学位論文（要約）

# Structural and functional analyses of nitrate transporters

# （硝酸イオン輸送体の構造・機能解析）

平成29年12月博士（理学）申請

東京大学大学院 理学系研究科

生物科学専攻

福田 昌弘

# Abstract

Nitrogen is an essential element in biology, required for the synthesis of DNA, RNA, and amino acids, and therefore is a fundamental nutrient for cellular metabolism. Most organisms utilize nitrate, nitrite and ammonium ions as inorganic nitrogen sources. Nitrate respiration and assimilation are the key processes in biological system and also in the nitrogen cycle of the ecosystem. The cellular nitrate uptake, the first step of nitrate respiration and assimilation, is mediated by nitrate transporters which are divided into three families; ATP-binding cassette-type (ABC) transporter superfamily, Nitrate transporter 1/Peptide transporter Family (NPF), and Nitrate/Nitrite Porter (NNP) family. Among these nitrate transporters, the members of the NNP family are most widely conserved from prokaryotes to eukaryotes and play a major role in cellular nitrate uptake, and also in allocation and storage of nitrate in higher plants. However, the structural information of NNP family transporters has remained elusive for more than 20 years since the discovery of the gene encoding NarK in NNP family from *Escherichia coli* in 1989. Hence, the purpose of this study is to elucidate the detailed molecular mechanism for nitrate uptake by NNP family transporters.

*Escherichia coli*, one of the most extensively studied prokaryotic organisms, has two NNP family proteins: NarK and NarU. Despite intense research for over two decades,

the precise molecular-level functions of NarK and NarU were not well-understood. Recently, the crystal structures of NarK and NarU were reported [Zheng *et al.*, *Nature* **497**, 647–651 (2013) and Yan *et al.*, *Cell Rep.* **3**, 716–723 (2013)]. These structures revealed the overall architecture of prokaryotic NNP family transporters and also provided the important insights into the key residues for substrate recognition. However, the molecular mechanism of NNP transporters, including the structural change and the driving force (or fueling mechanism) for nitrate transport, remains elusive.

In this study, to elucidate the dynamic mechanism of the substrate transport by NNP family, the crystal structures of NarK from *Escherichia coli* were determined in three different states in the transport cycle (nitrate-bound occluded, nitrate-bound inward-open and apo inward-open states) at the resolutions of 2.35–2.4 Å. These high-resolution structures represent the occluded and inward-open conformations, in contrast to the previously reported partially occluded and partially inward-open conformations. Furthermore, a novel nitrite influx assay system to detect the NarK activity *in vitro* using NarK-reconstituted proteo-liposomes was developed and the results demonstrated that NarK functions as a tightly-coupled nitrate/nitrite antiporter. The combination of X-ray crystallography, *in vitro* nitrite influx assay, genetic analyses and an *in silico* molecular dynamics simulation enabled us to reveal the novel mechanism of substrate transport by

NarK, which clearly explains how the transport of nitrate is tightly coupled with that of nitrite: the coupling between the substrate recognition and the opening/closing of the transport pathway is accomplished by the orchestrated movement of the three TM helices and the key residues Tyr263 and Arg305 in the substrate-binding site. These results will facilitate the precise understanding of the general molecular mechanisms of MFS secondary transporters, in addition to the nitrate/nitrite exchange mechanism of the NNP family.

これ以降については、５年以内に雑誌等で刊行予定のため、非公開。

# Contents

# Table of Abbreviations

| Abbreviation | Full name |
|---|---|
| ABC | ATP-binding cassette |
| APBS | adaptive Poisson-Boltzmann solver |
| ATP | adenosine triphosphate |
| CBB | Coomassie brilliant blue |
| CCD | charge coupled device |
| CM | chloramphenicol |
| DDM | *n*-Dodecyl β-D-maltopyranoside |
| DME | dimethyl ether |
| DNA | deoxyribonucleic acid |
| DTT | dithiothreitol |
| EC | evolutionary constraint |
| EDTA | ethylenediaminetetraacetic acid |
| EPR | electron paramagnetic resonance |
| EtBr | ethidium bromide |
| HATS | high-affinity nitrate transport system |
| IPTG | isopropyl β-D-thiogalactopyranoside |
| KM | kanamycin |
| LATS | low-affinity nitrate transport system |
| LB | Luria-Bertani |
| LCP | lipidic cubic phase |
| MFS | Major facilitator superfamily |
| NNP | nitrate nitrite porter |
| NPF | nitrate transporter 1/peptide transporter family |
| MeHgCl | methyl mercury chloride |
| MD | molecular dynamics |
| MES | 2-morpholinoethanesulfonic acid |
| MFS | major facilitator superfamily |
| MO | monoolein |
| MWCO | molecular weight cutoff |
| NBD | nucleotide-binding domain |
| Ni-NTA | nickel-nitrilotriacetic acid |
| NMR | nuclear magnetic resonance |
| PCR | polymerase chain reaction |

| | |
|---|---|
| PDB | Protein Data Bank |
| PEG | polyethylene glycol |
| SeMet | selenomethionine |
| PMSF | phenylmethylsulfonyl fluoride |
| SDS-PAGE | sodium dodecyl sulfate poly-acrylamide gel electrophoresis |
| SEC | size-exclusion chromatography |
| TB | terrific broth |
| TEV | tobacco etch virus |
| TM | transmembrane |
| Tris | tris(hydroxymethyl)aminomethane |
| WT | wild type |

# Table of Amino Acid Abbreviations

| Abbreviation | Full name |
|---|---|
| A, Ala | alanine |
| C, Cys | cysteine |
| D, Asp | aspartic acid |
| E, Glu | glutamic acid |
| F, Phe | phenylalanine |
| G, Gly | glycine |
| H, His | histidine |
| I, Ile | isoleucine |
| K, Lys | lysine |
| L, Leu | leucine |
| M, Met | methionine |
| N, Asn | asparagine |
| P, Pro | proline |
| Q, Gln | glutamine |
| R, Arg | arginine |
| S, Ser | serine |
| T, Thr | threonine |
| V, Val | valine |
| W, Trp | tryptophan |
| Y, Tyr | tyrosine |

# Chapter 1 General Introduction

## 1.1 The nitrogen cycle and biological relevance of nitrate

Nitrogen is an essential element for all organisms, and is present in numerous biological molecules, including DNA, RNA, and protein. In the biosphere, nitrogen undergoes a variety of oxidations and reductions, which produce a variety of nitrogen-containing compounds with oxidation states ranging from +V (as in nitrate: $NO_3^-$) to –III (as in ammonia: $NH_3$). The series of reactions constitute the biogeochemical cycle of nitrogen (N-cycle) that are performed in different ways in wide range of organisms including bacteria, archaea, fungi, and plants (**Figure 1-1**).

