Structural Mechanism Underlying the Plant Dual-affinity Nitrate Transceptor, CHL1

Ji Sun

A dissertation submitted in partial fulfillment
of the requirements for the degree of

Doctor of Philosophy

University of Washington

2013

Reading Committee:

Ning Zheng, Chair

Wenqing Xu

Williams Zagotta

Program Authorized to Offer Degree:

Department of Pharmacology

University of Washington

Abstract

Structural Mechanism Underlying the Plant Dual-affinity Nitrate Transceptor, CHL1

Ji Sun

Chair of the Supervisory Committee:

Prof. Ning Zheng

Department of Pharmacology

Nitrate is a key nutrient and signaling molecule for plant development and growth. It is the main source of inorganic nitrogen for plants grown in the aerobic soil condition. The growth of many species, including cultivated crops, is strictly dependent on the effective processes of nitrate uptake under frequent fluctuations of nitrate levels in soil. Nitrate is also an important signal molecule that affects plants' metabolism and development. In order to accommodate the changing nitrate concentrations in the soil, which can be varied by four orders of magnitude, plants have evolved a dual-affinity transporter system, NRT1 (low-affinity) and NRT2 (high-affinity) families. Previous studies have shown that *Arabidopsis thaliana* CHL1 (NRT1.1), which belongs to the low affinity family, functions as a dual-affinity nitrate transporter and sensor, whose high- and low-affinity states are switched upon phosphorylation of a key residue Thr101. However, the following two questions remain unknown: 1) What is the structural mechanism underlying the dual-affinity nature of CHL1; and 2) What is the molecular mechanism that governs the integration of nitrate transporter and sensor function in one protein,

CHL1. Therefore, I solved the crystal structure of CHL1 at atomic resolution, which reveals a face-to-face inward dimer with the key residue T101 buried in the dimer interface. In chapter 2, I confirm that functional CHL1 indeed dimerizes on the cell membrane with a cell-based fluorescence resonance energy transfer (FRET) assay, and further show that decoupling of the dimer by phosphorylation mimetic mutation on Thr101 is sufficient to convert CHL1 into a monophasic high-affinity transporter. In chapter 3, I revealed a potential secondary substrate-binding site on the intracellular side of CHL1, which may be important in nitrate induced signal transduction and regulation. Together with structural analysis and mutational assays, this study provides a structure basis for the working mechanism of transceptor.

TABLE OF CONTENTS

Page

List of Figures

List of Tables

Acknowledgements

This work would not have been possible without the guidance and support of my advisor Ning Zheng, which is reflected in every part of this dissertation. I would like to thank my collaborator John Bankston and William Zagotta. Without their contributions, this work won't be as complete and exciting. I also want to thank the Zheng lab and Xu lab for their advice and expertise, especially Wenqing Xu, Thomas Hinds, Jian Payandeh, Haibin Mao, Ti Li, Weiman Xing, Xiaobo Tang, Kenneth Garbutt, Nabiha Saifee, Laura Sheard, Heng Li, Vivian Tran, Zhizhi Wang and Zhihong Chen. I greatly appreciate the advices from all my committee members, including William Catterall, Richard Gardner, Wenqing Xu, and William Zagotta. At last, this dissertation would never have been completed without the encouragement and help from all my friends.

Dedication

For my parents.

CHAPTER 1. Introduction to CHL1 (NRT1.1)

Nitrogen is a fundamental building block of amino acids and nucleotides, which are essential constituents of all life. Although certain bacteria species are able to capture nitrogen from the atmosphere through nitrogen fixation, plants acquire nitrogen mainly by assimilating nitrate, the dominant mineral form of nitrogen in aerobic soil. Because animals are incapable of either nitrogen fixation or assimilation, they rely on other organisms, particularly plants, to obtain the essential element in the organic forms.

Nitrate is an essential element for plant growth, both as a primary nutrient in the nitrogen assimilation pathway and as an important signal for various aspects of plant development, including lateral root branching, flowering time, and seed dormancy (Alboresi et al., 2005; Castro Marin et al., 2011; Zhang and Forde, 2000). As sessile organisms, plants have developed sophisticated mechanisms to ensure appropriate adaptations to constantly changing environmental conditions. In order to accommodate the fluctuating concentrations of nitrate, plants evolved complicated nitrate transport systems.

1.1 Introduction to plant nitrate transport system

Nitrogen (N) can be available to plant roots in several different forms, but mainly in the forms of nitrate (NO_3^-) and ammonium (NH_4^+). In a typical aerobic agricultural soil, both nitrate and ammonium are present, but nitrate is usually the major form. However, nitrate is not only more abundant, its concentrations in soil are also more variable, which can differ by four orders of magnitude (from μM to >10 mM) (Xu et al., 2012). To cope with the variability of nitrate concentrations in soil, plant roots have developed really

sophisticated nitrate transporter systems with distinct kinetic properties. The high affinity transport system (HATS), which consists of members of the NRT2/NNP nitrate transporter family, drives nitrate uptake with K_M values in the μM range, while the low affinity transport system (LATS), constituted by the NRT1/PTR family of nitrate transporters, transports nitrate at mM concentrations. NRT1 and NRT2 together constitute a large NRT family in plants (Tsay et al., 2007). Take *Arabidopsis Thaliana* as an example; there are 53 NRT1 genes and 7 NRT2 genes in this family (Fig 1.1). Recent studies of these transporters have provided insights into the molecular mechanisms of nitrate uptake and allocation. Interestingly, several of these transporters also play versatile roles in nitrate sensing (Ho et al., 2009), plant development (Krouk et al., 2010), pathogen defense, and/or stress response (Camanes et al., 2012). In my thesis, I will focus on the genetically and biochemically best characterized nitrate transporter, CHL1 (also named as NRT1.1).

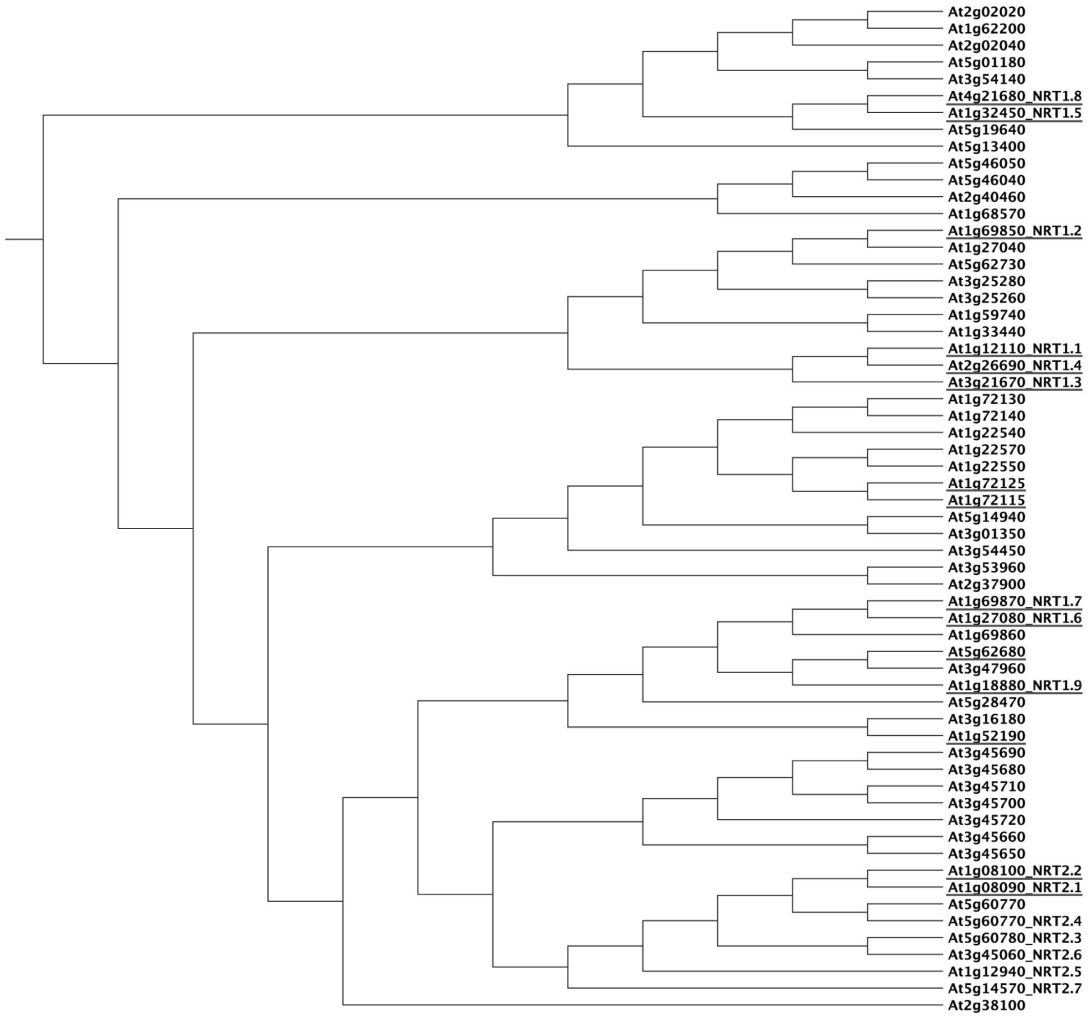

Figure 1.1 Phylogenetic tree of the *Arabidopsis* NRT family. Genes marked with underlines are characterized nitrate transporters. The phylogenetic tree was constructed using the neighbor-joining method and drawn in CLC Sequence Viewer 6.

1.2 CHL1 is a member of the Major Facilitator Superfamily (MFS)

Despite of the importance of nitrate, little was known about nitrate assimilation before the cloning of the *Arabidopsis chl1* gene in 1993 (Tsay et al., 1993). The Arabidopsis *chl1-1* mutant is resistant to the herbicide chlorate (ClO_3^-), a nitrate analog, but shows normal activity in nitrate reduction, suggesting a defect in nitrate transport. Cloning of the *chl1* gene led to the identification of its protein product CHL1, which was immediately proved to be a nitrate transporter.

CHL1 is characterized by 12 transmembrane domains and belongs to the Major Facilitator Superfamily (MFS) (Fig. 1.2a) (Law et al., 2008; Pao et al., 1998), which is a large, and diverse transporter family with over 10,000 sequenced members. Substrates for the MFS include nutrient compounds, metabolites, drugs, neurotransmitters, peptides, osmolites, etc. CHL1 shares sequence homology with a subfamily of MFS known as proton-coupled oligopeptide transporters (POT). Recent structural studies of several prokaryotic MFS members, including GlpT, LacY, EmrD, $PepT_{so}$ and FucP (Abramson et al., 2003; Dang et al., 2010; Huang et al., 2003; Newstead et al., 2011; Yin et al., 2006), have revealed important features of this superfamily of transporters, including their overall assembly, inward-facing, occluded, and outward-facing conformational states, as well as structural determinants for substrate specificity. A "rocker-switch" working model was proposed, suggesting that substrate might be fixed in space at its binding site, and alternating access involves a "rocker-switch" type movement of the N-ter and C-ter halves of the protein (Fig. 1.2b) (Dang et al., 2010).

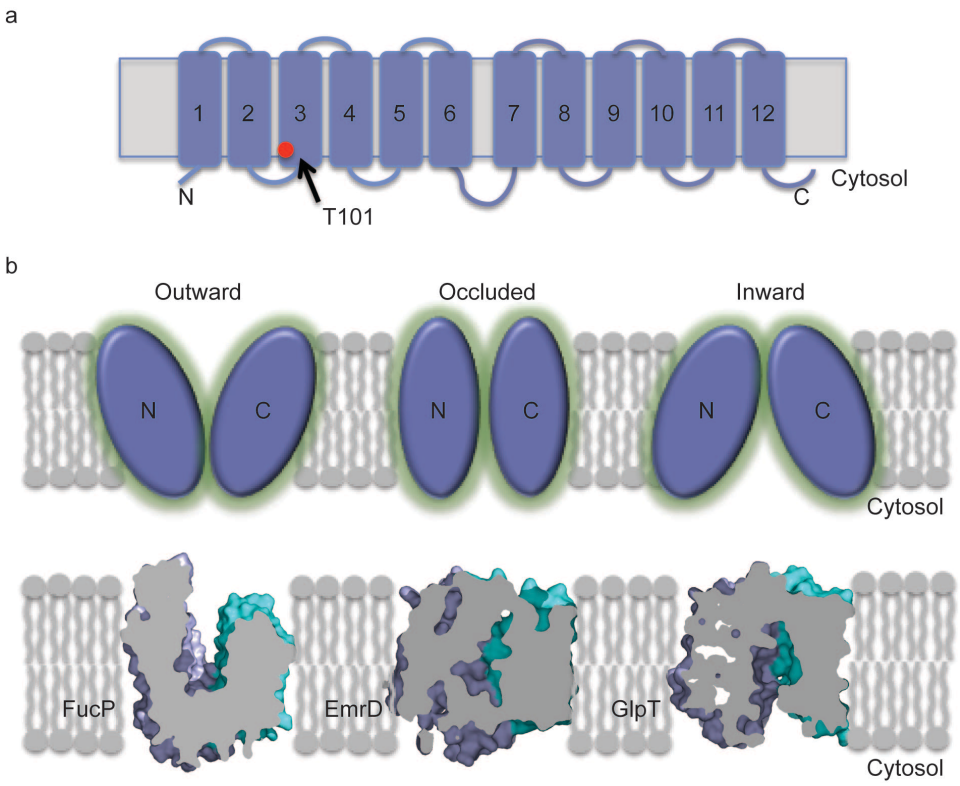

Figure 1.2 The "Rocker-Switch" model of MFS. a, The domain structure of CHL1. Transmembrane helix 1-6 (TM1-6) constitutes the N-ter domain of CHL1, and TM1-7 forms the C-ter domain. Red dot stands for the Thr101. **b,** The "Rocker-Switch" model of MFS with representative crystal structures of three different conformations. The substrate-binding pocket (between N-ter and C-ter domains) of MFS is either open toward outside (Outward), inside (Inward) or occluded from both sides (Occluded).

As a member of POT, $PepT_{so}$ shares 29% sequence identity with the N-terminal half of CHL1. The rest of the CHL1 protein, however, has a distinct and much longer sequence that is distant to the prokaryotic POT member. To understand how CHL1 functions as a nitrate transporter, structural characterization of the protein will be essential. Besides being a eukaryotic transporter, CHL1 is expected to have a unique architecture with special structural elements for its ligand specificity and proton coupling.

1.3 CHL1 is a dual-affinity transporter

CHL1 also distinguishes itself from most other MFS members by being a dual-affinity transporter (Huang et al., 1996; Liu et al., 1999; Liu and Tsay, 2003; Wang et al., 1998). Before the cloning of CHL1, nitrate uptake systems were categorized into two types according to their kinetic properties: low-affinity system (K_M > 0.5mM) and high-affinity system (0.01mM < K_M < 0.3mM). A series of studies have been performed to characterize the kinetic properties of CHL1, with both high- and low-affinity data reported by different groups. It is now clear that CHL1 is in fact a dual-affinity transporter, which functions at both low and high nitrate concentrations. Importantly, recent studies have shown that phosphorylation of a single residue, Thr101, is responsible for switching CHL1 from the low-affinity to the high-affinity mode (Fig. 1.2a) (Liu and Tsay, 2003).

In the nitrate uptake curve, CHL1 displays a biphasic kinetic pattern with a K_M of ~100 μM for the high-affinity mode and ~5 mM for the low-affinity mode (Liu and Tsay, 2003). Characterization of two CHL1 mutants, T101A and T101D, which mimic the

dephosphorylated and phosphorylated state of the protein, respectively, demonstrates that both CHL1 mutants only exhibit a monophasic curve. The high-affinity phase is absent in T101A, whereas the low-affinity phase is absent in T101D (Fig. 1.3). Therefore, T101 phosphorylation is required for the high affinity uptake of nitrate at low substrate concentration by CHL1, and dephosphorylation of T101 is required for low affinity uptake of nitrate at high substrate concentration. T101 phosphorylation appears to function as a switch, converting the transporter between two forms.

Why CHL1 was evolved to have the dual-affinity function while plant roots encode other nitrate transporters with different affinities (e.g. NRT1 and NRT2 families)? One explanation could be that plants lack mobility and have to respond rapidly to the changing environments. Plants are known to reprogram gene expression to survive different conditions. If a single transporter can function effectively with highly variable substrate concentration, it will leave plants enough time to undergo the necessary transcriptional and cellular changes to adapt. However, the molecular mechanism underlying how the phosphorylation modification on T101 can change the transporter affinity remains unclear.

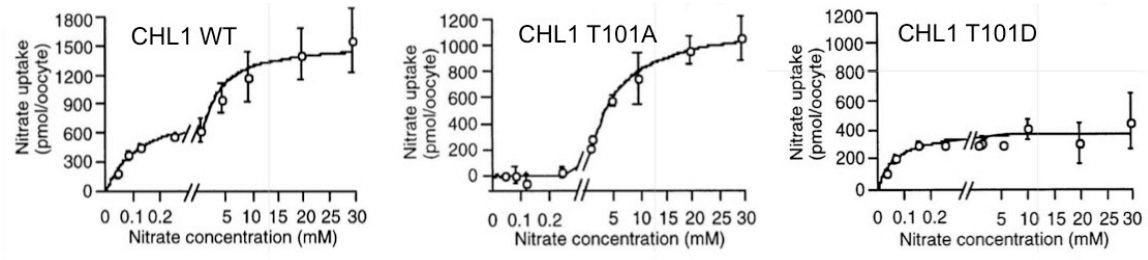

Figure 1.3 Kinetic analysis of nitrate uptake by wild type CHL1 and T101 mutants. The uptake assay is done in oocytes. Left, the uptake curve of wild type CHL1; middle, the uptake curve of CHL1_T101A, the unphosphorylated mutant; right, the uptake curve of CHL1_T101D, the phosphorylated mimetic mutant. (Adapted from Liu, KH *et al*. EMBO J, 2003).

1.4 CHL1 is a dual-affinity sensor

Nitrate is not only a nutrient but also a signaling molecule, which modulates many aspects of plant physiology and optimal nitrate acquisition (Guo et al., 2003; Ho et al., 2009; Munos et al., 2004). How do plants perceive different nitrate concentrations and initiate signaling events that ultimately regulate gene expression? Surprisingly, a 2009 groundbreaking work from the Tsay's group revealed that the nitrate transporter CHL1 also functions as a nitrate sensor in plant. In this study, Tsay et al. used NRT2.1 as a marker gene, whose transcriptional level changes as the nitrate concentration varies. As shown in Fig. 1.4, the expression level of NRT2.1 in wild type plants exhibits a biphasic pattern, responding to the nitrate concentration increases in the soil. In the CHL1 deletion mutant (*chl1-5*) background, however, the expression of the NRT2.1 marker gene drops to no nitrate baseline. CHL1, thus, must play a key role in nitrate signaling. Since CHL1 is a nitrate transporter, it remained highly possible that the unresponsive expression level of NRT2.1 in the CHL1 deletion mutant is mostly due to the failure of nitrate uptake by the mutant plant, which would naturally lead to defective nitrate sensing by some downstream sensors.

