Nodulin 26 like Intrinsic Proteins: Structurally Similar Membrane Channels with Diverse Functions in Plant Hypoxia Stress, Metalloid Nutrition & Toxicity

Zachary Beamer
University of Tennessee, Knoxville, zbeamer@vols.utk.edu

Follow this and additional works at: https://trace.tennessee.edu/utk_graddiss

Part of the Molecular Biology Commons, and the Plant Biology Commons

Recommended Citation
https://trace.tennessee.edu/utk_graddiss/7108

This Dissertation is brought to you for free and open access by the Graduate School at TRACE: Tennessee Research and Creative Exchange. It has been accepted for inclusion in Doctoral Dissertations by an authorized administrator of TRACE: Tennessee Research and Creative Exchange. For more information, please contact trace@utk.edu.
To the Graduate Council:

I am submitting herewith a dissertation written by Zachary Beamer entitled "Nodulin 26 like Intrinsc Proteins: Structurally Similar Membrane Channels with Diverse Functions in Plant Hypoxia Stress, Metalloid Nutrition & Toxicity." I have examined the final electronic copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Biochemistry and Cellular and Molecular Biology.

Dr. Daniel Roberts, Major Professor

We have read this dissertation and recommend its acceptance:

Dr. Brad Binder, Dr. Tessa Burch-Smith, Dr. Tarek Hewesi

Accepted for the Council:

Dixie L. Thompson

Vice Provost and Dean of the Graduate School

(Original signatures are on file with official student records.)
Nodulin 26 like Intrinsic Proteins: Structurally Similar Membrane Channels with Diverse Functions in Plant Hypoxia Stress, Metalloid Nutrition & Toxicity

A Dissertation Presented for the
Doctor of Philosophy
Degree
The University of Tennessee, Knoxville

Zachary Beamer
May 2022
Acknowledgements

I would first like to express my gratitude for my mentor, Dr. Dan Roberts. I was sitting in his BCMB401 class back in the fall of 2010 as an undergraduate when my interest in the subject of biochemistry was first realized. It seemed only fitting that I would later pursue a Ph.D. in biochemistry under the same man who started me down this path. Dr Roberts always allowed me the opportunity and freedom to pursue my research without stifling my ideas (no matter how half-baked they might have been). Although I struggled at times in my research, it was his hands-off mentorship style that has given me the confidence to be able to solve any problem that I may encounter in my future endeavors.

It would be remiss of me to fail to credit much of my Ph.D. success to the lab mates that I had over the years at UTK. Tian Li, Pratyush Routray, Ansul Lokdarshi all had instrumental roles in molding me as a scientist and a person. Tian demonstrated true passion in her research and was the embodiment of hard work. I'll never forget our first week working together, filled with great conversations and long hours spent doing oocyte assays. Pratyush has always been a reliable friend and served as a second mentor to me. He showed me that a bit of patience and thoughtful planning is a crucial component to being an effective researcher. Ansul was always a source of guidance when experiments were not working, or I needed to vent my problems. It was a laugh a minute when we were together, and I will cherish those memories. Additionally, I would also like to acknowledge all the talented and dedicated undergraduates that I had the pleasure to work with in the lab (Katey Gibson, Samantha McIntire, Clayton Nunn, Katie Ostrouchov, Maggie Spangler). I feel that our time together taught me as much about
what it takes to be a good mentor as I tried to teach them about research. I take great pride in knowing that all of you have moved on to successful careers and professional schools.

Finally, I would like to thank my parents for their unconditional support, without which, none of this would have been possible. Although I dreaded hearing the inevitable question of “So... when are you graduating?”, I knew that I count on them for anything I needed.
Abstract

Plant nodulin 26 intrinsic proteins are categorized into three groups (NIP I, II, and III) based on pore architecture. NIP II and III participate in metalloid nutrition, whilst the function of a third (NIP I) is less understood. Here we investigate the physiological function of one NIP I protein (*Arabidopsis thaliana NIP2;1*) as a lactic acid channel, and also explore the structural basis for metalloid and water permeability of NIP I and NIP II proteins in general. In addition, a strategy was developed for the purification and crystallization of soybean nodulin 26 as a step towards structure determination of a NIP I protein.

*NIP2;1* is specifically expressed in roots in response to low oxygen stress where it accumulates predominantly on the plasma membrane. Compared to wild type plants, *nip2;1* mutant plant roots over accumulate lactic acid, and show lower lactic acid efflux and acidification of rhizosphere during hypoxia. *nip2;1* plants survive hypoxia poorly compared to wild type plants, suggesting that lactic acid efflux by NIP2;1 prevents lactic acid toxicity. *Nip2;1* plants show altered expression of pyruvate and lactate metabolizing enzymes, implying regulation of fermentation by lactic acid accumulation in vivo.

The structural basis for Arabidopsis NIP I and NIP II permeability was investigated by biochemical assays, plant genetics, and molecular modeling. Compared to NIP II channels, boric acid is poorly permeated by some NIP I channels (NIP4;1) but not others (NIP1;1). Over expression analysis *in planta* shows that NIP I proteins have some ability to transport boric acid in plants, but not as well as NIP II proteins which are bona fide physiological boric acid channels. The primary distinction between NIP I and NIP II proteins is that the latter have lost their function as aquaporin water channels.
Modeling of NIP6;1 reveals a larger selectivity filter with space for five amino acids rather than the four seen in classical aquaporins and NIP I proteins. Potential models for how the NIP II pore accommodates boric acid, as well as hypotheses for gating of the NIP II pore to prevent water transport were generated by homology modeling and molecular dynamics of NIP6;1.
Table of Contents

CHAPTER I – Introduction .......................................................................................... 1
  Aquaporins .............................................................................................................. 2
  Conserved topology of the AQP superfamily and their basis of selectivity .......... 4
  Diversification of the AQP superfamily in terrestrial plants ............................. 9
  Plasma Membrane Intrinsic Proteins ................................................................. 12
  Tonoplast Intrinsic Proteins ............................................................................... 12
  Small Basic Intrinsic Proteins ............................................................................ 14
  Nodulin 26-like Intrinsic Proteins ..................................................................... 14
  Goals of thesis research ...................................................................................... 21

CHAPTER II - Materials and Methods ................................................................... 23
  Standard plant growth conditions ................................................................. 24
  Hypoxia survival phenotype analysis ........................................................... 24
  Photosynthetic efficiency measurement under hypoxia ................................. 25
  T-DNA insertion mutant nip2;1 and complementation lines ......................... 25
  Molecular cloning of NIPs for *Xenopus* and *nip5;1-1* complementation experiments ............................................................... 28
  Generation of Arabidopsis *nip5;1-1* complementation lines ......................... 30
  Limiting boron phenotype analysis and B uptake in plants ............................. 31
  RNA purification and quantitative PCR .......................................................... 32
  Histochemical and microscopy techniques ...................................................... 34
  Immunochemical techniques .......................................................................... 35
  Media acidification and L-lactate measurements ............................................ 36
  Expression and functional analyses in *Xenopus* oocytes .............................. 37
  Computational methods and protein modeling techniques ............................. 39
  Expression and purification of *Glycine max* Nodulin26 in *Pichia pastoris* .... 41
  Detergent screening for Nod26 crystallization .............................................. 44
  Nod26 crystallization ...................................................................................... 44

CHAPTER III - The Arabidopsis thaliana NIP2;1 Lactic Acid Channel promotes Plant Survival Under Low Oxygen Stress .............................................................. 46
  Background and Rationale: NIP2;1 shows a unique transport selectivity among the NIPs .............................................................. 47
NIP2;1 expression is induced early during hypoxia primarily in root tissue........ 48
NIP2;1 enhances plant survival under low oxygen conditions...............................51
NIP2;1 is expressed on the plasma membrane as well as on internal membranes
during hypoxia and reoxygenation recovery......................................................56
NIP2;1 participates in lactic acid efflux and media acidification during hypoxia..63
The loss of NIP2;1 function affects the expression of pyruvate and lactate metabolic
enzymes.................................................................................................................65
Summary and Conclusions Chapter III.................................................................67

CHAPTER IV - Structural Basis for water and metalloid specificity of NIP I and NIP II channels........................................................................................................69
Background and Rationale......................................................................................70
The NIP fold and selectivity filter deviates from water-selective aquaporins......71
NIP I and II and corresponding H2 mutants show distinct preference for metalloid
uptake in oocytes........................................................................................................80
Overexpression of NIP I proteins partially rescue the B deficiency phenotype of nip5;1
mutant seedlings.......................................................................................................85
The loss of aquaporin activity in NIP II pores: Models for pore gating in a water-tight
channel.....................................................................................................................91
Summary and Conclusions Chapter IV.................................................................97

CHAPTER V – Expression, purification and crystallography of with the NIP I channel archetype, soybean nodulin 26.................................................................100
Background and Impetus for structural studies.........................................................101
Optimization of nod26 expression, solubilization and purification for crystallization
trials.............................................................................................................................102
Initial crystallization of nod26 with various transport substrates.........................110
Attempts at crystallization optimization with other neopentyl glycol
detergents..................................................................................................................113
Summary and Conclusions Chapter V.................................................................118

CHAPTER VI – Discussion......................................................................................119
The NIP Structure and Function Conundrum: Same, same, but different.............120
NIPs as “Metalloidoporin” channels........................................................................120
NIP II protein selectivity for Boron........................................................................122
Water transport through NIP II proteins and role of the H2 position in the Ar/R
selectivity filter: potential gating and its importance..............................................124
NIP2;1-mediated lactic acid efflux promotes Arabidopsis survival during low oxygen stress

NIP2;1 expression during normoxia, hypoxia stress and recovery

NIP2;1 localization during hypoxia stress and recovery

Lactic acid and ethanolic fermentation pathways

Aquaporins as lactic acid channels in other plant and microbial systems

Resolving the NIP conundrum: The need for more structural data

List of References

Vita
List of Tables

Table II-1 Sequences of primers used for genotyping of nip2;1, NIP2;1 Promoter and CDS cloning and quantitative RT-PCR experiments ................................................................. 26

Table II-2 Sequences of primers used for cloning AQP and NIP CDS with C-terminal 3xGlycine linker, NIP ar/R mutagenesis and genotyping of nip5;1 complementation lines ........................................................................................................................................ 29

Table II-3 Sequences of primers used for cloning codon optimized nod26 with an N-terminal 6xHis tag and sequencing primers ........................................................................................................ 42

Table IV-1 Pairwise RMSD values of backbone C-alpha atoms in superposed structures for the Lsi1 open, Lsi closed and the AlphaFold NIP homology models ...................... 76

Table V-1 Sitting drop crystallization conditions for nod26 solubilized with the different detergents .................................................................................................................. 112
List of Figures

Figure I-1 Hourglass fold of MIP family member Aqp1........................................5
Figure I-2 Selectivity filter regions of AQP1 and GlpF........................................8
Figure I-3 The Arabidopsis MIP family’s phylogenetic tree.................................11
Figure I-4 Phylogenetic analysis and representative models of the selectivity filter
regions of Arabidopsis NIP pore subtypes............................................................19
Figure III-1 NIP2;1 expression in Arabidopsis seedlings in response to oxygen
deficit....................................................................................................................49
Figure III-2 Cellular localization of AtNIP2;1 promoter activity in the roots under
normoxic growth conditions..............................................................................50
Figure III-3 Characterization of nip2;1 T-DNA insertional mutant seedlings........52
Figure III-4 Effects of oxygen deprivation on the survival of nip2;1 T-DNA insertional
mutant seedlings.................................................................................................53
Figure III-5 Complementation of nip2;1 T-DNA mutant with a NIP2;1pro::NIP2;1-GFP
construct..............................................................................................................55
Figure III-6 NIP2;1-GFP expression in the roots of NIP2;1:GFP plants during hypoxia
and reoxygenation...............................................................................................57
Figure III-7 Subcellular localization of NIP2;1:GFP in the roots of hypoxia challenged
NIP2;1:GFP complementation seedlings.............................................................59
Figure III-8 Subcellular localization of NIP2;1-GFP in the roots of hypoxia challenged
seven-day-old NIP2;1:GFP complementation lines............................................60
Figure III-9 Comparison of surface and internal localization of NIP2;1-GFP during
hypoxia and recovery........................................................................................61
Figure III-10 Media acidification and lactic acid efflux in hypoxia-challenged
 nip2;1 ..........................................................................................64
Figure III-11 Effect of nip2;1 mutation on transcripts of pyruvate metabolism
enzymes..............................................................................................................66
Figure IV-1 Alignment and Alpha-fold modeling of Arabidopsis NIP I and II
proteins....................................................................................................................73
Figure IV-2 Representative models of the selectivity filter regions of NIP pore
subtypes................................................................................................................74
Figure IV-3 Effect of H2 tryptophan substitution in NIP I-like pore models........ 77
Figure IV-4 Boric Acid bound model of NIP6;1............................. 78
Figure IV-5 NIP-GFP expression in Xenopus oocytes and Arabidopsis thaliana transgenic plants................................................................. 81
Figure IV-6 Comparison of the metalloid hydroxide permeabilities of NIP I and II wild type proteins and H2 mutants...................................................... 82
Figure IV-7 The H2 mutant protein NIP5;1 A177W produces an inactive protein... 84
Figure IV-8 Comparison of Boric acid and arsenous acid permeabilities of four A. thaliana NIP proteins........................................................................ 86
Figure IV-9 Comparison of growth of wild type, nip5;1-1, and complementation lines under sufficient boric acid conditions........................................... 88
Figure IV-10 Complementation of nip5;1 growth and B uptake phenotypes by NIP1 wild type and H2 ar/R mutant constructs................................................. 89
Figure IV-11 Effect of NIP I protein overexpression on the B sensitive phenotype of nip5;1......................................................................................... 90
Figure IV-12 Comparison of the aquaporin activities of NIP I and II wild type proteins and H2 mutants................................................................. 93
Figure IV-13 Conserved gating loop residues in NIP structures...................... 94
Figure IV-14 NIP6;1 Selectivity Filter Arginine occupies two conformational states during MD simulations........................................................... 96
Figure V-1. Synthetic nod26 gene expression construct for Pichia pastoris........ 103
Figure V-2. Detergent solubilization and molecular weight aggregate screen for purified nod26.................................................................................. 106
Figure V-3 Purified nod26 conformation and particle size distribution............. 107
Figure V-4 Enhanced solubilization of nod26 by polyamines......................... 109
Figure V-5 Crystallization of nod26................................................................ 111
Figure V-6 Crystallization and diffraction of nod26-LMNG............................ 114
Figure V-7 Detergent screen of LMNG analogues........................................ 116
Figure V-8 Crystallization and diffraction of nod26-DMNG............................ 117
Figure VI-1 Comparison of hypoxia-induced expression of *NIP2;1* and lactate metabolizing enzyme transcripts *LDH* and *GOX3*…………………………………133

Figure VI-2 Phylogenetic analysis of Arabidopsis *NIP2;1*……………………………………139
CHAPTER I – Introduction
Water movement across biological membranes is facilitated by a superfamily of membrane protein channels called Aquaporins

Molecular water is a fundamental component of all living organisms. Within biological systems, water flux is critical for maintaining osmotic equilibrium, body temperature, turgor pressure generation, cell volume control, and osmotic stress adaptation (Vokes, 1987; Johansson et al., 2000). The plasma membrane is the primary barrier to bulk water flow between cells and their surroundings. Historically, scientists suggested that biological membranes are readily permeable to specific molecules such as water. However, early biophysical experiments provided evidence that certain biological membranes show elevated water permeable compared to others, implying that rapid transmembrane water movement could occur through a facilitated pathway rather than by simple diffusion. In support, it was observed that water movement into red blood cells occurs with a low Arrhenius activation energy typical of water diffusion in free solution (Paganelli and Solomon, 1957), suggesting a rapid pathway of water flow through an opening within the membrane.

Subsequently, research showed that mercurial compound such as HgCl₂ inhibit water permeability of red blood cell membranes in a manner that is reversed by mercaptan reducing agents (Macey and Farmer, 1970). Due to the fact that HgCl₂ has no effect on diffusive water movement through bare bilayers, it was concluded that red blood cells possess a proteinaceous “water channel” that aids in transmembrane water flow. This prompted a search for integral membrane channel proteins capable of
facilitating bulk water movement driven by osmotic and pressure gradients with minimal energy cost.

The laboratory of Peter Agre lab made the groundbreaking discovery of such a protein in the 1980s. They discovered a 28-kDa protein that co-purified with the 32-kDa subunit of human red blood cell Rh polypeptides (Agre et al., 1987; Saboori et al., 1988), and showed that it is an integral membrane channel protein component of red blood cell membranes (Smith and Agre, 1991). It was initially referred to as a channel-forming integral protein of 28-kDa or “CHIP28” (Preston and Agre, 1991). In his seminal analysis, Preston et al. (1992) demonstrated that Xenopus laevis oocytes injected with CHIP28 RNA demonstrated proteinaceous water transport activity with the hallmarks of a predicted protein water channel including a reduction in activation energy (Ea) compared to control oocytes, an 8-fold increase in water permeability (Pf), and sensitivity to mercurial compounds. The CHIP28 water channel was later renamed aquaporin-1 (AQP1) and is regarded as the first characterized integral membrane water channel. Peter Agre was awarded the Nobel Prize in Chemistry for his discovery in 2003.

Following the discovery of AQP1, additional members of the AQP family were discovered in all kingdoms of life including animals, yeast, bacteria, archaeabacteria, and plants (Maurel et al., 1993; Ecelbarger et al., 1995; Carbrey et al., 2001). The discovery that the bacterial glycerol facilitator GlpF has sequence similarities to other AQPs led to the discovery of a distinct action for these proteins in glycerol transport (Sweet et al., 1990). GlpF expressed in Xenopus laevis oocytes demonstrated rapid and selective
glycerol transport with similar features to AQP transport, including inhibition by mercurial compounds and low activation energy (Maurel et al., 1994).

From these early observations of AQP and GlpF, members of the aquaporin gene family were subsequently classified into two functional classes (Borgnia et al., 1999a). These include water specific aquaporin channel proteins, which are specialized for water transport, and multifunctional aquaglyceroporin channel proteins, which promote the passage of water in addition to glycerol and other uncharged solutes. For example, the thirteen human AQPs are classified as water-selective aquaporins subgroup includes (e.g., AQP0, AQP1, AQP2, AQP4, AQP5, AQP6, and AQP8) (King et al., 2004; Sorani et al., 2008), or aquaglyceroporins (e.g., AQP3, AQP7, AQP9, AQP10, AQP11, and AQP12A) (Rojek et al., 2008; Sorani et al., 2008). The discovery that abnormalities in aquaporins are also related with human illness and other aberrant water balance pathologies demonstrates the relevance of certain of these genes in transport physiology (Sorani et al., 2008). Aquaporin 2 (AQP2), for example, is implicated in vasopressin-dependent urine concentration (Deen et al., 1994), and AQP2 mutation is a significant cause of autosomal dominant nephrogenic diabetes insipidus (Mulders et al., 1998).

Conserved topology of the AQP superfamily and their basis of selectivity

As illustrated by the mammalian archetype, AQP 1 (Fig. 1-A), members of the aquaporin family have a similar oligomeric structure and membrane protein fold. AQP1 forms a homotetramer with each monomer consisting of six transmembrane helices (H1-6 from the N-terminus) that are connected by five loops (A-E), with the hydrophilic amino- and carboxyl termini exposed to the cytoplasmic side of the membrane (Fig. 1-
**Figure I-1 Hourglass fold of MIP family member Aqp1.**

A. Sideview of Aqp1 monomer (Sui et al. 2001, PDB 7CJS) with both the NPA Arginines shown. B. Arrangement of the Aqp1 tetramer as viewed from the extracellular side (left) and from the membrane plane (right).
Two of the loop sections (loop B and E) include highly conserved asparagine-proline-alanine (NPA) sequence motifs that are canonical features of the aquaporin superfamily. As noted below, these have significant hydrophobicity that were postulated to comprise part of the pore (Fig. I-1A). Another characteristic of the aquaporin fold is an obverse two-fold symmetry between the amino terminal half (helices 1, 2, and 3 with the first NPA motif) and carboxyl terminal half (helices 4, 5, and 6 with the other NPA motif) of the protein (Fig. I-1A). Jung et al. (1994) postulated that each AQP1 monomer creates a water channel pore and that the architecture resembles an hourglass with large cytosolic and extracellular vestibules that narrow in the center of the bilayer.

After the atomic structure of several AQPs was solved, the structural aspects of this protein superfamily became better understood. The first atomic structure of an aquaporin solved was that of rat AQP1 extracted from red blood cells, which provided a cryo electron crystallography structure of 3.8 resolution (Murata et al., 2000). Following this first electron crystallographic structure, many other AQP atomic structures have been solved by X-ray crystallography, including bovine AQP1 (Sui et al., 2001), E. coli GlpF (Fu et al., 2000; Tajkhorshid et al., 2002), lens specific AQP0 (Gonen et al., 2004; Harries et al., 2004), human AQP5 (Horsefield et al., 2008) and many others (Kreida and Törnroth-Horsefield, 2015).

These high-resolution AQP structures were astonishingly compatible with the hourglass model's early predictions. AQPs are homotetrameric with each monomer containing a single transport pore, according to the X-ray structures (Fig. I-1B). As predicted by the hourglass topology model, each monomer contains six transmembrane helices arranged in a right-handed tilted helical bundle (Fig. I-1A and B). Loops B and E
with the two canonical NPA motifs form two half helices (Loops B and E), which fold back into the pore to produce a seventh “pseudo” transmembrane helix. The result is a densely packed helical cluster that forms the transport pore.

Transport selectivity and differences between aquaporins and aquaglyceroporins is determined by the aromatic/arginine (ar/R) region, which produces the smallest pore constriction (Fu et al., 2000; Sui et al., 2001). The NPA motifs in loops B and E, converge at the center of the pore. The amide groups of the two Asn residues of the NPA motifs make hydrogen bonds with the water or glycerol molecules that traverse the channel. As a result of the water molecule dipole reorientation by this interaction, the hydrogen bonding potential between transported water molecules within the pore is limited. Disruption of a continuous hydrogen bond network between single file waters in the pore prevents potential proton flux via a “proton wire” by the Grotthuss mechanism (Fu et al., 2000; Sui et al., 2001; Cukierman, 2006). In water selective aquaporins, such as AQP1, the ar/R constriction is composed of four residues: phenylalanine 58 (Helix 2), histidine 182 (Helix 5), cysteine 191 (Loop E), and arginine 197 (Loop E) (Sui et al., 2001). These act as a size selectivity filter (2.8 Å in diameter; see Fig. I-2), enabling water to pass through while rejecting bigger solutes (Sui et al., 2001).

In addition, the side chains of H182 and R197, as well as the carbonyl backbone oxygen of C191, form energetically favorable hydrogen bond interactions with the transported water molecule at the ar/R. Furthermore, the nearby cysteine 189 appears to be the crucial location of mercury inhibition, providing a structural reason for why these compounds reversibly inhibit water transport (Preston et al., 1993; Zhang et al., 1993). Molecular dynamic (MD) simulation studies also indicate that there is a high
Figure I-2 Selectivity filter regions of AQP1 and GlpF. A. The selectivity filter residues of Aqp1 (Sui et al. 2001, 1J4N) and GlpF (Fu et al. 2000, 1LDA) are shown as viewed from the extracellular side. B. Comparison of selectivity filter residues by position for Aqp1 and GlpF (H2: helix 2, H5: helix 5, LE1: E loop residue, R: conserved arginine. C. Plot of the calculated pore radius along the Z-coordinate of both AQP1 (orange) and GlpF (green). The data were obtained by using MoleOnline 2.5 (Pravda et al., 2018) in the Channels mode. The position of the conserved NPA motifs that define the center of the pore as well as the ar/R selectivity filter region are indicated.
barrier for proton conduction at the NPA and ar/R constriction (de Groot et al., 2003), demonstrating that both regions are important determinants of water specific transport within aquaporins.

In the glycerol facilitator, the H182 and C191 residues of the AQP1 ar/R region are replaced by a glycine (G) and phenylalanine (F) respectively, and F58 is replaced with a conserved tryptophan residue (Fu et al., 2000). (Fig. I-2A and B). These modifications have a significant impact on the ar/R region of GlpF. First, they expand the width of the ar/R region of the pore to around 4Å, allowing bigger solutes such as glycerol to pass through (Fig. I-2C). Second, an aromatic surface produced by the glyceroporin specific tryptophan and phenylalanine residues increases the hydrophobicity of the selectivity filter area region to provide an interaction site for the carbon backbone of glycerol. Increased size and hydrophobicity allow for bulkier hydrophobic solutes such as glycerol to be accommodated while also resulting in decreased water permeability.

**Diversification of the AQP superfamily in terrestrial plants**

Since the initial discoveries of AQP1 and GlpF, thousands of AQP genes have been found in a variety of prokaryotic and eukaryotic organisms throughout all kingdoms of life, all of which have a conserved “hourglass” membrane topology and polypeptide structure (Johansson et al., 2000; Chaumont et al., 2005; Benga, 2009). While members of the aquaporin superfamily are widely distributed and found in virtually all living species, their diversity differs within each organism. Bacterial species such as *E. coli*, for example, reveals just two AQP genes: AqpZ and GlpF, an aquaporin and aquaglyceroporin, respectively (Heller et al., 1980; Borgnia et al., 1999a; Borgnia et al.,
1999b; Calamita, 2000). As previously stated, the human genome has genes encoding thirteen AQP genes, which can alternatively be classified as water-specific aquaporins or aquaglyceroporins. These bacterial and animal proteins can be segregated structurally into one of the two classical ar/R motifs, as illustrated in Fig. 1-2B. Terrestrial plants, on the other hand, have a substantially higher number of AQP genes with more diversified pore architectures and phylogeny (Wallace and Roberts, 2004; Bansal and Sankararamakrishnan, 2007). For example, 35 AQP family genes have been discovered in Arabidopsis (Johanson et al., 2001; Quigley et al., 2002), and subsequent analysis of many land plant genomes (reviewed in Laloux et al., 2018) reveals that all have over 30 AQP genes with some (e.g., poplar, (Gupta and Sankararamakrishnan, 2009)) as high as 60. Plant AQPs have been further divided into four phylogenetic subfamilies based on sequence homology and localization: plasma membrane intrinsic proteins (PIPs), tonoplast intrinsic proteins (TIPs), nodulin26-like intrinsic proteins (NIPs), and the small basic intrinsic proteins (SIPs) (Johansson et al., 2000; Chaumont et al., 2005; Maurel et al., 2008). Arabidopsis, for example, has 13 PIP genes, 10 TIP genes, 9 NIP genes, and 3 SIP genes among its 35 MIP genes (Fig. 1-3).