There are two different reductive pathways in the N-cycle: the dissimilatory and assimilatory pathways. The dissimilatory reactions are predominantly performed by prokaryotic species[1] and a part of fungi[2]. In this pathway, $NO_3^-$ is used as the respiratory electron acceptor under anaerobic conditions and is further reduced by several different reductases. This is called nitrate respiration. In the complete dissimilatory nitrate reduction pathway, $NO_3^-$ is reduced to nitrite ($NO_2^-$), nitric oxide (NO), nitrous oxide ($N_2O$), and finally to $N_2$. This complete pathway is called denitrification. The majority of denitrifying bacteria are facultative aerobic heterotrophs that switch from aerobic respiration to nitrate respiration in the absence of oxygen. However, the full process of denitrification is accomplished only by a part of prokaryotes such as *Paracoccus denitrificans* and *Pseudomonas aeruginosa.* In the alternative dissimilative pathway, $NO_3^-$ is rapidly reduced anaerobically via $NO_2^-$ to $NH_4^+$, which is then excreted; this process is also known as $NO_3^-$/ $NO_2^-$ ammonification. Most commensal bacteria including *Escherichia coli* do not denitrify nitrate but catalyze this short-circuit reaction in the nitrogen cycle[3]. It should be noted that although nitrate in tissues and plasma in mammalian cells are inert end product of endogenous and exogenous nitrogen oxides and cannot be further metabolized by themselves, the commensal bacteria reduces nitrate to nitrite, which can be formed to various biologically active compounds, including the nitrogen oxides that bring possible protective effects to the host[3].

Another reductive pathway is constituted by the assimilatory processes, such as nitrate assimilation and $N_2$ fixation. Although nitrogen is very abundant in the atmosphere, it is mostly present in the gaseous form ($N_2$), and therefore is not available to most organisms, except for diazotrophs[4,5]. Most organisms utilize nitrate, nitrite and ammonium ions as inorganic nitrogen sources. Among these, nitrate is one of the most widely used inorganic nitrogen sources for both prokaryotes and eukaryotes. Plants growing in aerobic soil absorb nitrate from the roots and assimilate it[6]. In this pathway, $NO_3^-$ is reduced via $NO_2^-$ to $NH_4^+$. These reactions are similar to the $NO_3^-$/$NO_2^-$

ammonification, however the reductases used in each pathway are different. The produced $NH_4^+$ is assimilated via the glutamine synthetase and glutamine-oxoglutarate aminotransferase (GS/GOGAT) pathway into glutamine/glutamate, which serve as substrates for transamination reaction to produce all the other proteinaceous amino acids.

In both assimilative and dissimilative nitrate reduction pathways in N-cycle, the first step is cellular uptake of nitrate, which is catalyzed by nitrate transporters.

## 1.2 Nitrate transporters

The nitrate transporters are divided into three groups: the ATP-binding cassette (ABC) transporter superfamily[7], the Nitrate transporter 1/Peptide transporter Family (NPF)[8,9], and Nitrate Nitrite Porter (NNP) family.

ABC transporters are classified as primary transporters. ATP hydrolysis on the nucleotide binding domain (NBD) drives conformational changes in the transmembrane domains, resulting in the alternating access from outside and inside of the cell for unidirectional transport across the lipid bilayer. Although ABC transporters are widely conserved from prokaryotes to eukaryotes including mammals, the ABC-type nitrate (or nitrite) transporters are only conserved in part of bacteria such as *Klebsiella* and cyanobacteria such as *Synechococcus* species. The ABC-type nitrate transporters show high affinity for nitrate ($K_m$ value is ~ 1 μM).

NPF family is widely conserved from prokaryotes to eukaryotes including mammals. This family shows the diverse ranges of substrate specificity and only the subfamily in plant species known as NRT1 functions as a nitrate transporter. NFP family belongs to the Major Facilitator Superfamily (MFS) that is one of the largest secondary transporter groups. Several crystal structures of NPF family has been reported[10–16] including NRT1.1 from *Arabidopsis thaliana* (AtNRT1.1) which have the dual affinity for nitrate. The previously reported structures revealed that NPF family basically adopts the canonical MFS-fold in which 12 TM helices are organized into two discreetly folded N- and C-terminal bundles displaying a two-fold pseudosymmetry. Some of the prokaryotic peptide transporters have additional two helices[10], but these are not seen in plant NRT1 nitrate transporters[12,13]. The crystal structures of AtNRT1.1[12,13] revealed its unique dual affinity mechanism; the transition of $K_m$ value is due to the increase in the rate of nitrate transport as a result of increasing of structural flexibility, which is caused by the dimer decoupling induced by phosphorylation.

The NNP family is the most prevalent nitrate transporter group conserved in archaea, bacteria, fungi, yeast, algae and plants[17,18]. Although both the NNP and NPF families belong to the MFS, they share low amino-acid sequence similarity and the structure and transport mechanism of NNP family were poorly understood when the author started this research project. The author will summarize about the NNP family transporters in the next section; what has been clarified by previous research, and what has not been clarified yet, focusing on bacteria and plants.

## 1.3 NNP family in bacteria

The bacterial NNP family transporters play an important role in the uptake of nitrate as a terminal electron acceptor for anaerobic respiration, and the extrusion of cytotoxic nitrite produced by anaerobic respiration[19]. In *Escherichia coli*, there are two NNP family transporters, NarK and NarU (**Figure 1-2**). NarK is highly expressed during anaerobic growth in the presence of nitrate, and plays a central role in nitrate uptake in anaerobic conditions, while NarU is constitutively expressed in low level as a “house-keeper” nitrate transporter and has limited advantage during severe nutrient (nitrate) starvation or slow growth[20]. NarK and NarU share 75% amino acid sequence similarity, suggesting they have a similar transport mechanism. In general, in the process of nitrate assimilation in bacteria, ABC transporters are postulated to be used as well as the secondary transporters in NNP family. In contrast, it is postulated that bacteria rely solely on secondary transporters for nitrate respiration under anaerobic conditions[21].

Despite the importance of the NNP family transporters, their molecular-level functions are still under debate. Several hypotheses for the functions of bacterial NNP family transporters have been proposed: nitrate/nitrite antiporter[22,23], nitrite/$H^+$ antiporter[19,24,25], nitrate/$H^+$ symporter[21,24,25], simple nitrite uniporter[21,24], and $Na^+$- and/or $K^+$-coupled nitrate and/or nitrite transporter[26] (**Figure 1-3**). The lack of facile nitrogen and oxygen radioisotopes has hampered the direct measurement of their transport activity using *in vitro* reconstituted systems, and thus has limited an understanding of the precise functions of the NNP family transporters.