Remarkably, the authors identified a second CHL1 mutant (*chl1-9*), which allows them to decouple CHL1's transporter and sensor functions. In the decoupled mutant *chl1-9* (carrying a missense mutation P492L), CHL1 loses its function as a transporter, but retains its ability to respond to nitrate concentration changes and to regulate the expression of downstream genes. In fact, its response curve for the marker gene, NRT2.1, expression perfectly matches the wild type (Fig. 1.4). All together, these results proved that CHL1 is able to regulate nitrate-related gene expression and function as a

nitrate sensor independent of its transporter activity. Thus, it acts as a *bona fide* nitrate receptor.

Although the sensor function of CHL1 is independent of its transport function, the two functions seem to share overlapping regulation mechanisms. The phosphorylation state of T101 in CHL1, which controls the dual-affinity transporter function, is also involved in regulating its receptor function. CHL1_T101D, the phosphorylation mimetic mutation, retains the high-affinity mode of the primary gene expression response, but is defective in the low-affinity mode. CHL1_T101A, on the other hand, shows no defect in neither the high- nor low-affinity phases. In fact, this non-phosphorylable mutant displays noticeably enhanced responses in both phases. Together, these results suggest that phosphorylation of T101 is dispensable for both high and low affinity nitrate sensing and/or signaling functions of CHL1, yet dephosphorylation of T101 is critical for the protein to sense and/or transduce the signal of high concentration of nitrate (i.e. low affinity sensing function). In other words, both the phosphorylated and non-phosphorylated forms of CHL1 can sense and signal low concentration of nitrate with high affinity, but only the non-phosphorylated form can sense and signal high concentration of nitrate with low affinity.

Tsay et al. also identified CIPK23 (CBL-Interacting Protein Kinase 23) as a protein kinase that can interact with and phosphorylate CHL1 *in vitro*. *Cipk23* defective mutant reduces the phosphorylation of CHL1 on T101 and has a similar phenotype as CHL1_T101A. Therefore, CIPK23 is thought to be responsible for the phosphorylation of CHL1 *in vivo* when nitrate concentration is low. An elusive phosphatase is expected

to dephosphorylate T101 at high nitrate concentration to convert CHL1 from a high affinity nitrate sensor into a low affinity sensor.

CHL1 functions both as a transporter and sensor, thus established the transceptor concept in plants (Gojon et al., 2011). Although, lots of genetic and biochemical studies have been done to characterized CHL1's properties, the answers to the following questions are still largely unknown. First, as a eukaryotic member of MFS-POT, is the overall structure of CHL1 similar to its prokaryotic homologs, like $PepT_{so}$ and $PepT_{St}$? And how substrate specificity is achieved in this transporter? Secondly, what is nature of a dual-affinity switch for CHL1? How can the posttranslational modification (phosphorylation) change the substrate affinity of a transporter? Third, what is working mechanism of a transceptor? How are two functions (transporter and sensor) incorporated in one protein? My thesis, combining multiple methods including X-ray crystallography, mutation studies, FRET assay and biochemical characterization, provides molecular insights into the first two questions and shed light on the working mechanism of transceptors.

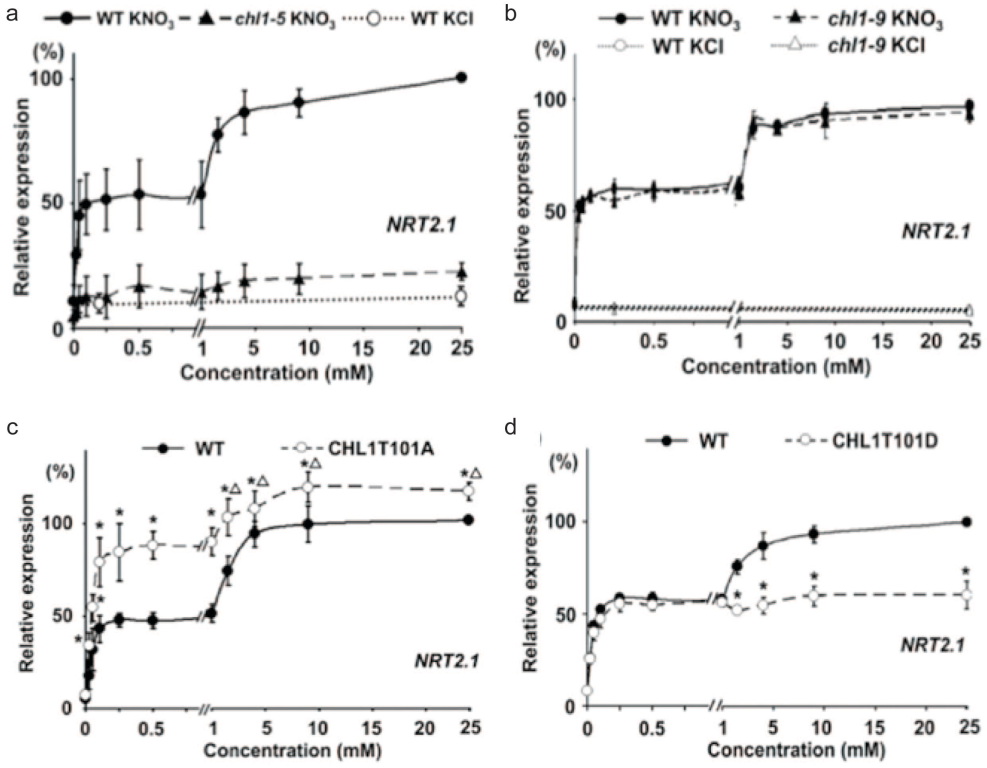

Fig. 1.4 CHL1 functions as a nitrate sensor and is regulated by phosphorylation. **a,** The relative expression level of NRT2.1 displays a biphasic curve as nitrate concentration increases in WT, while the NRT2.1 RNA level in the deletion mutant *chl1-5* is close to the no nitrate background. **b,** The transport-defective mutant *chl1-9* shows similar marker gene expression responses as WT. **c,** CHL1T101A enhances the both high- and low-affinity primary nitrate response; **d,** CHL1T101D has the high-affinity phase of the primary response. (Adapted from Ho, CH, *et al* Cell, 2009).

CHAPTER 2. Structure of a Plant MFS Transporter

Part of the material in this chapter is under publishing:

Sun, J., Bankston, J.R., Hinds, T.R., Payandeh, J., Zagotta, W.N., and Zheng, N. (2014), Crystal structure of a plant dual-affinity transporter. Nature.

2.1 Introduction

Members of the Major Facilitator Superfamily (MFS) constitute the largest superfamily of secondary active transporters (Pao et al., 1998). Its diverse members generally function as symporters or antiporters driven by proton or sodium gradients. Structures of about ten bacterial MFS transporters have been determined by two-dimensional and three-dimensional crystallography (Abramson et al., 2003; Dang et al., 2010; Doki et al., 2013; Huang et al., 2003; Newstead et al., 2011; Pedersen et al., 2013; Solcan et al., 2012; Sun et al., 2012; Yan et al., 2013; Yan, 2013; Yin et al., 2006; Zheng et al., 2013). On the basis of these structures, a 'rocker-switch' mechanism was proposed, suggesting that the two symmetry-related domains found in MFS transporters rock back and forth as 'banana-shaped' rigid bodies with the central substrate-binding site as the pivot point.

CHL1 shares sequence homology with a subfamily of MFS known as proton-coupled oligopeptide transporters (POT) (Pao et al., 1998). As a member of POT, $PepT_{so}$, whose crystal structure is available, shares 29% sequence identity with the N-terminal half of CHL1. The rest of the CHL1 protein, however, has a distinct and much longer sequence that is distant to the prokaryotic POT members. Besides being a eukaryotic transporter, CHL1 is expected to have a unique architecture with special structural elements for its ligand specificity and proton coupling. As a member of NRT1 family, structure study of CHL1 will shed lights on how plants assimilate nitrate, a key nutrient for plant growth and development.

2.2 Expression and characterization of CHL1 in vitro

The *Arabidopsis thaliana* CHL1 gene encodes a 590 amino acid protein, which is highly conserved among plant CHL1 orthologs (Fig 2.1), but not *Arabidopsis* NRT1 family members (Fig 2.2). To prepare for structural studies, I first verified that CHL1 construct still retains its characteristic dual-affinity nitrate transporter activity in the established *Xenopus* oocyte system as previously described (Fig 2.3) (Ho et al., 2009; Liu and Tsay, 2003).

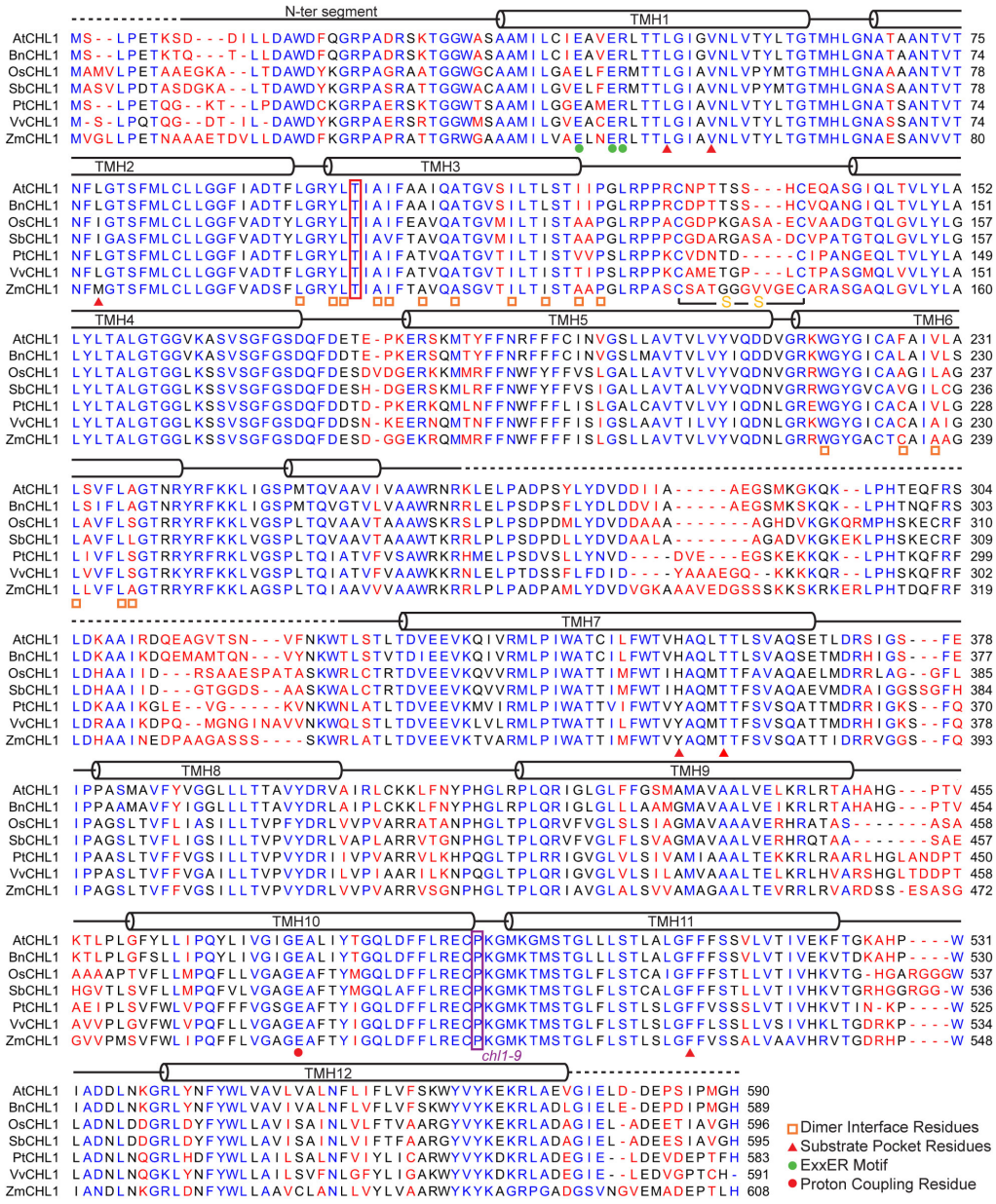

Figure 2.1 Sequence alignment of plant CHL1 orthologs. Alignment and secondary structure assignments of CHL1 orthologs from *Arabidopsis thaliana* (At), *Brassica napus* (Bn), *Oryza sativa* (Os), *Sorghum bicolor* (Sb), *Populus trichocarpa* (Pt), *Vitis vinifera* (Vv), and *Zea mays* (Zm). Strictly conserved residues are colored in blue. Green dots indicate the ExxER motif. Orange empty squares indicate dimer interface residues. Red triangles indicate residues in the substrate-binding pocket. Red dot indicates the energy-coupling residue. Dash lines represent the disordered region in the crystal structure.

Figure 2.2 Sequence alignment of *Arabidopsis* NRT1 family members. Alignment and secondary structure assignments of the NRT1 family members from *Arabidopsis thaliana*. Strictly conserved residues are colored in blue. The two residues potentially important for substrate binding are indicated by red triangle with green stroke. Energy coupling residue is indicated by red dot. Dash lines represent the disordered region in the crystal structure.

2.3 Overall structure of CHL1

The C-terminally tagged recombinant CHL1 protein was then overexpressed, solubilized and purified from *Spodoptera frugiperda* insect cells with DDM and crystallized in the presence of 10 mM $NaNO_3$. CHL1 crystals were grown using hanging-drop vapour diffusion method, using 2 µl protein sample mixed with 1 µl reservoir solution containing 100 mM Sodium Acetate, pH 4.5, 30% PEG300 and 3% MPD. Crystals of maximal sizes were obtained after 1 month. With combined phases from *Rosetta*-improved molecular replacement (DiMaio et al., 2011) and single wavelength anomalous diffraction, we determined and refined the CHL1 structure at 3.25 Å resolution (Table 2.1).

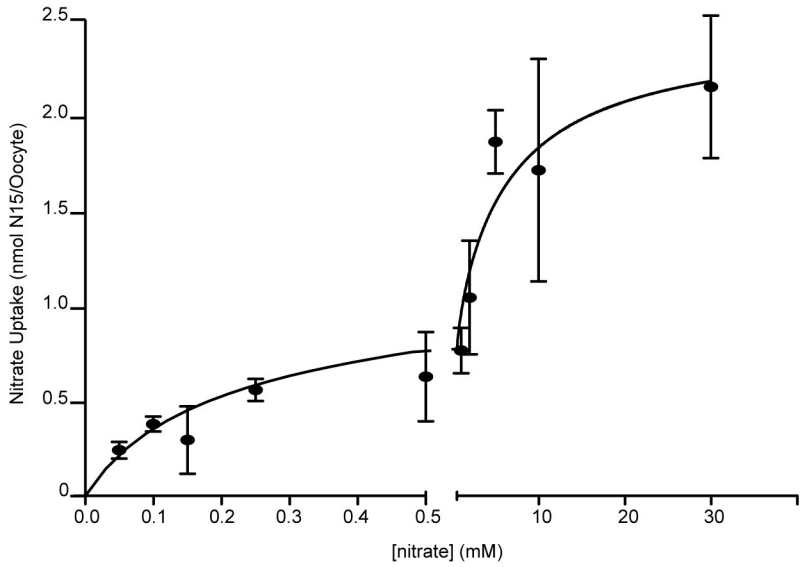

Figure 2.3 Nitrate uptake analysis of CHL1 in oocytes. Measurement of nitrate uptake by CHL1 was carried out in *Xenopus* oocytes in the presence of increasing concentrations of nitrate. Q-test was used to identify statistical outliers in data. All data points are mean ± s.d. of one experiment in quintuplicates or sextuplicates. The data were fit with a two site nonlinear binding curve using Prism.

	Native	Hg Derivative
Data collection		
Space group	C2221	C2221
Cell dimensions		
a, *b*, *c* (Å)	84.8, 188.5, 262.8	84.7, 189.9, 262.8
α, β, γ (°)	90, 90, 90	90, 90, 90
Resolution (Å)	50.0-3.2 (3.26-3.20)*	50-3.50 (3.56-3.50)*
R_{svm} or R_{merge}	0.07 (0.91)	0.08 (0.65)
I/s*I*	17.92 (2.05)	26.96 (1.84)
Completeness (%)	99.1 (94.6)	98.2 (96.5)
Redundancy	5.9 (4.9)	7.0 (5.5)
Refinement		
Resolution (Å)	3.25	
No. reflections	206925	
R_{work}/ R_{free}	0.239/0.305	
No. atoms	7995	
Protein	7947	
Ligand/ion	48	
Water		
B-factors	97.70	
Protein	97.60	
Ligand/ion	112.30	
Water		
R.m.s deviations		
Bond lengths (Å)	0.010	
Bond angles (°)	1.47	

Table 2.1 Summary of crystallographic data analysis. Data collection, phasing and refinement statistics.
*Highest resolution shell is shown in parenthesis.

CHL1 was crystallized in a C2221 space group, with two molecules in the asymmetric unit (ASU). The two copies of CHL1 in each ASU pack in a head-to-head manner with amino termini facing towards each other, and can be superimposed with a root mean square deviation of 0.9 Å over 504 Ca atoms, indicating a common overall structure (Fig. 2.4).

As predicted, the transporter adopts a typical MFS fold, which is characterized by 12 transmembrane helixes (TMHs) in a reverse "V" shape (Fig. 2.5). In detail, eight peripheral helices form a rectangular fence, which acts as a scaffold for the remaining four central helices. Whereas four of the peripheral helices, TMH3 and TMH6 in the N-terminal domain and TMH9 and TMH12 in the C-terminal domain, form the two narrow walls of the fence, the two wide walls are formed by TMH2 and TMH11 in the front, and TMH5 and TMH8 in the back. The four central helices (TMH1, TMH4, TMH7, and TMH10) line up and form the substrate tunnel. The N-terminal (TMH1-6) and C-terminal (TMH7-12) domains can be related by a pseudo 2-fold axis. The two parts are connected by a loop between TMH6 and TMH7, which is partial disordered and invisible in our crystal structure.