Plant AQPs are predicted to have a more diversified ar/R selectivity filter than animal and microbial AQPs, with at least eight different ar/R selective filter categories based on homology modeling (Wallace and Roberts, 2004). Furthermore, other model plant species are even more diversified in the ar/R groups, displaying an extra ninth group that is not present in Arabidopsis AQPs (Bansal and Sankararamakrishnan, 2007). Notably, most of these ar/R pore “signatures” differ from traditional aquaporin and
Figure I-3 The Arabidopsis MIP family's phylogenetic tree. The Arabidopsis genome's 35 full-length MIP subfamily members were aligned using the ClustalW method and constructed into a phylogenetic tree using the DNAstar software. To the right of the tree are the four plant MIP subgroups. Each subfamily's members are identified by their Arabidopsis Genome Initiative (AGI) number. The axis beneath the tree represents the number of amino acid substitutions. The subfamily designations are based on Johanson et al. (2001) nomenclature. Adapted from Choi, W.G. (2009) (https://trace.tennessee.edu/utk_graddiss/26).
aquaglyceroporin ar/Rs (Fig. I-2B) implying that plant AQP genes likely have activities other than water and glycerol transport.

**Plasma Membrane Intrinsic Proteins (PIPs)**

PIPs are predominantly located in the plasma membrane, where they aid in passive water transport (they are the plant AQP family most similar to classical animal aquaporins with respect to the ar/R signature). The first PIP subfamily member was found in Arabidopsis, and biochemical analysis showed that it was a water-selective aquaporin (Daniels et al., 1994). In Arabidopsis, the PIP subfamily consists of 13 members, which are separated phylogenetically into two groups: PIP1 and PIP2 (Fig. I-3) While both the PIP1 and PIP2 subgroups share a high degree of sequence identity, they have unique functional features as aquaporins (Chaumont et al., 2000). The PIP2 subgroup has been found to have lower water permeability when compared to the PIP1 subgroup's robust water permeability (Chaumont et al. 2000). It has also been observed that in order to obtain aquaporin functionality, PIP1 proteins must form heterotetramers with PIP2 monomers (Fetter et al., 2004; Zelazny et al., 2007). PIP3, a third PIP subgroup, has been discovered in the genome of the model moss *Physcomitrella patens* (Danielson and Johanson, 2008). PIPs are involved in a variety of plant activities that require bulk water transport, including response to drought, salinity, and flooding stress (Afzal et al., 2016).

**Tonoplast Intrinsic Proteins (TIPs)**

TIPs were the first AQP proteins that were discovered in plants, with the first characterized in *Phaseolus vulgaris* seeds as a 27kDa intrinsic tonoplast membrane protein (Johnson et al., 1989). Hydropathy plots and secondary structure studies
predicted that TIPs have the canonical AQP protein topology with six transmembrane domains and possessed sequence similarity to GlpF (Johnson et al., 1990). In Arabidopsis, the TIP subfamily has 11 members, all of which are present on the vacuole tonoplast membranes (Quigley et al., 2002). Analysis in *Xenopus* oocytes show that TIPs are robust water channels (Maurel et al., 1993; Maurel et al., 1995; Daniels et al., 1996; Gerbeau et al., 1999), and their presence in the tonoplast membrane is expected to allow rapid water exchange between the cytoplasm and the vacuole, an important process during cell expansion that is driven by turgor pressure.

While TIP subfamily proteins were initially identified as aquaporins (Maurel et al., 1993), however further research revealed that TIP proteins also transport a variety of other uncharged substrates in addition to water (Gerbeau et al., 1999; Klebl et al., 2003). Yeast complementation experiments have revealed that numerous Arabidopsis TIPs transport ammonium, methylammonium, urea or formamide permeation (Liu et al., 2003; Jahn et al., 2004a; Loqué et al., 2005; Kapilan et al., 2018). Additionally, TIPs expression in *Xenopus* oocytes show enhanced \( \text{NH}_4^+ / \text{NH}_3 \) permeability in a pH dependent manner (Jahn et al., 2004b; Holm et al., 2005; Loqué et al., 2005). Therefore, in addition to their function as water channels, TIPs may also aid in the movement of these types of molecules between the vacuolar and the cytoplasmic compartments. Collectively, TIP subfamily proteins are expected to be important not just in aquaporin-based osmoregulation activities, but also in the transport of nitrogenous substrates such as urea and ammonia (Liu et al., 2003; Jahn et al., 2004a; Loqué et al., 2005).
Small Basic Intrinsic Proteins (SIPs)

The small basic intrinsic protein subfamily has the least number of genes and diversity within the plant aquaporin gene family. The SIP subfamily was first discovered through bioinformatic database mining and phylogenetic analysis of plant genomes (Johanson and Gustavsson, 2002). The fundamental reason for their small protein size is because they have a truncated cytosolic N-terminal region in comparison to other plant aquaporins. Arabidopsis SIP proteins tagged with green fluorescent protein (GFP) showed preferential localization to the ER with minimal distribution to the plasma and vacuolar membranes, suggesting a localization and function distinct from TIP and PIP proteins (Ishikawa et al., 2005). Phylogenetic analysis of SIP subfamilies from various plant lineages (Chaumont et al. 2000, Johanson et al. 2001, Quigley et al. 2002) shows that these proteins are separated into two conserved subgroups (SIP1 and SIP2), implying that each of these subgroups may have a unique physiological role. Arabidopsis SIP1;1 and SIP1;2 have aquaporin activity based on permeability assays in yeast and proteoliposomes while SIP2;1 demonstrated only mild water permeability in reconstituted membrane vesicles (Ishikawa et al. 2005).

Nodulin 26-like Intrinsic Proteins (NIPs)

Nodulin 26-like intrinsic proteins represent a terrestrial plant-specific subfamily of the larger plant aquaporin superfamily that are named for the archetype of the family, soybean nodulin 26 (nod26). Nodulins are genes of leguminous plants that are specifically expressed during the formation of nitrogen-fixing symbioses with rhizobia soil bacteria. Nod26 was first identified in soybean by Fortin et al. (1987) as a component of the peribacteroid/symbiosome membrane that houses the
Bradyrhizobium japonicus bacteroid in nitrogen-fixing root nodules (Fortin et al., 1987). Sandal and Marcker (1988) noted the similarity to mammalian lens AQP0 and bacterial GlpF, and thus the protein was one of the first members of the aquaporin superfamily to be identified in plants (Sandal and Marcker, 1988). Weaver et al (1991) confirmed specific localization of the protein to the soybean symbiosome as the major protein component (15% of the protein mass).

Symbiosomes are specialized nitrogen-fixing symbiotic organelles created by legumes in coordination with endosymbiotic nitrogen-fixing Rhizobia bacteria (Oldroyd and Downie, 2008). The plant-derived symbiosome membrane contains a variety of specific channel and transport proteins, including several nodulins, that help establish and maintain the symbiosis (Udvardi and Poole, 2013). Transport analysis of nod26 in symbiosome and proteoliposome vesicles (Rivers et al., 1997)(Dean et al., 1999; Hwang et al., 2010; Niemietz and Tyerman), as well as in Xenopus oocytes (Rivers et al., 1997; Dean et al., 1999; Guenther et al., 2003) show that nod26 is a multifunctional aquaglyceroporin that transports water, glycerol and uncharged ammonia gas, accounting for these activities on the native symbiosome (Rivers et al., 1997; Dean et al., 1999). The symbiosome is the predominant organelle comprising much of the space within the mature soybean nodule core populated by the specialized nitrogen-fixing infected cells, and likely engages in cell volume management and osmotic buffering in response to changing metabolic and environmental conditions within the nodule (; Guenther et al., 2003; Roberts and Routray, 2017).

Prior to its functional characterization, nod26 was initially described as a phosphorylation substrate for a symbiosome membrane- calcium-dependent protein
kinases of the CDPK/CPK family (Weaver et al., 1991), making it one of the first endogenous CDPK targets identified in plants. CDPK phosphorylates nod26 specifically on serine-262 found in the cytosolic carboxyl terminal domain (Weaver and Roberts, 1992). The consensus phosphorylation site found in nod26 is generally conserved across proteins of the NIP-I pore family (Wallace et al. 2006) suggesting that this is a common site for regulation among this subfamily. The carboxyl terminus is a common site of phosphorylation for a variety of AQP proteins from both plants and animals (Hachez and Chaumont, 2010; Kreida and Törnroth-Horsefield, 2015; Maurel et al., 2015), and regulates function in a variety of ways, including channel gating (Johansson et al., 1998; Tornroth-Horsefield et al., 2006; Nyblom et al., 2009), trafficking, and tailored localization to certain membranes (van Balkom et al., 2002; Boursiac et al., 2008). Nod26 phosphorylation in nodules stimulates intrinsic water permeability activity in response to drought and salinity stress (Guenther et al. 2003), suggesting a possible role in increasing symbiosome water permeability in response to osmotic stress, perhaps assisting in infected cell volume regulation or transcellular water movement in the nodule.

In addition to transporting water and glycerol, nod26 also mediates NH$_3$ transport, which is the primary form of produced and effluxed from the symbiosome during symbiotic nitrogen fixation (Niemietz and Tyerman, 2000; Hwang et al., 2010). The primary symbiotic metabolic exchange associated with the nitrogen-fixing symbiosome membrane is the uptake of malate as a carbon source for rhizobia bacteroids to support nitrogen fixation, and the efflux of fixed nitrogen (NH$_3$/NH$_4^+$) to the plant cytosol for assimilation (Udvardi and Poole 2013). Analysis of symbiosome
membrane vesicles (Niemietz and Tyerman, 2000) and purified nod26 proteoliposomes (Hwang et al., 2010) show that nod26 accounts for the NH₃ efflux activity of the symbiosome membrane. Further, ammonia is the preferred substrate for the nod26 channel (P_{ammonia} is fivefold greater than P_f) and represents a potential conduit for NH₃ movement from the symbiosome (Hwang et al. 2010).

Strong support for this hypothetical function of nod26 came from the discovery that nod26 forms a complex with soybean nodule cytosolic glutamine synthetase (Masalkar et al., 2010). Cytosolic ATP-dependent glutamine synthetase (GS), which is a key component of the infected cell cytoplasm, catalyzes a critical step in nitrogen assimilation once ammonia effluxes from the symbiosome. GS binds to the symbiosome membrane (Masalkar et al., 2010; Clarke et al., 2014) via interacting with the carboxyl terminal domain of nod26 (Masalkar et al. 2010). The association of the major ammonia channel with the major ammonia assimilation enzyme is predicted to efficiently produce glutamine while preventing the accumulation of toxic levels of ammonia in the cytosolic compartment (Masalkar et al., 2010; Routray et al., 2015). Interestingly, the carboxyl terminal domains of other aquaporins (e.g., AQP0 and AQP2 in mammals) is a common site of binding of soluble regulatory proteins (Rose et al., 2006; Wang and Schey, 2011; Reichow et al., 2013; van Balkom et al., 2009)].

Since the initial discovery of nodulin 26, and the sequencing of multiple plant genomes, it has become obvious that “Nodulin 26-like” Intrinsic Proteins (NIPs) are widely distributed in all terrestrial plants, from mosses to angiosperms (Wallace et al., 2006; Roberts and Routray, 2017). The wealth of information regarding NIP sequences across many plant lineages has led to the recognition of structural and functional
diversity within the NIP family, as well as the emergence of even more transport functions associated with these proteins (reviewed in Roberts and Routray, 2017; Pommerrenig et al., 2015). Based on phylogeny and homology modeling of the nine NIP genes in Arabidopsis, Wallace and Roberts proposed classification of NIPs into two pore families based on the ar/R selectivity filter (Wallace and Roberts 2004). NIP subgroup I (NIP I) is made up of six genes that encode NIP1;1, NIP1;2, NIP2;1, NIP3;1, NIP4;1 and NIP4;2, all of which have the same ar/R selectivity filter as the NIP archetype, soybean nod26 (Fig. I-4). This filter consists of Trp at the H2 position, Ile or Val at the H5 position, and Ala and Arg at the LE1 and LE2 positions, respectively, and is highly similar to amphipathic ar/R of glyceroporins (Wallace and Roberts, 2005). NIP subgroup II (NIP II) in Arabidopsis has three genes: NIP5;1, NIP6;1, and NIP7;1. The fundamental difference between NIP I and NIP II ar/R selectivity filters is that NIP II proteins have an Ala residue in lieu of the NIP I-like Trp at the H2 location (Fig. I-4B). This leads to a larger pore aperture (Wallace and Roberts, 2005), resulting in NIP II proteins having a different substrate selectivity than NIP I (Wallace and Roberts 2004; Wallace and Roberts 2005). Investigation of the NIP subfamily in rice resulted in the discovery of a third NIP pore type that is widely distributed in the grasses but is absent in Arabidopsis (Bansal and Sankararamakrishnan, 2007). Rice has a a small subset of NIPs with a unique Ar/R structure, which has been designated as NIP subgroup III. The NIP III protein ar/R selectivity filter is made up of Gly(H2), Ser(H5), Gly(LE1), and Arg(LE2), resulting in a broader and more hydrophilic pore aperture compared to NIP I and NIP II (Mitani et al., 2008).
Figure I-4 Phylogenetic analysis and representative models of the selectivity filter regions of Arabidopsis NIP pore subtypes. A. Phylogenetic tree of the seven Arabidopsis thaliana NIP proteins as well as their pore subtype classification. NIPs were classified into two subgroups based on the the Ar/R selectivity filter formed by four amino acid residues (helix 2 H2, helix5 H5, loop E1 LE1, and loop E2 LE2). B. The ar/R selectivity filters of the NIP I (soybean nodulin 26) and NIP II (Arabidopsis NIP6;1) based on the homology models of Roberts and Routray (2017) are shown. The diagram below each structure indicates the consensus sequence for the NIP I and II pore subtypes based on phylogenetic analyses across plant species.
Due to the distinct ar/R regions, each of the three NIP subgroups has a different substrate selectivity. NIP I proteins have an ar/R amino acid composition comparable to nod26 (Wallace and Roberts, 2004), and they exhibit inherent substrate transport selectivity for water, glycerol, and ammonia (Roberts and Routray, 2017), as well as the capacity to be penetrated by toxic metalloids (e.g., arsenite and antimonite) (Kamiya et al., 2009; Kamiya and Fujiwara, 2009; Xu et al., 2015; Diehn et al., 2019). Other substrates, such as uncharged lactic acid is transported by Arabidopsis NIP 2;1, a root-specific NIP I protein that is induced by flooding and hypoxia (Choi and Roberts, 2007).

NIP II protein pore predictions show a greater pore aperture and a comparison of an NIP I protein (Nodulin 26) with an NIP II protein (NIP6;1) revealed that NIP II proteins are penetrated by bulkier solutes, such as urea, whereas NIP I proteins are not (Wallace and Roberts, 2005). NIP II proteins are also “water-tight,” with little to no aquaporin activity while still being able to flux glycerol and urea (Wallace and Roberts, 2005; Li et al., 2011; Routray et al., 2018). In terms of biological substrates, overwhelming genetic, cell biological, and biochemical evidence show that the major physiological function of these proteins is to serve as boric acid channels that cooperate with BOR transporters to mediate the uptake and distribution of boric acid as a nutrient under limiting conditions (Takano et al., 2006; Tanaka et al., 2008; reviewed by Takano et al., 2008; Miwa et al., 2010; Yoshinari and Takano, 2017).

Pore predictions for NIP III proteins suggest a larger and more hydrophilic pore than NIP I and II (Ma et al., 2006), and these proteins have been demonstrated to be penetrated by bulkier metalloid solutes such as germanic acid (GeOH4) and silicic acid (SiOH4) (Ma et al. 2006). Genetic evidence supports a role for NIP III proteins as silicic
acid channels that promote optimal growth and development, as well as resistance to abiotic and biotic stress in rice and other grasses (Ma et al. 2006; Chiba et al. 2009; Epstein 1999, Ma and Yamaji 2006). Site-directed mutagenesis investigations of the ar/R region in OsNIP2;1 (NIP III) and AtNIP5;1 (NIP II) offer information on the mechanism of substrate selectivity regulation (Mitani-Ueno et al.). The results show that the residue at the H5 position of the ar/R filter of both OsNIP2;1 and AtNIP5;1 plays an important role in the permeability of silicic acid and boric acid (Mitani-Ueno et al., 2011).

**Goals of thesis research**

Members of the aquaporin superfamily are present in all kingdoms of life, implying their importance in transporting biologically relevant solutes across membranes. As previously stated, many MIPs are capable of transporting uncharged solutes in addition to water. Plants have been discovered to have a greater number of MIP superfamily members that are more functionally varied than their animal and microbial counterparts. Plant MIPs have a more diversified combination of Ar/R amino acids, indicating unique functionality within the aquaporin superfamily. Within the MIP superfamily is a plant specific aquaporin subfamily, referred to as the NIPs. In Arabidopsis, the NIP proteins are divided into two groups: NIP I and NIP II, which have different Ar/R tetrad amino acids which are thought to be responsible for the variable substrate selectivities. We employed homology modeling, molecular dynamics simulations, site-directed mutagenesis, and functional analysis to examine the substrate selectivity of Arabidopsis NIP I and NIP II proteins. Additionally, we examined NIP2;1 biological function, cellular localization, and confirmed its function as a lactic acid efflux
protein in the hypoxia stress response in Arabidopsis. Finally, we instigated structural studies of the NIP I archetypal protein, soybean nodulin 26, using x-ray crystallography.
CHAPTER II - Materials and Methods
Standard plant growth conditions

*Arabidopsis thaliana* ecotype Columbia-0 was used in all experiments. Seeds were sterilized and stratified at 4 °C for 2 days, and were germinated as in (Choi and Roberts, 2007). Seedlings were grown vertically on in Murashige-Skoog (MS) media supplemented with 1% (w/v) sucrose and 0.8% (w/v) Phyto-agar (plantMedia) with a long day cycle of 16 hours of light (100 μmol m⁻² s⁻¹) and 8 hours of dark (LD conditions).

Hypoxia survival phenotype analysis

Hypoxia treatment was administered at the end of the light cycle by the argon-treatment protocol described by (Lokdarshi et al., 2016). For normoxic controls, seedlings were treated simultaneously under identical conditions except in the presence of air instead of argon gas. For reoxygenation, the seedlings were returned to normoxic LD conditions at the end of the hypoxia time course. Phenotype analysis for hypoxia survival was conducted by established stress/recovery protocols for Arabidopsis (Licausi et al., 2010; Sorenson and Bailey-Serres, 2014; Lokdarshi et al., 2016) with modifications. Seven days old seedlings grown vertically as described above were administered 8 hours of argon gas-induced hypoxia (treatment) or air (normoxia control) in darkness, and were returned to normoxic LD growth conditions. The survival frequency (the absence of chlorosis and shoot meristem death as described by Sorensen and Bailey-Serres, 2014) was scored after three days after return to normoxic conditions.
Photosynthetic efficiency measurement under hypoxia

The maximum quantum yield of photosystem II \( QY_{\text{max}} = F_v / F_m \) was measured with a FluorCam 800MF instrument (Photon Systems Instruments) by the general method of (Murchie and Lawson, 2013). Seven-day old seedlings were administered 8 hr hypoxia, and \( QY_{\text{max}} \) was measured at different recovery time points upon return to LD conditions. For the first time point (time=0), seedlings were removed from hypoxia and were subjected to a saturating pulse of 1800 µEin m\(^{-2}\)s\(^{-1}\) for 0.8 sec (\( F_m \)). Variable fluorescence \( (F_v) \) was calculated as the difference between \( F_o \) and \( F_m \) to calculate the maximum quantum yield \( [F_v/F_m] \). For subsequent measurements, seedlings were dark adapted for 2 min \( (F_0) \) prior to application of the saturating pulse and conducting measurements.

T-DNA insertion mutant \textit{nip2;1} and complementation lines

A sequence tagged T-DNA insertion line within the \textit{NIP2;1} gene (WiscDsLox233237_22k) from the WiscDs-Lox T-DNA collection (Woody et al., 2007) was obtained from Arabidopsis Biological Resource Center at the Ohio State University. The \textit{nip2;1} mutant was selected on MS media supplemented with 15 µg/mL Basta and was genotyped by a PCR-based genotyping protocol as described at (http://signal.salk.edu/tdnaprimers.2.html). For this purpose, genomic DNA was extracted from 2-week old seedlings using the \textit{Wizard} Genomic DNA purification kit (Promega) and was subjected to PCR analysis using two \textit{NIP2;1} gene specific primers and the left border T-DNA primer (Table II-1). The precise site of T-DNA insertion was verified by a cloning of the PCR product into the pCR2.1-TOPO vector (Invitrogen) followed by automated DNA sequencing. All DNA sequencing conducted in this study
Table II-1 Sequences of primers used for genotyping of *nip2.1*, **NIP2.1** Promoter & CDS cloning and quantitative RT-PCR experiments.

<table>
<thead>
<tr>
<th>NAME</th>
<th>SEQUENCE (5'-3')</th>
<th>Description</th>
<th>PURPOSE</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>A</strong></td>
<td><strong>AACCCGATCCATGGTTATTTAGGTTGC</strong></td>
<td>TSSA, Mf border primer</td>
<td><strong>Nip2.1</strong> genotyping primers</td>
</tr>
<tr>
<td><strong>A</strong></td>
<td><strong>GCGGGAGGATCCATAGGTTGC</strong></td>
<td><strong>NIP2.1</strong> gene specific primer</td>
<td></td>
</tr>
<tr>
<td><strong>NIP2.1</strong> <em>PromF3</em></td>
<td><strong>GAAGCGGATCCATGGTTATTTAGGTTGC</strong></td>
<td><strong>NIP2.1</strong> gene promoter specific forward primer</td>
<td>Cloning <strong>NIP2.1</strong> Promoter</td>
</tr>
<tr>
<td><strong>NIP2.1 PromR2_</strong></td>
<td><strong>GAAGCGGATCC</strong></td>
<td><strong>NIP2.1</strong> gene promoter specific reverse primer</td>
<td></td>
</tr>
<tr>
<td><strong>NIP2.1_XbaF1_</strong></td>
<td><strong>GAGCTCTAGA</strong></td>
<td><strong>NIP2.1</strong> coding region forward primer</td>
<td>Cloning <strong>NIP2.1</strong> CDS and qPCR verification of transgenic lines</td>
</tr>
<tr>
<td><strong>NIP2.1 XbaR2_</strong></td>
<td><strong>GAAGCGGATCC</strong></td>
<td><strong>NIP2.1</strong> coding region reverse primer</td>
<td></td>
</tr>
<tr>
<td><strong>NIP2.1 F_</strong></td>
<td><strong>GATATCCGGATCC</strong></td>
<td><strong>NIP2.1</strong> coding region reverse primer</td>
<td>Cloning <strong>NIP2.1</strong> CDS</td>
</tr>
<tr>
<td><strong>NIP2.1 R_</strong></td>
<td><strong>GAAGCGGATCC</strong></td>
<td>PCR primers to confirm <strong>NIP2.1</strong> construct</td>
<td></td>
</tr>
<tr>
<td><strong>NIP2.1 qPCR_F1_</strong></td>
<td><strong>GCTTACACTAATCCAGACACGAC</strong></td>
<td>Forward primer for <strong>NIP2.1</strong> qPCR</td>
<td><strong>qPCR</strong> target genes</td>
</tr>
<tr>
<td><strong>NIP2.1 qPCR_R1_</strong></td>
<td><strong>TGACCTGCTCTATTCATGCTG</strong></td>
<td>Reverse primer for <strong>NIP2.1</strong> qPCR</td>
<td></td>
</tr>
<tr>
<td><strong>LDH1 qPCR_F1_</strong></td>
<td><strong>TGACCTGCTCTATTCATGCTG</strong></td>
<td>Forward primer for <strong>LDH1</strong> qPCR</td>
<td></td>
</tr>
<tr>
<td><strong>LDH1 qPCR_R1_</strong></td>
<td><strong>TCCACGGTTTATAGCTGCTG</strong></td>
<td>Reverse primer for <strong>LDH1</strong> qPCR</td>
<td></td>
</tr>
<tr>
<td><strong>GOX3 qPCR_F1_</strong></td>
<td><strong>AAGAAGGCAGAAGACGACG</strong></td>
<td>Forward primer for <strong>GOX3</strong> qPCR</td>
<td></td>
</tr>
<tr>
<td><strong>GOX3 qPCR_R1_</strong></td>
<td><strong>TACGTTACAGACGACG</strong></td>
<td>Reverse primer for <strong>GOX3</strong> qPCR</td>
<td></td>
</tr>
<tr>
<td><strong>AMP00875 qPCR_F1_</strong></td>
<td><strong>CGATTATGGACATTACCGG</strong></td>
<td>Forward primer for <strong>POD1</strong> qPCR</td>
<td></td>
</tr>
<tr>
<td><strong>AMP00875 qPCR_R1_</strong></td>
<td><strong>TGCGAGGCGACGACG</strong></td>
<td>Reverse primer for <strong>POD1</strong> qPCR</td>
<td></td>
</tr>
<tr>
<td><strong>AHR1 qPCR_F1_</strong></td>
<td><strong>ATGCGGCGACG</strong></td>
<td>Forward primer for <strong>AHR1</strong> qPCR</td>
<td></td>
</tr>
<tr>
<td><strong>AHR1 qPCR_R1_</strong></td>
<td><strong>GGAGAGACAGACG</strong></td>
<td>Reverse primer for <strong>AHR1</strong> qPCR</td>
<td></td>
</tr>
<tr>
<td><strong>AT1g17980 qPCR_F1_</strong></td>
<td><strong>TTGACCTGCTCTATTCATGCTG</strong></td>
<td>Reverse primer for <strong>AT1g17980</strong> qPCR</td>
<td></td>
</tr>
<tr>
<td><strong>AT1g17980 qPCR_R1_</strong></td>
<td><strong>CGATGACGACGACG</strong></td>
<td>Reverse primer for <strong>AT1g17980</strong> qPCR</td>
<td></td>
</tr>
<tr>
<td><strong>UBQ10.1 qPCR_F1_</strong></td>
<td><strong>CAAGGCGGATCCCTTGGCGG</strong></td>
<td>Forward primer for <strong>UBQ10</strong> qPCR</td>
<td><strong>qPCR</strong> reference genes</td>
</tr>
<tr>
<td><strong>UBQ10.1 qPCR_R1_</strong></td>
<td><strong>TGACCTGCTCTATTCATGCTG</strong></td>
<td>Reverse primer for <strong>UBQ10</strong> qPCR</td>
<td></td>
</tr>
<tr>
<td><strong>F1 ACTIN 2 qPCR_F1_</strong></td>
<td><strong>CGAGCGGATCCCTTGGCGG</strong></td>
<td>Forward primer for <strong>ACTIN2</strong> qPCR</td>
<td></td>
</tr>
<tr>
<td><strong>F1 ACTIN 2 qPCR_R1_</strong></td>
<td><strong>TGACCTGCTCTATTCATGCTG</strong></td>
<td>Reverse primer for <strong>ACTIN2</strong> qPCR</td>
<td></td>
</tr>
</tbody>
</table>

Notes: For cloning primers, the lower case letters indicate gene specific sequences, while upper case letters indicate engineered restriction sites or other added sequences.
(including all DNA cloning constructs described below) was done by the Sanger sequencing method performed with a Perkin Elmer Applied Biosystems 373 DNA sequencer at the University of Tennessee Molecular Biology Resource Facility (Knoxville, TN). T4 homozygous mutant seedlings were used for all phenotyping and other analyses in this study.