## 1.4 5年以内に雑誌等で刊行予定のため、非公開。

## 1.5 Previously reported structures of nitrate transporters

Recently, the crystal structures of NarK and NarU from *Escherichia coli* were reported, at resolutions of 2.6–2.8 Å and 3.0–3.1 Å, respectively[26,27] (**Table 1-1**). These structures revealed that both NarK and NarU contain 12 transmembrane (TM) helices with pseudo two-fold symmetry axes, like other MFS transporters, and suggested that they transport substrates via an alternating-access mechanism[28], in which the transporter adopts several conformations, involving outward-open, occluded, and inward-open states, during the transport cycle (**Figure 1-4**). Furthermore, these structures provided insights into the substrate recognition[26,27]. However, the molecular mechanism of the NNP family transporters remains largely unknown, due to our insufficient understanding of their molecular-level functions and the lack of the atomic-resolution structures. Recently, the structures of NRT1.1 from *Arabidopsis thaliana*, in the NPF family, were reported[12,13] (**Table 1-1**). The substrate recognition mechanism of NPF is largely different from that of the NNP family, and the molecular mechanism of the NNP family remains elusive.

## 1.6 Major facilitator superfamily (MFS)

Major facilitator superfamily (MFS) is one of the largest and most extensively studied secondary transporter superfamily, which shows the diverse range of substrate specificity. NNP and NPF families are both belonging to this superfamily. Since the determination of the structure of a lactose permease LacY as the first MFS structure in 2003[29], the structures of various kinds of MFS transporters in various conformations in the transport cycle have been reported; *e.g.* GlpT from *Escherichia coli* in organophosphate/phosphate antiporter (OPA) family in wide-open inward-facing state[30], multidrug exporter EmrD in occluded state[31], several bacterial PepT homologues in Proton-coupled oligopeptide transporter (POT) family in inward-open state[10–16], Fucose transporter FucP in Fucose $H^+$ symporter (FHS) family in outward open-state[32], the high-affinity phosphate transporter PiPT in phosphate:$H^+$ symporter (PHS) family as the first eukaryotic MFS transporter in an inward-facing occluded state[33], bacterial and mammalian GLUTs in Sugar Porter (SP) family in both inward- and outward-open states[34–39], and bacterial Ferroportin homologue in both inward- and outward-open states[40] (**Figure 1-5**). Although the MFS transporters share the common topological motif containing 12 transmembrane (TM) helices with pseudo two-fold symmetry axes and adopt the alternating-access mechanism for substrate transport[28], the sequence similarities are quite low between the different MFS families (typically under 20%) and there are large variations in structures (**Figure 1-5**). These structural variations led to the variety of substrate-recognition and transport mechanism: *i.e.* symporter, antiporter, or uniporter.

## 1.7 Overview of this study

In this study, to understand the transport mechanism by NNP transporters, the author performed the structural analysis of the crystal structures of NarK, a NNP transporter from *Escherichia coli*.

In **Chapter 2**, the expression, purification, LCP crystallization, and X-ray diffraction analysis of NarK are described in detail. The author determined the crystal structures of NarK in three distinct conformations (*i.e.*, nitrate-bound occluded, nitrate-bound inward-open, and *apo* inward-open states) at 2.35-2.4 Å resolutions.

In **Chapter 3**, a structural comparison of NarK revealed the substrate recognition mechanism and the conformational change between occluded and inward-open states.

In **Chapter 4**, The author performed various kinds of functional analyses; the genetic complementation assays were performed under the supervision of Dr. Koichi Ito in the University of Tokyo, the proteo-liposome based nitrite influx assay under the supervision of Dr. Hironori Takeda in the University of Tokyo, and the molecular dynamics simulations under the supervision of Dr. Ryuichiro Ishitani in the University of Tokyo. Based on the structural and functional analyses, the nitrate/nitrite antiport mechanism of NarK is proposed. In this mechanism, the structural changes and nitrate/nitrite recognition are coupled.

In **Chapter 5**, ５年以内に雑誌等で刊行予定のため、非公開。

In **Chapter 6**, this study is summarized and further discussed.

## Figures and Tables of Chapter 1

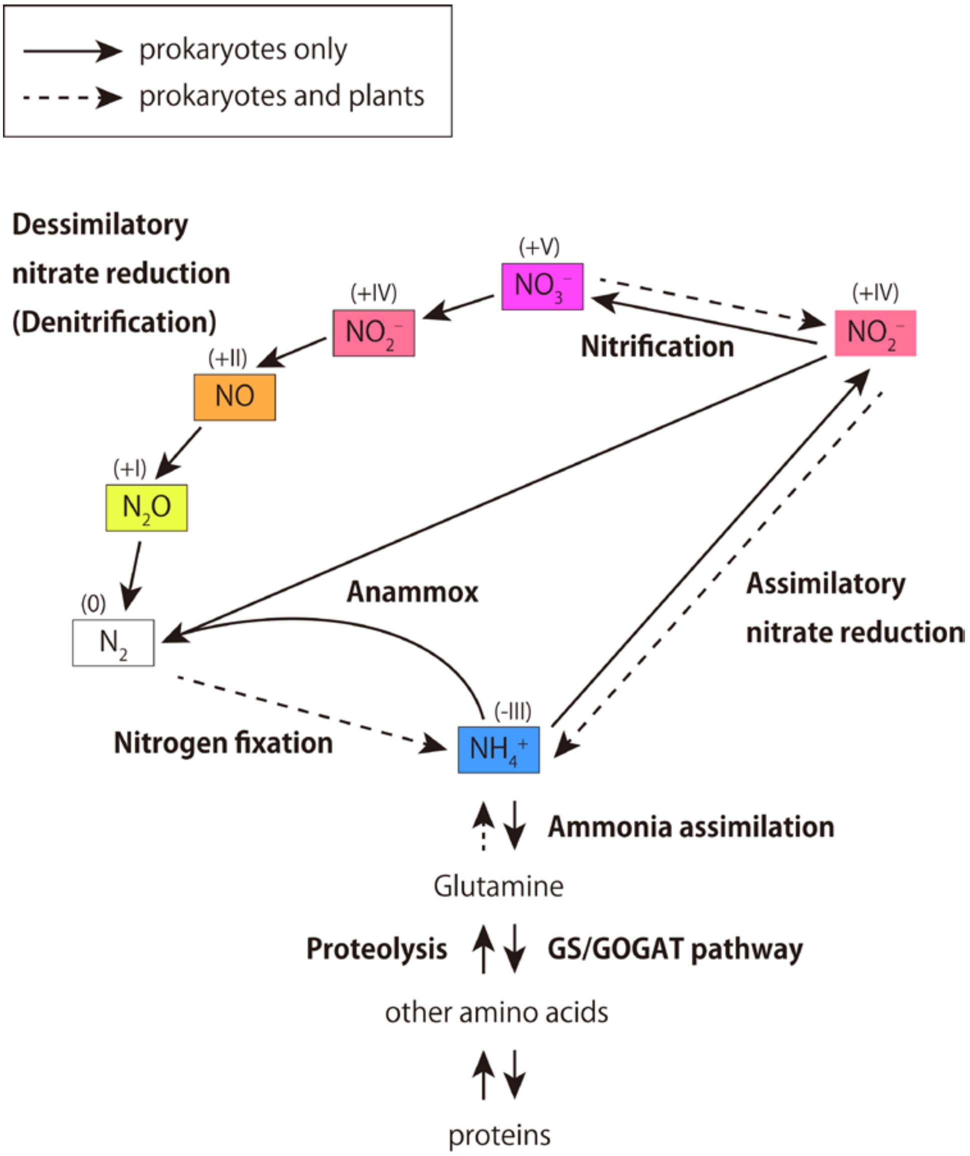


**Figure 1-1** **The nitrogen cycle.**

The schematic diagram of biogeochemical cycle of nitrogen (N-cycle). The processes seen only in prokaryotes are depicted by black solid arrows. The processes seen both in prokaryotes and plants are depicted by black dashed arrows. The oxidation numbers of each nitrogen-containing compound are also shown in parentheses.