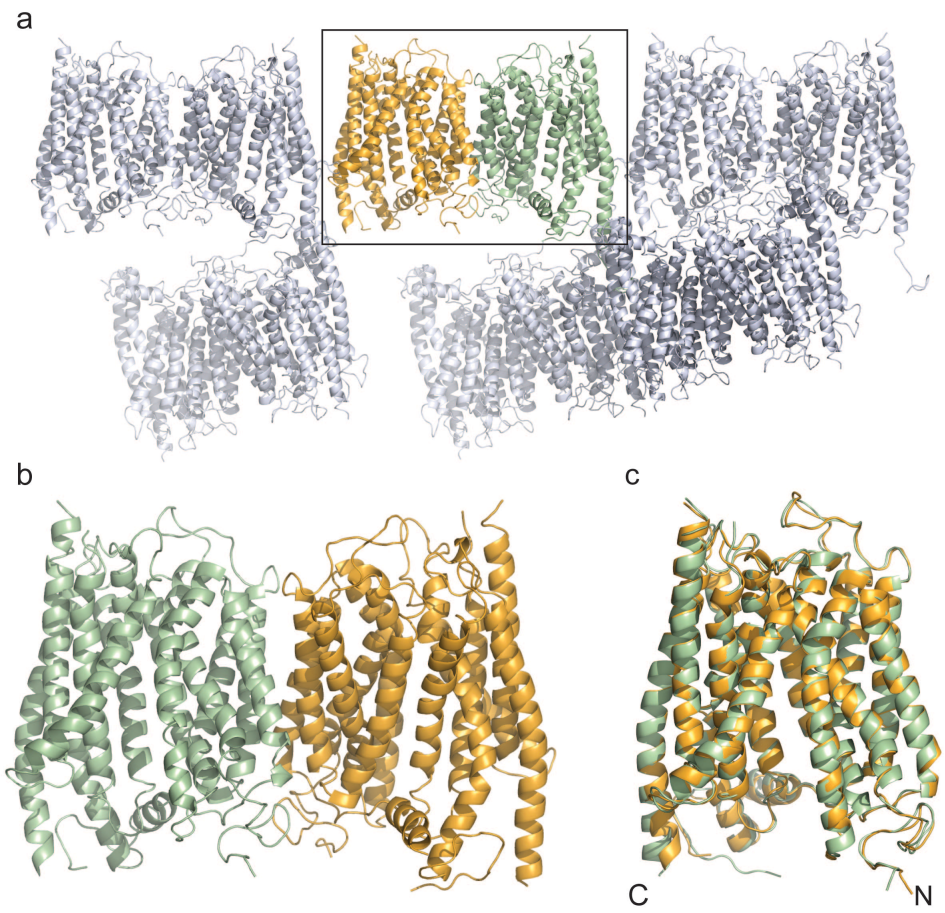

Figure 2.4 Crystal packing and overall structure of CHL1. a, Crystal packing of CHL1 in the space group of C2221. Each asymmetric unit contains two CHL1 molecules, colored in pale green and bright orange. **b,** The two copies of CHL1 in each asymmetric unit. **c,** Superimposition of the two CHL1 molecules in each asymmetric unit.

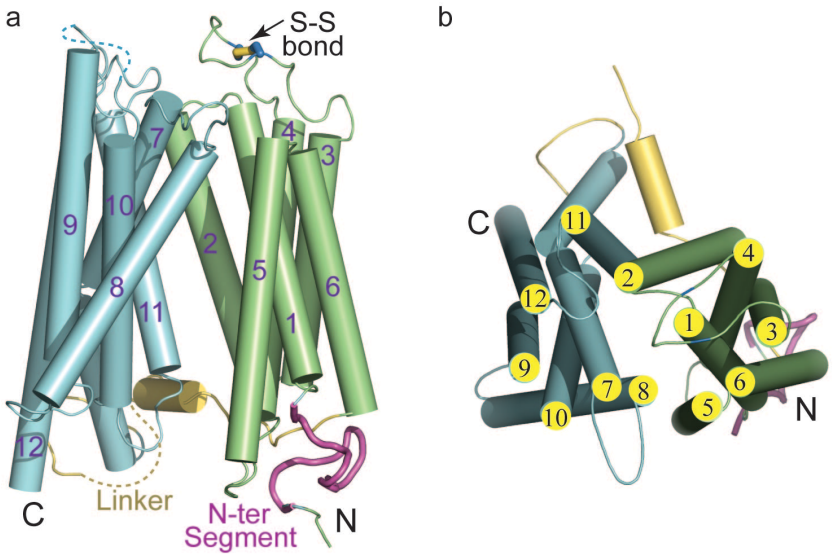

Figure 2.5 Helix arrangement of CHL1. a, Side view of CHL1 helix arrangement in the cylinder representation. The N-terminal and C-terminal domains, the N-terminal conserved segment, and the inter-domain linker are colored in pale green, cyan, magenta and yellow, respectively. The extracellular disulfide bond is indicated. **b,** Extracellular view of CHL1 helix arrangement.

Both CHL1 molecules in the asymmetric unit feature a closed extracellular gate and a central substrate-conducting path open towards the cytosol, suggesting that the transporter is captured in an inward-facing conformation as previously observed in the LacY, GlpT, $PepT_{St}$ structures (Fig. 2.6) (Abramson et al., 2003; Doki et al., 2013; Huang et al., 2003).

By comparing the structures of CHL1 and a bacterial peptide transporter, $PepT_{St}$, we confirm that eukaryotic and prokaryotic members of the NRT1/PTR family of MFS transporters share a similar overall architecture (Fig. 2.7). The plant nitrate transporter, nonetheless, has three unique and conserved structural elements. These include a well-structured N-terminal cytoplasmic segment, a disulfide bond-stabilized extracellular TMH3-TMH4 loop, and a partially ordered central linker sequence, which connects the two halves of the protein (Fig. 2.5). Consistent with the sequence divergence between the NRT1/PTR and NRT2/NNP nitrate transporter families, the structures of CHL1 and two bacterial NRT2/NNP family members, NarK and NarU (Yan et al., 2013; Zheng et al., 2013), share little common feature except the MFS fold (Fig. 2.7).

To sum up, we confirm that CHL1 shares an overall typical MFS fold with its prokaryotic homologues. However, due to the sequence differences, CHL1 has some unique structural features, which we will discuss further in the following chapters.

2.4 Substrate binding pocket

In both protomers of the refined CHL1 dimer structure, an island of strong electron density is present in the middle of the transport tunnel between the N-terminal and C-terminal domains (Fig. 2.8). Although its position is slightly shifted between the two

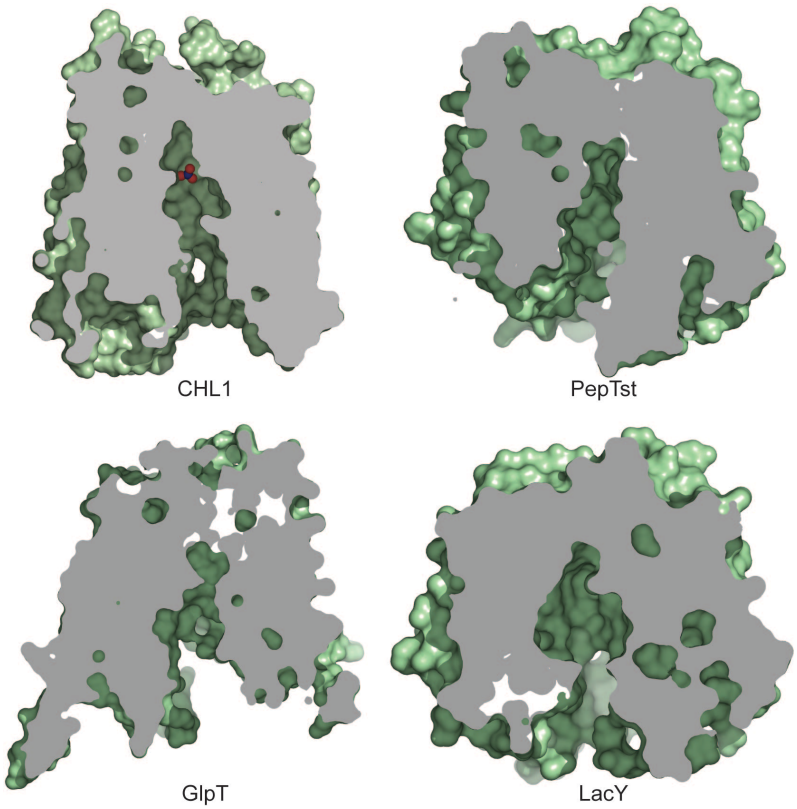

Figure 2.6 Cutaway view of CHL1. CHL1 is captured in an inward conformation with nitrate displayed in spheres. The cutaway views of PepTst, GlpT and LacY are also shown for comparison.

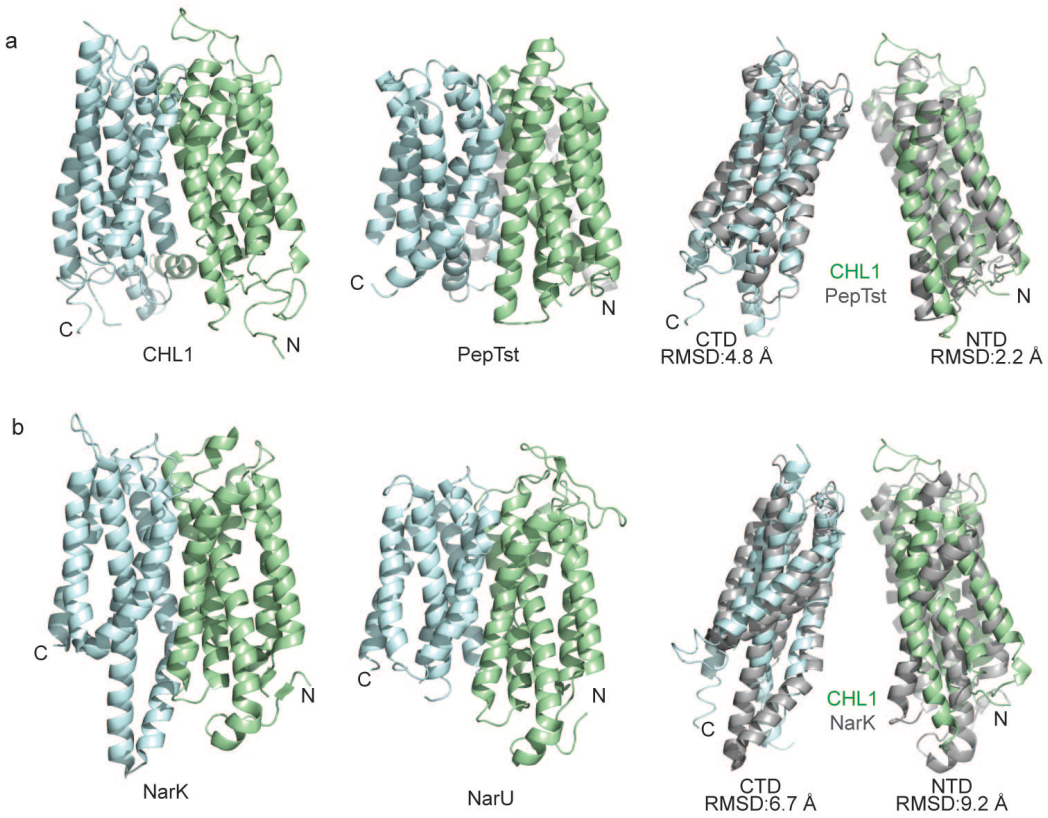

Figure 2.7 Structural comparison of CHL1 and other MSF transporters. a, Overall structural comparison of CHL1 and PepTst with their N-terminal and C-terminal domains (NTD & CTD) colored in pale green and cyan, respectively. Superposition of their NTDs and CTDs are shown separately. **b,** Comparisons between CHL1 and two bacterial NRT2 family nitrate transporters, NarK and NarU.

CHL1 molecules, the overall location of the density coincides with the substrate binding sites of other MFS structures, suggesting that it belongs to the substrate molecule, nitrate. Indeed, when nitrate was omitted from the cryo-protection buffer, this electron density completely disappeared from both CHL1 molecules (Fig. 2.9).

Distinct from the nitrate/nitrite binding sites of NarK and NarU, which coordinate the substrate(s) with two opposing conserved Arg residues (Yan et al., 2013; Zheng et al., 2013), the nitrate-binding pocket in CHL1 is predominantly formed by hydrophobic residues, including Leu49, Val53, Leu78, and Phe511 (Fig. 2.8). His356 on TMH7 is the only polar residue that is in close contact with nitrate, whose precise binding mode cannot be resolved due to the resolution limit of the structure. Based on its close proximity to the nitrate density and the crystallization condition (pH=4.5), His356 likely stabilizes the substrate in the pocket through a charge-charge interaction. Its side chain conformation, meanwhile, is supported by two nearby residues, Tyr388 and Glu476. Although Tyr388 and two other polar residues, Thr360 and Thr48, are also around the substrate, their hydroxyl groups do not appear to be at the optimal hydrogen bond distance (Fig. 2.8).

To validate the substrate-binding site, we mutated His356 and compared the nitrate uptake activities of the wild type and mutant transporters. In support of a critical role of His356 in substrate transport, its mutation to alanine completely abolished the transport activity of CHL1 (Fig. 2.10). A similar effect was also found when the His356-interacting residue, Glu476, was altered (Fig. 2.8 and Fig 2.10). Interestingly, His356 is not conserved among plant CHL1 orthologs and *Arabidopsis* NRT1 family members,

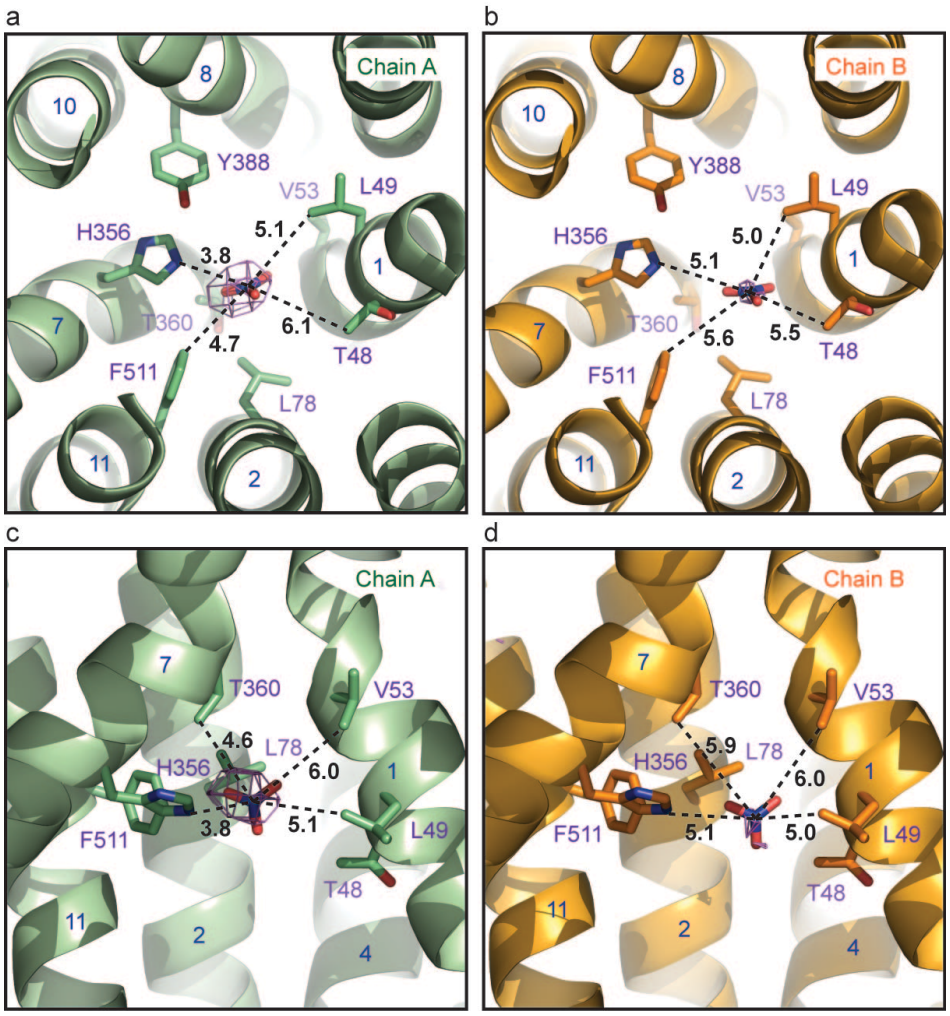

Figure 2.8 Substrate binding pocket of CHL1. a-b, Intracellular view of the substrate binding site in the two copies of CHL1 in each asymmetric unit. To compare the relative position of the substrate to its surrounding residues, distances between the nitrogen atom of the modeled nitrate and select amino acid atoms in its vicinity are shown in dash lines and indicated. Nitrate is shown in sticks with electron density countered at 4σ from a Fo-Fc map calculated without the substrate. TMHs are numbered. **c-d,** Side view of the substrate binding site.

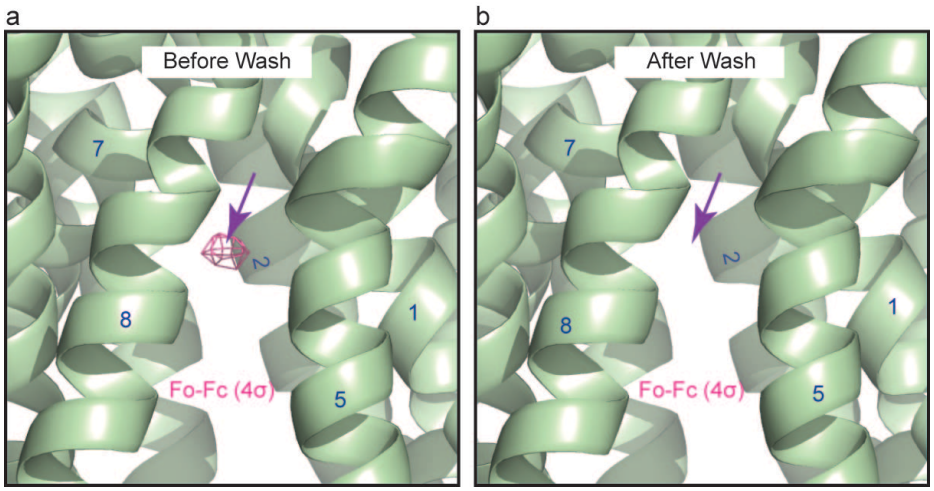

Figure 2.9 Nitrate binding pocket before and after wash with nitrate free buffer. Comparison of the putative substrate density between the CHL1 structures determined with a cryo-protectant solution containing 10 mM or 0 mM nitrate**. a,** CHL1 structure determined with 10mM nitrate. **b,** CHL1 structure determined without nitrate.

which harbor either a tyrosine or a hydrophobic amino acid (Leu/Met/Phe) at the equivalent position (Fig. 2.8). This key residue, nevertheless, has closely co-evolved with the adjacent residue, Phe511. Among all CHL1 orthologs and paralogs, a combination of a polar and a hydrophobic side chain has been generally maintained between the two residues, suggesting that one of them is responsible for specific nitrate binding (Fig. 2.2 and Fig. 2.10a). CHL1 is unique among all *Arabidopsis* NRT1 family members by featuring a histidine at the nitrate-binding pocket. This charged residue provides a plausible explanation for the high affinity nitrate uptake activity acquired by CHL1, which is otherwise a member of LATS. On the other hand, the replacement of the histidine residue by tyrosine in some of the plant CHL1 orthologs raises a question about their dual affinity transporter function (Fig 2.1).