For complementation of nip2;1, as well as to localize NIP2;1 protein by fluorescence microscopy, transgenic lines containing a construct consisting of a NIP2;1-GFP translational fusion under the control of the NIP2;1 promoter were generated in the nip2;1 background. The promoter region of the NIP2;1 gene (from the translational start site to a site 2000 bp upstream) was amplified by PCR from Arabidopsis genomic DNA with gene specific primers with added KpnI and AatII restriction sites (Supplemental Table II-1) to facilitate its insertion into the binary vector pKGW_RedRoot_OCSA replacing the ubiquitin promoter and dsRed CDS (Niyikiza et al., 2020). The modified destination vector was named pKGW_OCSA_NIP2;1Pro. The NIP2;1 coding sequence was amplified with gene-specific primers with XbaI and EcoRI sites (Table II-1) using a template of cDNA prepared from total RNA from 4 hr hypoxic Arabidopsis seedlings. The resulting product was cloned into the Gateway entry vector CD3-1822 (Wang et al., 2013) to generate a construct encoding NIP2;1 as an in-frame carboxyl-terminal fusion with GFP separated by a 3X Gly linker. The NIP2;1-GFP construct was then recombined into the pKGW_OCSA_NIP2;1Pro vector by using a gateway LR reaction (Invitrogen) to generate the final binary vector with NIP2;1.pro::NIP2;1-GFP. The constructs were sequenced and verified by using Snap Gene 4.2.11 software. Agrobacterium tumefaciens GV3101 (Koncz and Schell, 1986) was transformed with the
final construct by electroporation with a Bio-Rad Gene Pulser Xcell Electroporation system. Colonies carrying the correct construct were verified by PCR, and were used to transform Arabidopsis *nip2;1* plants by using floral dip method (Clough and Bent, 1998). Transgenic lines were selected on MS media supplemented with 25 μg/ml kanamycin, and were confirmed by PCR-based genotyping with transgene specific primers (Table II-1). Twelve transgenic lines were generated and the two with the highest *NIP2;1* transcript abundance with expression levels similar to WT under hypoxia stress (lines K and I) were selected for further analysis. T2 generation homozygous complementation lines were used for further studies.

**Molecular Cloning of NIPs for Xenopus and nip5;1-1 complementation experiments**

cDNA constructs with NIP coding sequences (CDS) were generated from total RNA isolated from total seedling (*NIP1;1, NIP5;1* and *NIP6;1*) or flowers (*NIP4;1*) as previously described (Wallace and Roberts, 2005; Tanaka et al., 2008; Li et al., 2011a). The CDS encoding the aquaporin control protein, *Acyrthosiphon pisum* AQP2, was generated as described in Wallace et al. (2012). GFP fusions of the NIP and AQP constructs were generated by the general approach described in Beamer et al. (2021). Briefly, the NIP or AQP CDS were amplified by using gene specific primers (Table II-2) and were cloned into the BioVector Gateway entry vector Fu28 (ABRC stock no: CD3-1822; Wang et al., 2013), which contains an eGFP CDS at the 3’ end of a multiple cloning site. The primers were designed with a linker of three glycines between the NIP and eGFP coding sequences. A gateway LR reaction (Invitrogen) was performed to transfer the CDS-eGFP construct from the entry clone to the Gateway destination.
Table II-2 Sequences of primers used for cloning AQP and NIP CDS with C-terminal 3xGlycine linker, NIP ar/R mutagenesis and genotyping of nip5.1 complementation lines.

<table>
<thead>
<tr>
<th>NAME</th>
<th>SEQUENCE (5'→3')</th>
<th>Description</th>
<th>PURPOSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>AQP5.1_BspE_F</td>
<td>GAGGCCAGATCTGAGAAGACGTTG</td>
<td>ASPE coding region forward primer</td>
<td>Cloning of AQP and NIP CDS sequences</td>
</tr>
<tr>
<td>AQP5.1_BspE_R</td>
<td>GAGGCCAGATCTGAGAAGACGTTG</td>
<td>ASPE coding region reverse primer</td>
<td>with C-terminal 3xGlycine linker. Reverse</td>
</tr>
<tr>
<td>NIP5.1_EcoRI_F</td>
<td>CAAGCTGATCGTGATCGAGAAGACGTTG</td>
<td>NIP5.1 coding region forward primer</td>
<td>primers used for PCR confirmation of</td>
</tr>
<tr>
<td>NIP5.1_EcoRI_R</td>
<td>CAAGCTGATCGTGATCGAGAAGACGTTG</td>
<td>NIP5.1 coding region reverse primer</td>
<td>transgenic lines</td>
</tr>
<tr>
<td>NIP5.1_SalI_F</td>
<td>CAAGCTGATCGTGATCGAGAAGACGTTG</td>
<td>NIP5.1 coding region forward primer</td>
<td></td>
</tr>
<tr>
<td>NIP5.1_SalI_R</td>
<td>CAAGCTGATCGTGATCGAGAAGACGTTG</td>
<td>NIP5.1 coding region reverse primer</td>
<td></td>
</tr>
<tr>
<td>NIP5.1_AvaI_F</td>
<td>CAAGCTGATCGTGATCGAGAAGACGTTG</td>
<td>NIP5.1 coding region forward primer</td>
<td></td>
</tr>
<tr>
<td>NIP5.1_AvaI_R</td>
<td>CAAGCTGATCGTGATCGAGAAGACGTTG</td>
<td>NIP5.1 coding region reverse primer</td>
<td></td>
</tr>
</tbody>
</table>

Notes: For cloning primers, the lower case letters indicate gene specific sequences, while upper case letters indicate engineered restriction sites or other added sequences.
vectors for *Xenopus laevis* expression (pT7TSGW) or for plant transformation for complementation experiments. For *Xenopus* experiments, the destination vector pT7TSGW was generated by incorporating the Gateway cassette system from pB7WG (from VIB-UGent Center for Plant Systems Biology, Ghent University, Belgium) into the *Xenopus laevis* expression plasmid pT7TS (Vincill et al., 2005) at the EcoRV site. Binary plant transformation vectors with the NIP-eGFP under the control of the CaM35S promoter were generated by using the Gateway destination vector pB7WG2 (Karimi et al., 2002).

Site-directed mutagenesis to generate alanine or tryptophan substitutions at the codon encoding the H2 residues in the NIP constructs was done by using the mutagenesis primers listed in Table II-2 and the Q5 Site-Directed Mutagenesis Kit (NEB) as described previously (Li et al., 2011). Flag-tagged *Xenopus* expression constructs were generated in the pXβG-ev-1 vector (Wallace and Roberts, 2005).

**Generation of Arabidopsis nip5;1-1 complementation lines**

The *nip5;1-1* T-DNA insertion line (Salk_122287C, Takano et al., 2006) was a kind gift of Professor Junpei Takano, Osaka Prefecture University. Transgenic complementation of homozygous *nip5;1-1* mutant lines was done with Arabidopsis NIP-GFP translational fusion constructs under the control of the 35S CaMV promoter by the general approach described in Beamer et al. (2021). *Agrobacterium tumefaciens* GV3101 strains carrying the the desired *CaM35S*<sub>pro</sub>-*CDS-eGFP* construct was used to transform *nip5;1-1* background using floral dip method (Clough and Bent, 1998). The transgenic lines were identified by initial screening and selection on MS media supplemented with 15 ug/mL BASTA. For confirmation and genotyping, genomic DNA
was extracted from 2-wk old seedlings by using the Wizard Genomic DNA purification kit (Promega) and was subjected to PCR analysis by using a primer set consisting of a specific CaM35S promoter sequence and a gene-specific primer (Table II-2). Seed from homozygous T2 generation complementation lines were used for all analyses in this study.

**Limiting boron phenotype analysis and B uptake in plants**

For seedling growth under limiting or sufficient boron conditions on plates, MGRL media (Fujiwara et al., 1992) supplemented with 1% (w/v) sucrose was treated O/N with 3 g/L of Amberlite IRA743 (Sigma) B-chelating resin with gentle shaking at 22°C. The resin was removed and boric acid was added to a final concentration of 1 μM (low) or 50 μM (sufficient) in addition to 1.5% (w/v) Phytagel (Sigma) before autoclaving. Seeds were soaked in a 8% (v/v) sodium hypochlorite solution with 0.05% (v/v) Tween-20 for 15 min and washed five times with sterilized ultrapure water. The sterilized seeds were sown in the MGRL medium with 1μM boric acid and subjected to a vernalization treatment at 4 °C for 2 d. The plates were grown under LD conditions. After 7 days the plants were screened for GFP signal using a wide-field epifluorescence microscope (DM6000 B; Leica) before images were taken using a DSLR camera (Canon Rebel XS) and seedling weight was recorded.

For growth in soil, seeds were sown on MGRL media (Fujiwara et al., 1992), were supplemented with 0.1% (w/v) sucrose, 30 ug/mL BASTA and 50 μM boric acid, were vernalized, and were grown under LD conditions. After 10 days the plants were transferred to soil and watered every 5 days with MGRL media supplemented with either 0.3 μM (low) or 30 μM (normal) boric acid. Images were taken at 29 d and 40 d
growth using a DSLR camera (Canon Rebel XS) before plant tissues were harvested for weight measurements and elemental analysis.

For the determination of boric acid uptake and content of plants, rosettes from 40-day old plants (n= 6 to 12 per treatment) were harvested and were extracted and analyzed by the general approach of (Diehn et al., 2019). Samples were digested in nitric acid at 70°C overnight. The B content was determined by ICP-MS using a Agilent 7500 cx instrument operating in the He collision mode (Spectroscopy and Biophysics Core, University of Nebraska, Lincoln). The samples were diluted 10 to 20-fold into the autosampler plates in 2% (v/v) HNO₃ and were supplemented with 50 ppb ⁷¹Ga as an internal standard. The instrument was operated in NoGas mode with a carrier flow rate of 1.0 mL/min, make-up flow at 0.1 mL/min, plasma gas at 15 L/min, auxiliary gas of 1 L/min. The samples were loaded with an ESI autosampler at a flow rate of 55 uL/min with 2% (v/v) HNO₃ as carrier solution.

**RNA purification and quantitative PCR**

Total RNA was isolated from plant tissues by grinding in liquid nitrogen followed by extraction with the PureLink Plant RNA Reagent (Invitrogen). RNA isolation and DNase treatment was carried out with a Direct-zol RNA MiniPrep Plus Kit by using the manufacturer’s protocol (Zymo Research). Q-PCR was performed on a Bio-Rad IQ5 real-time PCR detection system by using iTaq Universal SYBR Green One-Step RT-qPCR kit (Bio-Rad Laboratories) according to the manufacturer’s instructions with the following program: 50°C for 10 min, 95°C for 1 min, and 40 cycles of 95°C for 10s and 60°C for 30s. This amplification protocol was followed by an additional thermal denaturation cycle (65 °C to 95 °C with 0.5°C increments) was performed to generate
melting curves to validate the amplification specificity. All primer sets produced a single amplification product with the expected Tm. All gene-specific primers used in this study are listed in Table II-1. PCR amplification cycle efficiencies for all primer pairs were over 95%. Accession number of the genes used in this study are AT2g34390 (NIP2;1), AT4g33070 (PDC1), AT4g18360 (GOX3), AT1g17290 (AlaAT1), AT1g77120 (ADH1), AT3g18780 (ACTIN 2), AT4g05320 (UBIQUITIN 10), AT4g17260 (LDH1).

Quantitative expression analysis was calculated by the comparative threshold cycle (Ct) method (Pfaffl, 2001) modified to take into account two separate reference genes as described in Hellemans et al. (2007). In the present study, two transcripts that are commonly used as references for the hypoxia-induced gene expression, UBQ10 (Choi and Roberts, 2007; Giuntoli et al., 2014; R et al., 2018) and ACTIN2 (Loreti et al., 2020) were used. The stability of both reference genes was assessed using GeNorm software; both showing the M value (0.354) lower than the threshold of 1.5 suggesting their stable expression and suitability for the tested samples (Vandesompele et al., 2002). The relative expression of genes was calculated by normalizing the data to the geometric mean of relative quantity of the reference genes as described here.

Relative Expression = $2^{\Delta \text{Ct}_{\text{GOI}}} / \text{Geomean}[2^{\Delta \text{Ct}_{\text{ref}}}]$

Where $\Delta \text{Ct} = \text{Ct}_{\text{calibrator}} - \text{Ct}_{\text{sample}}$; GOI is the gene of interest; and Geomean refers to the geometric mean of $2^{\Delta \text{Ct}}$ of UBQ10 and ACTIN2. The specific calibrators are described in the figure legends. All data represent at least six determinations from three different biological replicates.
**Histochemical and microscopy techniques**

GUS staining and clearing of *Arabidopsis thaliana* lines with the NIP2;1\textsubscript{pro}:GUS reporter transgene was done by the protocol of (Choi and Roberts 2007), and stained tissues were imaged with a Leica MZ16FA microscope (Leica Microsystems). For analysis of NIP2;1 promoter activity in root cross-sections, GUS-stained roots were dehydrated in ethanol and embedded in Technovit 7100 resin by the manufacturer’s (Kulzer GmbH) protocol. Cross sections (2.5 µm thickness) were generated from the mature differentiated region of the primary root with a Reichert OMV3 microtome equipped with a glass knife, and were mounted in 50% (w/v) glycerol. Cross-sections were imaged with a Nikon ECLIPSE E600 microscope equipped with Micropublisher 3.3 and QCapture 2.60 software (QImaging corporation).

Epifluorescence imaging of NIP2;1-GFP seedlings were captured with an Axiosvert 200M microscope (Zeiss) equipped with filters for GFP fluorescence (Zeiss; filter set 38 HE) and a digital camera (Hamamatsu Orca-ER) controlled by the Openlab software (Improvision). Subcellular localization analysis under hypoxia and reoxygenation was done with a Leica SP8 white laser confocal microscope system at the Advanced Microscopy and Imaging Center at The University of Tennessee, Knoxville. To stain the plasma membrane, hypoxia-treated seedlings were removed from the plate and were incubated in 4 µM FM4-64 (Invitrogen) under hypoxic conditions in darkness for 10 min. For reoxygenation, seedlings were returned to aerobic conditions under light for 1 hr before staining and visualization. The 488-nm excitation filter was used, and the emission filter for detection was set to 495 to 550 for GFP, 580 to 650 nm for FM4-64 and 680-720 for chlorophyll. Confocal micrographs
were captured with the Leica LASX software and uncompressed images were exported and analyzed in ImageJ version 1.53a (Schneider et al., 2012) to adjust the brightness and contrast of images, and to generate merged images.

For GFP localization in nip5;1-1 line complemented with NIP-GFP constructs, 7 d old seedlings grown on MS media supplemented with 1% (w/v) sucrose under the growth conditions mentioned above, were stained with 4 μM FM4-64 for 10 min. Confocal imaging of the leaves of transgenic plants was performed with a Leica SP8 confocal microscope using the settings described above.

**Immunochemical techniques**

Anti-NIP2;1 antisera were produced against a synthetic peptide (GenScript) corresponding to the C-terminal sequence of NIP2;1 (CHKMLPSIQNAEPEFSKTGSSHKRV) following the immunization protocol of Guenther et al. (2003) with the exception that Titermax-Gold was substituted for Freund's adjuvants. Antibodies were affinity purified on peptide resins as described in Guenther et al. (2003).

For analysis of NIP2;1 protein in WT and nip2;1 mutant seedlings after hypoxia treatment, Arabidopsis root tissues from seedlings treated with 6 hr hypoxia or normoxic controls were extracted, and a membrane microsomal fraction was prepared as described by Ishikawa et al. (2005). Protein concentrations were determined by using the BCA assay (Pierce Biochemical). The SDS-poly acrylamide gel electrophoresis (PAGE) and Western blot analysis were performed using 10 μg protein from membrane microsomal fractions as previously described Guenther et al. (2003). For the analysis of NIP2;1-GFP expression in complementation lines, hypoxia and reoxygenation
treatments were conducted as described above, and samples were directly extracted into SDS-PAGE sample buffer (Laemmli, 1970) for Western blot analysis. The same approach was used for the analysis of NIP-GFP expression in 10 d old seedlings of nip5;1-1 complementation lines. For detection of NIP-GFP protein in oocytes, SDS-polyacrylamide gel electrophoresis (PAGE) and Western blot analyses were performed using 10 μg protein from oocyte lysates (Wallace and Roberts, 2005). Rabbit anti-GFP polyclonal antibodies (Abcam) were used for the detection of all NIP-GFP fusion proteins.

**Media acidification and L-lactate measurements**

Media acidification assays with the pH-sensitive indicator bromocresol purple were done by the method of (Silva et al., 2018) Seven-day-old WT and nip2;1 seedlings (20 seedlings per treatment) were transferred from MS media to 1.5% (w/v) agarose plates containing 60 mg/L bromocresol purple (Acros Organics) and 1 mM CaSO₄ in sterile distilled water with the pH adjusted to 5.7 using KOH. Seedlings were subjected to argon-induced hypoxia or air (normoxic controls) as described above, and the change in pH indicator color was assessed throughout hypoxia treatment. Images were captured with a DSLR camera (Canon Rebel XS).

To determine L-lactic acid/lactate concentration within Arabidopsis roots, vertically grown seedlings (63-72 seedlings per treatment replicate) were submerged in MS media and hypoxia treatment carried out over a 12 hr time course by flushing continuously with nitrogen gas. At each time point, samples were removed from treatment, roots were dissected, and were snap frozen in liquid nitrogen. The frozen tissue was ground in a mortar with a pestle and was extracted in two volumes of 1N
perchloric acid and was neutralized with potassium carbonate on ice prior to assay by the LDH method of (Bergmeyer and Bernt, 1974). NADH production was assayed by the change in absorbance at 340 nm, with background corrected from duplicate samples treated identically except for no added LDH enzyme.

To quantify the rate of L-lactic acid/lactate efflux from roots to the external media, a modified assay with increased sensitivity was used that utilizes bacterial lactate oxidase (Cell Biolabs). Seven-day-old seedlings (60-75 per treatment replicate) were weighed and transferred to each well of a 6-well plate with roots submerged in 2 mL of MS liquid media. Argon gas hypoxia was administered over an 8 hr time course, the lactate content of media aliquots was assayed hourly, and the rate of root lactate release was standardized to seedling fresh weight.

Expression and functional analyses in Xenopus oocytes

*Xenopus laevis* expression and functional analyses were performed as previously described (Wallace and Roberts, 2005; Tanaka et al., 2008; Li et al., 2011). NIP or AQP expression constructs were linearized by digestion with *Xba*I, and capped cRNA was generated by in vitro transcription by using the mMessage mMachine kit (Ambion). Stage V and VI *Xenopus* oocytes were collected surgically and defolliculated (Guenther et al., 2003) or were obtained from Ecocyte Bioscience (Austin, TX). Oocytes were microinjected with 46 nL of 1 ug/ul of cRNAs or with RNase-free water as a negative control using a “Nanoject” automatic injector (Drummond Scientific Co., Broomall, PA). The oocytes were cultured for 72 h in 96-well microtiter plates at 16 °C in standard Ringer’s solution (96 mM NaCl, 2 mM KCl, 5 mM MgCl₂, 5 mM HEPES-NaOH
pH 7.6, 0.6 mM CaCl$_2$, 190 mosmol/kg) supplemented with 100 ug/mL penicillin/streptomycin.

The boric permeability of the oocytes was determined as described previously (Tanaka et al., 2008) by incubation of groups of eight oocytes in standard Ringer’s solution supplemented with 2 mM boric acid. Assays were conducted at 16°C for 20 to 30 min, after which oocytes were washed five times on ice (1 ml per 8 oocytes per wash) with standard Ringer’s solution without boric acid followed by homogenization and overnight digestion at 65 °C in 100 uL of nitric acid. The uptake of arsenous acid was done identically except in the presence of 1 mM sodium arsenite instead of boric acid. The As or B content of the digests were determined by ICP-MS analysis as described above (Spectroscopy and Biophysics Core, University of Nebraska, Lincoln).

The osmotic water permeability ($P_f$) of the oocytes was measured by the standard swelling assay as described previously (Guenther and Roberts, 2000). Oocytes were transferred from standard Ringers to hypoosmotic dilute Ringer’s media (60 mosmol/kg) and the rate of oocyte swelling (($dV/V_o)/dt$) was determined from the cross-sectional area change determined by video microscopy. The $P_f$ was calculated from:

$$P_f = \frac{V_0 \frac{d(V)}{dt}}{S_0 (osm_{in} - osm_{bath})V_W \left( \frac{S_{real}}{S_{sphere}} \right)}$$

Where V is the volume at a specific time; $V_o$ and $S_0$ are the initial oocyte volume and cross sectional area; $osm_{in} - osm_{bath}$ is the osmotic gradient; $V_W$ is the partial molar
volume of water, and $S_{\text{real}}/S_{\text{sphere}}$ is a surface area correction constant that accounts for the topology of the oolemma (Rivers et al., 1997).

**Computational methods and protein modeling techniques**

The NIP structural models were obtained from the AlphaFold2 website, which is an AI system developed by DeepMind to predict a 3D structure of a protein from its amino acid sequence (Senior et al., 2020; Jumper et al., 2021; Pearce and Zhang, 2021). The mutant models were generated using the Molecular Operating Environment (MOE, 2022). The NIP structural models were refined by molecular dynamics simulation (MD) performed using Amber simulation engine (Case et al., 2005). The systems were hydrated using the water model TIP3P (Price and Brooks, 2004) in an octahedral box of 10 Å around the protein in each direction. The MD was performed on the proteins using Amber14 ff14SB (Maier et al., 2015) force field with a non-bonded cutoff of 10 Å using the Particle Mesh Ewald algorithm (Darden et al., 1993). The restraints applied during the simulation were taken from a previous study (Dutagaci et al., 2018), which reported a protocol for the structure refinement of membrane proteins with proper restraints.

All systems were initially minimized while applying restraints on water and ions with a 5 kcal/mol/Å$^2$ force constant. The system was energy minimized with 5000 steps of steepest descent followed by 5000 steps with the conjugate gradient method. The systems were further equilibrated using restraints on Cα and Cβ atoms with a force constant of 0.5 kcal/mol/Å$^2$. The systems were heated to 300 K, and 1000 MD steps were performed. The SHAKE algorithm (Miyamoto and Kollman, 1992) was used to constrain all bonds involving hydrogen in the simulations. MD production runs were performed at 300 K using the NPT ensemble and a 2 fs time step. The temperature was
fixed with the Langevin dynamics thermostat\(^9\) and the pressure was fixed with the Monte Carlo barostat (Åqvist et al., 2004). C\(\alpha\) atoms were restrained during the simulations with a force constant of 0.025 kcal/mol/Å\(^2\) to avoid large deviations from the initial structures. Thirty nanosecond production runs were performed for each system. The structures obtained during the last 10 ns of each of the simulation trajectories were segregated into four clusters based on the backbone atoms with the hierarchical agglomerate clustering algorithm present in the Cpptraj module (Roe and Cheatham, 2013). This method helps in dimension reduction and generates an ensemble of structures that also takes the local and global motion of the protein into account.

Comparison of the conformations and rmsd calculations of the structural models was done in MOE. The dimension of the transmembrane pores along the z-coordinate of protein models was calculated by using the MOLEonline Websever in the “Channels mode” (Probe Radius 5, Interior Threshold 1.5, Merged Pores On) for the open structures and “Pore mode” (Probe Radius 5, Interior Threshold 0.3) for the closed structures (Berka et al., 2012; Pravda et al., 2018). Models of metalloid hydroxides were constructed and minimized in MOE and the molecular volumes were calculated by using VEGA ZZ (Pedretti et al., 2021). The pdb coordinates for the open (7CJS, Saitoh et al., 2021) and the closed (7NL4, van den Berg et al., 2021) structures of the rice silicic acid channel Lsi1 were obtained from the NCBI Structure database, and the silicic acid bound Lsi1 structural model (Model 2) was taken from Saitoh et al. (2021). Boric acid docked structures of NIP6;1 were constructed by using MOE by using the selectivity filter waters of Lsi1 structure (Waters 1, 2, and 4 in 7CJS) as a template for orientation of the three hydroxyl groups of the boric acid molecule. The final boric acid-bound
NIP6;1 protein structure was energy minimized taking into account the trigonal planar geometry of the boric acid hydroxyl groups.

**Expression and purification of *Glycine max* Nodulin26 in *Pichia pastoris***

The OptimumGene™ codon optimization analysis system was used to codon optimize the *Glycine max* Nodulin 26 coding sequence (NP_001235870.1) for *Pichia pastoris* (GenScript,NJ). With the primers listed in Table II-3, optimized coding sequences were subcloned as a 6x N-terminal histidine tagged fusion into the *BamH*1 and *NotI* restriction sites of the pPIC3.5K expression vector. Expression constructs were transformed into *Pichia pastoris* according to the instructions in the Easyselect™ Pichia handbook (Life Technologies). Clones with multiple inserts were identified through selection on 1.75 mg/mL geneticin sulfate (MP Biomedicals) media and cultured for protein expression in IsoYeast media (Sigma). A single colony was inoculated into a 20 mL IsoYeast growth medium and cultured overnight at 30°C with shaking at 250 rpm. The seed culture was inoculated into 1 L IsoYeast growth media and grown under identical conditions until the OD600 exceeded 2.0. The cells were collected by centrifugation at 3000xg at 4°C in a Sorvall GS-3 rotor, and the pellet was resuspended to 250 mL IsoYeast expression media with 0.5% (v/v) methanol to induce expression. The culture was shaken at 28°C and supplemented with 0.5% (v/v) methanol every 24 hours, for 96 hours in total. Cell pellets (~10 grams) were resuspended in 25 mL of 20 mM HEPES–NaOH, pH 8.0, 150 mM NaCl, 0.5 mM Tris(2-carboxyethyl)phosphine (TCEP) supplemented with protease inhibitor tablets (Thermo Scientific) before being lysed by several passes through a French-Press Cell Disrupter (Thermo Scientific) at 4°C. The lysate was centrifuged at 7000xg for 45 minutes at 4°C, and the supernatant
**Table II-3** Sequences of primers used for cloning codon optimized nod26 with an N-terminal 6xHis tag and sequencing primers.