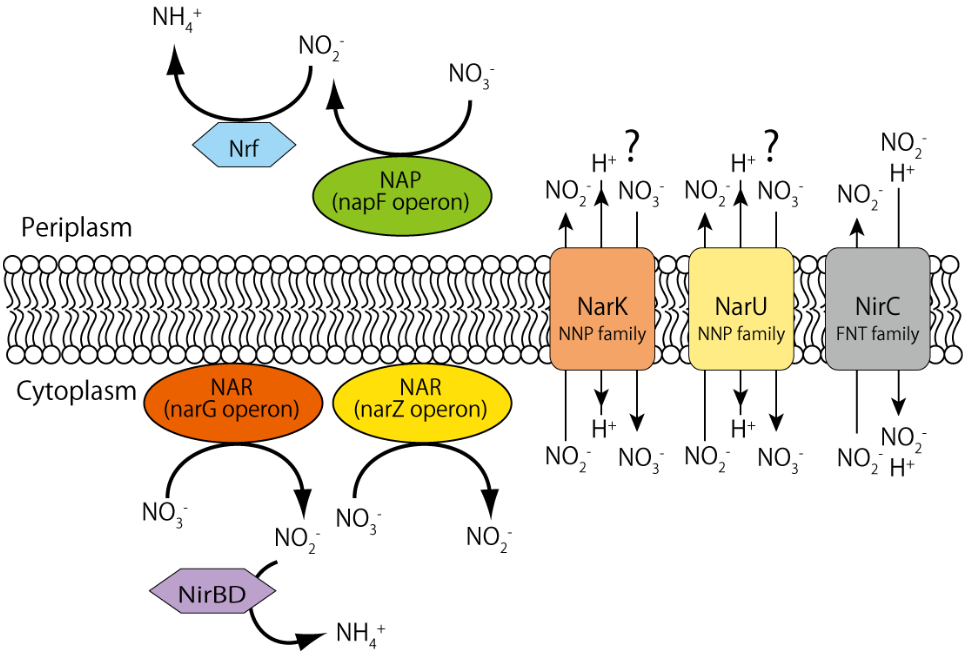


**Figure 1-2** **Nitrate transport and reduction system in *Escherichia coli.***

*E. coli* does not have an assimilatory nitrate reductase (known as Nas) in cytoplasm, but contains two membrane-associated nitrate reductases NarGHI and NarZYW encoded by the *narGHJI* and *narZYVW* operons, respectively. The third nitrate reductase is encoded by the *napFDAGHBC* operon and present in the periplasm as a soluble protein. In the cytoplasm, the NADH-dependent nitrite reductase NirBD reduces nitrite directly to ammonia. In the periplasm, this reaction known as $NO_3^-$/ $NO_2^-$ ammonification is also carried out by the cytochrome *c* nitrite reductase Nrf. The known nitrate transporters (NarK and NarU) and a nitrite channel (NirC) are also shown.

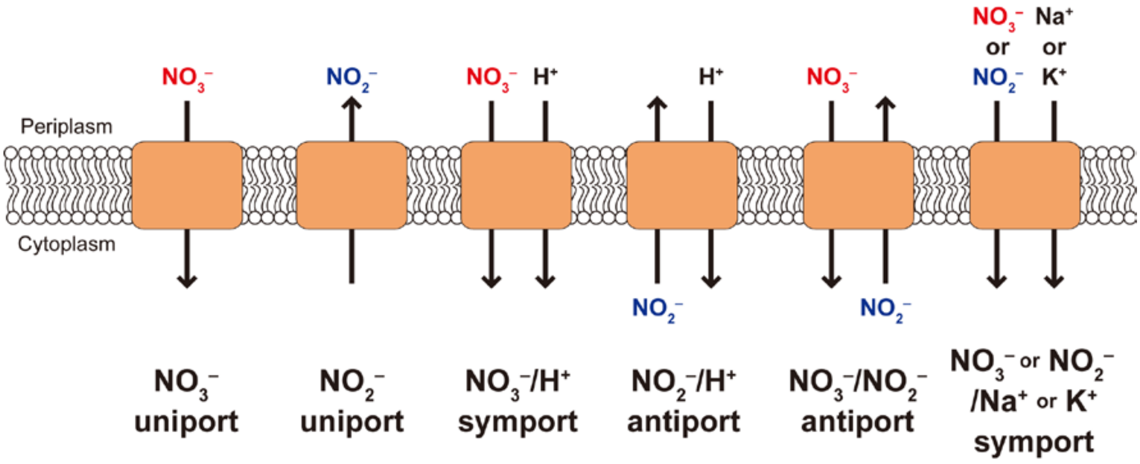


**Figure 1-3 Previously proposed transport models of NarK.**

The previously proposed substrate transport models of NarK are shown. The directions of substrate transport are depicted by black dashed arrows.