2.5 Energy coupling

In the $PepT_{St}$ structure (Doki et al., 2013), a conserved motif, ExxERFxYY, on TMH1 has been identified to play an important role in proton coupling. Part of this motif, ExxER, is also found in all plant CHL1 orthologs (Fig. 2.1). Together with the conserved residue, Lys164, this motif presents a cluster of interacting residues under the nitrate-binding pocket and facing toward the transport tunnel (Fig. 2.11b). Consistent with a key function in the symport cycle, alanine mutation of each of the four residues abrogated the transporter activity of CHL1 in the oocyte-based nitrate uptake assay (Fig. 2.10b). Surprisingly, these four residues have been simultaneously evolved into non-charged residues in two *Arabidopsis* NRT1 family members, AtNRT1.5 and AtNRT1.8 (Fig. 2.2).

Their documented pH-dependent nitrate transporter activities necessitate an alternative proton coupling mechanism.

Another potential amino acid, Glu476 on TMH10 may involve in proton coupling due to its charging nature. Transporter assays in oocytes indeed show its potential importance in the substrate transport cycle (Fig. 2.10b). However, this residue is also interacting with the key residue H356 (Fig. 2.11) that involves in nitrate binding. It is not sure whether the function disruption is due to deficiency in proton coupling or substrate binding.

2.6 Other structure elements

Outside the core MFS fold, CHL1 has evolved multiple structural elements that are potentially associated with its regulation and versatile functions. These include a well-structured and conserved N-terminal cytoplasmic segment, a disulfide bond-stabilized extracellular TMH3-TMH4 loop, and a partially ordered central linker sequence, which connects the two halves of the protein. We will talk about the N-terminal cytoplasmic segment and central linker sequence in Chapter 3. At the extracellular side above the gate, the C-terminal end of TMH1 is capped by a ~20 amino acids long loop, which links TMH3 and TMH4 and contains a conserved disulfide bond (Fig. 2.5a). Although the central sequence of this loop is highly variable among CHL1 orthologs (Fig. 2.2), elimination of the disulfide bond severely affected nitrate uptake in the oocyte system, implicating an important role of the extracellular loop in regulating the structure and activity of the transporter (Fig. 2.10b).

In the crystal structure, we also see a clear density between TMH5 and TMH8 (Fig. 2.12). We were able to place a DDM head group, maltose, in this density. The binding of DDM is clipped by TMH5 and TMH8 between N-ter and C-ter, and generates a side gate for entering the transporter tunnel.

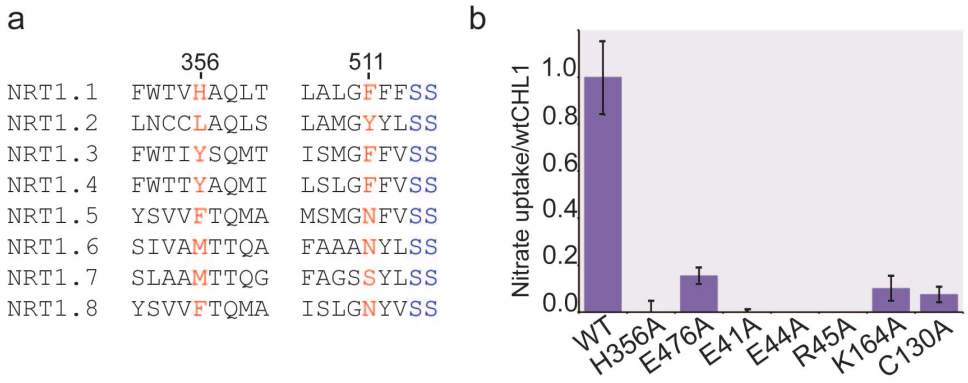

Figure 2.10 Substrate binding residues and uptake assay. a, Sequence alignment of eight NRT1 family members from *Arabidopsis thaliana* showing regions surrounding H356 and F511 of CHL1/NRT1.1, both of which are highlighted in red. Strictly conserved residues are colored in blue. **b,** Relative nitrate uptake activities of CHL1 mutants to the wild type protein measured in the *Xenopus* oocyte-based assay. All results are the mean ± s.d. of one experiment in quintuplicates or sextuplicates.

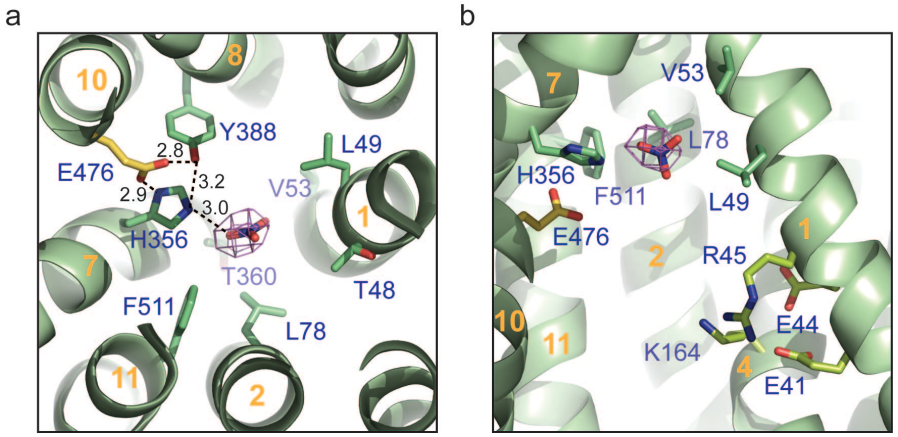

Figure 2.11 Substrate binding and energy coupling in CHL1. a, Intracellular view of the amino acids forming and surrounding the nitrate substrate binding pocket. Nitrate is shown in sticks together with electron density contoured at 4s from a Fo-Fc map calculated before the nitrate was modeled in. Dash lines indicate potential salt bridges hydrogen bonds with inter-atom distances labeled. Transmembrane helices are numbered (organge). E476 (yellow) is potentially important for energy coupling. **b,** Side view of the putative nitrate binding site and the transporter tunnel with the clustered ExxER motif and K164.

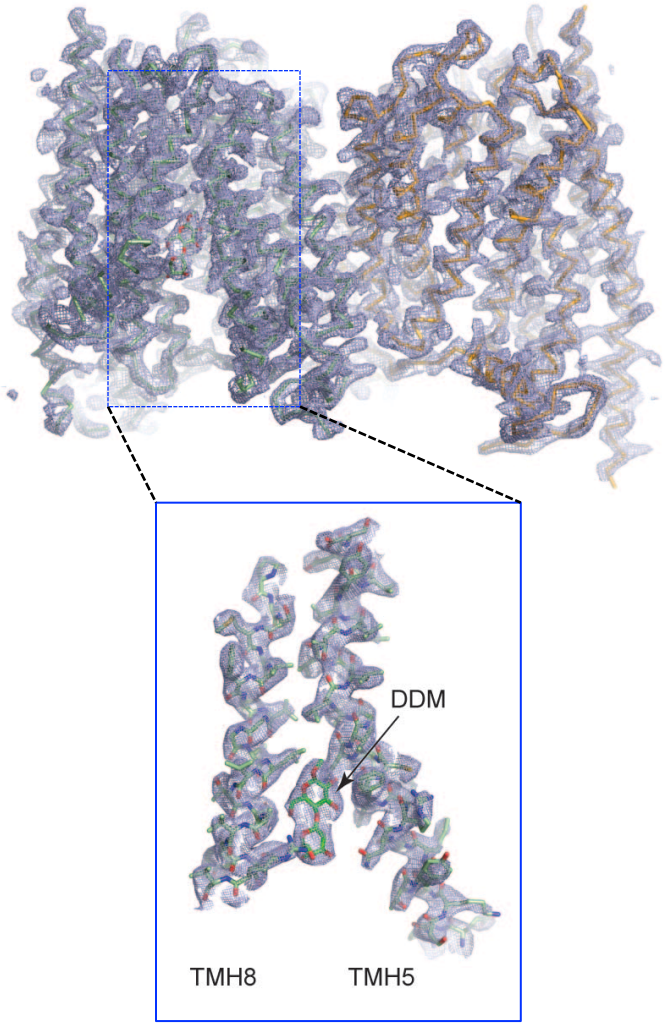

Figure 2.12 DDM binding. An island of density from the 2Fo-Fc map contoured at 1.5σ is assigned to the head group of DDM bound between TMH5 and TMH8.

CHAPTER 3. Phosphorylation Controlled Dimerization Switch

Part of the material in this chapter is under publishing:

Sun, J., Bankston, J.R., Hinds, T.R., Payandeh, J., Zagotta, W.N., and Zheng, N. (2014), Crystal structure of a plant dual-affinity transporter. Nature.

3.1 Introduction

The major facilitator superfamily (MFS) represents the largest group of secondary active membrane transporters, and its members transport a diverse range of substrates. MFS plays critical roles in lots of cellular process, like metabolites uptake, signal transduction, and drug resistance. In order to respond to the changing environment, MFS transporters are regulated through multiple mechanisms, including phosphorylation (Liu and Tsay, 2003), protein-protein interaction (Orsel et al., 2006) and oligomeric states (Veenhoff et al., 2001), which impact different aspects of the transporter function, including cellular localization, stabilization, substrate affinity, and energy coupling. Take oligomeric states as example, numerous transporters in the MFS family have been reported to exist as oligomers, including LacS, GLUT1, TetL and human RFC (hRFC) (Hou et al., 2010; Pessino et al., 1991; Safferling et al., 2003; Veenhoff et al., 2001). The oligomer state could potentially be pharmaceutically important. For example, the GLUT1 protein, a glucose transporter that is considered as a drug target for Type II diabetes, has been reported to form functionally coupled oligomers. The dominant negative effects have been observed in the hRFC transporter and LacS transporter as well. These suggest the importance of homo-oligomerization. Despite the information gained from recent structure studies, the molecular mechanism underlying how the posttranslational modifications and regulation manipulate transporters' function remains largely untouched.

The affinity of CHL1 is regulated by the phosphorylation modification on Thr101. However, the molecular mechanism underlying this regulation is unknown. More specifically, how the phosphorylation modification on Thr101 alters the affinity states of CHL1 remains largely unclear. In this chapter, we show that functional CHL1 indeed dimerizes on the cell membrane and the phosphomimetic mutation of Thr101 converts the protein into a monophasic high affinity transporter by structurally decoupling the dimer. Together with analyses of the substrate transport tunnel, our results establish a phosphorylation-controlled dimerization switch that allows CHL1 to uptake nitrate with two distinct affinity modes.

3.2 CHL1 dimer in the crystal

To date, crystal structures of more than ten MFS transporters have been determined, all of which are in the monomeric form. Prior to crystallization, the DDM-solubilized CHL1 protein was also isolated by size exclusion chromatography (SEC) in a monomeric state as determined by SEC-coupled multi-angle light scattering, differential refractive index, and ultraviolet absorption measurements (Fig. 3.1) (Slotboom et al., 2008). However, a closer examination of the two CHL1 molecules in the asymmetric unit reveals an unexpected dimeric arrangement, hinting at a possible biological dimer. More importantly, the key residue Thr101 whose posttranslational modification alters the affinity of CHL1 also locates at the dimer interface, making the interface more biologically significant if it is a physiological relevant dimer.

In the crystal, the two adjacent noncrystallographically related CHL1 molecules are juxtaposed in a side-to-side fashion with their N-terminal halves facing and interacting

with each other (Fig. 3.2). The intermolecular packing is predominantly mediated by TMH3 and TMH6, which are located at a peripheral edge of the canonical MFS fold. Although crystal contacts do not always reflect biological interactions, two prominent features of the crystallographic dimer arrangement support its physiological relevance.

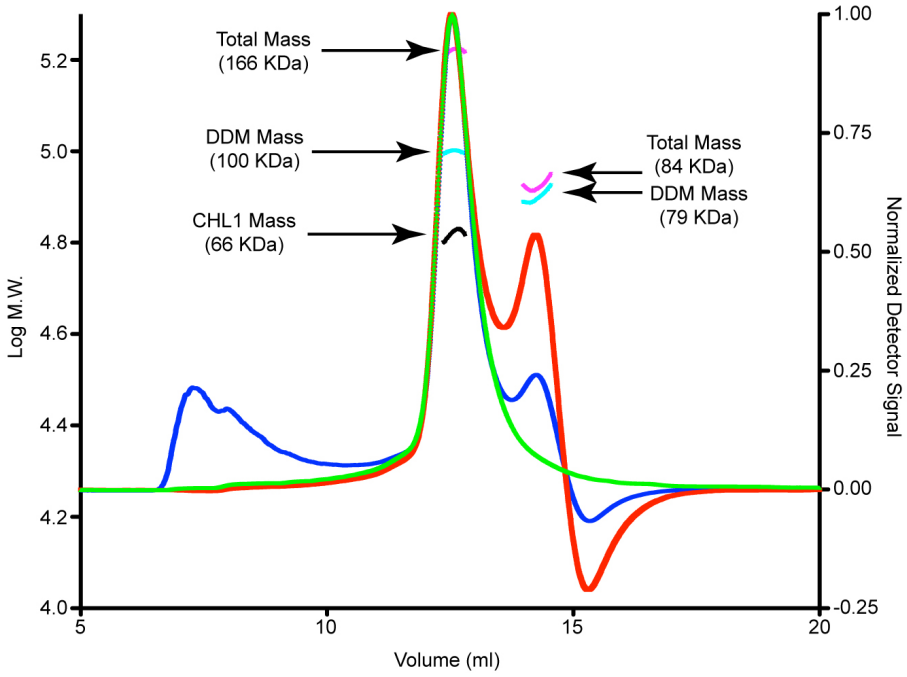

Figure 3.1. SEC-LS-RI-UV analysis of DDM-solubilized CHL1. Size exclusion chromatography elution profile of DDM-solubilized and purified CHL1 protein. Normalized light scattering signal from the 90° detector, UV absorption signal, and the refractive index signals are plotted in blue, green, and red lines, respectively. The highest peak contains CHL1 bound to DDM. The calculated masses of the protein-detergent micelle complex (magenta), DDM micelle (cyan), and the CHL1 protein (black) are shown. The complex contains about 196 detergent molecules (MW 510.62) and 1 CHL1 molecule (67 kDa). The second peak belongs to the detergent micelle with a mass of 79 kDa or 155 detergent molecules.

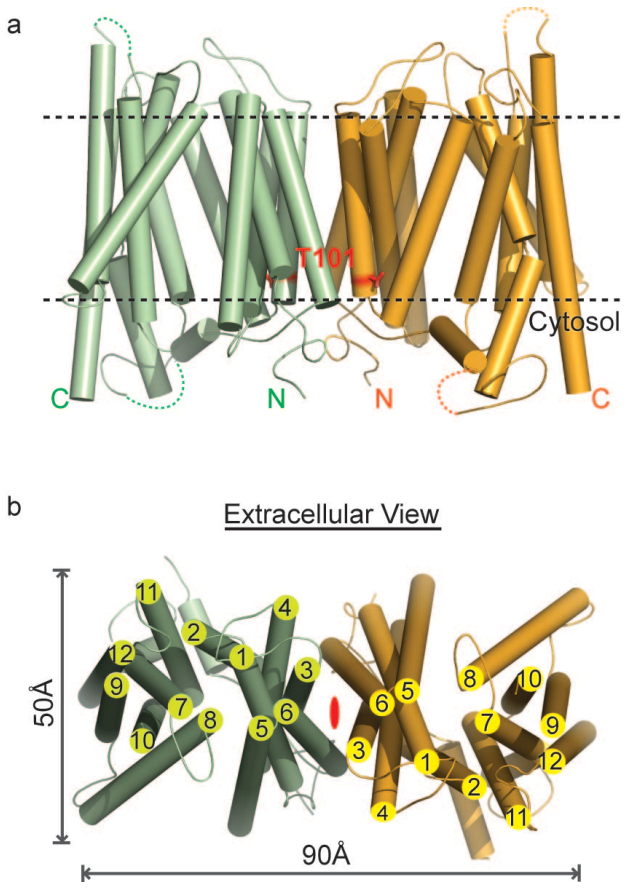

Figure 3.2. CHL1 dimer. a, Cylinder representation of the CHL1 dimer with Thr101 shown in red sticks. **b,** Extracellular view of the CHL1 dimer with a central 2-fold axis indicated in red.

First, the overall topology of the putative CHL1 dimer is perfectly compatible with its transporter function at the membrane (Fig. 3.2a). Within the dimer, the two protomers are related by a central two-fold symmetry axis perpendicular to the membrane plane, which allows both molecules to be simultaneously embedded in the membrane and span the entire lipid bilayer (Fig. 3.2). Second, the interface between the two CHL1 molecules is extensive and complementary, characteristic of biological interactions instead of non-specific crystal packing. The putative CHL1 dimer buries a total surface area of 2160 $Å^2$ and is constructed by a panel of ~36 hydrophobic residues spread along the two TMHs of each protomer. The detailed interaction is shown in Fig. 3.3. The close engagement of the pair of TMH3 and TMH6, which cross the membrane in opposite orientation, creates an anti-parallel four-helix bundle with its intracellular side tightly sealed (Fig. 3.5). Overall, the two inward-facing CHL1 molecules give rise to a putative "in-phase" dimer assembly, which is about 90 Å wide and 50 Å thick (Fig. 3.2). When viewed from the side, the two substrate-transporting tunnels are not in parallel with the central two-fold axis but slant at an ~15° angle in opposite directions (Fig. 3.4).

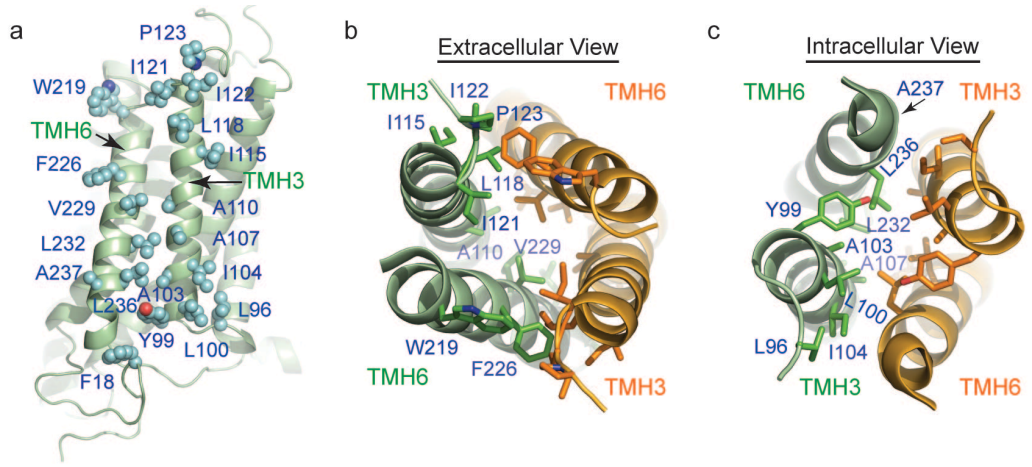

Figure 3.3 CHL1 dimer interface. a, The side chains of all dimer interface residues are shown in cyan spheres and labeled. **b-c,** Extracellular and intercellular views of the CHL1 dimer interface with TMH3 and TMH6 shown in ribbon and the side chains of interacting residues shown in stick. The interface residues of the chain colored in green are labeled.