<table>
<thead>
<tr>
<th>NAME</th>
<th>SEQUENCE (5'-3')</th>
<th>Description</th>
<th>PURPOSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ferri_SyntheseDeF</td>
<td>GAAAGACTGCGaCATCATGATGATGATGATGATGATGATGA</td>
<td>codon optimized nod26 forward primer with 6xHis tag insertion</td>
<td>Cloning of codon optimized with addition of N-terminal 6xHis tag</td>
</tr>
<tr>
<td>Synthesis_Rep_F</td>
<td>TCCCATCTCCTCCCTACCCTACCCTACCCTACCCTACCCT</td>
<td>codon optimized nod26 reverse primer</td>
<td></td>
</tr>
<tr>
<td>AOX_Pep_F</td>
<td>gcggggtcggggtcggggtcggggtcggggtcggggtcggg</td>
<td>Aldehyde oxidase [AOX1] primer forward primer</td>
<td></td>
</tr>
<tr>
<td>AOX_Pep_R</td>
<td>gcccgtccggccgtccggccgtccggccgtccggccgtccgg</td>
<td>Aldehyde oxidase [AOX1] primer reverse primer</td>
<td></td>
</tr>
</tbody>
</table>

Notes: For cloning primers, the lower case letters indicate gene-specific sequences, while upper case letters indicate engineered restriction sites or other added sequences.
fraction was collected and was then centrifuged at 200,000xg for 2 hours at 4°C. The membrane pellet fraction was resuspended in solubilization buffer (50 mg/mL) containing 20 mM HEPES–NaOH, pH 8.0, 150 mM NaCl, 5 mM imidazole, 5 mM spermidine (ACROS Organics) and 2% (w/v) Lauryl Maltose Neopentyl Glycol (LMNG) (Anatrace). The membrane pellet was solubilized by gentle agitation overnight at 4 °C. The mixture was centrifuged at 12,000xg for 20 minutes, and the supernatant fraction collected and applied to the Ni Sepharose High Performance resin (Cytiva) that had been pre-equilibrated in 10 mL of 20 mM HEPES–NaOH, pH 8.0, 150 mM NaCl, 10 mM imidazole and 0.01% (w/v) LMNG for 2 hrs at 4°C. The resin was washed with 200 column volumes of wash buffer (20 mM HEPES–NaOH, pH 8.0, 150 mM NaCl, 60 mM imidazole and 0.01% (w/v) LMNG) before eluting the histidine-tagged protein with 20 mM HEPES–NaOH, pH 8.0, 150 mM NaCl, 500 mM imidazole and 0.01% (w/v) LMNG. The elution fraction was concentrated on a Vivaspin sample concentrator (GE Healthcare) with a 50 kDa MW cutoff and chromatographed on Superdex 200 10/300 GL (GE Healthcare). To pre-equilibrate the column, two column volumes (48 mL) of 20 mM HEPES–NaOH, pH 8.0, 150 mM NaCl and 0.01% (w/v) LMNG were used. The chromatography was done in the same buffer as the equilibration. The elution of the protein from the column was monitored by the absorbance at 280 nm on the AKTA FPLC system utilizing the UNICORN software (GE Healthcare). The protein-containing fractions were combined and used for biophysical studies or concentrated to 6-10 mg/mL for crystallization trials using a Vivaspin sample concentrator (GE Healthcare) with a 50kDa MW cutoff filter. SDS-PAGE was used to assess protein purity, and the final product was stored at 4°C until crystallization trials were initiated.
Detergent screening for Nod26 crystallization

To assess the stability of purified nod26 solubilized in analogues of the LMNG detergent, nod26 was solubilized from *Pichia pastoris* membranes as described above using one of the following alternative detergents: Decyl Maltose Neopentyl Glycol (DMNG, Anatrace), Undecyl Maltose Neopentyl Glycol (UMNG, Anatrace), Octyl Glucose Neopentyl Glycol (OGNG, Anatrace) or n-Octyl-β-D-Glucopyranoside (OG, Anatrace). Following Ni\(^{2+}\)-NTA chromatography and size exclusion chromatography, the protein-detergent complexes were concentrated on a Vivaspin concentrator to 4-17 mg/mL and was stored at 4°C for one week to assess the stability of the complex to aggregation. Following an amended procedure of Gutmann et al. (2007), a 2 uL sample of the protein was removed prior to ultracentrifugation. The remaining protein solutions were ultracentrifuged at 200,000xg for 1 hour at 4 °C and a second 2 uL protein sample was collected. Comparison of the protein densities on SDS-PAGE was used to determine the extent of aggregation of the protein-detergent complexes.

**Nod26 crystallization**

The concentrated protein was ultracentrifuged at 100,000xg for 30 min to remove any aggregated protein prior to crystallization trials. Crystallization screening trials by sitting drop vapor diffusion (Newby et al., 2009) was initiated using the MemGold and MemGold2 screens (Molecular Dimensions) with a Pheonix crystallization robot (Art Robbins) at temperatures ranging from 4 to 25°C. Crystals generally appeared within three weeks and were imaged with a wide-field epifluorescence microscope (DM6000 B; Leica). Crystals were obtained directly from the initial trials or optimized by hanging and sitting drop vapor diffusion with larger drops (2-3 uL total volume). Crystals selected
for diffraction analysis were flash frozen, transported and analyzed at the Advanced Photon Source (Argonne, IL).
CHAPTER III - The Arabidopsis thaliana NIP2;1 Lactic Acid Channel promotes Plant Survival Under Low Oxygen Stress
Background and Rationale: NIP2;1 shows a unique transport selectivity among the NIPs

The biochemical investigation of Arabidopsis NIP2;1 revealed that it varies from other NIP I transport characteristics and shows preferential transport of lactic acid (Choi and Roberts, 2007). Lactic acid is a product of the fermentative processes used by plants to sustain energy generation under oxygen-limiting situations and other stress circumstances in which respiration is suppressed (Drew, 1997; Gibbs and Greenway, 2003). Lactic acid accumulation is one of the processes contributing to the cellular acidification observed under low oxygen stress (Davies et al., 1974; Roberts et al., 1984; Felle, 2005). The capacity of plant roots to efflux lactic acid to the rhizosphere is accompanied by hypoxia-induced fermentation (Xia and Saglio, 1992; Rivoal and Hanson, 1993; Xia and Roberts, 1994; Dolferus et al., 2008). This lactic acid/lactate efflux mechanism in maize (Zea mays) root tips coincides with reduced sensitivity to low oxygen stress from acclimation, and it is hypothesized to be an adaptive mechanism to limit cytosolic acidification or other harmful consequences of cellular lactic acid/lactate buildup (Xia and Saglio, 1992; Xia and Roberts, 1994).

The discovery that NIP2;1 is preferentially permeable to lactic acid together with its identification as a core-hypoxia gene product has led to the notion that it could control lactic acid efflux and/or compartmentation during the Arabidopsis response to low oxygen stress (Choi and Roberts, 2007). This hypothesis has yet to be rigorously
tested *in planta*. The objective of this research was to test this hypothetical function of NIP2;1 during the Arabidopsis hypoxia response by using T-DNA insertional *nip2;1* mutant lines and *NIP2;1-GFP* in complementation seedlings.

**NIP2;1 expression is induced early during hypoxia primarily in root tissue**

Quantitative PCR (qPCR) analysis of hypoxia treated Arabidopsis seedlings resulted in a rapid >1000-fold increase in *NIP2;1* transcript levels within two hours after the onset of anaerobiosis (Fig. III-1A). *NIP2;1* transcript levels then showed a sharp decline by 12 hr but still remained over 100-fold elevated relative to normoxic controls before leveling off at 24 hr (Fig. III-1A). While *NIP2;1* is predominantly a root transcript, hypoxia also induced *NIP2;1* expression in shoot tissues, but with a lower overall expression (40-fold relative to basal levels) compared to roots (Fig. III-1A). In addition, the time course of accumulation in shoots was delayed compared to roots and the expression did not peak until 12 hr treatment.

Similar patterns of *NIP2;1* expression were observed with two-week-old *NIP2;1* promoter::GUS fusion plants subjected to the same oxygen deprivation regime (Fig. III-1B). Analysis of the cellular localization of GUS staining under normoxic conditions showed that expression is principally limited to roots with little staining detected in shoot tissues (Fig. III-1B). Cross-sections of unstressed (normoxic) roots showed that GUS staining was principally observed in the cells of the stele (pericycle, phloem and procambium,) with little or no GUS signal apparent in endodermal, cortical and epidermal cells (Fig. III-1B, Fig. III-2). At 4 hr after the induction of hypoxia, root tissues showed an increase in intensity of the GUS staining in the stele as well as the appearance of the GUS signal in the cortex, epidermis and root hairs (Fig. III-1B).
Figure III-1 *NIP2;1* expression in Arabidopsis seedlings in response to oxygen deficit. **A.** Quantitative real-time RT-PCR (Q-PCR) analysis of *NIP2;1* transcripts in root and shoot tissues during a hypoxia time course of two-week old Arabidopsis seedlings. The ΔCt value of *NIP2;1* obtained from the 0 hr shoot sample was used as the calibrator for expression level normalization. Error bars indicate the SD of three biological replicates. **B.** GUS staining analysis of two-week old *NIP2;1pro::GUS* Arabidopsis seedlings subjected to the oxygen deprivation conditions as in panel A. Top panel are representative whole seedlings while the bottom panel shows root cross-sections. Scale bars are 1.0 mm for the top panel and 20 μm in the bottom panel. **C.** Close up of GUS-stained leaves from Panel B. Scale bar = 1.0 mm. Data obtained from from Choi, W.G. (2009) (https://trace.tennessee.edu/utk_gradiss/26) and adapted from Beamer et al. (2021).
Figure III-2 Cellular localization of AtNIP2;1 promoter activity in the roots under normoxic growth conditions. Two-week-old NIP2;1pro::GUS transgenic seedlings were grown under normoxic conditions and were stained for the reporter enzyme GUS. The GUS-stained roots were then transverse sectioned (2 μm thickness). Top, Full cross section of mature root zone. E, epidermis; c, cortex; en, endodermis; stele. Bottom, Close up the stele region with cell types indicated based on the morphological study of Dolan et al. (1993). c, cortex; en, endodermis; p, pericycle; x, metaxylem. Data obtained from Choi, W.G. (2009) (https://trace.tennessee.edu/utk_graddiss/26) and adapted from Beamer et al. (2021).
Similar to the Q-PCR result, this staining peaked at 4 hr post hypoxic treatment and then decreased by 12 hr post treatment (Fig. III-1B), although the expression at these later time points was still much higher than the basal expression in the unstressed roots (Fig. III-1B). In comparison to roots, increases in GUS expression in shoots were less acute and appeared more slowly and the expression was mainly restricted to the vascular tissues of leaves (Fig. III-1C).

**NIP2;1 enhances plant survival under low oxygen conditions**

Core hypoxia-response gene loss-of-function mutants generally result in reduced survival or increased sensitivity to low oxygen stress (Ismond et al., 2003; Kursteiner et al., 2003; Licausi et al., 2010; Giuntoli et al., 2014; Sorenson and Bailey-Serre, 2014; Lokdarshi et al., 2016). To investigate whether *NIP2;1* is necessary for hypoxia stress survival, a T-DNA insertion mutant (WiscDsLox233237_22K; referred to here as *nip2;1*) line was studied. The *nip2;1* line contains a T-DNA insertion within the promoter region (-30) between a cluster of Anaerobic Response Elements and the transcriptional start site (Fig. III-3A). Consistent with the position of this insertion, hypoxia treatment (4 hr) of *nip2;1* mutants resulted in poor expression of *NIP2;1* compared to WT (Fig. III-3B). Western blot analysis of WT roots with site-specific anti-*NIP2;1* antisera showed the hypoxia-induced appearance of an immunoreactive band that was not detected under the same conditions in *nip2;1* mutant roots (Fig. III-3C). Overall, the data confirm that this T-DNA insertional mutant shows a severe defect in *NIP2;1* expression and do not produce a detectable protein product.

To determine the effects of low oxygen stress on *nip2;1* plants, their growth and survival under normoxic and hypoxic conditions were compared (Fig. III-4). While WT
Figure III-3 Characterization of nip2;1 T-DNA insertional mutant seedlings. A. Schematic diagram showing the site of T-DNA insertion in the nip2;1 mutant line. The position of Anaerobic Response Elements (AREs) in the promoter region based on Olive et al. (1990) (cross hatched bar) or Dolferus et al. (1994) (open bars) relative to the site of T-DNA insertion is shown. B. Q-PCR results for NIP2;1 expression in the roots of seven-day-old wild type (Col-0) and nip2;1 during hypoxic treatment. The ΔCt value of NIP2;1 expression in normoxic wild type roots was used as a calibrator for expression normalization. Error bars show the SD of 3 biological replicates. C. Root extracts (10 μg protein/lane) were analyzed by Western blot with site-directed NIP2;1 antibodies. 0 hr, normoxic control; 6 hr, 6 hr hypoxia-treated plants. Data obtained from from Choi, W.G. (2009) (https://trace.tennessee.edu/utk_graddiss/26) and adapted from Beamer et al. (2021).
Figure III-4 Effects of oxygen deprivation on the survival of nip2;1 T-DNA insertional mutant seedlings. 

A. Seven-day-old, vertically grown seedlings of wild type (Col-0) and nip2;1 were subjected to a 8 hrs of argon gas treatment and were allowed to recover under normal oxygen conditions for 72 hours prior to assessing survival. 

B. PSII maximal quantum yield [QYmax (Fv/Fm)] was calculated from chlorophyll fluorescence analysis of seven-day-old wild type Columbia (WT) and nip2;1 mutant seedlings treated with 8 hr argon (hypoxia) or air control (normoxia). Error bars represent std. error of mean of five biological replicates. 

C. Box and whiskers plot showing the survival of seven-day-old wild type, nip2;1, and two NIP2;1-GFP complementation seedling lines to 8 hr argon treatment represented as % seedling survival. Each data point represents one biological replicate with the median value indicated in each box. Statistical significance was assessed by multiple comparisons by One way ANOVA analysis. (**** represents p<0.0001 for nip2;1 compared to wild type and both complementation lines; ns, not significant).
and nip2;1 seedlings showed little difference in growth under normoxic conditions (Fig. III-5), nip2;1 seedlings showed higher sensitivity to hypoxia treatment (Fig. III-4). After exposure to argon gas-induced hypoxia, followed by transfer back to normoxic conditions for recovery, nip2;1 seedlings exhibited a higher incidence of chlorosis and seedling death (Fig. III-4A). Comparison of the overall survival frequency of WT and nip2;1 seedlings showed that the mutant exhibited significantly poorer survival to hypoxic stress (Fig. III-4C).

The sensitivity of the nip2;1 mutant to hypoxia was further assessed by measuring the chlorophyll fluorescence properties and calculating the maximum potential quantum efficiency (Q_{max} or F_{v}/F_{m}) of photosystem (PS) II of nip2;1 and WT seedlings under normal and low oxygen stress conditions. Chlorophyll fluorescence measurements is a common technique used to assess the photosynthetic efficiency of PS II which is an index of the susceptibility of plants to different environmental stressors (Murchie and Lawson, 2013). Under normoxic conditions, WT and nip2;1 seedlings were not significantly different with F_{v}/F_{m} values that fell within the optimum range (0.78 - 0.8, Murchie and Lawson, 2013). This suggests that the lack of NIP2;1 expression in the mutant line does not exhibit any detectable adverse effect on this parameter under standard growth conditions. However, upon hypoxia treatment both WT and nip2;1 seedlings showed a steep reduction in photosynthetic efficiency with the F_{v}/F_{m} ratio declining to below 0.6 within four hours of the recovery period (Fig. III-4B). At all time points, the F_{v}/F_{m} ratio is lower for nip2;1 than WT. At later time points the quantum efficiency of PS II starts to recover for both wild type and nip2;1 seedlings (Fig. III-4B), but the nip2;1 F_{v}/F_{m} ratio remains significantly lower than WT after 24 hours of recovery.
Figure III-5 Complementation of *nip2;1* T-DNA mutant with a *NIP2;1pro::NIP2;1-GFP* construct. *Nip2;1* T-DNA mutants were transformed with a construct consisting of the *NIP2;1* open reading frame with an in-frame C-terminal placed under the control of the native *NIP2;1* promoter. Seven-day-old plants were subjected to 8 hr of Argon gas induced hypoxia (Argon) or air (Normoxic Control) and were returned to normal growth conditions for three days. Wild type, *A. thaliana* (Col-0), *nip2;1*, T-DNA insertional mutant of *NIP2;1*; *NIP2;1-GFP*, *nip2;1* plants complemented with *NIP2;1-GFP* translational fusion under the *NIP2;1* promoter. Scale bars represent 1 cm.
To confirm that the increased sensitivity of *nip2;1* plants to hypoxia is result of the loss of *NIP2;1* gene, two complementation lines containing a *NIP2;1pro::NIP2;1-GFP* transgene (*NIP2;1-GFP* plants) in the *nip2;1* background were analyzed (Fig. III-4C). Both complementation lines showed enhanced tolerance to hypoxia challenge compared to *nip2;1* seedlings, and were not statistically different to WT seedlings with respect to survival frequency (Fig. III-4C). Taken together, the *nip2;1* phenotype data indicate that *NIP2;1* is hypoxia core response protein that participates in the hypoxia adaptation response, and that reduction in expression of *NIP2;1* lowers the ability of Arabidopsis to survive this stress. One *NIP2;1-GFP* complementation line (line K) was chosen for further study.

**NIP2;1 is expressed on the plasma membrane as well as on internal membranes during hypoxia and reoxygenation recovery**

To investigate the dynamics of NIP2;1 expression and its subcellular localization, NIP2;1-GFP expression in the complementation line was analyzed. Similar to WT, Q-PCR analysis shows that the *NIP2;1-GFP* transgene transcripts are acutely induced by hypoxia, with a peak at 4 hr followed by a decline (Fig. III-6A). As has been documented with other core hypoxia transcripts (Branco-Price et al., 2008), reoxygenation resulted in a rapid decline of the transcript to basal levels. Analysis of NIP2;1-GFP protein accumulation during hypoxia and re-oxygenation was done by using epifluorescence microscopy and Western blot analysis at different time points of hypoxia stress and reoxygenation recovery (Fig. III-6B and C). Microscopy revealed that the GFP signal first appeared within two hours of the onset of hypoxia (Fig. III-6C), and increased as hypoxia proceeded. Return to normal oxygen conditions resulted in a substantial
Figure III-6 NIP2;1-GFP expression in the roots of NIP2;1-GFP plants during hypoxia and reoxygenation. Ten day old NIP2;1-GFP plants were subjected to an argon-induced hypoxia time course, with oxygen resupplied at hour 6. **A.** Q-PCR analysis of NIP2;1-GFP transcript accumulation in roots. Relative expression was standardized as described in Figure 2. **B.** **Left,** Anti-GFP western blot showing NIP2;1-GFP protein accumulation (upper panel). **Bottom panel,** Coomassie blue stained loading control gel. **Right,** The relative expression of protein based on densitometry of the Western blot is shown in the right. **C.** Representative epifluorescent images of NIP2;1-GFP plants at the indicated times of hypoxia treatment or reoxygenation. Scale = 50 μm.
increase in the fluorescent intensity at 30 min and remained throughout the recovery period (Fig. III-6C).

Western blot analysis show a similar pattern for protein accumulation as that observed with fluorescence microscopy with the protein levels increasing during hypoxia, and remaining elevated during reoxygenation (Fig. III-6B). Two bands were observed in Western blot analyses, a major band migrating as expected for the NIP2;1-GFP fusion, and a second minor band with a mobility similar to free GFP, likely representing a degradation product. While transcript levels decline rapidly to non-detectable levels (Fig. III-6A), the fluorescent intensity and Western blot analyses indicate the persistence of the NIP2;1 protein for hours after return to normal oxygen conditions (Fig. III-6B and C).

Different NIP proteins show varied subcellular localization, ranging from polarized plasma membrane localization for the root boric acid channel NIP5;1 (Wang et al., 2017) to specific localization on subcellular organelles such as the soybean symbiosome membrane protein, nodulin 26 (Weaver et al., 1991). The determination of NIP2;1 localization under native conditions in response to low oxygen stress is important to understand the potential path of lactic acid transport. To investigate its subcellular localization under native conditions during hypoxia and reoxygenation, the roots of NIP2;1-GFP seedlings were closely examined more closely by using confocal microscopy (Fig. III-7, 8 and 9).

Analysis of NIP2;1-GFP at 2 hr after the onset of hypoxia revealed accumulation of GFP fluorescence throughout the root tip, and the elongation and maturation zones (Fig. III-8A). Closer examination and co-localization analyses with FM4-64 shows that
Figure III-7 Subcellular localization of NIP2;1:GFP in the roots of hypoxia challenged NIP2;1:GFP complementation seedlings. A. Seven-day-old vertically grown NIP2;1-GFP seedlings were subjected to anaerobic stress by root submergence under argon gas treatment for 2 hr, and the appearance of NIP2;1-GFP was monitored by confocal fluorescence microscopy. NIP2;1-GFP (top panel), FM4-64 (middle panel) and merged (bottom panel). Scale bar = 20 μm. B. NIP2;1:GFP seedlings were subjected to argon gas treatment at the times indicated followed by return to normoxic conditions at 6 hr. DIC, differential interference contrast images. Bars = 50 μm.
Figure III-8 Subcellular localization of NIP2;1-GFP in the roots of hypoxia challenged seven-day-old NIP2;1-GFP complementation lines. A. Seven-day-old vertically grown NIP2;1-GFP seedlings were subjected to anaerobic stress by root submergence under argon gas treatment for 2 hr, and the appearance of NIP2;1-GFP (upper panel) and FM4-64 staining (lower panel) was monitored. Scale bar = 50 µm. B. Higher magnification micrograph showing NIP2;1-GFP (top panel) and FM4-64 (bottom panel). Scale bars = 50 µm. The longitudinal and transverse axes used for fluorescence intensity quantitation and polarity index analysis by ImageJ are shown in the figure. C. Intensity profile of GFP and FM4-64 along the axes lines in the figure B. D. Polarity index of NIP2;1-GFP expression in the plasma membrane at 2 hr hypoxia. The polarity index of NIP2;1 expression was calculated by comparing the normalized GFP intensity of apical/basal end of the membrane to its the radial side as described in the Materials and Methods. Values represent here is the means± SD (n = 10 cells from two different NIP2;1-GFP transgenic lines).
Figure III-9 Comparison of surface and internal localization of NIP2:1-GFP during hypoxia and recovery. \textit{NIP2:1-GFP} complementation line K was subjected to hypoxia and reoxygenation as shown in Fig. 5 and the relative fluorescence intensity distribution of FM4-64 and GFP fluorescence was imaged and quantitated along the longitudinal (left column) and radial (right column) axes (see Fig. S3 and Materials and Methods for details).
NIP2;1-GFP is predominantly localized to the plasma membrane, although a lower level of signal within internal structures is also apparent (Fig. III-7A, 8B and 8C). Simple examination of the confocal micrograph images suggested that the intensity of the NIP2;1-GFP signal differs between the apical and basal end of the cells (Fig. III-7A). Some NIPs, such as the boric acid channel NIP5;1, have been documented to have polarized localization to the plasma membrane which aids in the directional flow of substrate across the root (Wang et al., 2017). To determine if a similar situation is apparent for NIP2;1, the relative distribution of NIP2;1-GFP fluorescence signal across the radial and longitudinal axes of multiple root cells was quantified and standardized to the FM4-64 plasma membrane signal by the method of (Wakuta et al., 2015; Wang et al., 2017). The analysis indicated that NIP2;1 is evenly distributed across the plasma membrane, and that the ratio of NIP2;1/FM4-64 staining did not differ significantly from unity (polarity indices of 1.17 for radial distribution and 0.97 for longitudinal distribution, Fig. III-8D), suggesting that NIP2;1 does not show polarized expression on selected surfaces of the plasma membrane.

As hypoxia proceeds, and more protein accumulates, there is stronger accumulation of NIP2;1-GFP signal internally, although plasma membrane localization is also still apparent (Fig. III-7B, 6 hr time point; Fig. III-9). Following 1 hr of reoxygenation, the predominant localization of NIP2;1 on the plasma membrane appears again (Fig. III-9). Overall, while the data show some changes in the degree of surface vs. internal localization of NIP2;1 during the hypoxia/recovery response, the protein shows strong localization to the plasma membrane at all phases of hypoxia and
early recovery, where it presumably mediates the transport of substrate from the cell into the apoplastic space.

**NIP2;1 participates in lactic acid efflux and media acidification during hypoxia**

Based on the specificity of NIP2;1 as a lactic acid channel from biochemical analyses (Choi and Roberts, 2007), its localization in part to the surface of root cells, and the observation that hypoxia triggers release of lactic acid/lactate from roots into the external media (Xia and Saglio, 1992; Dolferus et al., 2008; Engqvist et al., 2015), it is hypothesized that NIP2;1 may participate in the excretion of lactic acid during the low oxygen stress response *in planta*. In support of this, in comparison to wild type seedlings, *nip2;1* seedlings accumulated significantly higher levels of lactic acid/lactate level within root tissues during a hypoxia time course (Fig. III-10B), consistent with the inability to efflux this fermentation end product.

To test this hypothesis further, the media pH and rate of lactic acid/lactate efflux from the roots of WT and *nip2;1* seedlings challenged with hypoxia were examined. To compare the hypoxia-induced acidification of the external medium, 10 day-old WT and *nip2;1* seedlings were subjected to argon gas treatment on media containing the pH sensitive dye, bromocresol purple (Fig. III-10A). Upon transfer from normoxic to hypoxic conditions, WT seedlings showed significant yellowing of the media surrounding the roots, suggesting the decrease in the pH and acidification of the media, while *nip2;1* plants show no difference in bromocresol purple staining between normoxic and hypoxic conditions (Fig. III-10A).