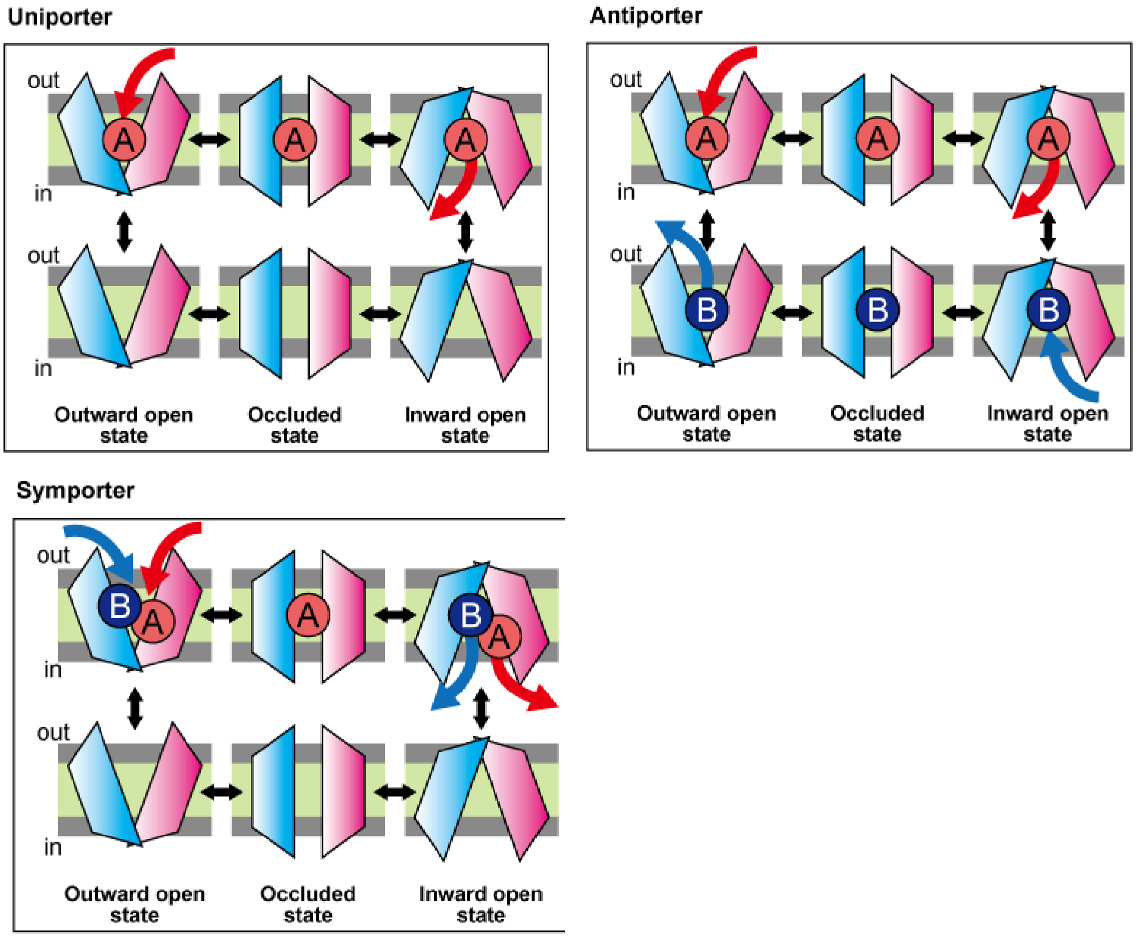


**Figure 1-4 Alternating access transport model.**

The substrate transport mechanism of uniporter, antiporter, and symporter. All of these transport cycles are accomplished by the conformational changes between an extracellular- and an intracellular-facing state via occluded state. In this alternating access mechanism, the substrate-binding site typically located in the central region of the transporter is accessible to only one side of the membrane at a time.

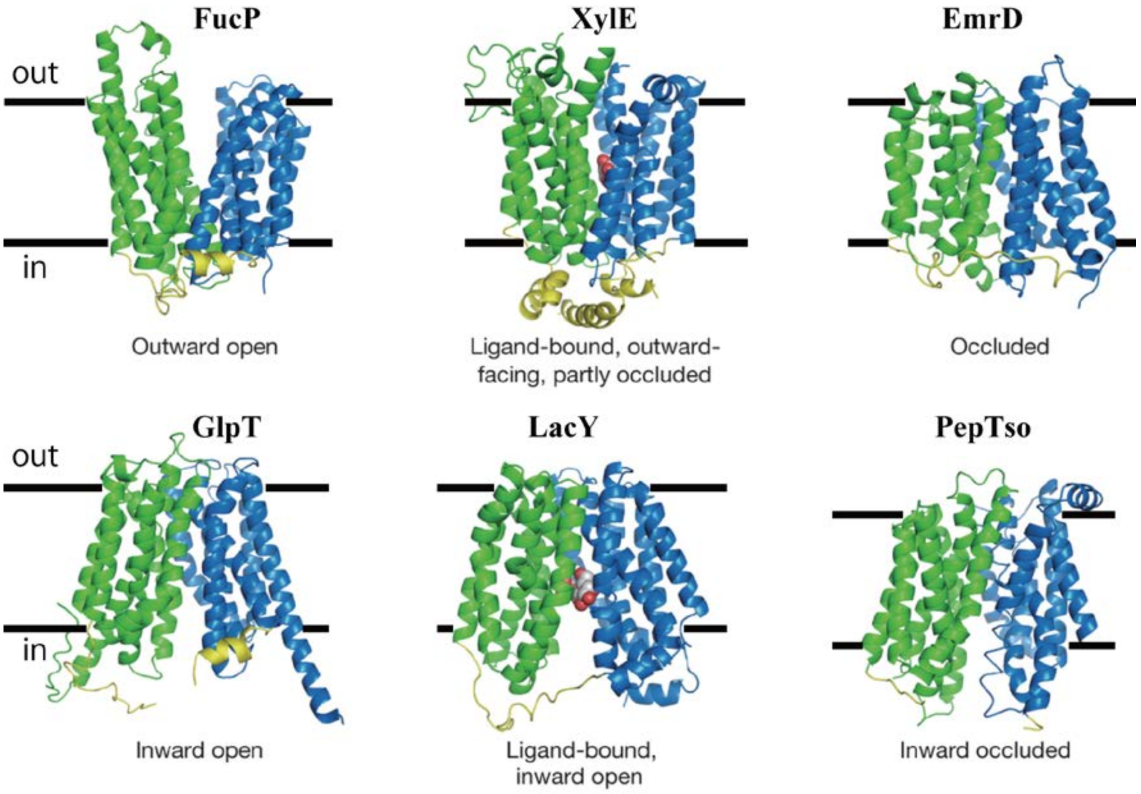


**Figure 1-5 Structures of various MFS transporters were reported in various conformations.**

The structures of MFS transporters have been determined in various conformations including inward-facing, outward-facing, and many other intermediate conformations. Although all the MFS share the canonical MFS fold composed of 12 transmembrane helices, their sequence similarities are quite low (~20%) and there are large variations in structures. These structural variations led to the wide range of substrate specificity. The Figures are edited from Sun, L., Zeng, X., Yan, C., Sun, X., Gong, X., Rao, Y., & Yan, N. (2012). Crystal structure of a bacterial homologue of glucose transporters GLUT1-4. *Nature*, **490**(7420), 361–6.