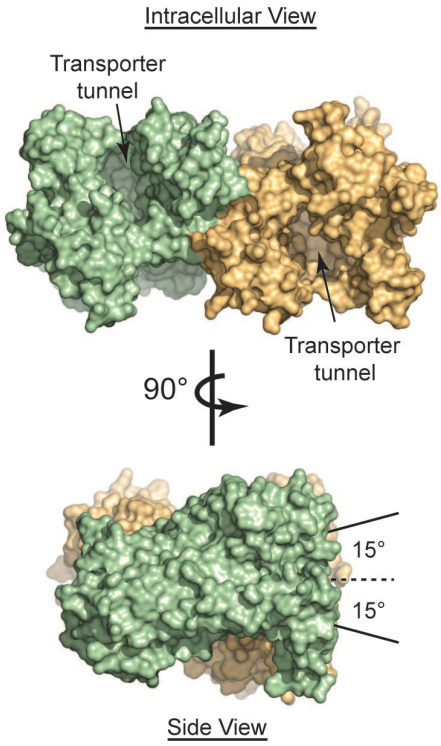

Figure 3.4 Transporter tunnel angle with two orthogonal views of the CHL1 dimer in surface representation. The dash line represents the central 2-fold axis. The two solid lines indicate the approximate directions of the transporter tunnels in the two protomers of the dimer.

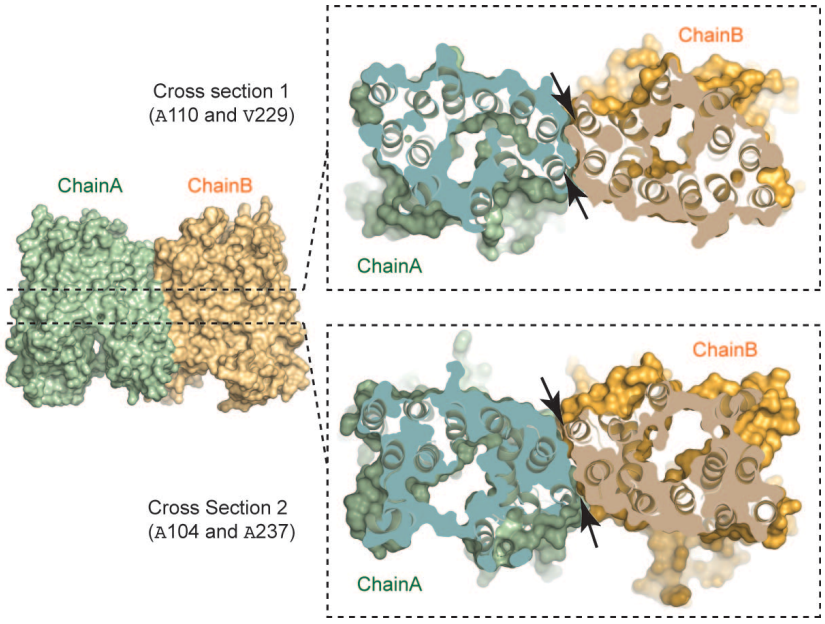

Figure 3.5. Shape complementarity of the TMH3-TMH6 four-helix bundle. Two representative cross section views of the CHL1 dimer interface that are parallel to the membrane. The top cross section goes through the plane defined by Ala110 and Val229 in the two CHL1 protomers. The bottom cross section goes through Ala104 and Ala237. The interface is indicated by black arrows.

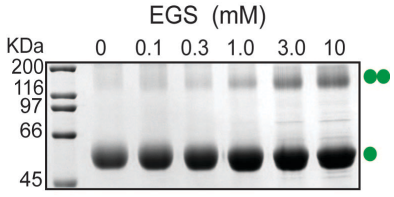

Figure 3.6 Crosslinking of CHL1 in increasing concentrations of EGS. The monomeric and dimeric species of CHL1 are indicated by single and double green circles.

3.3 Functional dimerization of CHL1

To dissect the biological relevance of the CHL1 dimer observed in the crystal, we first used a crosslinking experiment to assess the potential of detergent solubilized CHL1 to dimerize in solution. Despite its low efficiency, an amine reactive crosslinker (ethylene glycol bis-succinimidylsuccinate (EGS), about 16 Å in length) was able to crosslink CHL1 in a concentration dependent manner (Fig. 3.6). The crosslinked products migrated on SDS-PAGE with a size corresponding to a CHL1 dimer, indicating that DDM-solubilized CHL1 is capable of forming a weak and transient dimer in a membrane-free environment.

Solubilization by DDM allows the purification of CHL1 but might also interfere with its dimer formation. We next performed Förster resonance energy transfer (FRET) spectroscopy experiments with the nitrate transporter expressed on the membrane of *Xenopus* oocytes (Taraska and Zagotta, 2010; Zheng et al., 2002). FRET occurs between two fluorophores in close proximity, which is a non-invasive technique to monitor formation of multi-protein complexes or homo-oligomers so that it can be used to detect protein interactions in the cell. More importantly, it enables us to examine the oligomeric state of CHL1 in the same lipid environment where its dual-affinity transporter activity has been detected. We separately fused the N-terminus of two CHL1 constructs with the mCerulean variant of cyan fluorescent protein (mCFP) and the mCitrine variant of yellow fluorescent protein (mYFP), which constitute a FRET pair with an R_0 of 50 Å for 50% energy transfer efficiency (Fig. 3.8). In the structure of the putative CHL1 dimer, the N-terminal ends of the two CHL1 molecules are about 42 Å away from each other (Fig.

3.7). Therefore, FRET is expected to occur if CHL1 dimerizes on the membrane in the same fashion as seen in the crystal structure.

As shown in Fig. 3.8b, strong FRET signal measured by a spectrum-based approach was indeed detected between the co-expressed mCFP-CHL1 and mYFP-CHL1 fusion proteins, but not in the negative control between mCFP-HCN and mYFP-CHL1. This result strongly suggests that the plant dual-affinity nitrate transporter can form a homo-oligomer. Together with the crystal structure and crosslinking experiment, we conclude that CHL1 forms as home-dimer not only in the crystal but also in the oocyte cellular membrane.

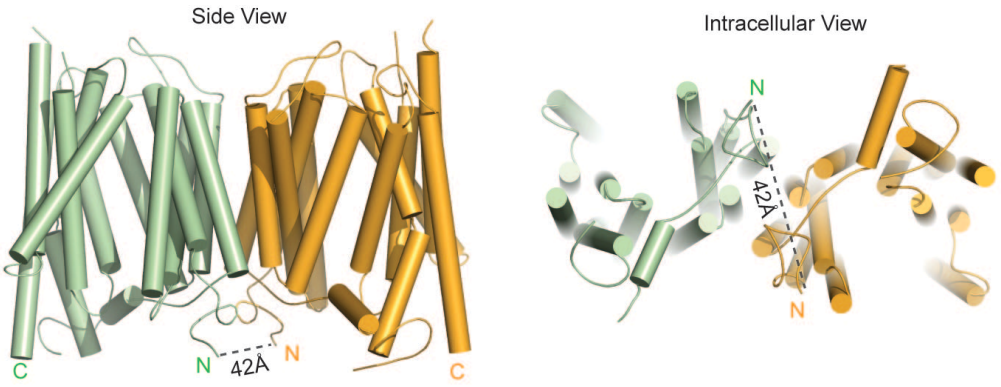

Figure 3.7 Spatial relationship of the N-termini in the two CHL1 protomers. The N-termini of the two CHL1 protomers in the crystal structure are about 42 Å away apart and are shown in two direction orientations.

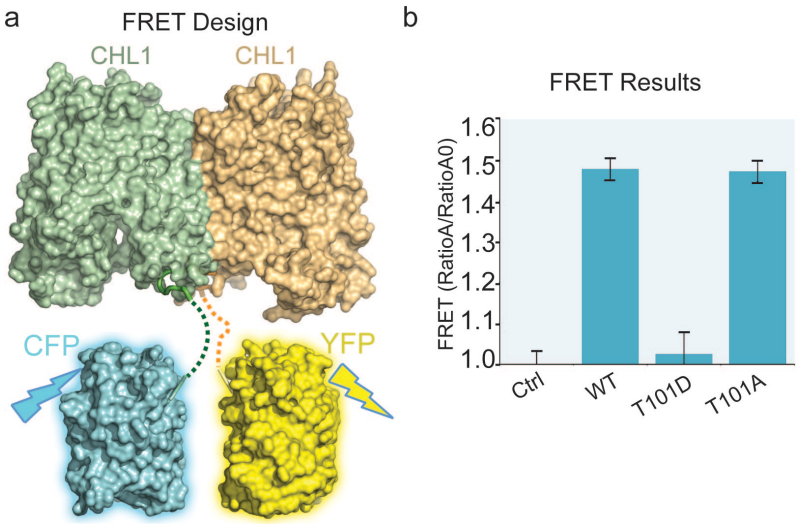

Figure 3.8 FRET experiments. a, The design of FRET assay for CHL1 dimerization. CHL1 and the fluorescence proteins are displayed in surface representation at the same scale. Dash lines indicate the eleven amino acids long linker between the fluorescence proteins and the structurally resolved CHL1 N-terminus. **b,** FRET measurements of wild type and mutant CHL1. A pair of fusion protein, mCFP-HCN and mYFP-CHL1, was used as negative control.

3.4 Thr101 phosphorylation as a dimerization switch

The phosphorylation site residue, Thr101, is strictly conserved among plant CHL1 orthologs and represents one of the hallmarks of the dual-affinity nitrate transporter (Fig. 2.1). In the CHL1 structure, Thr101 is located at the N-terminal end of TMH3 and is entirely buried in a hydrophobic pocket formed among TMH2, TMH3, and THM4 (Fig. 3.9a). Strikingly, this pocket is directly adjacent to the dimer interface with one of its walls demarcated by three hydrophobic interface residues, Leu96, Leu100, and Ile104 (Fig. 3.9b). While the side chain of Thr101 is unmodified in the crystal, its phosphorylation is expected to induce major conformational changes in its vicinity and have a direct impact to the dimer interface. This structural clue prompted us to postulate that the formation of the CHL1 dimer might be determined by the phosphorylation status of Thr101 and the two distinct affinity states of the transporter might be enabled by the difference in its oligomerization state.

To test this hypothesis, we compared the wild type CHL1 with the phosphomimetic mutant, T101D, and the phosphorylation defective mutant, T101A, in the oocyte-based FRET experiments (Fig. 3.8b). Consistent with the prediction, the phosphomimetic mutant, T101D, which has been previously shown to bear a monophasic high affinity nitrate transporter activity (Liu and Tsay, 2003), almost completely lose the FRET signal, indicating a spatial separation of the two N-terminal domains in the CHL1 dimer. By contrast, the phosphorylation defective mutant, T101A, which is known to transport nitrate in the low affinity state, generated a robust FRET signal as the wild type protein. The matching levels of FRET between wild type CHL1 and the T101A mutant indicates that the wild type protein is mostly in the unmodified form under the nitrate-free

condition of FRET measurement. This is in agreement with previous studies showing that phosphorylated CHL1 only started to accumulate when plants were exposed to nitrate. Together, these results not only confirm the functional relevance of the CHL1 dimer, but also suggest a dimerization-based switching mechanism for the dual affinity nitrate transporter — unmodified CHL1 forms a structurally coupled homo-dimer and functions as a low-affinity transporter, whereas phosphorylated CHL1 undergoes dimer decoupling and adopts a high-affinity state.

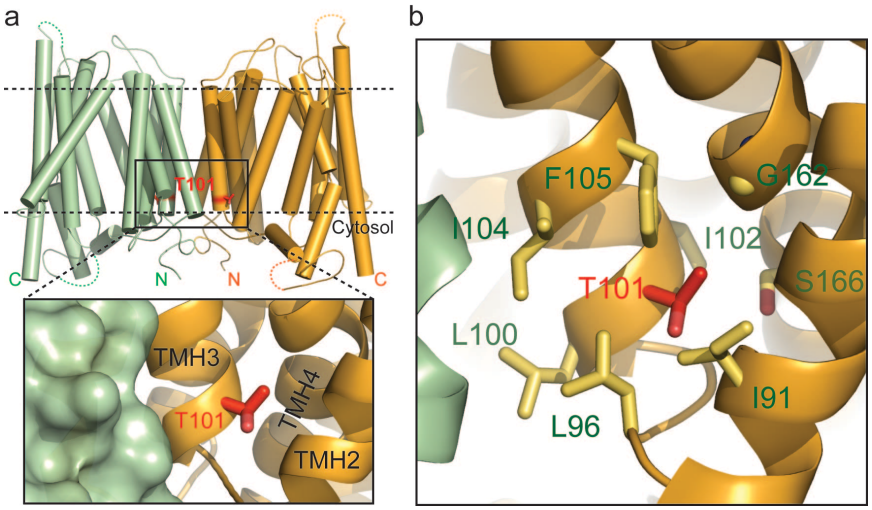

Figure 3.9 T101 and dimer. a, A close-up view of Thr101 at the CHL1 dimer interface. **b,** Thr101-interacting residues. The side chains of Thr101-interacting residues, which form a hydrophobic pocket, are shown in yellow sticks.

3.5 Extramembrane structural elements

Beside the disulfide bond we discussed in Chapter 2, the N-terminal segment and the ~80 amino acids long central linker sequence of CHL1 also contains regions that are conserved among its orthologs (Fig 2.1).

Although the majority of the central linker is disordered in the crystal, its N-terminal region forms a stable amphipathic α-helix, which sticks out of the MFS fold and is predicted to lie on top of the lipid bilayer (Fig. 3.2). Pointing at opposite directions, the linker helices of the two CHL1 protomers present yet another pair of potential protein docking sites that are removed from the dimer interface.

More interestingly, CHL1 protein has an ~30 amino acids long N-terminal cytoplasmic segment, which is highly conserved among its plant orthologs but absent in other NRT1 family members (Fig. 2.1). In the crystal, this sequence adopts a well-ordered loop structure and packs tightly against the C-terminal end of TMH6 and the domain linker sequence following it. This intracellular segment is not extensively involved in the dimer interface, but creates a pronounced cleft between the two protomers of the CHL1 dimer (Fig. 3.10). Facing to the cytoplasm, this well-defined cleft is lined by strictly conserved yet solvent-exposed amino acids, indicative of a putative two-fold symmetric protein-protein interaction site, whose function can be coupled to the CHL1 dimer formation.

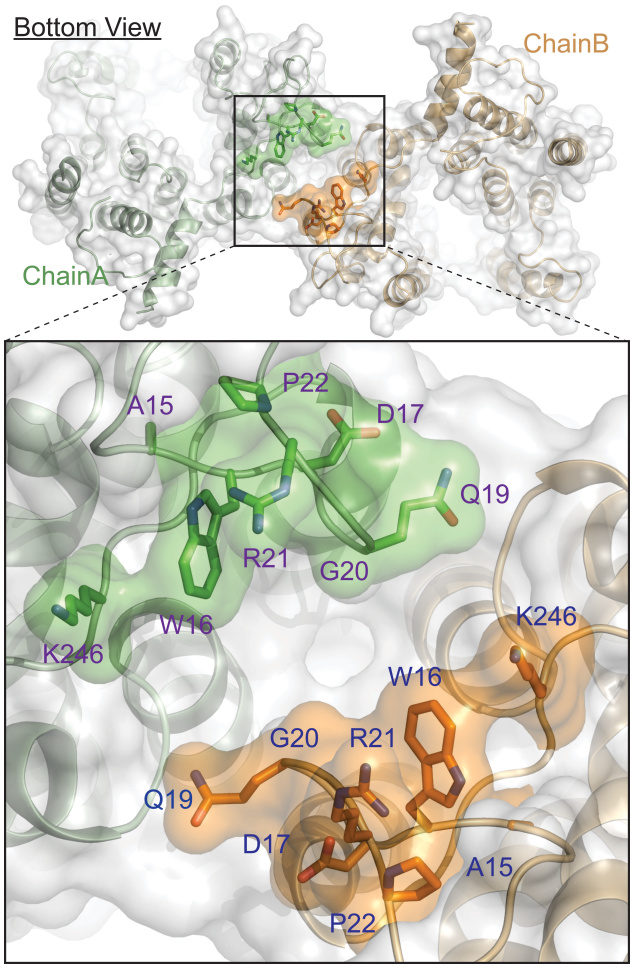

Figure 3.10 The bottom cleft. Overall and close-up views of the N-terminal segments of CHL1 within the dimer are shown. The cleft forming residues, which are strictly conserved in the CHL1 orthologs, are shown in sticks and labeled.

3.6 Working model and discussion

The crystal structure of CHL1 reveals a biologically relevant dimer, whose dynamic coupling and decoupling is controlled by the phosphorylation of a single residue near the dimer interface. Since the same post-translational modification switches the mode of action of the dual-affinity transporter, we propose that dimer assembly and disassembly enables CHL1 to toggle between the low-affinity and high-affinity states. To function as a dual-affinity transporter, CHL1 could have two separate substrate-binding sites with different affinities or a single substrate-binding site with two affinity states. Although we cannot rule out the possibility that CHL1 has a second high affinity substrate-binding site, which is masked in the current structure, the dimerization-based switching mechanism can conceivably create two affinity states with one physical substrate-binding site.

In this model (Fig. 3.11), structural engagement of two protomers at the interface allosterically regulates the affinity of substrate binding at the central transport tunnel. When the outside nitrate concentration changes, the nitrate sensor senses the difference, and then the phosphorylation state of CHL1 is altered. This posttranslational modification further affects the dimer coupling within CHL1, and finally changes the substrate affinity. Whether dimer disassembly itself is sufficient to shift CHL1 into the high-affinity mode awaits future analysis of CHL1 mutants that are defective in both dimerization and phosphorylation.

Previous structural studies of several MFS members have established a “rocker-switch” mechanism of substrate transport, in which the transporters cycle through outward-facing, occluded, and inward-facing conformations. The structure of the dimeric unmodified CHL1 reveals a buried Thr101 phosphorylation site in the inward-facing

conformation. Thr101 phosphorylation, therefore, likely occurs when the dimeric transporter adopts either the outward-facing or occluded conformation. It is, however, equally possible that the unmodified dimeric transporter is in fast equilibrium with the monomeric form, which is susceptible for phosphorylation.

Transporter oligomerization and phosphorylation have been implicated in the proper functions of a number of MFS members, such as LacS, GLUT1, TetL, hRFC, and rOCT1 (Hou et al., 2010; Mehrens et al., 2000; Pessino et al., 1991; Safferling et al., 2003; Veenhoff et al., 2001). The crystal structure of the CHL1 dimer not only establishes a structural framework for understanding its dual-affinity nitrate transporter activities, but also reveals for the first time how protein oligomerization and post-translational modification can synergistically expand the functional capacity of a MFS transporter, revealing a structurally unprecedented mechanism of membrane transporter regulation.