Further comparison of hypoxia-challenged WT and *nip2;1* seedlings reveal that WT roots show a significantly higher rate of lactic acid/lactate release into the medium
Figure III-10 Media acidification and lactic acid efflux in hypoxia-challenged nip2;1. A. Ten day-old nip2;1 and wild type (Col-0) seedlings were transferred to pH indicator plates containing bromocresol purple, and were subjected to 8 h treatment of hypoxia induced by argon gas (Argon). Air indicates a normoxic control. A color change from purple to yellow indicates a decrease in the pH of the environment. Scale bar = 1 cm. B. Quantitation of lactic acid/lactate in seedling roots based on enzymatic analysis during hypoxia. Values are the average of three biological replicates at each time point, with the error bars showing the SD. C. The rate of media lactic acid/lactate release from the seedlings of indicated lines during a hypoxia time course. Each value represents an individual determination from samples taken over an 8 hr hypoxia time course as described in the Materials and Methods with the error bars showing the SD. Asterisks in panels B and C indicate statistically significant differences based on One way ANOVA analysis with multiple comparisons (panel B) or a paired Student’s t-test analysis (panel C).
compared with the roots of *nip2;1* plants (Fig. III-10C). Assay of a *NIP2;1-GFP* complementation line shows that the rate of lactic acid/lactate efflux is restored to wild type levels (Fig. III-10C), verifying that the reduction in lactic acid efflux in *nip2;1* plants is due to the loss of NIP2;1 protein. The results, combined with previous functional studies (Choi and Roberts, 2007), suggest that NIP2;1 participates in lactic acid transport, homeostasis, and efflux from roots during low oxygen stress.

**The loss of *NIP2;1* function affects the expression of pyruvate and lactate metabolic enzymes**

Anaerobic metabolism of pyruvate during oxygen limitation in plants occurs through three conserved pathways: lactic acid fermentation, ethanolic fermentation and alanine synthesis (Fig. III-11A). While all three pathways use pyruvate as a substrate, lactic acid and ethanolic fermentation regenerate NAD⁺, whereas alanine synthesis serves as a mechanism to store nitrogen and carbon for reoxygenation (Sato et al., 2002; Ricoult et al., 2005). Genes that encode enzymes in these pathways (such as *ADH, PDC, LDH*, and *AlaAT*) are among the “core hypoxia response” genes that are induced in Arabidopsis roots during hypoxia (Mustroph et al., 2009; Lee et al., 2011; Mustroph et al., 2014). Conversely, L-lactate produced via LDH is proposed to be converted back to pyruvate in peroxisomes (Fig. III-11A) by the root-specific glycolate oxidase3 (GOX3) enzyme (Engqvist et al., 2015). Unlike other members of this enzyme family that participate in the metabolism of glycolate, GOX3 is specific for L-lactate and utilizes oxygen as an electron acceptor to oxidize lactate, producing hydrogen peroxide and pyruvate as end products (Engqvist et al., 2015). GOX3 is not a hypoxia-induced transcript and is rather proposed to regulate the concentration of lactate in a coordinate
Figure III-11 Effect of nip2;1 mutation on transcripts of pyruvate metabolism enzymes. A. Scheme showing the principal pathways of pyruvate and lactate metabolism during fermentation. B. Q-PCR analysis of the indicated transcripts in the roots of ~50 seven-day-old wild type seedlings grown under normoxic conditions (black bars) or in response to 4 hr of argon-induced hypoxia (gray bars). The data are normalized to the transcript levels of GOX3 under normoxic conditions. C. Comparison of the hypoxia-induced changes of selected transcripts from the roots of wild type and nip2;1 seedlings. The data are normalized to the expression levels of the indicated transcript under normoxic conditions. Error bars in panels B and C show the SD, and asterisks indicate statistically significant differences based on an unpaired Student's t-test analysis. ns indicates p-values >0.05.
fashion with LDH under aerobic conditions (Engqvist et al., 2015; Maurino and Engqvist, 2015).

Q-PCR analysis shows that hypoxia response transcripts (ADH1, PDC1, LDH, and AlaAT1) show induction in both WT and nip2;1 mutant root tissues during 4 hr of argon gas (Fig. III-11B and C). However, closer analysis of WT and nip2;1 roots show different levels of selective transcripts. While LDH and PDC1 transcript levels show no statistical differences between WT and nip2;1 roots, ADH1 and AlaAT1 are significantly elevated in nip2;1 compared to WT roots (Fig. III-11C). Conversely, GOX3, which is expressed at the same level under normal and hypoxic conditions in WT roots, is substantially reduced in hypoxic nip2;1 roots. Overall, the data suggest that the alterations in lactic acid homeostasis within the nip2;1 mutant affect the expression of enzymes in pyruvate and lactate metabolic pathways, with transcripts encoding fermentation enzymes in ethanol (ADH1) and alanine (AlaAT1) producing pathways enhanced, whereas the transcript that encodes the lactic acid metabolizing enzyme GOX3 is suppressed.

Summary and Conclusions Chapter III

NIP2;1 is a core hypoxia gene that encodes a member of the “Nodulin 26-like Intrinsic Protein” (NIP) subgroup of the aquaporin superfamily of membrane channel proteins. Under normal growth, NIP2;1 expression is limited to the “anoxia core” region of the root stele, but shows substantial induction in response to low oxygen stress (as high as 1000-fold by 2-4 hr of hypoxia challenge), and accumulates in all root tissues. During hypoxia, NIP2;1-GFP, accumulates predominantly on the plasma membrane by 2 hr, is distributed between the plasma membrane and internal membranes during sustained
hypoxia, and remains elevated in root tissues through 4 hrs of reoxygenation recovery. T-DNA insertional mutant nip2;1 plants show elevation of lactic acid within root tissues in response to hypoxia challenge, and reduced efflux of lactic acid as well as reduced acidification of the external medium compared to wild type plants. Together with previous biochemical evidence demonstrating that NIP2;1 has lactic acid channel activity, the present work supports the hypothesis that the protein facilitates the release of cellular lactic acid from the cytosol to the apoplastic space as part of an eventual efflux to the rhizosphere to prevent lactic acid toxicity. In support of this, nip2;1 plants show poorer survival to argon-induced hypoxia stress. Nip2;1 mutant plants also show elevated expression of the ethanolic fermentation transcript ADH1 and the core hypoxia-induced transcript AlaAT1, as well as reduced expression of the lactic acid metabolic enzyme GOX3, suggesting that the altered efflux of lactic acid through NIP2;1 regulates other pyruvate and lactate metabolism pathways.
CHAPTER IV - Structural Basis for water and metalloid specificity of NIP I and NIP II channels
Background and Rationale

The evolution of land plants led to an expansion and diversification of the aquaporin gene family, the acquisition of modified pore structures compared to canonical aquaporin channels, and additional transport functions beyond water transport (Wallace and Roberts, 2004; Ludewig and Dynowski, 2009). “Nodulin 26-like intrinsic proteins” (NIPs) represent a land plant-specific subfamily of aquaporin-like proteins that are phylogenetically and structurally organized into three broad families, NIP I, II and III, that have distinct Ar/R signatures and transport selectivities (Wallace and Roberts, 2004; Mitani et al., 2008; Rougé and Barre, 2008; Ludewig and Dynowski, 2009; Liu and Zhu, 2010; Roberts and Routray, 2017). Transport activities for NIPs include classical substrates such as water and glycerol, as well as ammonia, $\text{H}_2\text{O}_2$, formamide, urea, lactic acid, and various metalloid hydroxides (Dean et al., 1999; Weig and Jakob, 2000; Ma et al., 2006; Takano et al., 2006; Tanaka et al., 2008; Kamiya et al., 2009; Kamiya and Fujiwara, 2009; Pommerrenig et al., 2015; Xu et al., 2015; Roberts and Routray, 2017). NIP III proteins are widely distributed among the Graminaceae, and overwhelming biochemical and genetic evidence show that they principally function as facilitators of silicic acid ($\text{Si}[\text{OH}]_4$) that promotes optimal growth and development, as well as resistance to abiotic and biotic stress (Epstein, 1999; Ma et al., 2006; Ma and Yamaji, 2006; Chiba et al., 2009). NIP II proteins are boric acid channels that play a critical role in facilitating uptake of this micronutrient under B limiting conditions (Takano et al., 2006; Tanaka et al., 2008; Routray et al., 2018). In contrast, the biological and transport roles of NIP I are less defined, although myriad functions have been postulated, ranging from metalloid permeability (Kamiya et al.
to aquaporin and ammoniaporin transport during symbiosis (Niemietz and Tyerman, 2000; Hwang et al., 2010; Masalkar et al., 2010), to lactic acid efflux during flooding stress (Choi and Roberts, 2007; Beamer et al., 2021). Each pore subfamily possesses signature amino acid compositions within their ar/R selectivity filters that are postulated determine to these disparate substrate specificities and biological functions (Froger et al., 1998; Savage et al., 2003; Hove and Bhave, 2011). Insight into the unique pore features and potential regulatory features of the NIP family has recently emerged with the solution of two atomic resolution crystal structures of an open and closed conformation of the OsNIP2;1/Lsi1 silicic acid channel (a NIP III protein) from rice (Saitoh et al., 2021; van den Berg et al., 2021). Structural information for NIP I and II proteins is not yet available.

The objective of this study was to examine the solute selectivity and the function of signature ar/R residues of NIP Type I and Type II proteins from Arabidopsis by evaluating the transport of three key NIP solutes – water, boric acid and arsenite. Further, the role of H2 residue of the ar/R in providing boric acid selectivity of NIP proteins was tested by using in planta complementation approaches with various constructs. Finally, structural modeling and molecular dynamics were used to evaluate representative NIP I and II protein models based on the properties of the NIP III Lsi1/OsNIP2;1 silicic acid channel.

The NIP fold and selectivity filter deviates from water-selective aquaporins

Insight into the NIP structure and how it has diverged from classical water specific aquaporins has come from the recent solution of the rice silicic acid transporter.
Lsi1 structure at atomic resolution (van den Berg et al., 2021; Saitoh et al., 2021). Lsi1 is a NIP III protein that contains the signature sequence of silicic acid channels with a wide and hydrophilic ar/R with a consensus sequence of G-S-G-R [H2, H5, LE1, LE2] (Fig. IV-1 and Fig IV-2). This property leads to three essential features that distinguish NIP III from other aquaporin and aquaporin-like proteins. First, the invariant serine at H5 is unique to NIP III proteins and provides a hydrogen bond donor to transported substrate. Second, the open nature of the NIP III ar/R allows the insertion of a fifth residue, a conserved threonine, from helix 1 into the selectivity filter, providing an additional ligand for transported substrates and bound water (Fig. IV-2A). Third, unlike water specific aquaporins, which narrow to the diameter of a single water molecule at the ar/R allowing only single file water transport, Lsi1 is highly hydrated (Fig. IV-1) with three waters within the ar/R selectivity filter (blue in Fig. IV-2), as well as two additional bound waters (green in Fig. IV-2). The NIP III/Si(OH)4 model and MD simulations (Saitoh et al., 2021; van den Berg et al., 2021) show how these three features account for silicic acid transport by providing steric accommodation and multiple hydrogen bond contacts with the serine side chain, as well as with the two bound waters as Si(OH)4 traverses the pore (Fig. IV-2).

To investigate the comparative pore properties of NIP I and NIP II proteins in light of the features of the high resolution Lsi1 structure, structural models were generated from representative Arabidopsis NIPs (NIP1;1 and NIP4;1 representing NIP I channels, and NIP5;1 and 6;1 representing NIP II channels). Initial structures were obtained from the AlphaFold Protein Structure Database (Jumper et al., 2021), and were refined by short MD simulations. On each simulation trajectory, hierarchal clustering was
Figure IV-1 Alignment and Alpha-fold modeling of Arabidopsis NIP I and II proteins. A. Phylogenetic tree of the seven Arabidopsis thaliana NIP proteins as well as their pore subtype classification. B. Sideview of Ls1 monomer (Saitoh et. al. 2021, PDB 7CJS) with pore water molecules (light blue and pink spheres) and selectivity filter Arg shown. C. Alignment of Arabidopsis NIP1;1, NIP2;1, NIP4;1 & NIP6;1 with Oryza sativa NIP2;1 (Ls1). The five selectivity residues of Ls1 are denoted by H1, H2, H5, LE, and R (conserved arg at LE3). The residues responsible for the intracellular gating in the Ls1 crystal structure (van der Berg et. al. 2021) are denoted by Loop B (G1) and Loop D (G2) boxes. Helices are colored based on the Ls1 structure (panel B). D. Ribbon structures of the indicated alpha-fold model structures with the same color scheme and the position of the ar/R arg side chain highlighted. The predicted position of a disulfide bond between loop C and B in NIP1;1 is indicated. E. Superimposed structures of NIP6;1 homology model obtained from AlphaFold (red) with the open (dark green, PDB 7CJS) and closed structures (light green, PDB 7N4L) of Ls1.
Figure IV-2 Representative models of the selectivity filter regions of NIP pore subtypes. A. Left, diagram of the ar/R selectivity filter for classical water specific aquaporins viewed down the pore axis from the extracellular vestibule with the ar/R nomenclature of Wallace and Roberts (H1: helix 1, H2: helix 2, H5, helix 5, LE1: E loop residue, R, conserved arginine). Right, Aqy1, the selectivity filter of yeast aquaporin Aqy1. B. Right, Lsi1, the rice NIP III silicic acid permease viewed from the same orientation with waters shown as aqua (transported waters) or green (tightly bound waters) spheres. The waters are numbered based on the nomenclature of Saitoh et al., 2021. Right, Lsi1 with silicic acid bound at the ar/R selectivity filter generated by QM/MM (Saitoh et al., 2021). C. The ar/R selectivity filters of the Lsi1 (NIP III) and the NIP6;1 (NIP II) and NIP4;1 (NIP I) homology models (this study) shown with the predicted positions of the Lsi1 ar/R bound waters. The diagram below each structure indicates the consensus sequence for the indicated NIP I, II, and III pore subtype based on phylogenetic analyses across plant species (Supplemental Fig. S1), Roberts and Routray, 2017.)
performed to generate four representative structures which take into account both local and global motions of the protein. Structural models in each cluster show the conservation of the core hourglass topology of the aquaporin fold including the conservation of the six transmembrane alpha-helices, two NPA half helices, and the position of the interhelical loops (Fig IV-1 C and D). The overall backbone conformation within each of the four clusters is small with RMSD ranging from 1 to 1.5 Å. The predicted backbone conformations of the models superimpose well with the recently solved 1.8 Å crystal structure of the rice Lsi1 NIP III silicic acid channel (Saitoh et al., 2021), with the highest similarity observed with the NIP4;1 and NIP6;1 models (Table IV-1, Figure IV-2A). These were analyzed further as examples of NIP I and II protein pores.

Examination of the NIP II structure, as illustrated by the NIP6;1 model, shows a more hydrophobic ar/R selectivity region compared to Lsi1 (Fig IV-2C). Unlike the invariant serine residue found at the H5 position in NIP III channels, NIP6;1 possess a branched isoleucine residue at this position. The invariant glycine residues found at H2 and LE1 in NIP III are occupied by alanines in NIP6;1. As a result the NIP6;1 pore, similar to Lsi1, is still wider than classical aquaporins (Fig. IV-2B). This allows access of the H1 residue (glycine in NIP6;1) to the selectivity filter, and the residues involved in the binding and positioning of bound waters (Cys 39 and Ala 110 in NIP6;1, Fig. IV-3A) involved in metalloid transport in Lsi1 are conserved in the NIP 6;1 model, and the predicted NIP6;1 pore model would accommodate this feature of the NIP III structure (Fig. IV-2C and Fig. IV-4). The potential positions of the silicic acid hydroxide interaction site in Lsi1 were determined based on the positions of the selectivity filter
Table IV-1 Pairwise RMSD values of backbone C-alpha atoms in superposed structures for the Lsi1 open (Saitoh et al., 2021, pdb 7CJS), Lsi closed (van den Berg et al., 2021 pdb 7NL4) and the AlphaFold NIP homology models.

<table>
<thead>
<tr>
<th>Chains</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>1:7CJS</td>
<td></td>
<td>1.89</td>
<td>2.24</td>
<td>1.64</td>
<td>2.11</td>
<td>1.64</td>
</tr>
<tr>
<td>2:7NL4</td>
<td>1.89</td>
<td></td>
<td>2.65</td>
<td>2.22</td>
<td>2.58</td>
<td>2.38</td>
</tr>
<tr>
<td>3:NIP1_1</td>
<td>2.24</td>
<td>2.65</td>
<td></td>
<td>1.88</td>
<td>2.59</td>
<td>2.15</td>
</tr>
<tr>
<td>4:NIP4_1</td>
<td>1.64</td>
<td>2.22</td>
<td>1.88</td>
<td></td>
<td>2.14</td>
<td>1.69</td>
</tr>
<tr>
<td>5:NIP5_1</td>
<td>2.11</td>
<td>2.58</td>
<td>2.59</td>
<td>2.14</td>
<td></td>
<td>1.83</td>
</tr>
<tr>
<td>6:NIP6_1</td>
<td>1.64</td>
<td>2.38</td>
<td>2.15</td>
<td>1.69</td>
<td>1.83</td>
<td></td>
</tr>
</tbody>
</table>
Figure IV-3 Effect of H2 tryptophan substitution in NIP I-like pore models. A. The position of immobile waters 3 and 9 in the Lsi1 structure and the NIP6;1 and NIP6;1 A119W viewed from the membrane plane (upper panel) and extracellular vestibule (lower panel). B. Comparison of NIP6;1 (NIP II) and NIP4;1 (NIP1) selectivity filter models (viewed perpendicular to the pore axis from the extracellular vestibule with waters bound as in Fig. 1C.)
Figure IV-4 Boric Acid bound model of NIP6;1 A. Left, the Lsi1 structure with Si(OH)$_4$ bound at the selectivity filter taken from the QM/MM model of Saitoh et al. (2021) with molecular surfaces of substrate and ar/R residues shown. Right, corresponding model for NIP6;1 based on boric acid docking to the ar/R selectivity filter. The position of the five ar/R residues and the immobile (green) and selectivity filter bound (blue) waters are shown. B. Left, hydrogen bond contacts between Lsi1/silicic acid (Saitoh et al., 2021), and the position of comparable residues in the NIP6;1/boric acid model.
associated water molecules (Saitoh et al., 2021; Fig. IV-2B). By using a similar approach with the NIP6;1 model, a potential interaction site for boric acid was investigated. Compared to the tetrahedral silicic acid molecule, boric acid is a smaller molecule with three hydroxyl groups oriented in a trigonal and planar arrangement. A model of boric acid docked into the NIP6;1 selectivity filter is shown in Fig. IV-4. Similar to the Lsi1 Si model, potential hydrogen bond interactions with some selectivity filter residues, notably the conserved arginine and H1 (Gly 96) and LE1 (Ala 247) positions, are retained in NIP6;1. However, the replacement of the conserved Lsi1 Ser with Ile 238 in NIP6;1 at the H5 position would cause a steric restriction of the pore as well as a loss of hydrogen bond donor capability, consistent with the critical role of serine in providing silicic acid specificity (Mitani-Ueno et al., 2011; Saitoh et al., 2021).

The NIP4;1 model illustrates that NIP I protein pores are even more constricted and hydrophobic (Fig. IV-2C and Fig. IV-3). Like all NIP I pores (Wallace and Roberts, 2004; Roberts and Routray, 2017), NIP4;1 possesses a conserved tryptophan at the H2 position compared to the invariant glycine in Lsi1. To accommodate this bulkier residue, the H1 position is an invariant glycine in NIP I pores. The result is a smaller ar/R with four residues instead of five and the predicted loss in the ability to accommodate the bound waters (water 3 and 9) and one of transported waters (water 2) from the Lsi1 channel (Fig. IV-2B). This is predicted to result in the orientation of remaining selectivity filter waters in a single file arrangement (Fig IV-2C). In this regard, NIP4;1 resembles microbial and animal aquaglyceroporins that have similar amphipathic ar/R selectivity filter properties with a narrow ar/R constriction that permits single file transport of water and solutes.
**NIP I and II and corresponding H2 mutants show distinct preference for metalloid uptake in oocytes**

One prediction from modeling results of NIP6;1 and NIP4;1 is that the nature of the H2 residue determines the size and characteristics of the selectivity filter of NIP I and II proteins. Consistent with this prediction, early results with a series of test amide substrates show that NIP I proteins (e.g., soybean nodulin 26) exclude test solutes that permeate NIP II proteins (Wallace and Roberts, 2005). To determine whether this residue affects the comparative permeability of physiological metalloid hydroxide nutrients (B) and toxins (As), the boric acid and arsenous acid permeability of NIP I and II proteins, and their corresponding H2 mutants, were evaluated in *Xenopus* oocytes by direct uptake analyses and ICP-MS quantitation.

For this purpose, two NIP I proteins (NIP4;1 and NIP1;1) and two NIP II proteins (NIP6;1 and NIP5;1) were expressed as GFP fusions in *Xenopus* oocytes to allow analysis of expression levels by Western blot (Fig. IV-5). *Xenopus* oocytes that express NIP II and NIP I proteins showed 20-fold greater As(OH)$_3$ uptake rates compared to negative control oocytes, with no statistical differences in the permeability properties of NIPs from either pore subclass (Fig. IV-6A). Further, substitution of an alanine for tryptophan in the NIP I proteins, which would be predicted to increase the aperture of the ar/R, results in no significant differences As(OH)$_3$ uptake compared to wild type NIP I controls. Conversely, the substitution of a tryptophan for alanine in the NIP6;1 ar/R, which would be predicted to restrict the ar/R diameter, also showed no effect on As(OH)$_3$ uptake rate (Fig. IV-6A). However, NIP5;1A117W showed As(OH)$_3$ uptake rates that were indistinguishable from negative control oocytes, reflecting the loss of
Figure IV-5 NIP-GFP expression in *Xenopus oocytes* and *Arabidopsis thaliana* transgenic plants. A. Anti-GFP Western blot showing NIP-GFP protein in oocytes injected with the indicated cRNAs (upper panel). Bottom panel, Coomassie blue stained loading control gel. DEPC represents negative control oocytes injected with sterile DEPC instead of cRNA. ApAQP2-GFP represents oocytes injected with B. Anti-GFP Western blot from extracts of 10 d old Arabidopsis seedlings (upper panel). Bottom panel, Coomassie blue stained loading control gel. C. Leaves from seven-day old transgenic *nip5;1-1* Arabidopsis seedlings complemented with the indicated construct were dissected, stained with FM4-64, and were imaged by confocal microscopy. GFP (first panel), FM4-64 (second panel), merged images (third panel) and DIC (fourth panel). Scale bar = 20 µm.
Figure IV-6 Comparison of the metalloid hydroxide permeabilities of NIP I and II wild type proteins and H2 mutants. A. Structural models of various metalloid hydroxides that permeate NIP channels illustrating their disparate geometries and sizes. B. Oocytes expressing GFP fusions of the indicated NIP or NIP mutants were incubated with 1 mM sodium arsenite for 30 min and As uptake was determined by ICP-MS. Values represent means ± SD (n = 4 pools of eight oocytes per pool, error bars indicate SD). DEPC represents negative control oocytes. C. Oocytes injected with the indicated NIP construct were subjected to boric acid (2 mM) uptake analyses and B content was determined by ICP-MS. D. Standardization of B uptake rates as a fraction of As uptake under the same conditions. The background (DEPC) oocyte uptake rate of B or As was subtracted, and the NIP-mediated B uptake is expressed as a fraction of the As uptake (n = 7 pools of 8 oocytes per pool, error bars show SD). Statistical significance in B, C, and D was determined by One-way ANOVA with a different letter indicating statistical significance (p<0.05). In panel C and D, ****, p<0.0001
transport function. Subsequent analysis (Fig. IV-7) showed that NIP5;1A117W is impermeable to all substrates tested, and genetically is unable to complement nip5;1-1 boron sensitive mutant in Arabidopsis screens. Therefore, for reasons that are not apparent, the NIP5;1A117W mutation produces an inactive channel and was not pursued further. Nevertheless, the data suggest that NIP I and II proteins have indistinguishable permeability for arsenous acid.

Next, NIP4;1 (wild type and alanine to tryptophan mutant) and NIP6;1 (wild type and tryptophan to alanine mutant) were analyzed for their permeability to boric acid. As shown in Fig. IV-6A, boric acid and arsenous acid are trihydroxylated Lewis acids that exist in a neutral form at physiological pH. They have a similar molecular volume but differ substantially in geometry and electronic properties. Boric acid is a planar trigonal molecule, whereas As(OH)3 is a trigonal pyramid with more restricted bond angles (Fig. IV-6A).

Since NIPs show indistinguishable arsenous acid permeability, for comparative purposes the boric acid permeability was normalized to the arsenous acid uptake. This allows the elimination of nonspecific background uptake of the two substrates as well as elimination of any differences related to slight variations in protein expression. Wild type NIP6;1 shows robust transport of both metalloids with a preference for B over As (fractional B/As permeability of 1.32) (Fig. IV-6 D). In contrast, NIP4;1 was a poorer boric acid transporter with a 4-fold lower permeability (B/As = 0.32) (Fig. IV-6 D). In contrast, the NIP4;1 W82A mutant has substantially higher boric acid permeability (Fig. 3C), with a B/As permeability (1.28) statistically indistinguishable from the NIP6;1 boric
Figure IV-7 The H2 mutant protein NIP5;1 A177W produces an inactive protein. 
A. Comparison of the boric acid and water permeability of NIP5;1 A177W to positive control (NIP5;1 and the aphid aquaporin ApAQP2, respectively) and negative control (DEPC oocytes). Statistical comparison was done by One Way ANOVA analysis. B. Complementation analysis of the *nip5;1-1* B sensitive phenotype with wild type and mutant NIP5;1-GFP constructs. Plants represent 29-day old wild type (Col-0), *nip5;1-1* or *nip5;1-1* complemented with the indicated *NIP5;1-GFP* constructs driven by the 35S promoter were grown under low (1 µM) or normal (30 µM) boric acid conditions. Scale bars are 2 cm.
acid channel. Conversely, the substitution of the H2 alanine in NIP6;1 with a tryptophan residue (NIP6;1 A119W) results in a selective reduction in boric acid permeability (B/As = 0.69, Fig. IV-6 D).