**Table 1-1 Previously reported structures of nitrate transporters.**

| Name | Organisms | Super family | Family | Resolution (Å) | Bound substrate | PDB ID | Reference |
|---|---|---|---|---|---|---|---|
| NarU | *Escherichia coli* | MFS | NNP | 3.0 Å<br>3.1 Å | None<br>$NO_3^-$ | 4IU9<br>4IU8 | Yan *et al*., ***Cell reports***, 2013 |
| NarK | *Escherichia coli* | MFS | NNP | 2.6 Å<br>2.8 Å | None<br>$NO_2^-$ | 4JR9<br>4JRE | Zheng *et al*., ***Nature***, 2013 |
| NRT1.1 | *Arabidopsis thaliana* | MFS | PNF | 3.7 Å<br>3.7 Å | None<br>$NO_3^-$ | 5A2N<br>5A2O | Parker *et al*., ***Nature***, 2014 |
|  |  |  |  | 3.25 Å | $NO_3^-$ | 4OH3 | Sun *et al*., ***Nature***, 2014 |

# Chapter 2 X-ray crystallographic analysis of NarK, a bacterial nitrate transporter from *Escherichia coli*

## 2.1 Materials and Methods

### 2.1.1 Plasmid construction

The *narK* (GI: 945783, UniProtKB: P10903) and *narU* (UniProtKB: P37758) genes were cloned from the *Escherichia coli* str. K-12 substr. MG1655 genome (ATCC 47076) provided by RIKEN BRC (Ibraki, Japan) and subcloned into a pET-modified vector[41]. This modified vector contains a C-terminal tobacco etch virus (TEV) protease cleavage site (ENLYFQG) followed by a GFP-His$_8$-tag (**Figure 2-1**). The resulting constructs were designated as NarK WT and NarU WT. The secondary structures and disordered regions of NarK and NarU were predicted using the program PSIPRED (http://bioinf.cs.ucl.ac.uk/psipred/) and the program DISOPRED2[42], respectively (**Figure 2-2**). The macromolecule-production information and DNA primers for cloning and creating mutants are summarized in **Table 2-1** and **Table 2-2**, respectively.

### 2.1.2 Small-scale expression of NarK and NarU

The NarK WT and NarU WT plasmids were introduced into *Escherichia coli* strain C41(DE3) Δ*acrB* cells. The transformed cells were grown in 3 mL of Luria-Bertani (LB) medium containing 50 µg mL$^{-1}$ ampicillin at 37°C to an absorbance at 600 nm

($Abs_{600}$) of 0.4-0.6, and expression was induced with 0.5 mM isopropyl-β-D-thiogalactoside (IPTG) at 20˚C for 24 hrs. The cells were centrifuged at 8,000 × g for 1 min, resuspended in buffer A’ and disrupted by sonication using Bioruptor UCW-310 (Cosmo Bio). The buffer compositions are summarized in **Table 2-3**. After removal of the debris by centrifugation at 12,000 × g, for 30 min at 4˚C, the supernatant was ultracentrifuged at 138,000 × g for 20 min at 4˚C to pellet the membrane fraction. The membrane fraction was resuspended in buffer A’, and solubilized at the rotator at 4˚C for 1.5 hrs. After removal of the insolubilized pellet by ultracentrifugation at 138,000 × g for 20 min at 4˚C, the homogeneity and expression level of the GFP-fusion protein samples were assessed by Fluorescence-detection Size Exclusion Chromatography (FSEC) using the HPLC equipped with the high sensitive fluorescence detector (Shimadzu). The buffer composition used for FSEC is as same as that of buffer G except that CHS was not added.

### 2.1.3 Large-scale expression and purification of native NarK

The NarK WT was expressed as described in **2.2.1** except that the culture scale was 2.5 L. The typical purification method is described below. The cells were centrifuged at 4,500 × g for 10 min, resuspended in buffer A and disrupted by three passages through a Microfluidizer Processor (Microfluidics) at 15,000 p.s.i. After removal of the debris by centrifugation at 12,000 × g, for 30 min at 4˚C, the supernatant was ultracentrifuged at

138,000 × g for 1 hr at 4˚C to pellet the membrane fraction. The membrane fraction was resuspended in buffer A, flash-frozen by liquid nitrogen, and stored at −80ºC until use. The membrane fraction was solubilized in buffer B for 2 hrs at 4ºC. Insoluble materials were removed by ultracentrifugation at 138,000 × g for 30 min at 4˚C, and the supernatant was mixed with 10 mL Ni-NTA Superflow resin (Qiagen) equilibrated with buffer C. After binding for 2 hrs at 4ºC, the mixture was loaded into an Econo-column (Bio-Rad), and the flow-through fraction was discarded. The resin was washed with ten column volumes of buffer D, and the protein sample was eluted with buffer E. To cleave the GFP-His$_8$-tag, C-terminus His$_6$-tagged tobacco etch virus (TEV) protease S219V mutant (produced in-house) was added to the eluted fraction in a 10:1 (w/w) protein:TEV protease ratio. During the overnight cleavage reaction at 4ºC, the solution was dialyzed against buffer F. The sample was mixed with 10 mL Ni-NTA superflow resin again for 1 hr at 4ºC, to remove the TEV protease. The flow-through fraction containing NarK was collected, concentrated by an Amicon Ultra centrifugal filter (50 kDa molecular weight cutoff, Millipore), and further purified on a HiLoad 16/600 Superdex 200 pg column (GE Healthcare), equilibrated with buffer G. The peak fractions were concentrated to approximately 18-30 mg mL$^{-1}$ by an Amicon Ultra filter (50 kDa molecular weight cutoff, Millipore), and ultracentrifuged at 138,000 × g for 20 min at 4˚C to remove the precipitation. The sample was flash-frozen by liquid nitrogen, and stored at −80ºC until use. The purity of the protein sample was assessed by SDS-PAGE (**Figure 2-3**). The concentrations of the protein samples in each step were calculated based on the $Abs_{280}$ values measured by the NanoDrop™ 2000/2000c Spectrophotometers (Thermo Fisher Scientific). The purity, homogeneity, and thermo-stability of the protein samples were assessed by FSEC.

### 2.1.4 Crystallization of native NarK

For the lipidic cubic phase (LCP) crystallization trials, the protein solution was mixed with monoolein (Nu-Chek Prep) at a 4:6 w/w ratio using coupled syringe devices, and the 30-80 nL protein-laden LCP droplets were overlaid with 1 μL precipitant solution on both 96-well plastic sandwich plates (SWISSCI) by the crystallization robot, mosquito LCP (TTP LabTech), and glass sandwich plates by the repeating dispenser (Hamilton). Initial crystallization screening was performed at 20ºC using the in-house-produced grid-screening crystallization kits as reservoir solutions. For crystallization optimization, StockOptions Salt and Additive Screen (Hampton Research) were added to the reservoir solutions, in addition to the optimization of the pH and the concentrations of precipitants and salts. The crystals of NarK were grown at 20˚C, and three different types of crystals (Crystal I, II, and III) were obtained in reservoir solutions I, II, and III, respectively (reservoir solution I: 27% PEG500DME, 100 mM MES-NaOH (pH 6.5), and 100 mM $NH_4NO_3$; reservoir solution II: 36% PEG400, 100 mM MES-NaOH (pH 6.0), 10 mM $Zn(OAc)_2$, and 3% sucrose; reservoir solution III: 31% PEG500DME, 100 mM MES-NaOH (pH 5.8), and 100 mM $NH_4NO_3$) (**Figure 2-5**). The protein concentrations were 30, 18, and 18 mg $mL^{-1}$ for Crystal I, II, and III, respectively. All crystals grew to full size in 2–3 weeks. The crystallization conditions are summarized in **Table 2-4**.