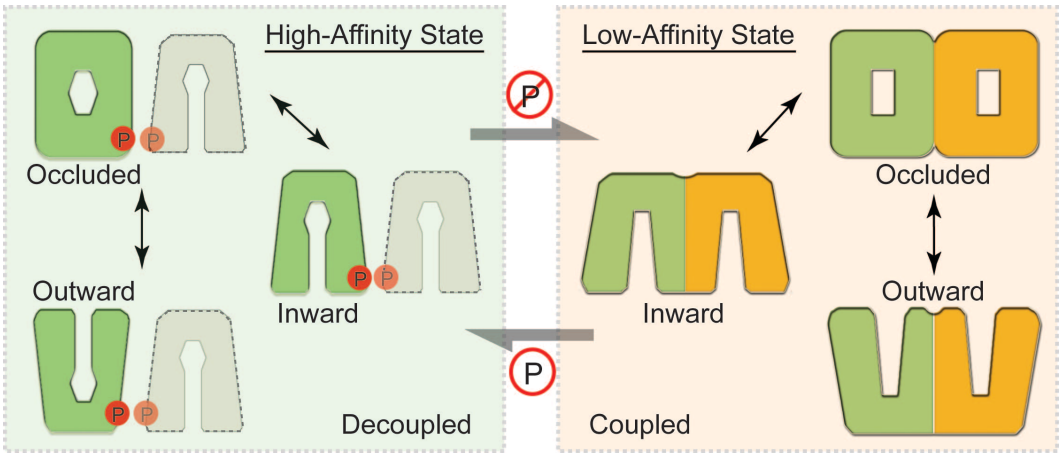

Figure 3.11 The working model. The phosphorylation-controlled dimerization state of the dual-affinity nitrate transporter dictates its mode of action. The non-phosphorylated and structurally coupled CHL1 dimer functions as an "in-phase" homodimeric low-affinity nitrate transporter (right). Once phosphorylated, the CHL1 dimer is decoupled, and each molecule functions as an independent high-affinity nitrate transporter (left). Different shapes of the putative substrate-binding site at the central transport tunnel reflect their differential nitrate binding properties, which are allosterically regulated by dimer interaction and possibly phosphorylation

CHAPTER 4. Structural Studies on CHL1 as a Nitrate Transceptor

4.1 Introduction

Membrane proteins like CHL1 that can fulfill a dual nutrient transporter/sensor function, have been named transceptors, and the studies of CHL1 has established a transceptor paradigm in plants (Gojon et al., 2011). Although transceptors has been widely found in bacteria, yeast and animals, little is known about the mechanisms underpinning their functions (Mouillon and Persson, 2006; Pinilla et al., 2011; Schwoppe et al., 2003; Thevelein and Voordeckers, 2009). A major issue of transceptors is the relationship between their transport and signaling activities. Does the sensor use the same substrate-binding site as the transporter does? Or have transceptors developed a totally new substrate-binding site with a new signaling function independent of its transport activity? Do the conformation changes associated with substrate transport play a role in signal transduction?

Recent advances in transceptor studies had shed some light on these questions. In the case of CHL1, a decoupled mutant has been found, which is defective in transporting substrates but not sensing the substrate signal (Ho et al., 2009). This indicates that the complete transport of a substrate is not necessary for signaling. The bacterial Glc6P transceptor, UhpC, on the other hand, has been documented to have specific mutations that can cause defects in signaling but not transport (Schwoppe et al., 2003), suggesting that its signaling function requires structural determinants beyond what are essential for the transport function. Together, these observations seem to favor a scenario, where the transport and sensor functions of the transceptors might involve two completely separate

mechanisms. Although such a scenario is easier to comprehend, it is still tempting to speculate that some common structural elements might be shared by the two activities. This is particularly true in the case of CHL1, whose transporter and sensor activities show similar sensitivity to substrate concentration and are both under regulation by the phosphorylation state of a key residue, Thr101.

4.2 Working models for CHL1 transceptor function

What are the possible mechanisms by which CHL1 functions as a dual-affinity nitrate transceptor? If we temporarily ignore its receptor function, one might think that Thr101 phosphorylation simply converts the transporter into two structural configurations, each capable of adopting different conformational states for transporting the substrate. These two structural configurations could share the same substrate-binding site(s), but with different substrate transport kinetic properties. In this simple model (Fig. 4.1), the kinase and the phosphatase controlling the phosphorylation state of the transporter can be regulated by a separate cellular component that can sense the substrate concentration. Since the CHL1 transporter itself is the sensor, one can modify the model by adding nitrate-sensing sites (high affinity and low affinity) in addition to the primary substrate-binding site(s) of the transporter. The high and low affinity sensing sites, upon binding nitrate, can allosterically regulate the binding and/or activation of kinase and phosphatase, respectively, thereby putting the transporter into different phosphorylation states and structural configurations. Through the same kinase and phosphatase or other mechanisms, these sensing sites could also initiate signal transduction that would eventually alter gene expression. How can this model explain the requirement of Thr101

dephosphorylation for CHL1 to sense and signal high concentration of nitrate with low affinity? One possibility is that it is not the sensor function of CHL1 but the signal transducing function that requires the de-phosphorylation state of T101. In essence, this model evokes a complete separation of the receptor and transporter functions of CHL1, but demands nitrate binding sites other than the ones involved in substrate transport.

Would it be possible that the sensor sites actually overlap or even equal to the substrate-binding sites? Going back to the initial model, in which the receptor function of the transporter is ignored, the two structural configurations dictated by Thr101 phosphorylation could actually have distinct substrate-binding sites with different affinities to nitrate (Fig. 4.2). It is possible that in the absence of nitrate, free CHL1 is in a "ground state" configuration and is constitutively phosphorylated and dephosphorylated at Thr101, keeping both the high and low affinity substrate-binding sites dynamically accessible. At low nitrate concentration, nitrate binding to the high affinity substrate-binding site prevents the phosphatase from binding via an allosteric mechanism, thereby fixing the transporter in the phosphorylated structural configuration (which can still undertake several conformational states necessary for transporting substrate). At high nitrate concentration, nitrate binding to the low affinity substrate-binding site facilitates phosphatase binding and maintains the transporter in the dephosphorylated structural configuration (which also can still undertake several conformational states necessary for transporting substrate with a different kinetic property). In such a scenario, certain specific conformations of the transporter dictated by nitrate binding, might serve as the initial signal for downstream transcription regulation. This model attributes the sensor and transporter function to the same or overlapping nitrate-binding sites, yet requires at

least two distinct substrate-binding sites in the transport path. Without any structural information, the above two models can only serve as a starting point for mechanistic understanding of the transceptor. Obviously, alternative models that combine the elements of the two models and perhaps additional components might also be equally attractive to explain the intricate functions of CHL1.

Yet, the structure of CHL1 added one level of complexity to problem, which is dimerization state. If we dissect the sensor function into two parts, nitrate sensing and signal transduction. Will the dimer be essential to either of the two parts or both? To completely understand these questions, several things need to be addressed. First, what proteins does CHL1 interact with to transduce the signal? Secondly, what is the sensing mechanism? To be more specifically, how could CHL1 sense the nitrate signal? We proposed two working models in this chapter, and the key difference lies in the amount of nitrate-binding sites in CHL1. Since the CHL1 structure was solved at 3.25 Å resolution, it will be difficult for us to see the nitrate directly in the density map. As a result, we have to use alternative approaches to solve this problem.

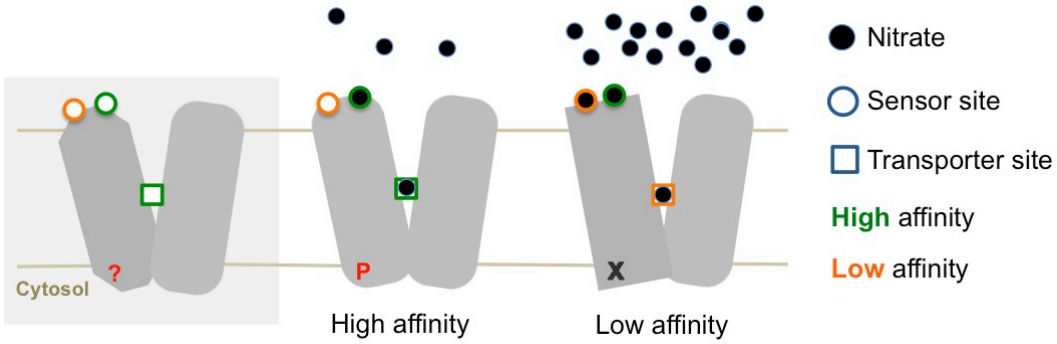

Figure 4.1 A model of CHL1 with spatially separate transporter and sensor function sites.

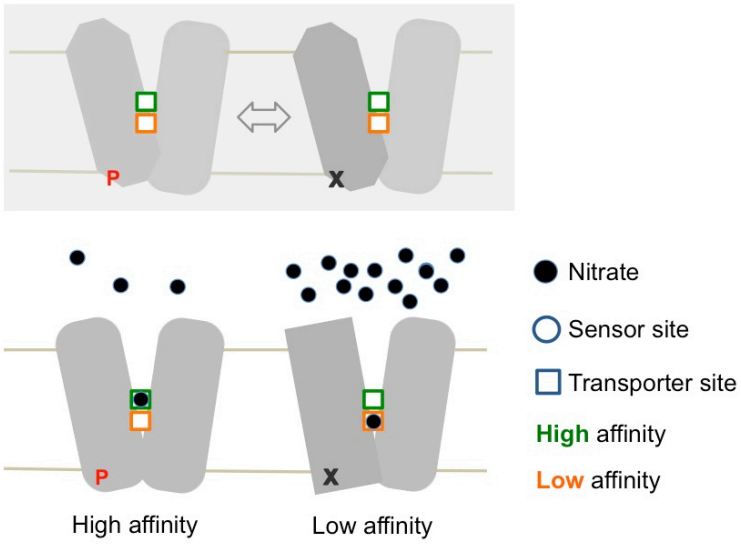

Figure 4.2 An alternative model of CHL1 transceptor with shared sensor and transporter sites.

4.3 A second nitrate-binding site

Locating the expected multiple nitrate-binding sites is one of the most important tasks in our crystal structure determination efforts, since it is the main differences between the transceptor function models we propose. As expected, we were not able to directly visualize the nitrate ions in the structure due to the relative low resolution. In crystallography, we take the advantage of the bromide anomalous signal to solve this problem with minimized ambiguity.

Bromate (BrO_3^-) is structurally similar to nitrate (NO_3^-) and chlorate (ClO_3^-), both of which can be transported by CHL1 (Tsay et al., 1993), so it is expected that bromate will bind to CHL1 in a similar way as nitrate does. Practically, we crystallized CHL1 as previously described, and the nitrate in crystals was washed away in nitrate-free buffer during crystal harvesting. 200mM sodium bromate was then added to replace nitrate. At last, data was collected at the peak wavelength of bromide. Once phase information was obtained using molecular replacement, bromate was located by anomalous difference Fourier calculation.

Interestingly, I did get bromate anomalous signal at a position outside the transporter tunnel. The detail surrounding residues are mapped in Fig. 4.3. The potential secondary binding site locates at the gap between two conserved loops, L2-3 (loop between TMH2 and TMH3) and L6-7 (the central loop between TMH6 and TMH7). The backbones of the following residues, Arg98 (L2-3), Tyr99 (L2-3), Lys245 (L6-7) and Lys246 (L6-7) are within hydrophilic interaction distance with bromate.

It is of great interest to identify a secondary substrate-binding site. Firstly, the possible binding pocket formed by Arg98, Tyr99, Lys245 and Lys246 is extremely

conserved in all CHL1 orthologs, but not within NRT1 family. The differences in conservation suggest that the secondary binding site may play a role in the sensor function of CHL1 but not transporter function. Secondly, this binding site is quite close to the phosphorylation site, Thr101. The spatial approximation indicates a potential relationship between phosphorylation modification and nitrate binding in the second site. In this scenario, it is likely that binding of nitrate in the second site may affect the phosphorylation of Thr101, or the Thr101 phosphorylation will disrupt this binding site. Thirdly, the binding of nitrate may also contribute to the bottom cleft forming (Fig. 3.10), since this binding site is also spatially close to the bottom cleft, and shares a common conserved residue, Lys246. The binding of nitrate in this site might also affect the docking of potential kinases or phosphatase.

The reliability of this binding site, however, needs to be further confirmed. In our soaking process, I introduced 200mM sodium bromate, which is extremely high for a substrate. In crystallography, anions of heavy halogens (Br^- and I^-) are used as phasing elements at similar concentration (Dauter and Dauter, 2007). As a result, the observation of the secondary bromate-binding site could be experimental artifact. However, the following two points give me confidence that what the site is highly likely true. First, bromate is structurally similar but not identical to nitrate. The four atoms in nitrate are on a plane, while they form pyramid in bromate. This may explain the low affinity between CHL1 and bromate. Second, there are only one fixed bromate site seen in each protomer, which indicates some specificity in the binding.

To sum up, I identified a potential secondary substrate-binding site, which could be of great importance to the sensor function. Yet, further studies, especially *in vivo* experiments in plants, will be needed to validate our discovery.

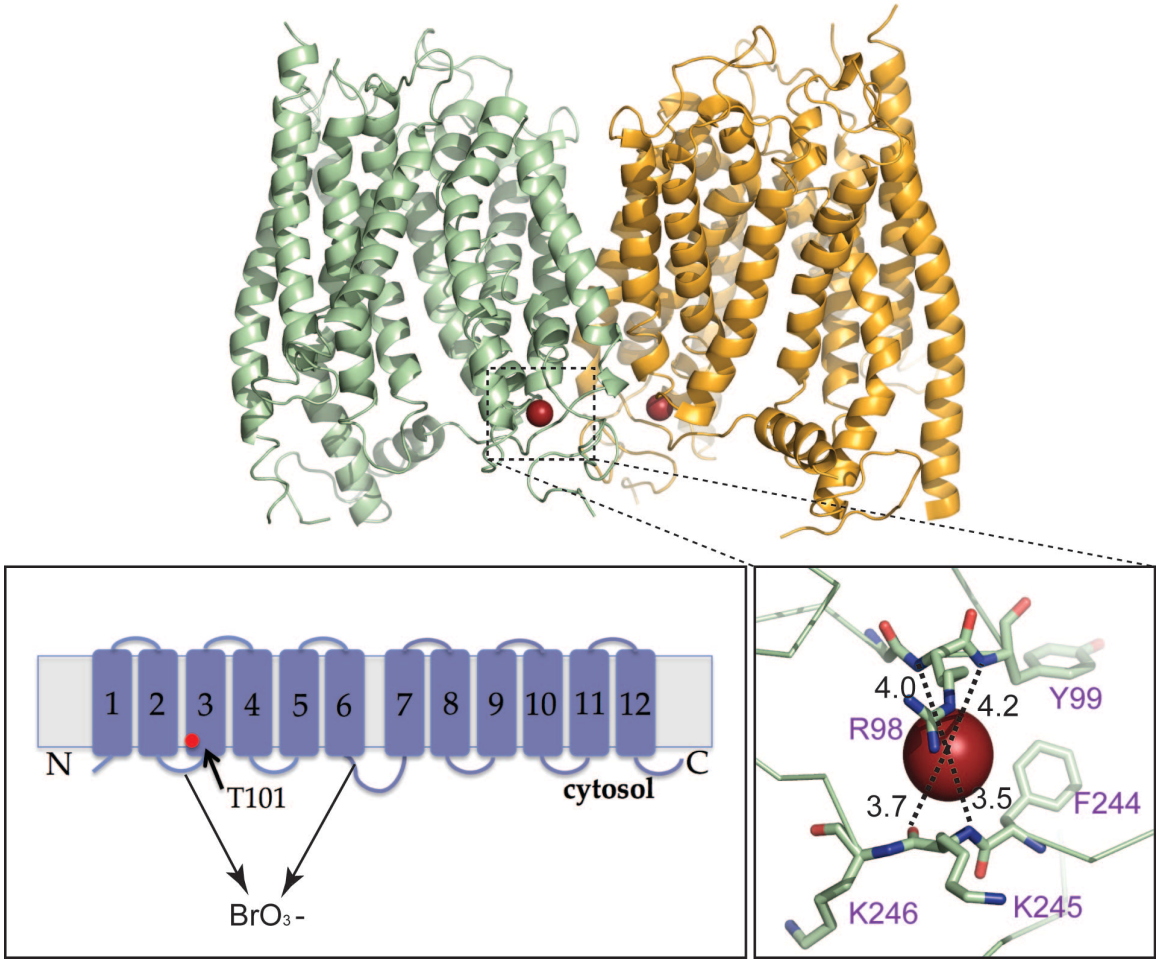

Figure 4.3 The position of a potential secondary substrate binding site. Top, the overall structure of CHL1 with bromate bound. **Bottom left**, cartoon view of the bromate-binding site. Bottom right, a close-up view of the local environment between bromate and CHL1. The distances are indicated by dash lines. Bromates are shown as spheres, colored in red.

4.4 A transceptor paradigm in plants

Transceptors have been identified in multiple systems, including *Escherichia coli (*UhpC*)* (Schwoppe et al., 2003), yeast (Gap1 and Mep2) (Van Zeebroeck et al., 2009; Van Zeebroeck et al., 2011), plants (CHL1) (Ho et al., 2009) and mammals (SNAT2) (Van Zeebroeck et al., 2009). They play essential roles in nutrients uptake as well as signaling. The study of CHL1 established a transceptor paradigm in plants, providing a new model system for study of this type of proteins.

The discovery of transporter-related receptors, now called transceptors, has raised interesting questions with respect to their functioning mechanism and significance for evolutionary origin of the receptors (Thevelein and Voordeckers, 2009). An interesting hypothesis has been proposed considering to their evolutionary role, suggesting that receptors for chemical signals may have been derived from nutrient transporters. The hypothesis was raised, based on two observations, 1) the existence of functionally related intermediates: transceptors and nontransporting transceptors; 2) the transceptor signaling also requires for a ligand-induced conformational change, which is similar to regular receptors.

As to the functioning mechanism of transceptors, a major issue is the relationship between transport and signaling. Does the sensor use the same substrate-binding site as the transporter, which state of the transport cycle plays a role in signaling? Our structural studies provide a plausible secondary binding site, which needs to be further validated. Yet, it opens a totally new gate of investigating the transceptors with this intracellular secondary substrate-binding site.

CHAPTER 5. Conclusions and Discussion

The crystal structure of CHL1 reveals a biologically relevant dimer, whose dynamic coupling and decoupling is controlled by the phosphorylation of a single residue near the dimer interface. Since the same post-translational modification switches the mode of action of the dual-affinity transporter, we propose that dimer assembly and disassembly enables CHL1 to toggle between the low-affinity and high-affinity states. To function as a dual-affinity transporter, CHL1 could have two separate substrate-binding sites with different affinities or a single substrate-binding site with two affinity states. Although we cannot rule out the possibility that CHL1 has a second high affinity substrate-binding site, which is masked in the current structure, the dimerization-based switching mechanism can conceivably create two affinity states with one physical substrate-binding site. Structural engagement of two protomers at the interface allosterically regulates the affinity of substrate binding at the central transport tunnel. Whether dimer disassembly itself is sufficient to shift CHL1 into the high-affinity mode awaits future analysis of CHL1 mutants that are defective in both dimerization and phosphorylation.