Overall, this shows that the presence of a smaller alanine residue at the H2 position enhances the selective boric acid permeability in the NIP6;1 and NIP4;1 channels, and that this property of NIP II channels could underpin their established function as boric acid permeases in higher plants (Mitani-Ueno et al., 2011; Pommerrenig et al., 2020). Previous work with the NIP I protein soybean nodulin 26 (Roberts and Routray, 2017) also suggests poor permeability to boric acid and its function as a ammonia-aquaporin has been postulated as its principal function. However, analysis of NIP1;1 shows that poor boric acid permeability is not a property shared by all NIP I pores (Fig. IV-8). Unlike NIP4;1, NIP 1;1 shows equally high permeability to both boric acid and arsenous acid, and the substitution of an alanine for tryptophan (NIP1;1 W74A) does not enhance boric acid permeability (Fig. IV-8). Thus, other unidentified pore structural determinants in addition to the H2 selectivity filter residue also contribute to metalloid preference and selectivity among selected NIP I channels.

**Overexpression of NIP I proteins partially rescue the B deficiency phenotype of nip5;1 mutant seedlings**

To evaluate further the ability of NIPI proteins as transporters of boric acid in plants, their ability to complement the low B sensitivity phenotype of the *nip5;1-1* T-DNA lines (Takano et al., 2006) was assessed. Transgenic plants were generated that express the C-terminally tagged GFP fusions of various Arabidopsis NIPs and their H2 mutants.
Figure IV-8 Comparison of Boric acid and arsenous acid permeabilities of four A. thaliana NIP proteins. A. The indicated NIP proteins were expressed in as GFP fusion proteins in Xenopus oocytes and were assayed for boric acid and arsenous acid uptake. The background (DEPC) oocyte uptake rate of B or As was subtracted, and the NIP-mediated B uptake is expressed as a fraction of the As uptake (n = 40 to 56 oocytes per construct). Statistical significance was determined by One-way ANOVA with a different letter indicating statistical significance (p<0.001). B. Comparison of boric acid permeabilities of NIP1;1 and the H2 mutant NIP1;1 W94A.
under the control of the 35S promoter were generated in the \textit{nip5;1-1} knockout background ($35S_{pro}:NIP$ transgenic plants). Western blot analysis of the seedlings of T2 transgenic lines showed the robust expression of each NIP-GFP fusion protein (Fig. IV-5B), and confocal microscopy showed co-localization of the GFP signal with the fluorescent plasma membrane marker FM4-64 (Fig. IV-5C).

To determine whether the NIP1 transgenes complemented the \textit{nip5;1} phenotype, growth under sufficient (50 μM) and limiting boric acid (1 μM) was compared by the approach of Fujiwara et al. (1992). Under sufficient boric acid conditions all plants, regardless of genotype showed normal growth (Fig. IV-9 and Fig. IV-10A), and were indistinguishable from wild type Col-0 plants. This is consistent with previous observations of \textit{nip5;1-1} plants, and further indicates that overexpression of the NIP transgenes does not affect plant growth under these conditions. Under limiting B conditions, \textit{nip5;1-1} plants showed defective growth that was complemented to different degrees by the various NIP I protein transgenes (Fig. IV-10 and Fig. IV-11). Seven-day old seedlings from complementation lines expressing wild type \textit{NIP1:1-GFP}, \textit{NIP1;1 W94A} and wild type \textit{NIP4;1-GFP} constructs showed significantly longer root lengths and higher seedling fresh weights compared to \textit{nip5;1-1} control plants. However, these constructs did not complement growth defects as well as the \textit{NIP5;1} complementation line and showed significantly reduced root length and seedling weights compared to Col-0 controls (Fig. IV-11). \textit{NIP1:1-GFP} was the most effective in complementing \textit{nip5;1-1} phenotype, consistent with its enhanced boric acid permeability. Surprisingly, the \textit{NIP4;1W94A} construct, which showed high boric acid permeability in Xenopus
Figure IV-9 Comparison of growth of wild type, *nip5;1-1*, and complementation lines under sufficient boric acid conditions. Seedling fresh weight (A) and primary root lengths (B) of 7-day old seedlings grown under non-limiting boric acid conditions (50 μM) conditions (mean and SD shown as a scatter plot of the data).
Figure IV-10  Complementation of nip5;1 growth and B uptake phenotypes by NIP1 wild type and H2 ar/R mutant constructs. A. Representative plants (Col-0, nip5;1-1 and nip5;1-1 complemented with the indicated NIP-GFP fusion transgene grown for 40 days under limiting (low B) or sufficient (high B) boric acid conditions. Scale bar = 2 cm. B. Fresh weight of rosette leaves (each data point represents a separate determination of a pool of dissected leaves) collected from plants grown under low B conditions as in Panel A. C. Plants were grown under low B conditions, and B uptake analysis was done as described in the Materials and Methods. For panels B and C, statistical analysis was performed by One-way ANOVA with different letters indicating statistical significance (p<0.05).
Figure IV-11 Effect of NIP I protein overexpression on the B sensitive phenotype of nip5;1. A. Representative seven-day old Arabidopsis seedlings growth on limiting (1 μM) boric acid media. Col-0, wild type control; nip5;1-1, nip5;1-1 plants without complementation; complementation, nip5;1-1 plants complemented with the indicated constructs (complementation line 1). B. Comparison of the primary root lengths of seven-day old Arabidopsis plants cultured as in panel A as measured in ImageJ software and represented as a scatter plot (mean and SD). Each datapoint represents a single seedling. The results of experiments from two representative complementation lines are shown. C. Seedlings grown as in panel A were collected in pools and fresh weights were recorded. The mean weight of the Col-0 control value was used to normalize the seedling weight values (each data point is a single pool). Statistical significance (P<0.05) was determined by One-way ANOVA analysis with different letters indicating statistically significant difference.
The partial ability of NIP1 proteins to complement \textit{nip5;1} growth and B-uptake deficiencies is supported by the analysis of plants the bolting and reproductive phase (Fig. IV-10). Complementation lines with \textit{NIP1;1} and \textit{NIP4;1} constructs showed restored rosette leaf growth compared to \textit{nip5;1-1} controls (Fig. IV-10B), but showed a delay in bolting compared to Col-0 plants (Fig. IV-10A). To determine whether the NIP complementation lines show enhanced boron accumulation, B uptake analysis was done by the method of Takano et al. (2006). While all B uptake values were significantly below Col-0 positive controls, all type 1 NIP complementation lines showed significantly enhanced B uptake and incorporation into rosette leaves compared to \textit{nip5;1-1} negative controls (Fig. IV-10C). In general, both NIP4;1 wild type and H2 mutant were less effective than the corresponding NIP1;1 constructs, both in restoring growth as well as in promoting the B uptake (Fig. IV-10B and C).

**The loss of aquaporin activity in NIP II pores: Models for pore gating in a water tight channel**

The solution of the NIP III Lsi1 structure reveals that the pore is unusually hydrophilic and hydrated compared to other members of the aquaporin family (Saitoh et al., 2021), with 16 water molecules occupying each monomeric channel (Fig. IV-1B). Consistent with this observation, NIP III possesses strong aquaporin activity in addition to its ability to transport silicic acid and other metalloids (citations). NIP I proteins also have aquaporin activity (Rivers et al., 1997; Dean et al., 1999), but the NIP II proteins are largely refractory to water transport (Wallace and Roberts, 2005; Takano et al., 2006).
2006; Tanaka et al., 2007; Li et al., 2011; Katsuhara et al., 2014; Routray et al., 2017; Diehn et al., 2019). Consistent with previous observations, all NIP I proteins (NIP1;1, NIP 4;1, and the canonical NIP I aqua-ammoniaporin protein soybean nodulin 26) show aquaporin activity based on oocyte swelling assays in hypotonic media whereas both NIP II proteins lacked detectable aquaporin activity and were statistically indistinguishable from negative control oocytes (Fig. IV-12). This difference between NIP I and NIP II proteins is controlled by the nature of the H2 residue in the selectivity filter. Substitution of the canonical H2 tryptophan in both NIP1;1 and 4;1 with the NIP II-like alanine residue (NIP1;1W94A, NIP4;1W82A, nodulin 26 W77A) significantly diminishes or abolishes water permeability (Fig. IV-12). Conversely, substitution of the tryptophan for alanine in NIP6;1 results in strong gain of function aquaporin activity (Fig. IV-12D).

The NIP II pore NIP6;1 model shows a wide selectivity filter that is able to accommodate the three transported waters as well as two bound waters found in the Lsi1 structure (Fig. IV-2C). Further, analysis of the pore diameter in the NIP 6;1 model along the transmembrane z-coordinate reveals a typical aquaporin-like pore architecture with a width that should be able to accommodate water movement. Why then is NIP6;1 unable to transport water unlike NIP I and NIP III proteins? One potential clue comes form a comparison of the two recently solved structures of Lsi1 that represent both an open conformation (pore filled with waters, Saitoh et al. 2021) and a closed structure (van den Berg et al., 2021), (Fig. IV-13A). The closed structure was solved in complex with Cd^{2+} which chelates histidine residue 120 (loop B) and histidine 46 between tetramers in the crystal lattice. The nature of the closed structure is a reorientation of
Figure IV-12 Comparison of the aquaporin activities of NIP I and II wild type proteins and H2 mutants. A. and B. Comparison of the osmotic water permeability of oocytes injected with cRNA encoding flag-tagged wild type Arabidopsis NIP6;1 and NIP4;1 or soybean nodule 26, or corresponding H2 ar/R mutants (NIP6;1 A119W, NIP4;1 W82A, Nod26W77A). The osmotic water permeability (Pₜ) was determined from the rate of oocyte swelling upon incubation in a hypoosmotic Ringer’s solution (Wallace and Roberts, 2005). Values represent means ± SD (n = 8-24 oocytes for each sample). DEPC, negative control oocytes injected with sterile DEPC water. ApAQP2, positive control oocytes injected with the flag-tagged Acrithosiphon pisum AQP2 aquaporin. C and D. Comparison of osmotic water permeability of NIP I proteins (NIP1;1 and NIP4;1) or NIP II proteins (NIP5;1 and 6;1) and their corresponding H2 mutants. Oocytes (n=16-20) were injected with cRNA encoding the indicated protein with an in-frame C-terminal GFP tag. Statistical significance for each data set was determined by One-way ANOVA with a different letter indicating statistical significance (p<0.05). Asterisks indicate statistical significance between wild type and H2 mutant pairs for each NIP subtype determined by pairwise t-test (*, p<0.05; **, p<0.01; ***, p<0.001; ****, p<0.0001).
Figure IV-13 Conserved gating loop residues in NIP structures A. The Lsi1 silicic acid channel structure shown in the open conformation with 16 pore water molecules (Saitoh et al., 2021, pdb 7CJS) and in the closed conformation obtained with Cd$^{2+}$ (van den Berg et al., 2021 pdb 7NL4). B. View of the closed structure from the cytosolic vestibule with showing the interactions between arginine and aspartate groups that stabilize the closed state (van den Berg et al., 2021). C. Top, Models for the open and closed states for NIP6;1. Bottom Plot of the calculated pore diameter along the Z-pore coordinate of the two NIP6;1 models shown in panel C. The data were obtained by using MoleOnline 2.5 (Pravda et al., 2018) in the Channels mode. The position of the conserved NPA motifs that define the center of the pore as well as the ar/R selectivity filter region are indicated.
loop D so that it forms electrostatic/hydrogen bond interactions between conserved R189 and D187 residues in loop D and a R119 residue in loop B (Fig. IV-13B). MD simulations showed spontaneous reorientation of loop D leading to closed and open conformations that are occupied during the MD trajectory (van den Berg et al., 2021).

Examination of the sequences of NIP I and NIP II proteins (Fig. IV-1C) shows high conservation of the loop D and loop B residues involved in Lsi1 gating. A comparison of NIP6;1 models generated based on the open and closed structures shows that protein could occupy an open or closed state stabilized by similar interactions observed in the closed Lsi1 structure. Such a structure would effectively restrict the pore diameter to less than that of a water molecule (Fig. IV-13C) which would prevent water movement. If NIP II proteins preferentially occupy a closed conformation, this could explain the disparate water permeability properties. Notably, during their MD simulations of Si(OH)4 permeation of Lsi1, interaction of the substrate with Loop D and Loop B residues was observed (van den Berg et al., 2021). If a similar mechanism exists in NIP6;1, the channel could remain closed until bound by substrate which could promote and stabilize an open conformation.

Examination of the four structural clusters of the NIP6;1 model generated by MD simulations revealed two different orientations (termed “up” and “down” based on orientation relative to the pore axis) of the side chain of the conserved selectivity filter Arg253 (Fig. IV-14). The up configuration is stabilized by and Arg253. In the “up” conformation, the arginine is oriented towards the extracellular side of the membrane, parallel to the pore axis, and is stabilized by a hydrogen bond with the backbone carbonyl of a conserved glycine (Gly186) in the central C-loop (Fig. IV-14). This is
Figure IV-14 NIP6;1 Selectivity Filter Arginine occupies two conformational states during MD simulations. A. Up and down orientation of the selectivity filter arginine (R253) in NIP6;1 homology models identified during MD simulations. Gly 186 is a conserved glycine residue found in the central C-loop of NIP proteins and other aquaporins that stabilizes the arginine in an “up” configuration that maximizes pore diameter. The down conformation is stabilized by a conserved threonine in NIP II proteins within transmembrane alpha helix 1 (see Fig. S1C). B. Comparison of the selectivity filter arginine orientation in the NIP6;1 “down” conformation and the NIP6;1A119W mutant homology model. The down configuration is not observed during MD simulations of the mutant model, presumably due to the steric constraints of the bulkier tryptophan sidechain.
similar to the orientation of the homologous arginine in Lsi1 (Saitoh et al., 2021) and most aquaporin structures, and represents an open configuration. In the “down” configuration, Arg253 hydrogen bonds with Thr97, a conserved residue of NIP II proteins found in transmembrane helix 1 (Fig. IV-14). Distinct up and down arginine rotamers have been observed in the structure of the bacterial aquaporin AQP Z (Jiang et al., 2006), and have been proposed to affect water permeation of the pore based on MD simulations (Xin et al., 2011).

Unlike wild type NIP6;1, the structural clusters of NIP I models and NIP6;1A119W mutants show the absence of distinct rotameric states with the selectivity filter arginine in the “up” configuration throughout the MD simulation trajectory. This could be due to the steric constraints of the bulkier tryptophan side chain (Fig. IV-14B). Thus, similar to AQP Z (Jiang et al., 2006) and other gated aquaporins (e.g., AQP0 [Gonen et al., 2005]), an alternative explanation for the lack of water permeability in NIP II pores could be a “pore pinching” mechanism due to arginine movement in the pore. Arginine rotamers were also predicted for the atypical NIP II protein NIP7;1 that is regulated by an unusual pore tyrosine unique to this subclass (Li et al., 2011).

Summary and Conclusions Chapter IV

The emergence of the NIP channel family in land plants led to their diversification into three broad families categorized based on their predicted pore structures. Two of these families (NIP II and III) have been linked to metalloid nutrition whereas the third (NIP I) has diverse functions ranging from lactic acid transport (previous chapter) to ammonia and water permeability (Roberts and Routray, 2017) to toxic arsenous acid permeability (Kamiya et al., 2009). Here we investigated the comparative metalloid and
water permeability of Arabidopsis NIP I and NIP II channels. All channels tested showed high and indistinguishable permeability to the toxic metalloid As(OH)₃, which is also readily fluxed by NIP III channels which account for As accumulation from toxic soils in rice (Ma et al., 2008). However, we find that boric acid, which differs from As(OH)₃ in molecular geometry and other chemical properties, shows more restricted permeability in some NIP I channels (NIP4;1) but not others (NIP1;1). The major difference between NIP I and NIP II proteins is that the latter are completely impermeable to water, and thus have evolved to lose their aquaporin activity. By site directed mutagenesis, we confirm that the residue that confers this property is the H2 residue of the selectivity filter (tryptophan in NIP I and alanine or glycine in NIP II).

Representative members of each pore family (NIP 4;1 and NIP 6;1) were modeled based on the high-resolution structure of the NIP III pore Lsi1 (Saitoh et al., 2021; van den Berg et al, 2021), to attempt to elucidate how these structures could account for difference in water and metalloid permeability. Modeling of NIP6;1 shows that similar to Lsi1, it forms a larger selectivity filter with room for five amino acids as opposed to the four that is typical of canonical aquaporins. Modeling of the NIP4;1 channel shows that the tryptophan side chain characteristic of NIP I proteins results in a more classical arrangement of four amphipathic amino acids in the selectivity filter similar to bacterial and animal aquaglyceroporins. Additionally, both NIP6;1 and NIP4;1 pores are more hydrophobic compared to the Lsi1 channel. Based on available closed and open structures of Lsi1, and the selectivity filter dynamics of the NIP6;1 model, potential hypotheses for gating of the NIP II pore were generated. How these models could
explain differences in substrate and water permeability are discussed in the Discussion, Chapter VI.
CHAPTER V – Expression, purification and crystallography of with the NIP I channel archetype, soybean nodulin 26
Background and Impetus for structural studies

As discussed throughout this dissertation, the Nodulin 26 Intrinsic Protein family (NIP) represents a plant-specific class of the aquaporin superfamily that has diverged and diversified from the water-specific aquaporin and glycerol transporting glyceroporin paradigm (Roberts and Routray, 2017). These unique structural features have imparted the ability to transport multiple substrates related to fundamental processes in plant nutrition and plant stress biology (water, glycerol, lactic acid, NH₃ and metalloids). As noted in the previous chapter, this diversity of function is linked to three “pore” signatures within the NIP canonical ar/R selectivity filter.

While the recent solution of the NIP III atomic structure (van den Berg et al., 2021; Saitoh et al., 2021) provides insight into the conserved features of the family that confer some elements of substrate specificity, other structural features of the pore and the regulatory domains of some family members remain unresolved. An example are several unresolved questions regarding the archetypal member of the NIP family, the NIP I protein nodulin 26. This protein is unique to the nitrogen fixing symbioses formed between legumes and rhizobia soil bacteria, and is proposed to serve as an aqua-ammoniaporin that is regulated by posttranslational modification via a calcium dependent protein kinase. Phosphorylation exhibits subtle effects on substrate selectivity, enhancing its aquaporin activity (Guenther et al., 2003) while suppressing its ammoniaporin activity (Niemietz and Tyerman, 2000) through an unresolved mechanism. To add to the complexity of the potential biological function of soybean nod26, it has been found to form a complex with glutamine synthetase (Masalkar et al., 2010). Glutamine synthetase is the major ammonia assimilatory enzyme in nitrogen
fixing root nodules, and its association with the nod26 ammoniaporin may have important metabolic significance with respect to nitrogen fixation and assimilation. While molecular modeling, site directed mutagenesis, biophysical and functional analyses have led to hypotheses about the basis for their unique transport properties, a rigorous test requires structural analysis at atomic resolution. Structural determination of nod26, the archetype of the NIP subfamily, by X-ray crystallization would clarify the structural and mechanistic basis for the activities of nod26 and related proteins. The work described in this chapter was undertaken to purify and crystallize soybean nodulin 26 as the first step towards elucidating the structure of a NIP I protein to evaluate the structural basis for these activities.

**Optimization of nod26 expression, solubilization and purification for crystallization trials**

To carry out structural determination by X-ray crystallography, high levels of purified nod26 are required. Therefore, the first aim of this project was to produce milligram quantities of soluble, active, folded and monodispersed preparations of soybean nodule nod26. Previous work by Hwang et al. (2010) has shown that soybean nod26 can be expressed and purified as an amino-terminal His6-tagged protein in *Pichia pastoris*. The recombinant protein is completely functional and retain the properties of the native protein when reconstituted into proteoliposomes (Hwang et al., 2010). While the work of Hwang et al. (2010) was adequate for functional analyses in liposomes, a modified approach was needed to obtain milligram amounts of protein for crystallization trials. For this purpose, two approaches were initially taken: 1. *Pichia pastoris* was transformed with multiple codon-optimized nod26 synthetic gene inserts (Fig. V-1) to
Figure V-1. Synthetic nod26 gene expression construct for *Pichia pastoris*. Schematic representation of the codon optimized synthetic gene encoding soybean nodulin 26 cloned into the pPIC3.5K *Pichia pastoris* expression plasmid under the methanol inducible AOX1 promoter. Cloning and sequencing primers denoted in purple text.
increase the levels of protein expression; and 2. A company (Tni2, Allele Biotechnology) was contracted to produce baculovirus-infected insect cells designed to express high levels of nod26.

Membrane protein crystallization is famously difficult as compared to soluble proteins, owing to the fact that membrane proteins are removed from their original phospholipid environment and shifted to detergents or membrane mimetics. The ability of detergents to maintain the stability and function of the protein during solubilization, purification, and crystallization is perhaps the significant barrier to membrane protein crystallography (Seddon et al., 2004). Complicating this further is the fact that the selection of a suitable detergent is inherently empirical, and a detergent with the optimal behavior is expected to emerge mostly from variations in the protein's characteristics. Some proteins with extensive hydrophobic domains and modest hydrophilic domains are far more prone to aggregation. Membrane protein complexes containing numerous quaternary structures, on the other hand, are considerably more prone to suffer from subunit dissociation, which leads to denaturation. As a result, the efficiency of detergent stabilization would be reliant on protein susceptibility to aggregation or loss of tertiary/quaternary structure (i.e., denaturation), the two primary sources of membrane protein instability in aqueous solutions (Cho et al., 2015). The optimal detergent extracts all of the membrane protein target from the membrane, preserves the original fold of the protein, and forms a protein detergent complex that is stable throughout purification and crystallization.

In membrane protein crystallography, a number of detergent classes have been utilized, however some, such as maltosides and glucosides, have been used more
frequently than others (Raman et al., 2006; Newstead et al., 2008). The majority of these detergents are available in a variety of hydrocarbon chain lengths, allowing for fine-tuning of the protein-detergent complex (PDC) size and stability (Garavito et al., 1996). To find suitable detergent and solubilization conditions, three detergents from this class that are commonly used for membrane protein solubilization and stabilization were selected for initial trials: 1. Undecylmaltoside (UDM); 2. Dodecylmaltoside; and 3. Lauryl Maltose Neopentyl Glycol (LMNG) (Fig. V-2). LMNG is a member of the neopentyl glycol class of detergents which is a novel, low critical micelle concentration (CMC) class of detergents designed specifically for membrane protein solubilization and crystallization due to its branched structure with two nonpolar tails that confer greater stability and less aggregation to solubilized membrane proteins (Chae et al., 2010).

Protein extracts under the three conditions were purified by Ni\(^{2+}\)-chelate chromatography and the homogeneity of the nodulin 26 protein were evaluated by size exclusion chromatography on Superdex-200. Nod26 forms a homotetramer in vivo and should have a size distribution coinciding with a single oligomeric state (tetramer) for crystallization. Examination of the molecular exclusion profiles of the three detergent solubilized preparations show that only one, LMNG, yields nod26 that migrates at the expected size (Fig. V-1). In contrast, nod26 shows the presence of high molecular weight protein aggregates in UDM and DDM-solubilized preparations that migrate in the void volume. The small amount of aggregated protein present in the LMNG solubilized preparation was removed by a second run on Superdex 200. The final product shows a single tetrameric peak by molecular exclusion analysis (Fig. V-3C). Further, circular dichroism spectroscopy shows a characteristic α-helical protein which we would expect
Figure V-2. Detergent solubilization and molecular weight aggregate screen for purified nod26. A. Structures of the test detergents used for solubilization. B. FPLC Superdex-200 Molecular exclusion chromatography profiles of Ni-chelate purified P. pastoris Nod26 from membrane samples solubilized with Dodecylmaltoside (DDM), Undecylmaltoside (UDM), and Lauryl Maltose Neopentyl Glycol (LMNG). The elution position of the excluded void volume and the expected position of the nodulin 26 tetramer (116kDa) is shown. C. SDS-PAGE gel of nod26 expressed in Pichia and Ni-NTA purification from LMNG solubilization. The position of the monomeric and dimeric species (29 and 58 kDa) are indicated. Protein was visualized with Coomassie blue staining.
Figure V-3 Purified nod26 conformation and particle size distribution. A. Far UV circular dichroism spectrum of final nod26 preparation. Void refers to the aggregated fractions collected during molecular exclusion purification B. Size distribution profile for final nod26 preparation determined by quasi-elastic light scattering. C. FPLC Superdex-200 Molecular exclusion chromatography profile of final nod26 preparation.
from a member of the aquaporin superfamily (Fig. V-3A). The presence of aggregate and oligomeric states was also evaluated by using dynamic light scattering. The final preparation is free of aggregation and in a monodispersed, single molecular species (Fig. V-3B). LMNG solubilization of insect cell membranes followed by Ni\textsuperscript{2+}-chelate chromatography and molecular exclusion purification yields a similar monodisperse, tetrameric preparation of nod26 (not shown).

One potential drawback to the use of the LMNG detergent is its limited ability to extract nod26 from cell membranes. To produce and purify the mg amounts of nod26 needed for crystallography screens, we examined additives that would provide a significant increase in solubilization and recovery of nod26. Polyamines, such as spermidine, have been used as additives for enhancing the solubilization of membrane proteins (Yasui et al., 2010). It is proposed that the moderate hydrophobic and cationic character of polyamines allow these molecules to interact with the negatively charged lipid bilayers and thus aid in the solubilization of membrane proteins. nod26 membrane fractions solubilized with LMNG with a range of spermidine concentrations were tested to determine the optimal conditions for solubilization of the protein based on anti-histidine tag Western blot (Fig. V-4). The maximum yield of nod26 was obtained with 5 mM spermidine, and this was adopted in the final solubilization and purification strategy. The optimized protocol was experimentally determined using both insect cell and yeast expression systems and allowed for purification of milligram amounts of nod26 for initial crystallization screening.
Figure V-4 Enhanced solubilization of nod26 by polyamines. Total membranes of *Pichia pastoris* expressing nod26 were solubilized with a range of Spermidine concentrations. Densitometry analysis of anti-His Western Blot displays maximal solubility of nod26 at 5mM Spermidine.
Initial crystallization of nod26 with various transport substrates

Using a Phoenix crystallization robot from Art Robbins Inc., nod26 was screened against 192 different conditions shown to be promising for α-helical membrane proteins (Molecular Dimensions, Memgold) using the sitting-drop, vapor diffusion method (Newby et al., 2009). Crystals appeared for several of the conditions and grew to full size in approximately 3 weeks (Table V-1, Fig. V-5). Due to the higher cost and increased level of nod26 aggregation of insect cell expression, subsequent crystallization trials were carried out using nod26 produced in P. pastoris.