### 2.1.5 X-ray data collection from native NarK crystals and data processing

All the diffraction data were collected by the helical data collection method from the single crystal using a micro-focus X-ray beam at SPring-8 beamline BL32XU[43]. To reduce radiation damage, all X-ray diffraction experiments were performed under nitrogen gas stream at 100 K. Loop-harvested crystals were identified by raster scanning.

A diffraction data set was collected from crystal I (described below), using a micro-focus X-ray beam (a 1μm width × a 15 μm height) at a wavelength of 1.0000 Å with an oscillation range of 180˚ (1.0˚ per image), an exposure time of 1.0 s per image and the camera distance of 240 mm. The diffracted X-ray photons were detected by an MX225HS charge coupled device (CCD) detector (Rayonix).

A diffraction data set was collected from crystal II (described below), using a micro-focus X-ray beam (a 1μm width × a 15 μm height) at a wavelength of 1.0000 Å with an oscillation range of 180˚ (1.0˚ per image), an exposure time of 1.0 s per image and the camera distance of 280 mm. The diffracted X-ray photons were detected by an MX225HS charge coupled device (CCD) detector (Rayonix).

A diffraction data set was collected from crystal III (described below), using a micro-focus X-ray beam (a 1μm width × a 15 μm height) at a wavelength of 1.0000 Å with an oscillation range of 68.40˚ (0.60˚ per image) and further 33.25˚ (0.35˚ per image), an exposure time of 1.0 s per image and the camera distance of 255 mm. The diffracted X-ray photons were detected by an MX225HS charge coupled device (CCD) detector (Rayonix).

X-ray diffraction data sets were indexed, integrated and scaled by the programs XDS[44], DIALS[45] and AIMLESS[46]. The data processing statistics are summarized in **Table 2-5**.

### 2.1.6 Trials of molecular replacement

Using the native data from the Crystal II collected in December 2011, the author tried to determine the structure of NarK by molecular replacement using the program Phaser. As the search models for molecular replacement, the following five MFS structures available at that time were used; LacY (PDB ID:2CFG), GlpT (PDB ID:1PW4), EmrD (PDB ID:2GFP), FucP (PDB ID:3O7Q), PepTso (PDB ID:2XUT). The amino acid sequence identity calculated by several different algorithms are shown in **Figure 2-6**.

### 2.1.7 Expression, purification, crystallization, and X-ray data collection of Selenomethionine (SeMet) derivative NarK

The NarK WT plasmid constructed in **2.1.1** was introduced into *Escherichia coli* strains C41 (DE3) Met- containing pRare plasmid, which requires methionine in the medium for growth (Ito *et al.*, unpublished). The expression conditions were optimized by using FSEC as described in **2.1.2**. SeMet derivative NarK was expressed in large scale and purified as described in **2.1.3** except the points as follows; (1) The SeMet core medium (Nihon Pharmaceutical) containing 50 μg mL$^{-1}$ ampicillin, 1% D-Glucose, 1 mM $MgSO_4$, 0.015 mM $FeSO_4$, 0.167 mM L-selenomethionine (L-SeMet), and 1% MEM Vitamin Solution (×100) (Wako) were used for culture. (2) The cells were cultured for 30 hrs at 25˚C after induction. (3) TEV protease was added to the eluted fraction in a 5:1 (w:w) protein:TEV protease ratio. The purified SeMet NarK samples (**Figure 2-8**) were concentrated to 22 mg/mL and crystallized by LCP method by using mosquito LCP crystallization robot as described in **2.1.4**. The tiny crystals were obtained under several conditions, and optimization of crystallization conditions led to needle-shaped crystals (70 µm × 5 µm × 5 µm) in the reservoir solutions containing 30% PEG500DME, 100 mM HEPES-NaOH pH7.0, 10 mM $Zn(OAc)_2$, and 50 mM $HCOONH_4$. SeMet NarK crystals grew to full size in 12 h (**Figure 2-8**).

Before the data collection, XAFS spectrums were measured to determine the proper wave length to collect the x-ray diffraction data from the SeMet NarK crystals. A diffraction data set was collected from the single crystal by the helical data collection method at SPring-8 beamline BL32XU[43], using a micro-focus X-ray beam (a 1μm width × a 15 μm height) at a wavelength of 0.97940 Å with an oscillation range of 360˚ (2.0˚ per image), an exposure time of 1.0 s per image and the camera distance of 300 mm. The diffracted X-ray photons were detected by an MX225HS charge coupled device (CCD) detector (Rayonix). The data processing statistics are summarized in **Table 2-5**. The SeMet crystal diffracted to 3.2 Å (**Figure 2-8**). The author tried to determine the structure by using the Single wavelength Anomalous Dispersion (SAD) method using this SeMet data set, however the anomalous signal was too low to determine the structure.

### 2.1.8 Trials of co-crystallization with heavy atom-containing reagents

The various heavy atom containing reagents shown in **Table 2-6** were dissolved in DMSO, added to the crystallization protein samples at the final concentrations of 1-10 mM, incubated at 20˚C for 3-20 hours, and crystallized by the LCP method by using the mosquito LCP crystallization robot as described in **2.1.4 (Figure 2-9)**. Although some crystals were obtained by co-crystallization, data sets were not collected because of the crystal size and low reproducibility.

### 2.1.9 Soaking with heavy atom reagents

#### 2.1.9.1 Heavy atom-binding assay

Prior to soaking the Hg-containing reagents into the native NarK crystals, the binding capacity of each Hg-containing reagent were assessed by the Fluorescence Detection of Heavy Atom Labeling (FD-HAL) competition assay (**Figure 2-10)**. First, the protein samples were incubated with various mercury-containing reagents at 20˚C, and then mixed with a fluorescence reagent tetramethylrhodamine-5-maleimide (TMRM), which can be bound to the accessible (surface-exposed) cysteine residues of the target protein by alkylation reaction. After the reaction, all the samples were analyzed by SDS-PAGE and the TMRM fluorescence signals were detected from the gel. If the mercury-containing reagent binds tightly to the cysteine residues of the target protein, TMRM cannot attack to the target, and therefore the fluorescence signal is not detected.

#### 2.1.9.2 Soaking with heavy atom reagents

The heavy atom-derivative crystals were prepared by the soaking method. After the native crystal III was grown on glass sandwich plates to the full size, about 15 × 40 ×

80 μm, the overlaid crystallization solution (1 μL) was replaced by 1-1.5 μL of the solution supplemented with a slightly lower concentration of PEG500DME and 4 mM $CH_3HgCl$. The crystal was carefully incubated at 20˚C for 5 hours in the highly humid atmosphere in order to prevent the phase transition in the LCP droplet. The crystals were harvested using MicroMounts (MiTeGen) and LithoLoops (Protein Wave), and were flash-cooled in liquid nitrogen.