Previous structural studies of several MFS members have established a "rocker-switch" mechanism of substrate transport, in which the transporters cycle through outward-facing, occluded, and inward-facing conformation. The structure of the dimeric unmodified CHL1 reveals a buried Thr101 phosphorylation site in the inward-facing conformation. Thr101 phosphorylation, therefore, likely occurs when the dimeric transporter adopts either the outward-facing or occluded conformation. It is, however,

equally possible that the unmodified dimeric transporter is in fast equilibrium with the monomeric form, which is susceptible for phosphorylation.

Despite the structural information revealed in our studies, questions remain as to how CHL1 senses nitrate and transduces the signal. Our structure reveals a potential secondary nitrate-binding site besides the substrate site in transport tunnel. However, whether the site is functionally relevant needs to be further validated. It is also possible that another substrate binding site also exists which is excluded from the inward-facing conformation. Previous studies have suggested that the high-affinity nitrate sensor function of CHL1 is independent of the phosphorylation status of Thr101, whereas its low-affinity sensor function relies on dephosphorylation of Thr101. It is possible that low-affinity nitrate sensing and signaling function of CHL1 is mediated by dimerization, which is analogous to the functional mechanism of many common cell surface receptors. The high-affinity sensor function of CHL1, on the other hand, might involve an entirely different mechanism to be revealed.

Transporter oligomerization and phosphorylation have been implicated in the proper functions of a number of MFS members, such as LacS, GLUT1, TetL, hRFC, and rOCT1 (Hou et al., 2010; Mehrens et al., 2000; Pessino et al., 1991; Safferling et al., 2003; Veenhoff et al., 2001). The crystal structure of the CHL1 dimer not only establishes a structural framework for understanding its dual-affinity nitrate transporter and receptor activities, but also reveals for the first time how protein oligomerization and post-translational modification can synergistically expand the functional capacity of a MFS transporter.

CHL1 has also been reported to transport a plant hormone, auxin (Krouk et al., 2010). The study links the nutrient-sensing and transport function of CHL1 to the regulation of the distribution of the plant hormone auxin, which, among other effects, controls the formation and outgrowth of lateral roots. Yet, being a transporter for both nitrate and auxin, how the substrate specificity is achieved in CHL1 is unknown. Our structural studies on CHL1 with nitrate will be a starting point to answer this question.

Overall, being the most well characterized nitrate transporter in plants, CHL1 plays key roles in both nitrate uptake and signaling. Understanding the working mechanism and regulation of CHL1 will potentially help to increase nitrogen usage efficiency (NUE) and reduce nitrogen pollution. Our structural studies provide a molecular platform for understanding the working mechanism of CHL1.

APPENDIX A. Materials and Methods

Protein expression and purification

The full-length *Arabidopsis thaliana* CHL1 was cloned into pFastBac vector with a C-terminal 8xHis tag linked by a thrombin cut site. Recombinant baculovirus was generated using Bac-to-Bac system (Invitrogen) (King, 1992) (Payandeh et al., 2011; Xing et al., 2013). Monolayer High Five insect cells were infected for protein expression. Cells were harvested 72 hours after infection, washed and lysed with a buffer containing 20 mM Tris-HCl, pH 8.0, 200 mM NaCl, 10 mM $NaNO_3$ (Buffer A). CHL1 was subsequently solubilized by DDM (Anatrace) at a final concentration of 1.5% with 1-2 incubation at 4 °C. Solubilized CHL1 was separated from the insoluble fraction by high-speed ultra-centrifugation (35,000rpm) for 1hr and applied to a Ni-NTA gravity column (GE Healthcare). The bound CHL1 protein was then washed and eluted in the presence of 0.02% DDM with Buffer A supplemented with 20 mM and 200 mM imidazole, respectively. The 8xHis tag was cleaved by thrombin overnight and the CHL1 protein was further purified by size exclusion chromatography with a Superdex 200 column (GE Healthcare) in Buffer A supplemented with 0.02% DDM. The peak fractions containing CHL1 were pooled and the protein was concentrated to 10 mg/ml using a 50 kDa MWCO centrifugal device (Ambion).

Crystallization, data collection and structure determination

The CHL1 crystals were grown at 4 °C by the hanging-drop vapour diffusion method, using 2 µl protein sample mixed with 1 µl reservoir solution containing 100 mM Sodium

Acetate, pH 4.5, 30% PEG300 and 3% MPD. Crystals of maximal sizes were obtained after 1 month. 7.5% ethylene glycol was added during crystal harvest and data collection as cryo-protectant. The heavy atom derivative crystals were obtained by soaking in the presence of 1 mM ethyl *mercury* thiosalicylate (EMTS) for 2 hours. All data sets were collected at the BL8.2.1 and BL8.2.2 beamlines at the Advanced Light Source. The single anomalous dispersion (SAD) data set was collected near the mercury absorption edge (λ = 1.008 Å). X-ray diffraction data were integrated and scaled with HKL2000 package (Otwinowski and Minor, 1997) and further processed with the CCP4 package (Collaborative Computational Project, 1994). Improved molecular replacement by *Rosetta* (DiMaio et al., 2011) combined with SAD was used to determine the initial phase using PHENIX (Adams et al., 2010) with a 3.5 Å mercury derivative data set. Initial structural models were built, refined and rebuilt using COOT (Emsley et al., 2010), O (Jones et al., 1991) and CNS (Brunger, 2007; Brunger et al., 1998). The final model was built and refined with a native data set of 3.25 Å resolution. All the structure model figures in the paper were prepared using PyMol (Schrödinger, 2010).

Bromate soaking experiment

After full size CHL1 crystals were obtained, the well buffer without nitrate was used to wash out the nitrate in the crystal by 30min soaking. Then CHL1 crystals were further soaked in the well buffer containing 50mM, 100mM, 150mM and 200mM bromate step by step for 15min each. At last, crystals were harvested as native crystals with 7.5% ethylene glycol in the presence of 200mM bromate. Anomalous data was collected at bromate peak at the wavelength of 0.923 Å. Then the data was processed using a way

similar to the native crytals, and bromate was located by anomalous difference Fourier calculation using Phenix (Adams et al., 2010).

Transporter assay

The transporter assays in oocytes were carried out as previously described (Ho et al., 2009; Liu and Tsay, 2003; Tsay et al., 1993). Briefly, genes were constructed into pGEMHE vector and the plasmids were then linearized by NheI (NEB). cRNAs were prepared using mMessage mMachine T7 Ultra Kit (Ambion) and diluted to a final concentration of 2 μg/ml. 50nl cRNA was manually injected into each oocyte. Oocytes were then incubated at 16°C in regular Barth's Solution supplemented with gentamicin. After 3 days, transporter assays were carried out by incubating oocytes in the following solution: 230 mM Mannitol, 0.3 mM $CaCl_2$, 10 mM MES-TRIS, pH 5.5 for 2 hours with N15 labeled $K^{15}NO_3$ (Sigma) supplemented when required at room temperature. The oocytes were rinsed 5 times with ND96 buffer after incubation and then dried individually at 80 °C for 24 hours in tin capsules (ALPHA). The retained ^{15}N was analyzed on a continuous flow isotope ratio mass spectrometry (IRMS) coupled with a nitrogen elemental analyzer (UC Davis Stable Isotope Facility). CHL1 mutants were constructed using Quikchange methods (Zoller and Smith, 1984). For uptake assays, the background was subtracted from the final uptake data, which were subsequently normalized to the wild type protein. Every data point in the experiments was measured and averaged on more than 5 oocytes. Curves were fit in Prism if needed.

FRET assay

The mCerulean variant of cyan fluorescent protein and mCitrine variant of yellow fluorescent protein were individually fused to the N-terminus of CHL1 with a three-alanine linker. Then cRNAs were prepared using the same vector and kit in the transporter assay and injected into oocytes. FRET signal was measured after incubation for 3 days using the method described previously (Zheng et al., 2002). Briefly, emission spectra of mCerulean and mCitrine were collected using laser excitation at 458 and 488 nm, respectively, and an emission window of 3.2 nm. FRET was calculated using a spectrum-based approach to remove contaminations caused by donor emission and direct excitation of the acceptor fluorophore. mCerulean spectra were collected from oocytes expressing mCerulean-tagged CHL1 transporters alone and then subtracted from spectra taken from oocytes expressing both mCerulean and mCitrine tagged transporters. The resulting extracted mCitrine emission spectrum, F_{458}, contains two components: one caused by direct excitation, $F^{\text{direct}}{}_{458}$, and the other by FRET, $F^{\text{FRET}}{}_{458}$. F_{458} was normalized by total mCitrine emission excited directly with 488nm light (F_{488}). The resulting ratio, termed RatioA, can be expressed as:

$$RatioA = \frac{F458}{F488} = \frac{F_{458}^{direct}}{F_{488}} + \frac{F_{458}^{FRET}}{F_{488}}$$

The direct excitation component, $\frac{F_{458}^{direct}}{F_{488}}$, termed RatioA_0, was experimentally determined with oocytes expressing mCitrine-tagged transporters only. In order to compare FRET across mutant and WT transporters, the ratio between RatioA and RatioA_0, termed FRET Ratio (FR), is calculated as follows:

$$FR = \frac{RatioA}{RatioA_0} = 1 + \frac{F_{458}^{FRET}}{F_{458}^{direct}}$$

Images were analyzed in ImageJ (National Institute of Health) by drawing regions of interest around the fluorescent membranes and using the measure stack feature. The data were then analyzed using programs written in MATLAB (MathWorks, Natick, NJ).

Crosslink

Purified CHL1 protein was dialyzed in PBS buffer, pH 7.4 supplemented with 0.01% DDM for 24 hours with one buffer exchange. The protein was then concentrated to 5 mg/ml, and 2 µl ethylene glycol bis-succinimidylsuccinate (EGS) (Pierce) was added into 18 µl protein sample at increasing concentrations (0, 0.1 mM, 0.3 mM, 1 mM, 3 mM, and 10 mM). The reaction was carried out at room temperature for 30 minutes and stopped by a 50 mM Tris-HCl buffer, pH 8.0. 1/5 of the reaction solution was analyzed by SDS-PAGE.

SEC-LS-RI-UV measurement

The SEC-MAL system consisted of a P900 HPLC pump (GE), a UV-2077 detector (Jasco), a Tri Star Mini Dawn light scattering instrument (Wyatt), and an Opti Lab T-Rex refractive index instrument (Wyatt) (Slotboom et al., 2008). 20 µL of purified and DDM-solubilized CHL1 (5 mg/ml) was injected into a Superdex 200 (10/300GL) gel filtration column and eluted isocratically at 0.5 ml/min in a buffer containing 20 mM Tris, 200 mM NaCl, 5mM $NaNO_3$, 0.02% DDM, pH 8.0. The extinction coefficient of CHL1 at 280 nm was calculated from the amino acid sequence (E=1.268 ml•mg^{-1}•cm^{-1}). DDM has no

absorbance at 280 nm. The specific refractive index of CHL1 and DDM was assumed to be 0.186 g/ml and 0.147ml/g, respectively. Data collection and analysis was performed with Astra 6 software (Wyatt). Total molecular mass and individual masses of the protein and the detergent were determined with Astra6 software using protein conjugate analysis. Both peak over lap and peak broadening were corrected with Astra 6 software. The SEC-MAL system was pre-calibrated with BSA.

APPENDIX B. Membrane Protein Crystallization

Introduction

Membrane proteins have unique complexities in structural studies due to their lipid-soluble nature. Since the first membrane protein structure was determined at atomic resolution in 1985 (Deisenhofer et al., 1985), lots of techniques and protocols have been developed for membrane structure studies (Andrell and Tate, 2013; Caffrey, 2003; Gonen et al., 2005; Kawate and Gouaux, 2006; Ujwal and Bowie, 2011; Zhao and Wu, 2012). However, structure determination for membrane proteins remains a major challenge as a result of the difficulties associated with generating milligram quantities of pure and monodisperse membrane protein, as well as getting well-diffracted crystals.

In the past few years, our lab accumulated some experiences in membrane protein structure study after successfully determining atomic resolution structures of a prokaryotic voltage-gate sodium channel, Na_vAb and a eukaryotic nitrate transporter, CHL1. Here, I will introduce my general protocol for membrane protein crystallization, established through my study on CHL1 and literature reading (Kawate and Gouaux, 2006; Newby et al., 2009; Newstead et al., 2007). This protocol described here is aimed at meeting the general requirements for crystallization.

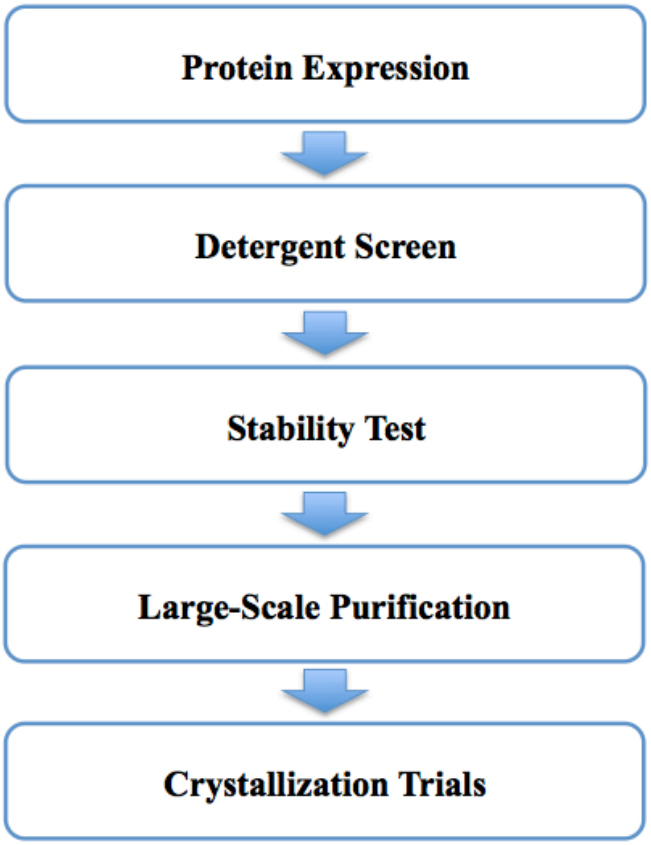

Figure B1 The workflow for membrane protein purification.

The overall workflow

The protocol will include the following major steps: protein expression, detergent screen, stability test, large-scale purification, and crystallization. The brief flow chart is shown in Fig. B1. The pathway to the final well-diffracted crystals is inherently an iterative one, and in every step, information will be obtained to optimize the previous or guide the next steps.

Protein expression

The protein expression part includes: 1) construct design and cloning, 2) expression test in a suitable expression system, and 3) yield estimation.

1) Construct design and cloning

Since membrane protein expression is usually difficult, designing a well-expressing construct will be of great importance. For a specific membrane protein, several constructs will be made in order to get optimal expression. These constructs usually include 3-5 homologs of the target protein, truncation of unconserved loops (usually N-ter and C-ter). In my experimental sets, a C-ter or N-ter His-tag will be added for affinity purification purpose.

2) Expression system

Escherichia coli is traditionally the system of choice for structural biologists to heterologously produce proteins for X-ray studies. It is inexpensive, easy to manipulate (e.g., transform, culture and harvest protein from) and is capable of producing milligram quantities of protein from culture sizes on the order of tens of liters. It has also been extremely successful in expressing prokaryotic membrane proteins (White, 2009).

However, when it comes to eukaryotic targets, alternative systems, like yeast (Long et al., 2005), insect cells (Jasti et al., 2007; Payandeh et al., 2011), and mammalian cells (Reeves et al., 2002), will be used for producing large amount of proteins. In our lab, insect cells are used extensively for membrane protein production. A brief protocol is described previously in the Appendix A (King, 1992).

3) Yield estimation

In order to estimate and compare the yields between different constructs, a relative hush detergent, Fos-choline 12 (FC12) will be used to extract the membrane protein, which is then followed by a standard Ni-NTA purification step. At last, the results will be compared on SDS-PAGE. The construct with highest yield will be further used for detergent screen. Meanwhile, constructs with the potential of easily generating milligram protein will also be applied into the pipeline if necessary.

Detergent screen and stability test

For purification and crystallization, membrane proteins need to be extracted from the lipid membrane, in which they were expressed, with detergents. The identification of the detergent most suitable for a particular protein target is an empirical process (Garavito et al., 1996; Linke, 2009). Detergent screen and stability test are two steps intertwined and carried out together.

Our detergent screen for extraction and stabilization was done in a two-step manner. In the first step, I take the detergent head group as the variable and tested 4-6 representative detergents, for example, FC12 (Fos-Choline 12), LDAO (Lauryldimethylamine-N-Oxide), DDM (n-Dodecyl-β-D-Maltoside) and $C_{12}E_8$ (Fig. B2).

These detergents have different head group structures but the same tail length (12 $–CH_2$-groups). Size exclusion chromatography elution profile was used as the criterion for judging whether a detergent is suitable or not for the target. Briefly, recombinant protein is extracted from cell lysates with the chosen detergents individually, followed by Ni-NTA purification in the presence of the same detergent. Proteins eluted from the metal affinity columns are subsequently applied to a superdex-200 gel filtration column without any detergent exchange event. If a non-aggregation peak of the target protein elution is found with any of the detergents, I will further test the stability of the target protein in those specific detergents. After taking samples from the non-aggregated protein-detergent complexes isolated by size exclusion chromatography and leaving them at 4 ˚C for a week, I apply the samples back to a gel filtration column. If the samples still form a monodispersed peak on size exclusion column in certain detergent, then the detergent will be further used in large-scale purification and crystallization. If none of them turns out to be successful, more detergents with different head groups will be chosen to do the same screen.

As the second step of our test, we fixed the head group of the detergent and screened the length of the tail group, which has a variable number of alkyl group, ranging from 10 to 13. Take maltoside family as example, I choose n-Decyl-β-D-Maltoside ($D_{10}M$), n-Undecyl-β-D-Maltoside ($U_{11}DM$), n-Dodecyl-β-D-Maltoside ($D_{12}DM$, same as DDM), n-Tridecyl-β-D-Maltoside ($T_{13}DM$) and also 6-Cyclohexyl-1-Hexyl-β-D-Maltoside (Cymal-6). $D_{10}M$, $U_{11}DM$, DDM, and $T_{13}DM$ each have an alkyl tail with an increasing length, while Cymal-6 has a tail similar to DDM, except that its last six-alkyl chain forms

a cyclohexane structure. If any of these members show any promising result, the detergent will also used for crystallization.

Large-scale purification

Large-scale purification is usually done in two-step manner, Ni-NTA purification and size exclusion column. At the end of the purification, I will use the following two criteria to judge whether the yield and purity of the protein sample is good enough for crystallization, 1) protein amount should be more than 2mg, and 2) purity should be better than 90%. If either of the criteria is not satisfied, the expression and purification step will be optimized.