Due to the desire to crystallize nod26 with different substrates in the channel, the crystals were generated and categorized by the substrate that would be expected to be in the pore (Fig. V-5). Successful diffraction with different substrates present in the crystallization condition could elucidate the interactions necessary for the selectivity of the multifunctional channel. As an alternative approach, substrates could be ‘soaked-in’ for conditions that diffract well by addition of the substrate to the cryoprotectant solution.

While crystals were readily observed under multiple conditions in our Memgold screens, most were small and not well shaped and would not be expected to give the high-resolution x-ray diffraction needed for structural analysis. The quality of protein crystals is a crucial criterion for determining structure via X-ray crystallography (Hou et al., 2015). Indeed, having high-quality crystals is a requirement for obtaining high-resolution diffraction data. This is the first step before advancing to data refinement and the construction of atomic models by Molecular Replacement or other phasing methods (isomorphous replacement). In general, the crystal quality of proteins is determined by the degree of packing order within the crystals, which is associated with the regularity of
<table>
<thead>
<tr>
<th>Expression system</th>
<th>Nodulin 26 water</th>
<th>Nodulin 26 glycerol</th>
<th>Nodulin 26 ammonia</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>P. pastorus</em></td>
<td><img src="image" alt="Image" /></td>
<td><img src="image" alt="Image" /></td>
<td><img src="image" alt="Image" /></td>
</tr>
<tr>
<td></td>
<td><img src="image" alt="Image" /></td>
<td><img src="image" alt="Image" /></td>
<td><img src="image" alt="Image" /></td>
</tr>
<tr>
<td>Insect Cell baculovirus</td>
<td><img src="image" alt="Image" /></td>
<td><img src="image" alt="Image" /></td>
<td><img src="image" alt="Image" /></td>
</tr>
</tbody>
</table>

**Figure V-5 Crystallization of nod26.** Representative crystals obtained from recombinant nod26 from the indicated expression systems in the presence of various substrates. Multiple pictures in a given cell indicate multiple crystallization conditions. The red bar represents 0.1mm.
Table V-1: Siting drop crystallization conditions for nod26 solubilized with the indicated detergent. Conditions in red produced crystals with near atomic resolution diffraction at the APS synchrotron.

### Nod26-LMNG Crystallization Conditions

<table>
<thead>
<tr>
<th>Buffer</th>
<th>Salt 1</th>
<th>Salt 2</th>
<th>Precipitant</th>
<th>Temperature [Celsius]</th>
</tr>
</thead>
<tbody>
<tr>
<td>★ 0.1 M Tris (pH 8.4)</td>
<td>0.05 M Calcium chloride dihydrate</td>
<td>0.05 M Barium chloride dihydrate</td>
<td>28% PEG 400</td>
<td>17</td>
</tr>
<tr>
<td>0.1 M Tris (pH 8.0)</td>
<td>0.1 M Magnesium chloride hexahydrate</td>
<td>0.1 M Sodium chloride</td>
<td>25% PEG 400</td>
<td>17</td>
</tr>
<tr>
<td>0.1 M Tris (pH 7.5)</td>
<td>0.1 M Potassium chloride none</td>
<td>10% PEG 400</td>
<td>17</td>
<td></td>
</tr>
<tr>
<td>0.1 M HEPEs (pH 6.8)</td>
<td>0.2 M Calcium chloride dihydrate none</td>
<td>25% PEG 400</td>
<td>17</td>
<td></td>
</tr>
<tr>
<td>0.1 M Sodium citrate (pH 3.5)</td>
<td>0.05 M Sodium chloride none</td>
<td>25% PEG 400</td>
<td>17</td>
<td></td>
</tr>
<tr>
<td>0.05 M MES (pH 7.3)</td>
<td>0.1 M Sodium chloride 0.15 M Ammonium Sulfate</td>
<td>22% PEG 1000</td>
<td>17</td>
<td></td>
</tr>
<tr>
<td>★ 0.1 M Tris (pH 7.3)</td>
<td>0.2 M Ammonium formate none</td>
<td>27% Pentaerythritol ethoxylate (15/4 EO/DH)</td>
<td>17</td>
<td></td>
</tr>
</tbody>
</table>

### Nod26-DMNG Crystallization Conditions

<table>
<thead>
<tr>
<th>Buffer</th>
<th>Salt 1</th>
<th>Salt 2</th>
<th>Precipitant</th>
<th>Temperature [Celsius]</th>
</tr>
</thead>
<tbody>
<tr>
<td>★ 0.1M MES (pH 6.5)</td>
<td>0.09 M Magnesium chloride none</td>
<td>28% PEG 400</td>
<td>22</td>
<td></td>
</tr>
<tr>
<td>0.05 MES (pH 6.5)</td>
<td>0.1 M Sodium chloride 0.15 M Ammonium sulfate</td>
<td>19% PEG 1000</td>
<td>4 &amp; 14</td>
<td></td>
</tr>
<tr>
<td>0.1 M Tris (pH 6.5)</td>
<td>0.1 M Calcium chloride dihydrate none</td>
<td>13% PEG 2000 MWFE</td>
<td>22</td>
<td></td>
</tr>
<tr>
<td>0.08 M Sodium phosphate (pH 6.2)</td>
<td>0.02 M Sodium citrate tribasic dihydrate none</td>
<td>18% PEG 2000</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>0.1 M Tris (pH 8.3)</td>
<td>0.05 M Calcium chloride dihydrate 0.05 M Barium chloride dihydrate</td>
<td>32% PEG 400</td>
<td>22</td>
<td></td>
</tr>
<tr>
<td>★ 0.1 M Tris (pH 7.5)</td>
<td>0.1 M Magnesium chloride hexahydrate none</td>
<td>22% PEG 400</td>
<td>14</td>
<td></td>
</tr>
<tr>
<td>0.4 M Ammonium acetate (pH 8.0)</td>
<td>none none</td>
<td>13% PEG 2000 MWFE</td>
<td>14</td>
<td></td>
</tr>
</tbody>
</table>

### Nod26-OGNG Crystallization Conditions

<table>
<thead>
<tr>
<th>Buffer</th>
<th>Salt 1</th>
<th>Salt 2</th>
<th>Precipitant</th>
<th>Temperature [Celsius]</th>
</tr>
</thead>
<tbody>
<tr>
<td>MES (pH 6.5)</td>
<td>0.3 M Lithium sulfate none</td>
<td>25% PEG 400</td>
<td>25</td>
<td></td>
</tr>
<tr>
<td>0.1 M Tris (pH 8.0)</td>
<td>0.22 M Sodium citrate tribasic dihydrate none</td>
<td>25% PEG 400</td>
<td>17 &amp; 25</td>
<td></td>
</tr>
<tr>
<td>0.1 M Sodium citrate (pH 6.0)</td>
<td>0.05 M Sodium chloride 0.02 M Magnesium chloride hexahydrate</td>
<td>22% PEG 400</td>
<td>17 &amp; 25</td>
<td></td>
</tr>
<tr>
<td>0.1 M Tris (pH 8.0)</td>
<td>0.1 M Sodium chloride 0.325M Sodium acetate trihydrate</td>
<td>21% PEG 400</td>
<td>17 &amp; 25</td>
<td></td>
</tr>
<tr>
<td>0.1 M Glycine (pH 9.3)</td>
<td>0.1 M Lithium sulfate none</td>
<td>30% PEG 400</td>
<td>17 &amp; 25</td>
<td></td>
</tr>
<tr>
<td>0.05 M Tris (pH 8.0)</td>
<td>0.05 M Sodium sulfate 0.05 M Lithium sulfate</td>
<td>35% PEG 400</td>
<td>17 &amp; 25</td>
<td></td>
</tr>
<tr>
<td>0.05 M glycine (pH 9.5)</td>
<td>0.1 M Sodium chloride none</td>
<td>33% PEG 400</td>
<td>25</td>
<td></td>
</tr>
<tr>
<td>0.1 M Tris (pH 8.0)</td>
<td>0.15 M Sodium chloride none</td>
<td>13% PEG 6000</td>
<td>17 &amp; 25</td>
<td></td>
</tr>
</tbody>
</table>

Crystals produced in the well conditions marked with star (★) symbols showed diffraction at a resolution less than 4.5 angstroms at the APS synchrotron.
the packing units, i.e. the protein's purity and uniformity. Obtaining high-quality protein crystals is a well-known bottleneck to protein structure determination and remains a difficult challenge.

Therefore, from the conditions that yielded the initial crystal 'hits', we narrowed down and refined the pH, precipitant concentrations and temperature conditions that allowed the growth a fewer number of larger, well-shaped crystals. From these refined screens, 10 crystals were tested for x-ray diffraction properties in initial experiments at the Advanced Photon Source (APS) synchrotron (Argonne, IL). While most crystals did not yield useful diffraction data, one condition diffracted up to a resolution of ~4.5 angstroms (Fig. V-6). Further optimization of the crystal forming conditions (additives, temperature, etc.) was attempted to increase crystal quality to approach the 3 angstrom diffraction limit needed to pursue structural modeling. However, this represented the best results obtained from a wide variety of crystal forming conditions in LMNG detergent, and in most cases the crystals did not diffract or diffracted with low resolution (~7 angstroms).

**Attempts at crystallization optimization with other neopentyl glycol detergents**

Recent studies have elucidated the characteristics of the neopentyl glycol (NG) family of nonionic detergents, of which LMNG is a member. The data shows that LMNG creates a micelle in solution that is significantly larger than the more commonly used Dodecylmaltoside (DDM) detergent and that the micelle size for neopentyl glycol detergents decreases significantly with decreasing acyl chain length (Cho et al., 2015). Although the detergent of choice is determined by a variety of factors, great effort should be made to screen for crystals in shorter chain detergents, since these are more
Figure V-6 Crystallization and diffraction of nod26-LMNG. **A.** Representative nod26-LMNG crystal from one of the conditions tested for diffraction at APS. **B.** Diffraction to 4.5Å from crystal in A. **C.** Detergent concentration and condition used to for crystallization of A.
likely to diffract to a higher resolution (Sonoda et al., 2010). In light of our results, it was hypothesized that the large size of the nod26-LMNG micelle was hindering strong intermolecular contacts in the crystal resulting in poor organization and low resolution x-ray diffraction.

To address this issue, we screened a number of neopentyl glycol detergent analogues of LMNG that share the low CMC (high hydrophobicity) characteristic but have shorter alkyl chain lengths, thus reducing the size of the detergent micelle. Experimental evidence has shown that the neopentyl glycol detergents Octyl Glucose Neopentyl Glycol (OGNG), as well as and Decyl Maltose Neopentyl Glycol (DMNG) are suitable alternatives that produce significantly smaller micelles than LMNG in solution (Cho et al., 2015). These detergents have generated promising results for improving x-ray diffraction resolution for other members of the aquaporin family (Frick et al., 2014).

We examined the ability of these neopentyl glycol detergents, as well as the classical low CMC detergent octyl glucoside, to solubilize nodulin 26 in a stable non-aggregated state (Fig. V-7). DMNG was the most effective in producing predominantly tetrameric nod26 in a stable state. Nodulin 26 solubilized and purified in DMNG was subsequently retested for crystal growth using Memgold & Memgold 2 kits (Molecular Dimensions). Several large and well-formed crystals were obtained (Fig. V-8A and B) and were analyzed at the APS synchrotron. Most did not yield diffraction data, but a crystal generated in 0.1 M MES-NaOH, pH 6.5, 30 mM MgCl₂, 28% PEG₄₀₀ generated a diffraction pattern with a resolution of 4 angstroms. A closer examination of the diffraction pattern shows that this level of resolution is limited to a small area of the diffraction grid, and this suggests strong diffraction anisotropy with this resolution is
Figure V-7 Detergent screen of LMNG analogues. A. FPLC Superdex-200 Molecular exclusion chromatography profile of nod26 solubilized with Decyl Maltose Neopentyl Glycol (MNG-C8), Undecyl Maltose Neopentyl Glycol (MNG-C9), Octyl Glucose Neopentyl Glycol (OGNG) or Octyl glucoside (OG) detergents. B. Measure of protein-detergent complex stability by comparison of protein densities on SDS-PAGE gel after a one week incubation period. Hydrodynamic radius of micelles for each detergent from Cho et. al. (2015) are listed in parenthesis.
Figure V-8 Crystallization and diffraction of nod26-DMNG. A. Representative crystals obtained from recombinant Nod26 solubilized in DMNG. B. Representative Nod26-DMNG crystal from one of the conditions tested for diffraction at APS. C. Diffraction to 4 Å from crystal in A. D. Detergent concentration and condition used to for crystallization of B.
directionally limited (Rupp, 2009). Anisotropic diffraction indicates that the crystal packing within the unit cell is not uniform, a situation that would prevent data refinement and atomic model building even if high resolution diffraction is obtained (Rupp, 2009).

**Summary and Conclusions Chapter V**

Nod26 easily crystallizes, and screens frequently give beautiful, small crystal pyramids under many conditions. A main challenge in this project is that while several crystallization conditions yield “pretty crystals”, it has been difficult to identify conditions that lead to improved x-ray diffraction. By using a battery of systematic crystallography screens employing hundreds of conditions with varying additives and other environmental conditions (e.g., pH and temperature), we were able to generate several larger, apparently more organized nod 26 crystals that we examined at the APS Synchrotron. In total, over ninety nod26 crystals were subjected to x-ray diffraction tests at APS. Unfortunately, the majority of these crystals either provided no diffraction, low resolution diffraction (<7 Å), or anisotropic diffraction at subatomic resolution (the best result was 4 Å). These results underscore the inherent difficulty of the elucidation of membrane protein structures. Potential reasons for this outcome as well as possible future experimental approaches to obtain high resolution diffracting nod26 crystals are discussed in Chapter VI.
CHAPTER VI – Discussion
The NIP Structure and Function Conundrum: Same, same, but different...

NIPs have been classified into three pore families with “signature” amino acids at the predicted ar/R selectivity filter based on the conserved aquaporin hourglass fold and molecular modeling. However, this classification scheme is complicated by the fact that some transport properties are remarkably shared between pore families with quite distinct ar/R selectivity regions, while in other cases NIPs with similar ar/R regions exhibit significant diversity in biochemical activity and biological function. It is clear that there is more to the determination of specificity and function beyond the canonical aquaporin ar/R, an assertion that is becoming clear with each new structure that is elucidated within the aquaporin family. From an evolutionary, genetic, and biophysical perspective, categorization and elucidation of biological functions of the diverse plant-specific NIP family remains a challenging endeavor. In this work, genetic and biochemical approaches were used to provide insight into the functions and structural properties of NIP I and NIP II proteins from Arabidopsis.

NIPs as “Metalloidoporin” channels

Since the groundbreaking work of Takano et al (2006) and Ma et al. (2006) that demonstrated NIP function in boric acid and silicic acid nutrition respectively, it has come to be generally accepted and supported by biochemical, biophysical and genetic evidence that a major biological function for the NIP family is to serve as permeases for uncharged metalloid hydroxides (Pommerrenig et al., 2015). This include nutrients such
as boric acid and orthosilicic acid, as well as toxic hydroxides of As (III), Ge and Sb. These compounds form weak Lewis acid structures with differing geometries ranging from trigonal planar (B[OH]₃), to trigonal pyramidal (As[OH]₃ and Sb[OH]₃) to tetrahedral (Si[OH]₄ and Ge[OH]₄) molecules. They exist in an uncharged state at physiological pHs.

NIPs transport these compounds with differing selectivity. The most selective are the NIP III pores which are specific for the tetrahydroxylated Si(OH)₄ and Ge(OH)₄ (Ma et al. 2006; Mitani-Ueno et al. 2011). As discussed in Chapter IV, the NIP III proteins contain an exceptionally wide and hydrophilic channel with a conserved five amino acid selectivity filter that accommodates the bulky Si(OH)₄ molecule both sterically as well as enthalpically through several hydrogen bond contacts (Saitoh et al., 2021; van den Berg et al., 2021). Quantum mechanical calculations and MD simulations of Lsi1 support a model (Saitoh et al., 2021) where the selectivity filter amino acids, two tightly bound waters (waters 3 and 9, see Fig. IV-2), and additional pore side chains and backbone carbonyl groups, provide hydrogen bonds to eight silicic acid molecules that move single file through the NIP III pore.

NIP I and II pores differ from NIP III in three major respects that prevent permeation by Si(OH)₄. First, NIP I and II pores have a hydrophobic substitution (valine or isoleucine) for serine at the H5 position as well as a glycine for threonine at the H1 position (Fig. IV-2C). This would reduce the hydrophilicity and hydrogen bonding capability of the selectivity filter. Second, the conserved glycines at H2 and LE₁ in NIP III are replaced with larger sidechains in NIP I and NIP II which would restrict the diameter (particularly for NIP I) of the pore. As a result the larger Si(OH)₄ (74 Å³) would
be sterically restricted and would be unable to form optimal hydrogen bond contacts compared to smaller As(OH)$_3$ or B(OH)$_3$ (52 to 54 Å$^3$).

In contrast to Si(OH)$_4$, As(OH)$_3$ is a promiscuous substrate that is transported by isoforms of all three functional NIP sub-classes (Bienert et al., 2008; Isayenkov and Maathuis, 2008; Ma et al., 2008; Kamiya et al., 2009; Kamiya and Fujiwara, 2009; Mitani-Ueno et al., 2011; Hayes et al., 2013; Katsuhara et al., 2014; Xu et al., 2015). This is consistent with our observations that show that NIP I and NIP II proteins from Arabidopsis are equally permeable to As(OH)$_3$ and that substitutions at the H2 position (tryptophan or alanine) have no effect on transport of this substrate. Similar results were found with NIPs from barley (Hayes et al., 2013) and rice (Mitani-Ueno et al. 2011).

Beyond NIP proteins, it is clear that As(OH)$_3$ promiscuity extends to other plant and non-plant aquaglyceroporins (Bhattacharjee et al., 2008; Bienert et al., 2008). If an aquaporin fluxes glycerol, usually it will flux As(OH)$_3$. In the case of NIP proteins, As permeability has potential agricultural importance since NIPs that are highly expressed in planta (e.g., NIP1;1 and NIP3;1 in Arabidopsis, Kamiya et al., 2009; Xu et al., 2015; Lsi1 in rice, Ma et al., 2008) facilitate the uptake and translocation of As from soils.

**NIP II protein selectivity for Boron**

Based on transport specificity, localization, and regulation, the NIP II subgroup appears to be specialized for boric acid uptake and transport to critical sink tissues under conditions of boric acid limitation to support the formation of the rhamnogalacturonan II (RG-II) pectin cell wall (Takano et al., 2008; Miwa et al., 2010; Yoshinari and Takano, 2017). By using the structural models for NIP6;1 and the hydrated structure of Lsi1 as a template, a model for the pore of NIP6;1 with boric acid
docked into the selectivity filter was generated (Fig. IV-4). Boric acid is a planar trigonal molecule (120° O-B-O bond angle), and within the minimized NIP6;1 structural model, the three boric acid hydroxyl groups are positioned to hydrogen bond to the conserved arginine and backbone carbonyls or side chains from the H1 (glycine 96), LE1 (Ala 247 and adjacent residue Ser 248). The small residue at H2 (Ala 119) allows for access of the H1 residue in the selectivity filter, and also provides room for one of the bound waters (water 3) found in the Lsi1 structure.

In the NIP4;1 structure (Fig. IV-2 and IV-3), the bulky tryptophan side chain would prevent the approach of the H1 residue, the positioning of water 3, as well as sterically restrict the pore at one of the proposed positions of the boric acid molecule docked in the ar/R region (occupied by water 2 in Lsi1 and NIP6;1 models). This prediction fits with the observation that boric acid is more poorly permeated in NIP4;1, and that the permeability of the protein for this metalloid increases four-fold with a substitution of alanine for tryptophan at the H2 position. Conversely, the substitution of a tryptophan for alanine in NIP6;1 has an opposing effect, causing a two-fold reduction in permeability.

This model does not explain why As(OH)₃, which has a molecular volume that is similar to boric acid, permeates NIP4;1 and NIP6;1 equally. The answer may lie in the distinct molecular geometry and greater hydrophobicity of As(OH)₃ compared to boric acid. As(OH)₃ is a trigonal pyramid with more restricted bond angles (O-As-O 97° based on X-ray absorption spectroscopy and Density Functional Theory calculations, (Ramírez-Solís et al., 2004)). As a result, unlike boric acid, the three hydroxyl groups in As(OH)₃ are clustered on one side of the molecule rather than in a trigonal planar array.
(see Fig IV-6A), while the lone pair of electrons occupies the opposing side. Based on quantum calculations and thermodynamic comparisons, As(OH)$_3$ is predicted to be an amphipathic molecule with a hydrophobic side (lone pair) and a hydrophilic side (clustered hydroxyl groups), not unlike the common aquaporin substrate glycerol. This arrangement could allow permeation of As(OH)$_3$ through multiple aquaporin and aquaglyceroporin channels (Porquet and Filella, 2007; Hernández-Cobos et al., 2010). NIP I and NIP II proteins are also equally permeable to glycerol (Wallace and Roberts, 2006).

The findings with NIP4;1 support previous work with the NIP I protein soybean nodulin 26 which has restricted ability to transport boric acid compared to glycerol and other test solutes (Wallace and Roberts, 2005; Roberts and Routray, 2017). However, our findings with another NIP I protein, Arabidopsis NIP1;1, showed that it is effective at transporting boric acid and that substitutions at the H2 position did not enhance this ability. This finding further underscores that there is more to pore specificity than the sequence of the ar/R selectivity filter, and that differences in transport selectivity for various metalloids can exist with each NIP pore family.

As a final note, we find that overexpression of NIP I proteins partially complements the boric acid sensitive phenotype of nip5;1-1. While this finding supports the contention that these proteins are permeable to boric acid to some level, there is lack of genetic and physiological data to firmly support a physiological role for boric acid transport in planta.

**Water transport through NIP II proteins and role of the H2 position in the Ar/R selectivity filter: potential gating and its importance**
Phylogenetically, the NIP II proteins in dicots are segregated into three subclades with preferential expression in roots (NIP5;1), shoots (NIP6;1), and flowers (NIP7;1) (Routray et al., 2018). Overwhelming genetic and biochemical evidence shows that all three subtypes of NIP II proteins lack aquaporin activity, and flux boric acid in a largely “water-tight” manner (Takano et al., 2006; Tanaka et al., 2008; Li et al., 2011; Routray et al., 2018). Why would aquaporin activity be restricted in NIP II proteins? NIP II proteins are plasma membrane localized, boric acid channels that coordinate uptake and distribution of this critical nutrient in collaboration with the BOR membrane transporters (Takano et al., 2006; Takano et al., 2010; Uehara et al., 2014; Wang et al., 2017). In the case of NIP II proteins, the ability to readily transport boric acid in a water-tight manner may be necessary to prevent disruption of cell turgor and cell volume regulation while facilitating the movement of this solute across the plasma membrane. Unlike animal cells, the water permeability of the plant plasma membrane is tightly and coordinately regulated. The plant plasma membrane water permeability is determined by PIP aquaporins, which are themselves regulated tightly by multiple mechanisms including genetic regulation, reversible trafficking between internal membranes and the cell surface, as well as gating at the protein level through through hetero-oligomerization, phosphorylation, pH and divalent cations (Chaumont and Tyerman, 2014; Maurel et al., 2015; Santoni, 2017; Takano et al., 2017). The osmotic water permeability of plant plasma membranes is often 100-times lower than tonoplast membrane, which is necessary for cell volume regulation and cytoplasmic buffering (Maurel et al. 1997).
The reduction or loss of aquaporin activity in NIP II channels compared to NIP I proteins is due solely to the H2 residue in the ar/R selectivity filter. We show here that the substitution of the bulky NIP I H2 residue (tryptophan) for the conserved NIP II H2 residue (alanine) actually opens the channel to water transport, while the reciprocal substitution of alanine for tryptophan has the opposite effect on NIP I channels. This is counterintuitive based on our comparative models and predicted pore apertures of the NIP I and NIP II pores, and it is unclear as to why the wider NIP II pores do not transport water.

The NIP6;1 model shows that, similar to the hydrated structure of Lsi1 solved by Saitoh et al. (2021) that enhanced width of the pore could accommodate five water molecules (Fig. IV-2). High permeability aquaporins (e.g., mammalian AQP) transport water in a single file manner in which strategically placed hydrogen bond donors and acceptors interact with water molecules along the length of the pore (Hub et al., 2009). Based on transition state analysis of aquaporin water permeability across aquaporins, a hypothesis was proposed that pores with large diameters similar to NIP II proteins that lack the single file arrangement of water may be unable to optimize water organization at the ar/R region resulting in low water permeability (Wallace & Roberts, 2005). However, this hypothesis cannot explain why Lsi1, which has a wider pore than NIP6;1 and contains multiple water molecules within the ar/R, has substantial aquaporin activity when expressed in Xenopus oocytes (Mitani et al., 2008). Instead based on this study, we propose two alternative models for NIP6;1 gating of water transport based on “capping” and “pinching” mechanisms that have been found in aquaporins (Hedfalk et al., 2006).
The capping model is based on the finding that Lsi1 exists in alternative open and closed conformations that are stabilized by interactions between loop D and loop B on the cytosolic side of the membrane (Fig. IV-13). Interestingly, the features of loop D that control the gate are conserved across all NIP protein subclasses suggesting this is a common feature of the subfamily. MD simulations of Lsi1 suggest that open and closed conformations of the gate occur spontaneously. In the case on NIP6;1, one possibility is that the closed conformation is favored resulting in a closed pore. Since loop D is proposed to interact with transported metalloids in Lsi1 based on Si(OH)₄ permeation MD trajectories (van den Berg et al., 2021), one possible model is that binding of the metalloid substrate to the gate could stabilize the open conformation enabling metalloid transport.

Similar regulation of gating in loop D of other members of the aquaporin family has been documented (Kreida and Tornroth-Horsefield, 2015). In plants, a notable example is provided by the SoPIP2;1 protein in which loop D forms a cap over the pore aperture (Tornroth-Horsefield et al., 2006). Regulation and channel opening of SoPIP2;1 occurs through several mechanisms including pH, phosphorylation and calcium ion interaction (Tornroth-Horsefield et al., 2006). Phosphorylation of the NIP I protein soybean nodulin 26 triggers enhanced aquaporin activity in nitrogen fixing nodules in response to osmotic stress, suggesting potential for gating of this activity.