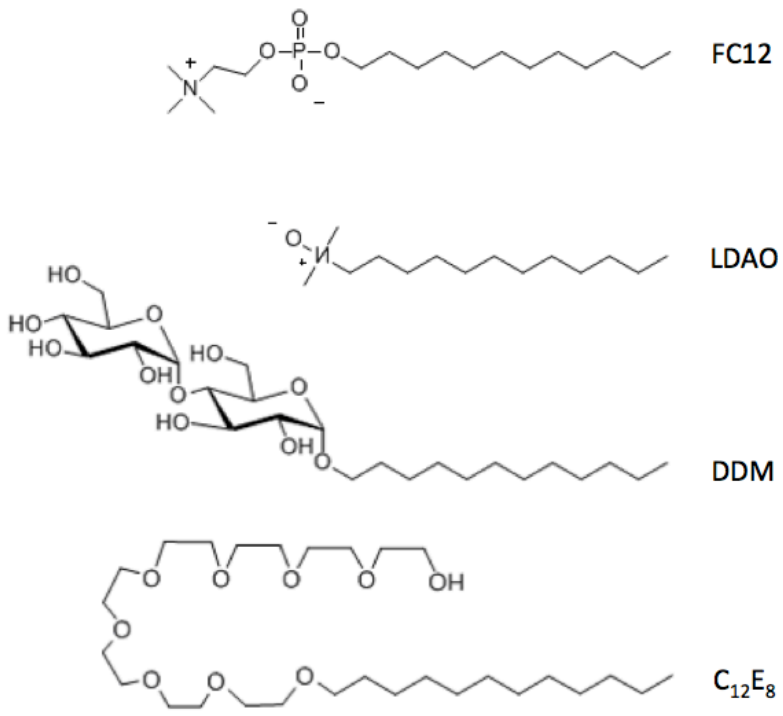

Figure B2 Detergent structure. Structures of the four detergents that have the same tail length (12 $–CH_2-$ group) but different head groups.

Crystallization

High-throughput sparse matrix screens will be carried out by a Mosquito nanodrop liquid handling robot (TTP Labtech). According to previous literatures in the membrane protein crystallization, lots of membrane proteins are crystallized in polyethylene glycol (PEG) bias conditions, so we performed crystal screen using a number of commercially available (Hampton, Molecular Dimensions) and in-house screen kits that mostly contain conditions with low molecule weight PEGs (PEG300, PEG400, and PEG1000) as major precipitants. Each protein will be tested in ~600-800 conditions for crystallization. If no promising crystal hits are observed, some other crystallization techniques like lipid bicelle methods will be used (Ujwal and Bowie, 2011). Alternatively, another construct will be applied into this whole process for crystallization.

It should be understood, however, that there is a tremendous amount of attrition that occurs along the pathway to membrane protein crystallization. That is to say, a particular target may fail at any stage, i.e., protein expression, solubilization, purification or crystallization. As a result, some specific target could fail at any of the above steps, so troubleshooting has to be done. The final success is also largely dependent on the protein characterization, so the protocol described here serves as a general guide for systematic studies of membrane proteins.

Bibliography

Abramson, J., Smirnova, I., Kasho, V., Verner, G., Kaback, H.R., and Iwata, S. (2003). Structure and mechanism of the lactose permease of Escherichia coli. Science *301*, 610-615.

Adams, P.D., Afonine, P.V., Bunkoczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.W., Kapral, G.J., Grosse-Kunstleve, R.W.*, et al.* (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta crystallographica Section D, Biological crystallography *66*, 213-221.

Alboresi, A., Gestin, C., Leydecker, M.T., Bedu, M., Meyer, C., and Truong, H.N. (2005). Nitrate, a signal relieving seed dormancy in Arabidopsis. Plant, cell & environment *28*, 500-512.

Andrell, J., and Tate, C.G. (2013). Overexpression of membrane proteins in mammalian cells for structural studies. Molecular membrane biology *30*, 52-63.

Brunger, A.T. (2007). Version 1.2 of the Crystallography and NMR system. Nature protocols *2*, 2728-2733.

Brunger, A.T., Adams, P.D., Clore, G.M., DeLano, W.L., Gros, P., Grosse-Kunstleve, R.W., Jiang, J.S., Kuszewski, J., Nilges, M., Pannu, N.S.*, et al.* (1998). Crystallography & NMR system: A new software suite for macromolecular structure determination. Acta crystallographica Section D, Biological crystallography *54*, 905-921.

Caffrey, M. (2003). Membrane protein crystallization. Journal of structural biology *142*, 108-132.

Camanes, G., Pastor, V., Cerezo, M., Garcia-Andrade, J., Vicedo, B., Garcia-Agustin, P., and Flors, V. (2012). A deletion in NRT2.1 attenuates Pseudomonas syringae-induced hormonal perturbation, resulting in primed plant defenses. Plant physiology *158*, 1054-1066.

Castro Marin, I., Loef, I., Bartetzko, L., Searle, I., Coupland, G., Stitt, M., and Osuna, D. (2011). Nitrate regulates floral induction in Arabidopsis, acting independently of light, gibberellin and autonomous pathways. Planta *233*, 539-552.

Collaborative Computational Project, N. (1994). The CCP4 suite: programs for protein crystallography. Acta crystallographica Section D, Biological crystallography *50*, 760-763.

Dang, S., Sun, L., Huang, Y., Lu, F., Liu, Y., Gong, H., Wang, J., and Yan, N. (2010). Structure of a fucose transporter in an outward-open conformation. Nature *467*, 734-738.

Dauter, M., and Dauter, Z. (2007). Phase determination using halide ions. Methods in molecular biology *364*, 149-158.

Deisenhofer, J., Epp, O., Miki, K., Huber, R., and Michel, H. (1985). Structure of the protein subunits in the photosynthetic reaction centre of Rhodopseudomonas viridis at 3A resolution. Nature *318*, 618-624.

DiMaio, F., Terwilliger, T.C., Read, R.J., Wlodawer, A., Oberdorfer, G., Wagner, U., Valkov, E., Alon, A., Fass, D., Axelrod, H.L.*, et al.* (2011). Improved molecular replacement by density- and energy-guided protein structure optimization. Nature *473*, 540-543.

Doki, S., Kato, H.E., Solcan, N., Iwaki, M., Koyama, M., Hattori, M., Iwase, N., Tsukazaki, T., Sugita, Y., Kandori, H.*, et al.* (2013). Structural basis for dynamic mechanism of proton-coupled symport by the peptide transporter POT. Proceedings of the National Academy of Sciences of the United States of America *110*, 11343-11348.

Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of Coot. Acta crystallographica Section D, Biological crystallography *66*, 486-501.

Garavito, R.M., Picot, D., and Loll, P.J. (1996). Strategies for crystallizing membrane proteins. Journal of bioenergetics and biomembranes *28*, 13-27.

Gojon, A., Krouk, G., Perrine-Walker, F., and Laugier, E. (2011). Nitrate transceptor(s) in plants. Journal of experimental botany *62*, 2299-2308.

Gonen, T., Cheng, Y., Sliz, P., Hiroaki, Y., Fujiyoshi, Y., Harrison, S.C., and Walz, T. (2005). Lipid-protein interactions in double-layered two-dimensional AQP0 crystals. Nature *438*, 633-638.

Guo, F.Q., Young, J., and Crawford, N.M. (2003). The nitrate transporter AtNRT1.1 (CHL1) functions in stomatal opening and contributes to drought susceptibility in Arabidopsis. The Plant cell *15*, 107-117.

Ho, C.H., Lin, S.H., Hu, H.C., and Tsay, Y.F. (2009). CHL1 functions as a nitrate sensor in plants. Cell *138*, 1184-1194.

Hou, Z., Cherian, C., Drews, J., Wu, J., and Matherly, L.H. (2010). Identification of the minimal functional unit of the homo-oligomeric human reduced folate carrier. The Journal of biological chemistry *285*, 4732-4740.

Huang, N.C., Chiang, C.S., Crawford, N.M., and Tsay, Y.F. (1996). CHL1 encodes a component of the low-affinity nitrate uptake system in Arabidopsis and shows cell type-specific expression in roots. The Plant cell *8*, 2183-2191.

Huang, Y., Lemieux, M.J., Song, J., Auer, M., and Wang, D.N. (2003). Structure and mechanism of the glycerol-3-phosphate transporter from Escherichia coli. Science *301*, 616-620.

Jasti, J., Furukawa, H., Gonzales, E.B., and Gouaux, E. (2007). Structure of acid-sensing ion channel 1 at 1.9 A resolution and low pH. Nature *449*, 316-323.

Jones, T.A., Zou, J.Y., Cowan, S.W., and Kjeldgaard, M. (1991). Improved methods for building protein models in electron density maps and the location of errors in these models. Acta crystallographica Section A, Foundations of crystallography *47 (Pt 2)*, 110-119.

Kawate, T., and Gouaux, E. (2006). Fluorescence-detection size-exclusion chromatography for precrystallization screening of integral membrane proteins. Structure *14*, 673-681.

King, L.A. (1992). Baculovirus Expression System : A Laboratory Guide, 1st edn (Springer).

Krouk, G., Lacombe, B., Bielach, A., Perrine-Walker, F., Malinska, K., Mounier, E., Hoyerova, K., Tillard, P., Leon, S., Ljung, K.*, et al.* (2010). Nitrate-regulated auxin transport by NRT1.1 defines a mechanism for nutrient sensing in plants. Developmental cell *18*, 927-937.

Law, C.J., Maloney, P.C., and Wang, D.N. (2008). Ins and outs of major facilitator superfamily antiporters. Annual review of microbiology *62*, 289-305.

Linke, D. (2009). Detergents: an overview. Methods in enzymology *463*, 603-617.

Liu, K.H., Huang, C.Y., and Tsay, Y.F. (1999). CHL1 is a dual-affinity nitrate transporter of Arabidopsis involved in multiple phases of nitrate uptake. The Plant cell *11*, 865-874.

Liu, K.H., and Tsay, Y.F. (2003). Switching between the two action modes of the dual-affinity nitrate transporter CHL1 by phosphorylation. The EMBO journal *22*, 1005-1013.

Long, S.B., Campbell, E.B., and Mackinnon, R. (2005). Crystal structure of a mammalian voltage-dependent Shaker family K+ channel. Science *309*, 897-903.

Mehrens, T., Lelleck, S., Cetinkaya, I., Knollmann, M., Hohage, H., Gorboulev, V., Boknik, P., Koepsell, H., and Schlatter, E. (2000). The affinity of the organic cation transporter rOCT1 is increased by protein kinase C-dependent phosphorylation. Journal of the American Society of Nephrology : JASN *11*, 1216-1224.

Mouillon, J.M., and Persson, B.L. (2006). New aspects on phosphate sensing and signalling in Saccharomyces cerevisiae. FEMS yeast research *6*, 171-176.

Munos, S., Cazettes, C., Fizames, C., Gaymard, F., Tillard, P., Lepetit, M., Lejay, L., and Gojon, A. (2004). Transcript profiling in the chl1-5 mutant of Arabidopsis reveals a role of the nitrate transporter NRT1.1 in the regulation of another nitrate transporter, NRT2.1. The Plant cell *16*, 2433-2447.

Newby, Z.E., O'Connell, J.D., 3rd, Gruswitz, F., Hays, F.A., Harries, W.E., Harwood, I.M., Ho, J.D., Lee, J.K., Savage, D.F., Miercke, L.J.*, et al.* (2009). A general protocol for the crystallization of membrane proteins for X-ray structural investigation. Nature protocols *4*, 619-637.

Newstead, S., Drew, D., Cameron, A.D., Postis, V.L., Xia, X., Fowler, P.W., Ingram, J.C., Carpenter, E.P., Sansom, M.S., McPherson, M.J*., et al.* (2011). Crystal structure of a prokaryotic homologue of the mammalian oligopeptide-proton symporters, PepT1 and PepT2. The EMBO journal *30*, 417-426.

Newstead, S., Kim, H., von Heijne, G., Iwata, S., and Drew, D. (2007). High-throughput fluorescent-based optimization of eukaryotic membrane protein overexpression and purification in Saccharomyces cerevisiae. Proceedings of the National Academy of Sciences of the United States of America *104*, 13936-13941.

Orsel, M., Chopin, F., Leleu, O., Smith, S.J., Krapp, A., Daniel-Vedele, F., and Miller, A.J. (2006). Characterization of a two-component high-affinity nitrate uptake system in Arabidopsis. Physiology and protein-protein interaction. Plant physiology *142*, 1304-1317.

Otwinowski, Z., and Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. Method Enzymol *276*, 307-326.

Pao, S.S., Paulsen, I.T., and Saier, M.H., Jr. (1998). Major facilitator superfamily. Microbiology and molecular biology reviews : MMBR *62*, 1-34.

Payandeh, J., Scheuer, T., Zheng, N., and Catterall, W.A. (2011). The crystal structure of a voltage-gated sodium channel. Nature *475*, 353-358.

Pedersen, B.P., Kumar, H., Waight, A.B., Risenmay, A.J., Roe-Zurz, Z., Chau, B.H., Schlessinger, A., Bonomi, M., Harries, W., Sali, A*., et al.* (2013). Crystal structure of a eukaryotic phosphate transporter. Nature *496*, 533-536.

Pessino, A., Hebert, D.N., Woon, C.W., Harrison, S.A., Clancy, B.M., Buxton, J.M., Carruthers, A., and Czech, M.P. (1991). Evidence that functional erythrocyte-type glucose transporters are oligomers. The Journal of biological chemistry *266*, 20213-20217.

Pinilla, J., Aledo, J.C., Cwiklinski, E., Hyde, R., Taylor, P.M., and Hundal, H.S. (2011). SNAT2 transceptor signalling via mTOR: a role in cell growth and proliferation? Frontiers in bioscience *3*, 1289-1299.

Reeves, P.J., Callewaert, N., Contreras, R., and Khorana, H.G. (2002). Structure and function in rhodopsin: high-level expression of rhodopsin with restricted and homogeneous N-glycosylation by a tetracycline-inducible N-acetylglucosaminyltransferase I-negative HEK293S stable mammalian cell line. Proceedings of the National Academy of Sciences of the United States of America *99*, 13419-13424.

Safferling, M., Griffith, H., Jin, J., Sharp, J., De Jesus, M., Ng, C., Krulwich, T.A., and Wang, D.N. (2003). TetL tetracycline efflux protein from Bacillus subtilis is a dimer in the membrane and in detergent solution. Biochemistry *42*, 13969-13976.

Schrödinger, L. (2010). The PyMOL Molecular Graphics System.

Schwoppe, C., Winkler, H.H., and Neuhaus, H.E. (2003). Connection of transport and sensing by UhpC, the sensor for external glucose-6-phosphate in Escherichia coli. European journal of biochemistry / FEBS *270*, 1450-1457.

Slotboom, D.J., Duurkens, R.H., Olieman, K., and Erkens, G.B. (2008). Static light scattering to characterize membrane proteins in detergent solution. Methods *46*, 73-82.

Solcan, N., Kwok, J., Fowler, P.W., Cameron, A.D., Drew, D., Iwata, S., and Newstead, S. (2012). Alternating access mechanism in the POT family of oligopeptide transporters. The EMBO journal *31*, 3411-3421.

Sun, L., Zeng, X., Yan, C., Sun, X., Gong, X., Rao, Y., and Yan, N. (2012). Crystal structure of a bacterial homologue of glucose transporters GLUT1-4. Nature *490*, 361-366.

Taraska, J.W., and Zagotta, W.N. (2010). Fluorescence applications in molecular neurobiology. Neuron *66*, 170-189.

Thevelein, J.M., and Voordeckers, K. (2009). Functioning and evolutionary significance of nutrient transceptors. Molecular biology and evolution *26*, 2407-2414.

Tsay, Y.F., Chiu, C.C., Tsai, C.B., Ho, C.H., and Hsu, P.K. (2007). Nitrate transporters and peptide transporters. FEBS letters *581*, 2290-2300.

Tsay, Y.F., Schroeder, J.I., Feldmann, K.A., and Crawford, N.M. (1993). The herbicide sensitivity gene CHL1 of Arabidopsis encodes a nitrate-inducible nitrate transporter. Cell *72*, 705-713.

Ujwal, R., and Bowie, J.U. (2011). Crystallizing membrane proteins using lipidic bicelles. Methods *55*, 337-341.

Van Zeebroeck, G., Bonini, B.M., Versele, M., and Thevelein, J.M. (2009). Transport and signaling via the amino acid binding site of the yeast Gap1 amino acid transceptor. Nature chemical biology *5*, 45-52.

Van Zeebroeck, G., Kimpe, M., Vandormael, P., and Thevelein, J.M. (2011). A split-ubiquitin two-hybrid screen for proteins physically interacting with the yeast amino acid transceptor Gap1 and ammonium transceptor Mep2. PloS one *6*, e24275.

Veenhoff, L.M., Heuberger, E.H., and Poolman, B. (2001). The lactose transport protein is a cooperative dimer with two sugar translocation pathways. The EMBO journal *20*, 3056-3062.

Wang, R., Liu, D., and Crawford, N.M. (1998). The Arabidopsis CHL1 protein plays a major role in high-affinity nitrate uptake. Proceedings of the National Academy of Sciences of the United States of America *95*, 15134-15139.

White, S.H. (2009). Biophysical dissection of membrane proteins. Nature *459*, 344-346.

Xing, W., Busino, L., Hinds, T.R., Marionni, S.T., Saifee, N.H., Bush, M.F., Pagano, M., and Zheng, N. (2013). SCF(FBXL3) ubiquitin ligase targets cryptochromes at their cofactor pocket. Nature *496*, 64-68.

Xu, G., Fan, X., and Miller, A.J. (2012). Plant nitrogen assimilation and use efficiency. Annual review of plant biology *63*, 153-182.

Yan, H., Huang, W., Yan, C., Gong, X., Jiang, S., Zhao, Y., Wang, J., and Shi, Y. (2013). Structure and mechanism of a nitrate transporter. Cell reports *3*, 716-723.

Yan, N. (2013). Structural investigation of the proton-coupled secondary transporters. Current opinion in structural biology *23*, 483-491.

Yin, Y., He, X., Szewczyk, P., Nguyen, T., and Chang, G. (2006). Structure of the multidrug transporter EmrD from Escherichia coli. Science *312*, 741-744.

Zhang, H., and Forde, B.G. (2000). Regulation of Arabidopsis root development by nitrate availability. Journal of experimental botany *51*, 51-59.

Zhao, Q., and Wu, B.L. (2012). Ice breaking in GPCR structural biology. Acta pharmacologica Sinica *33*, 324-334.

Zheng, H., Wisedchaisri, G., and Gonen, T. (2013). Crystal structure of a nitrate/nitrite exchanger. Nature *497*, 647-651.

Zheng, J., Trudeau, M.C., and Zagotta, W.N. (2002). Rod cyclic nucleotide-gated channels have a stoichiometry of three CNGA1 subunits and one CNGB1 subunit. Neuron *36*, 891-896.

Zoller, M.J., and Smith, M. (1984). Oligonucleotide-directed mutagenesis: a simple method using two oligonucleotide primers and a single-stranded DNA template. DNA *3*, 479-488.