A second type of regulation that is found within the pore of certain aquaporins is “pinching” or restriction of the pore aperture through the movement of amino acid side chains. Examples of this type of regulation are apparent in the E. coli AqpZ aquaporin (Jiang et al., 2006), mammalian AQP0 (Gonen et al., 2005), and the pollen specific NIP
protein NIP7;1 (Li et al., 2011). Based on our MD simulations, this model predicts alternative rotameric states of the conserved ar/R arginine in the NIP6;1 pore due to the wider aperture. A similar case in AqpZ shows that the ar/R arginine rotameric state could open or close the pore to water movement (Xin et al., 2011). Whether either of these models, or a combination of both, apply to NIP I or II proteins will require further detailed structural analysis.

NIP2;1-mediated lactic acid efflux promotes Arabidopsis survival during low oxygen stress

While there is a considerable amount of evidence that NIP II proteins act as boric acid channels and NIP III proteins as silicic acid channels, the biological role of the NIP I group is less obvious. From an evolutionary standpoint, the NIP I group is the most recent to emerge and is exclusive to angiosperms (Roberts and Routray, 2017). While it is obvious that NIP I proteins contribute to arsenite mobility in plants, this is most likely not their biological role. While the work in this study, as well as others (Diehn et al., 2019) support modest boric acid permeability as a function for NIP I proteins, the lack of genetic evidence makes a function in boron uptake less plausible. NIP I expression in specific developmental or stress responses (e.g., symbiosis, pollen, and waterlogging stress) provides intriguing leads to biological function, but genetic evidence (and possibly bioinformatic and system-based approaches) is required to provide more incisive support for these hypothetical roles.

Previous biophysical and biochemical analysis of Arabidopsis NIP2;1 in Xenopus oocytes (Choi and Roberts, 2007) indicate that it is an outlier from other classical NIP proteins and is impermeable to water and all traditional NIP solute substrates, and
instead shows specific bidirectional permeability to lactic acid. The present study, cellular, genetic, and physiological evidence, indicate that the aquaporin-like NIP2;1 assists in lactic acid efflux from Arabidopsis roots.

In response to low oxygen conditions resulting from flooding or submergence stress, plants switch to anaerobic fermentation pathways to maintain glycolytic flux and energy production. In Arabidopsis, enzymes of ethanolic (ADH and PDC) and lactic acid (LDH) fermentation pathways are necessary for optimal survival to low oxygen stress (Ellis et al., 1999; Ismond et al., 2003; Kursteiner et al., 2003; Dolferus et al., 2008). Root levels of lactic acid/lactate increase 14-fold during the first two hours of hypoxia challenge (Mustroph et al., 2014) suggesting that lactic acid fermentation is induced during the initial stages of hypoxia. As lactic acid/lactate accumulates, hypoxia-stressed Arabidopsis plants (Dolferus et al., 2008), similar to other plant lineages (Xia and Saglio, 1992; Rivoal and Hanson, 1993; Xia and Roberts, 1994), release this fermentation end product to the external media/rhizosphere. Lactic acid efflux mechanisms in plant roots may assist in mitigating cellular acidification or other toxic effects of lactic acid accumulation during anaerobic stress (Xia and Roberts, 1994; Gibbs and Greenway, 2003).

To alleviate potential cellular acidification from lactic acid accumulation, the pathways for the efflux of lactic acid must transport either the protonated form (lactic acid) or co-transport lactate with a proton [reviewed in (Greenway and Gibbs, 2003)]. In animal cells, lactate is effluxed or taken up by members of the SLC16 subgroup of the major facilitator superfamily known as Monocarboxylate Transporters or MCTs (Counillon et al., 2016; Sun et al., 2017). These symporters co-transport lactate with $H^+$
in a bidirectional fashion. They participate in the efflux of excess lactic acid during anaerobic fermentation, and also serve as an uptake mechanism for lactate from the serum for further metabolism (Sun et al., 2017). Land plants lack members of the SLC16/MCT transporter family, and the molecular identity of the transporters or channels that mediate the efflux of lactic acid/lactate produced during anaerobic fermentation have remained unclear. Several observations in the present work provide strong support that lactic acid transport and efflux is the biological function of NIP2;1. First, the expression of NIP2;1 in response to hypoxia coincides with the appearance of lactic acid/lactate in the external medium; second, genetic mutation of the NIP2;1 via T-DNA insertion results in the reduction of lactic acid efflux from hypoxic roots into the external medium and a concomitant increase in the accumulation of lactic acid/lactate within root tissue; and third, nip2;1 mutants show reduction in the acidification of the media surrounding hypoxic roots.

nip2;1 mutant seedlings show poorer survival to argon-induced low oxygen stress compared to WT, presumably because of the over accumulation of toxic levels of lactic acid due to a reduced ability to efflux this end product from roots. As noted above, cytosolic lactic acid generation would increase the acid load of the cytosol that could contribute to acidosis (Davies et al., 1974; Roberts et al., 1984; Felle, 2005). Additionally, the accumulation of lactic acid/lactate could also contribute to reduced glycolytic flux by affecting NAD+ regeneration by altering the equilibrium of the LDH reaction, or potentially through product feedback inhibition mechanisms. For example, recent studies in yeast and mammals show that over accumulation of lactate leads to the production of the toxic side product 2-phospholactate catalyzed by pyruvate kinase.
This side product of lactate blocks the production of fructose-2-6 bisphosphate, leading to the inhibition of the key glycolytic enzyme phosphofructokinase-1 (Collard et al., 2016). The production of similar toxic lactate metabolites side products could conceivably occur in plant tissues as well.

**NIP2;1 expression during normoxia, hypoxia stress and recovery**

NIP2;1 expression is predominately limited to root tissues with a precise pattern of transcript and protein expression during normoxic, hypoxic, and reoxygenation conditions. Under normoxic conditions, *NIP2;1* promoter activity is restricted to cells within the stele of the mature root. The cells of the stele are hypoxic even under well aerated growth conditions due to the low rate of lateral oxygen diffusion across the mature differentiated root (Gibbs and Greenway, 2003; Armstrong et al., 2019). “Anoxic cores” in the root stele are proposed to aid in hypoxia sensing and acclimation, potentially by the communication of low oxygen or energy signals (ethylene, metabolites, low pH, and Ca$^{2+}$) between hypoxic and well-aerated cells (Armstrong et al., 2019). The roots of *nip2;1* mutants show increased accumulation of lactic acid/lactate under normoxic conditions. This suggests that LDH is active in anoxic core tissues, even under aerobic conditions, and that NIP2;1 basal expression is necessary for preventing lactic acid/lactate accumulation.

*NIP2;1* Q-PCR and GUS data show the characteristics of a core hypoxia response transcript, with acute induction of *NIP2;1* expression in roots within 1 hr of the initiation of hypoxia stress, followed by a peak and eventual decline to a reduced but elevated steady state level that is sustained during hypoxia. Interestingly, examination of the cell-specific translatome atlas based on the work of (Mustroph et al., 2009) shows
that the expression *NIP2;1* during hypoxia parallels the expression of the two lactate metabolizing enzyme transcripts, *LDH* and *GOX3* (Fig. VI-1). All three transcripts are predominantly, if not exclusively, expressed in root tissue (Dolferus et al., 2008; Mustroph et al., 2014; Engqvist et al., 2015), and accumulate to the highest levels in the root cortex, as well as to high levels within the epidermal and vascular tissues, but are absent or poorly expressed in leaf tissues. The root-predominant expression pattern of *NIP2;1, LDH1* and *GOX3* is consistent with the distinct properties of lactate metabolism in roots and shoots during low oxygen stress (Ellis et al., 1999; Mustroph et al., 2014). Indeed, lactic acid fermentation and accumulation is predominantly restricted to root tissues during hypoxia in Arabidopsis (Mustroph et al., 2014). Based on the model of (Engqvist et al., 2015), these three gene products are proposed to coordinate root lactic acid/lactate homeostasis through its production (LDH), its recovery back to pyruvate (GOX3), and the excretion of lactic acid from the cell when it is overproduced during low oxygen stress (*NIP2;1*).

Similar to other hypoxia-induced genes (Branco-Price et al., 2008), reoxygenation results in suppression of *NIP2;1* mRNA expression and a return to low basal levels within two hours of recovery. In contrast, *NIP2;1* protein levels increase during early reoxygenation and remain elevated for several hours post recovery, suggesting that the activity of the protein is also required during recovery. In addition to excretion of lactic acid to the media, Arabidopsis roots take up L-lactate from the media and metabolize it (Dolferus et al., 2008; Engqvist et al., 2015).
**Figure VI-1 Comparison of hypoxia-induced expression of NIP2;1 and lactate metabolizing enzyme transcripts LDH and GOX3**

The figures and heat maps were obtained from the University of California Riverside’s Cell Type Specific Arabidopsis eFP Browser based on the cell-specific transcriptome and transcriptome described in Mustroph et al., 2009 and available online at http://efp.ucr.edu/. *Left*, The data represent the cell specific expression of indicated transcripts in the indicated root cell types. *Right*, comparison of total mRNA and translated, polysome associated mRNA for the indicated transcripts.
Since NIP2;1 mediates the bidirectional flux of lactic acid (Choi and Roberts, 2007), it could also assist in the recovery of excreted lactic acid to trigger its metabolism to pyruvate and entry into the TCA cycle as part of the replenishment of TCA cycle intermediates that takes place during post-anoxic recovery (Branco-Price et al., 2008; Tsai et al., 2014; Yeung et al., 2019).

**NIP2;1 localization during hypoxia stress and recovery**

Previous work with NIP2;1 expression in heterologous systems with non-native promoters under unstressed conditions showed conflicting results, with localization to the plasma membrane (Choi and Roberts, 2007), internal membranes (Mizutani et al., 2006), or a mixture of both locations (Wang et al., 2017) observed in different experiments. By using the complementation lines with a NIP2;1-GFP transgene under the native promoter, we were able to establish more clearly the localization of the protein during hypoxia stress and recovery. The work shows substantial, if not exclusive, localization of NIP2;1-GFP to the plasma membrane, both during hypoxia as well as during the first hours of reoxygenation recovery. This observation suggests that the protein is involved in the efflux of lactic acid from the site of production (cytosol) to the apoplastic space. Additional details on the subsequent pathway of lactic acid movement, including the participation of other transporters, that ultimately lead to the directional release of toxic fermentation product from the root to the media remain undetermined. By analogy to boric acid homeostasis, which involves the collaboration of NIP channels and BOR transporters (Takano et al., 2008), NIP2;1 may be part of a larger transport network that coordinates lactic acid efflux and release.
The processes that govern spatio-temporal localization of NIP2;1 to internal compartments vs. the plasma membrane remains an open question. The trafficking of plant aquaporins to various target membranes through endocytic and redistribution pathways is regulated based on metabolic need and stress physiology [reviewed by (Chevalier and Chaumont, 2015; Takano et al., 2017)]. For example, PIP2;1 aquaporins are dynamically cycled between the internal membranes and the plasma membrane (Li et al., 2011b), with regulation via phosphorylation (Boursiac et al., 2008; Prak et al., 2008) or other factors (Santoni, 2017; Takano et al., 2017) leading to preferential surface expression or internalization, which regulates the hydraulic conductivity of the cell. The dual localization of NIP2;1 could reflect a similar dynamic distribution and trafficking between internal membranes and the cell surface to regulate lactic acid efflux. In the case of some plant (Prak et al., 2008) as well as mammalian (Noda and Sasaki, 2006) aquaporins, preferential trafficking to the plasma membrane is controlled by the phosphorylation of serine within the cytosolic carboxyl terminal domain. Proteins of the NIP I subgroup are phosphorylated on a homologous serine within the carboxyl terminal domain (Wallace et al., 2006; Santoni, 2017), which is catalyzed by CDPK/CPK kinases (Weaver et al., 1991). This phosphorylation motif is conserved in NIP2;1 (Ser 278). In addition, phosphoproteomic analysis reveals that NIP2;1 is also phosphorylated in the N-terminal domain at Ser 5 by an unidentified protein kinase (Vialaret et al., 2014). Whether phosphorylation, or other regulatory factors, control trafficking or distribution of NIP2;1 in response to hypoxia or recovery signals to regulate lactic efflux or uptake merits further investigation.
Lactic acid and ethanolic fermentation pathways

Ethanolic fermentation through the PDC-catalyzed decarboxylation of pyruvate followed by subsequent production of ethanol from acetaldehyde via ADH is proposed to be the major anaerobic catabolism pathway (Gibbs and Greenway, 2003). However, lactic acid fermentation is also carried out in most plant species, and in many cases may precede, and regulate the transition to ethanolic fermentation (Gibbs and Greenway, 2003). The reason for initial reliance on lactic acid fermentation during hypoxia prior to a shift to ethanolic fermentation is not clear (Gibbs and Greenway, 2003). However, this pathway, unlike ethanolic fermentation, could allow recovery of the fermentation end product. Additionally, lactate production via LDH occurs under aerobic conditions in response to other abiotic and biotic stresses that could affect energy metabolism (Dolferus et al., 2008; Maurino and Engqvist, 2015). In animal systems, the physiological role of lactic acid/lactate transcends serving as an end product for anaerobic glucose metabolism, and its larger role as a metabolic regulator has emerged, including G-protein signalling as well as transcriptional regulation through histone modification (Latham et al., 2012; Sun et al., 2017; Zhang et al., 2019).

The possibility of lactic acid/lactate as a signal in plant systems remains largely unexplored. Nevertheless, there is evidence that the balance between the two fermentation pathways is regulated. For example, in classical studies of maize root tips, plants subjected to hypoxia stress initially engage in lactic acid fermentation followed by a switch to primarily ethanolic fermentation that is proposed to be driven by cellular acidification by lactic acid accumulation or other means that results in subsequent pH-dependent activation ethanolic fermentation [the “pH stat” model (Davies et al., 1974;
In the case of Arabidopsis, overexpression of *LDH* results in an increase in the activities of some ethanol fermentation enzymes (Dolferus et al., 2008), suggesting that increased lactic acid fermentation induces this alternative fermentation pathway. Conversely, *adh1* null plants induce higher levels of lactic acid production to apparently compensate for reduced flux through the ethanolic fermentation pathway (Ismond et al., 2003). In *nip2;1* mutants, the accumulation of higher tissue lactic acid/lactate may produce an effect similar to *LDH1* overexpression. Higher transcript levels encoding enzymes within alternative pathways of the metabolism of pyruvate (e.g., *ADH1* and *AlaAT1*) may be an adaptive response to the accumulation of lactic acid in *nip2;1* roots.

The reason for the selective reduction of *GOX3* inhibition in hypoxic *nip2;1* roots is less clear. As pointed out by Engqvist et al. (2015), the proposed role of this enzyme is to convert lactate back to pyruvate within the peroxisome which would serve to reduce lactic acid levels within the cell. Notably, however, this conversion occurs with the production of a reactive oxygen end product (hydrogen peroxide). ROS production is a major contributor to reoxygenation stress and is associated with poor tolerance to hypoxia and recovery (Yeung et al., 2019). If cytosolic lactic acid/lactate levels are elevated in *nip2;1* mutants, *GOX3* (which uses oxygen as a co-substrate) could trigger greater ROS production upon reoxygenation.

**Aquaporins as lactic acid channels in other plant and microbial systems**

In addition to Arabidopsis NIP2;1, select aquaporins with lactic acid permeability and efflux function have been described in other systems. For example, the Lactobacillales, which produce large quantities of lactic acid through fermentation,
possess isoforms of the glycerol facilitator encoded by the *GlpF1* and *GlpF4* that facilitate lactic acid efflux (Bienert et al., 2013). The human trematode pathogen, *Schistosoma mansoni*, which performs lactic acid fermentation during the pathogenic part of its life cycle, possesses a lactic acid permeable plasma membrane aquaporin SmAQP that is proposed to release this end product (Faghiri et al., 2010). More pertinent to the present study, recent work (Mateluna et al., 2018) has identified other NIP I proteins, PruavNIP1;1 and PrucxmNIP1;1, that are induced during low oxygen stress in hypoxia-tolerant *Prunus* root stocks, and which are proposed to be lactic acid permeable proteins based on yeast lactate auxotroph assays. These observations suggest that a subset of the NIP I family may have a biological function in lactic acid efflux.

From a phylogenetic perspective, the NIP2;1 subfamily is mostly restricted to plant lineages of the Brassicaceae family (Fig. VI-2). Since Arabidopsis NIP2;1 retains the ar/R pore constriction properties of other lactic acid-impermeable NIP aquaglyceroporins (e.g., nodulin 26, Choi and Roberts, 2007), the substrate preference for lactic acid over other substrates likely include pore structural features besides the canonical selectivity filter. Additionally, some closely related NIP2;1 orthologs in other Brassicaceae lineages such as *Brassica napus*, readily transport metalloids such as hydroxides of arsenite and boron, and these transport functions cannot be excluded for other NIP2;1 subfamily members (Diehn et al., 2019). Additional analyses are required to determine whether NIP2;1 orthologs in Brassicaceae are core hypoxia-response genes and share the lactic acid permeability properties of the Arabidopsis NIP2;1.
Figure VI-2 Phylogenetic analysis of Arabidopsis NIP2;1. The phylogenetic tree was created using NIP2;1 as the candidate to search its orthologs and paralogs from different plant species using the PhyloGenes webpage available online at http://www.phylogenesis.org/ (Zhang et al., 2020). The green circle represents the speciation node, brown circles shows duplication, blue diamond represents subfamily node and grey triangle represents the collapsed node that contains phylogeny of several genes mentioned near it. For constraint of the space these distance relatives are collapsed. NIP2;1 is highlighted in red and its close orthologs and the clade is demarcated with rectangle.
Resolving the NIP conundrum: The need for more structural data

The multiple atomic-resolution structures of aquaporins reveal that while the hourglass fold is a conserved feature, each new structure provides unexpected surprises, additional unforeseen pore constrictions, and unique structural characteristics that impart distinct functions to each protein. This is particularly critical for the various members of the NIP family studied here, many of which have unique substrate specificity as well as regulatory features for which a structural mechanism is lacking. A major step forward in the last year is the publication of the first atomic resolution crystal structure of a NIP protein, Lsi1, from the NIP III pore subclass (Saitoh et al., 2021; van den Berg et al., 2021). As discussed above, we have used this structure to build refined models for NIP I and NIP II pores, but additional structures of representatives of this class is essential and critical for providing mechanistic knowledge beyond conventional homology modeling. To this end, we attempted to generate diffraction quality crystals of the NIP I archetype, the symbiotic aqua-ammonia porin nodulin 26 from *Glycine max*.

Membrane protein structure elucidation by x-ray crystallography is a daunting undertaking that is inherently empirical and difficult for a number of reasons. First, the production of large quantities of membrane protein in heterologous systems in an active and folded state is a challenge. Second, membrane proteins need to be solubilized from their original phospholipid environment and purified in active, folded, and homogenous monodisperse state that is stable for crystallography screens. Third, conditions need to be found that produce membrane protein crystals that are isomorphous, well organized, and diffract to a resolution (<3Å) that allows refinement and model construction. Membrane protein crystallographers face a number of challenges beyond that
associated with soluble proteins. Protein-free micelles, for example, can disrupt protein–protein interactions and lower crystallization success rates. Furthermore, detergents and membrane mimetics cover the majority of the membrane protein (hydrophobic region) while leaving only a limited surface area (loops and hydrophilic region) for crystal contact formation. This is particularly a problem with highly hydrophobic proteins such as nodulin 26 in which most of the protein mass is imbedded within the detergent micelle. Further, crystals generated from detergent-solubilized proteins are frequently linked with low-resolution diffraction or crystallographic defects such as anisotropy or twinning (overlapping diffraction patterns from two closely linked but heterogeneous crystals) that confound refinement and model building (Kermani et al. 2021).

We were able to overcome the expression problem by using a codon-optimized synthetic gene and expression in the methanotrophic yeast *Pichia pastoris*. Further, we were able to solubilize and purify full length nod26 with detergents having low CMCs like LMNG (0.001 wt%), DMNG (0.0034 wt%) and OGNG (0.058 wt%). This novel family of sugar-based neopentyl glycol (NG) amphiphiles was discovered to be capable of promoting protein stability and crystallization with other membrane proteins due to their chemical composition and low CMCs (Chae et al., 2010; Rasmussen et al., 2011a; Rosenbaum et al., 2011; Chung et al., 2012). Given that detergent CMC values decrease with increasing detergent alkyl chain length (i.e., detergent hydrophobicity), the correlation observed for the nod26 between detergent CMC value and detergent stabilization efficacy suggests that a detergent with high hydrophobicity could effectively maintain the native structure of this channel protein. The increased PDC stability
coinciding with lower CMC detergents indicates that the primary means of solubilized nod26 loss is through aggregation rather than denaturation.

Often, while detergents are successful at protein solubilization/purification, they are not ideal for crystallization applications. This was the case for nod26. One critical factor to consider is the size of the detergent micelle. Detergents that form large micelles tend to envelop the hydrophilic surfaces of proteins, limiting the development of crystal contacts, thus preventing the formation of well diffracting crystals (Lisa et al., 2006). In our case, the use of the LMNG detergent produced the most stable PDC with nod26 of all the detergents tested. Interestingly, LMNG micelles appear to be significantly larger than those of the other NG analogues, with a hydrodynamic radius more than double that of micelles formed by DDM, a traditional detergent utilized for membrane protein studies (Cho et al., 2015). This property could explain why most of the nod26-LMNG crystals failed to diffract x-rays. The inability of the nod26-LMNG complexes to associate into tightly-packed, repeating units due to the disproportionally large size of the detergent micelle was deemed a likely cause of the poor diffraction properties of the crystals.

To address this issue, we screened a number of analogues of LMNG that share the low CMC (high hydrophobicity) characteristic, but have shorter alkyl chain lengths that reduce the size of the detergent micelle. This approach was effectively employed by Frick et al. (2014) for improving x-ray diffraction resolution for other members of the aquaporin family. Substitution of DMNG for LMNG (8 carbon alkyl chain vs 10 carbon alkyl chain) showed promise for improving the diffraction of solubilized nod26. DMNG has a detergent micelle with a hydrodynamic radius that is less than half the size of
LMNG micelles while maintaining a high degree of protein stability over time. We were able to obtain nod26-DMNG crystals under many different conditions and temperatures by using the sitting drop vapor diffusion method. With this condition, we achieved the best results with x-ray diffraction to an apparent resolution of ~4.0 angstroms. However, the diffraction pattern showed severe anisotropy indicating that the crystal organization was not uniform and that the data would not be useful for further refinement and model building.

Additional avenues for improving diffraction remain to be explored. The data collected suggests that crystals grown in the conditions showing diffraction could be grown to a larger size (using larger drop volumes) which may allow for more constructive interference and thus a higher resolution/more complete diffraction. Additionally, the incorporation of various additives to the crystallization conditions may improve the quality and size of the crystals. These additives have the ability to change the micelle character and improve the packing/stability of the nod26 crystals. Additionally, a truncated version of the protein could be created to remove the first 33 of the N-terminal residues of nod26 since they are not expected to be visible in the crystal structure (hypothesized from multiple alignments with other aquaporin structures) and may be interfering with the ability of the protein to pack in crystals to diffract to higher resolution. The recent articles elucidating the structures of the NIP type III, Lsi1, protein required limited proteolysis (van den Berg et al., 2021) or N and C-terminal deletion along with several points mutations (Saitoh et al., 2021) was necessary to alleviate issues with poor crystal diffraction due to anisotropy and lattice translocation defects, respectively. Lastly, there are synchrotron beamline facilities like the GM/AC@APS
micro-focus beamline (Argonne National Laboratory, Argonne, IL) that use micro-beams of x-rays to facilitate the acquisition of high-resolution structures. These beamlines use tunable microbeams to “scan” the crystal for areas of optimal diffraction and have been applied to difficult projects with heterogeneous crystals, resulting in the identification of many high impact structures (Rasmussen et al., 2007; Fleishman et al., 2011; Rasmussen et al., 2011b; Wei et al., 2012).

As a final consideration, recent breakthroughs in cryo-electron microscopy (cryo-EM) have transformed the area of membrane protein crystallography and has enabled the structure of biological macromolecules to be solved at near atomic resolution. Cryo-EM, adds a number of desirable, novel aspects to structural biology that were not previously available with existing methods of high-resolution structure determination (Cheng et al., 2017). The approach enables macromolecules to be investigated in more "native," i.e., physiologically relevant buffer conditions, rather than only the circumstances under which the molecules may crystallize. Off-pathway conformations will not be selected for by crystal packing forces, hence the observed conformations are likely to be functionally meaningful. Furthermore, structures for macromolecules in two or more functional states that are in equilibrium with one another can be determined. Finally, the approach requires only a few microliters of sample at concentrations as low as tens of nanomolar. Taken together, these advances make it possible to solve structures for macromolecules that were previously too difficult to crystallize.
List of References


Benga G (2009) Water channel proteins (later called aquaporins) and relatives: past, present, and future. IUBMB Life 61: 112–33


Chiba Y, Mitani N, Yamaji N, Ma JF (2009) HvLsi1 is a silicon influx transporter in barley. The Plant Journal 57: 810–818


Hove RM, Bhave M (2011) Plant aquaporins with non-aqua functions: deciphering the signature sequences. Plant Molecular Biology 75: 413–430


Klebl F, Wolf M, Sauer N (2003) A defect in the yeast plasma membrane urea transporter Dur3p is complemented by CpNIP1, a Nod26-like protein from zucchini (Cucurbita pepo L.), and by Arabidopsis thaliana delta-TIP or gamma-TIP. FEBS Lett 547: 69–74


altered cellular priorities during hypoxia in Arabidopsis. Proceedings of the National Academy of Sciences 106: 18843–18848


Altered Localization of NIP5;1, a Plasma Membrane Boric Acid Channel, Reveals the Requirement for d-Galactose in Endomembrane Organization. Plant and Cell Physiology 55: 704–714


VITA

Zachary Gardner Beamer was born in Knoxville, Tennessee on May 1, 1986. After graduating high school, Zach enrolled in the College of Arts and Sciences at the University of Tennessee where he obtained a Bachelor of Science in both Biological Sciences and Chemistry in the spring of 2012. Following his interest in biochemistry, he pursued a Ph.D. in the Biochemistry & Cellular and Molecular Biology Department at the University of Tennessee in 2013. Zach served as a Graduate Teaching Assistant during his tenure at UTK from August 2013 to May 2022.

Zach joined the laboratory of Dr. Dan Roberts in June 2014 and began his dissertation work focusing on the NIP subfamily of membrane protein channels. During his dissertation work, he characterized the biological role of NIP2;1 in Arabidopsis and additionally examined the more general transport properties of NIP I and NIP II proteins in Arabidopsis using various biochemical and computer modeling techniques.