#### 2.1.9.3 X-ray data collection from Hg-soaked NarK crystal and data processing

A diffraction data set was collected from $CH_3HgCl$-soaked crystal III (described below), using a micro-focus X-ray beam (a 1μm width × a 15 μm height) at a wavelength of 1.0000 Å with an oscillation range of 240˚ (1.5˚ per image), an exposure time of 1.0 s per image and the camera distance of 320 mm. The diffracted X-ray photons were detected by an MX225HS charge coupled device (CCD) detector (Rayonix). The data processing statistics are summarized in **Table 2-5**.

### 2.1.10 Phase determination, model building, and structure refinement

The structure was determined by the single isomorphous replacement with anomalous scattering (SIRAS) method, using the native and $CH_3HgCl$-soaked crystals.

Six Hg atom sites were identified with the program SHELXD[47,48]. The initial phases were calculated using the program SHARP[49], followed by solvent flattening with SOLOMON[50] and 2-fold non-crystallographic symmetry (NCS) averaging using DM[51]. The initial model structure of NarK was built with the program Phyre2[52], using the glycerol-3-phosphate transporter GlpT structure (PDB accession code: 1PW4)[30] as a template. The model was further built manually using COOT[53,54] and refined using PHENIX[55,56]. Using this Crystal III structure as a search model, the Crystal I and II structures were determined by the molecular replacement method, using the program PHASER[57]. Crystals I and II contain one molecule in each asymmetric unit, while Crystal III contains two molecules (Mol A and B) in each asymmetric unit. Mol A and B in Crystal III are essentially the same (r.m.s.d. of 0.38 Å for all Cα atoms), and thus Mol A is mainly used for the following discussion. The structure refinement statistics are summarized in **Table 2-5**.

## 2.2 Results

### 2.2.1 Plasmid construction and expression check by FSEC

In general, the expression level and stability of the protein sample are critical factors for structural analysis. In particular, membrane proteins are more difficult to express, purify and crystallize compared to water-soluble proteins. Recently the high throughput and efficient precrystallization screening strategy based on the fluorescence-detection size-exclusion chromatography (FSEC)[58]. In this method, the target membrane protein is expressed in a GFP-fusion construction, and the SEC profile of the resulting fusion protein is monitored by fluorescence spectroscopy. This FSEC method allows for evaluation of the expression level and monodispersity of the target protein rapidly using small amount (typically nanograms to micrograms) of unpurified material from whole-cell extracts.

The target proteins in this research, NarK and NarU, are both belonging to one of the largest secondary transporter superfamily, Major Facilitator Superfamily (MFS) which contains 12 transmembrane helices. As shown in **Figure 2-2**, both the C- and N-terminus were predicted to be in cytoplasmic side by using the program TOPCONS. The C-terminus GFP-fused construction was chosen for expression screening.

The full-length NarK and NarU were overexpressed in 3 mL *E. coli* cells in the cleavable C-terminus GFP-His-tag-fused construction. After solubilizing the membrane fractions, the crude lysates were analyzed by GFP-FSEC. As shown in **Figure 2-3**, although the peak profiles of both NarK and NarU were sharp and symmetrical, the peak shape of NarK was more closed to the ideal Gaussian shape, suggesting that NarK was highly monodisperse in GFP-fusion form. The peak profiles of NarK and NarU were not changed 1 week after the solubilization, suggesting their high stability in the detergent solution. Moreover, the GFP-fluorescence peak intensity of NarK was six-fold greater than that of NarU. These data led to the conclusion that NarK was a good candidate for structural analysis.

### 2.2.2 Expression and Purification of NarK

The full-length NarK in the cleavable C-terminus GFP-His-tag-fused construction was overexpressed in 2.5 L *E. coli* cells and purified by using Ni-NTA affinity chromatography and size exclusion chromatography (SEC). The purified protein sample was high purity and the SEC chromatogram peak showed monodispersity (**Figure 2-4**). The final yield was about 5 mg from 1 L *E. coli* culture.

### 2.2.3 Crystallization of NarK

An LCP crystallization screening of NarK was performed. Initially, tiny crystals were obtained under various conditions using sandwich plates. To improve the crystal size and quality, the purification condition was optimized. The key points were replacement of detergent from DDM to DDM+CHS and adding the nitrate into the buffer in all the purification step. The thermo-stability of NarK was dramatically increased by adding nitrate in the purification buffer. Furthermore, the protein concentration for crystallization was also the critical point. As a result of optimization efforts, including the screening of additives using StockOptions Salt and Additive Screen kits, three different crystals (Crystal I, II, III) were obtained from the different conditions (**Figure 2-5**). Crystal I and II are quite similar rod-shaped forms but the crystallization conditions are different. Crystal I and II appeared in a day and grew to approximate dimensions of 10 × 10 × 100-200 μm in a week. In contrast, Crystal III is a plate-shaped form appeared in a week and grew to approximate dimensions of 10 × 30 × 60-80 μm in three weeks. The Crystal I and III showed high sensitivity to pH of the crystallization conditions.

### 2.2.4 Molecular replacement

After the first native data set was collected from the Crystal II in December 2011, the author tried to determine the structure by molecular replacement using the available structures of MFS transporters. The best result was obtained when the poly-alanine model of GlpT was used as a search model (**Figure 2-6**). However, the *R* factors were too high to model and refine the structure, probably due to the relatively low sequence identity with other MFS transporters, strongly suggesting the necessity of experimentally phasing for structure determination.

### 2.2.5 Soaking of heavy atom reagents into LCP crystals

The result of molecular replacement strongly suggested that experimentally phasing is needed for structure determination. Although the SeMet derivative crystals and co-crystal with several heavy atom reagents were obtained but the phasing problem could not be not solved in any case. The author tried to apply the soaking method on the LCP crystals, although there was only one success report for phase determination by LCP-soaking method before the structure determination of NarK by our team. Mercury is one of the most popular heavy atoms used for experimental phasing because of its strong phasing power and proper wave length for detecting anomalous signal. Therefore, among the various heavy atoms that has been used to determine the protein structure, the mercury-containing reagents were chosen. Prior to the soaking trials, the binding capacity of each Hg-containing reagent were assessed by the Fluorescence Detection of Heavy Atom Labeling (FD-HAL) competition assay (**Figure 2-10)**. For reproducibly control the soaking condition, it was quite important to prevent drying up of the LCP droplet. The Crystal III was mainly used for soaking trials as it is well-diffracted and highly reproducible crystal form. It was technically difficult to soak the heavy atom-containing reagents into LCP crystals on the glass plate (**Figure 2-10)**, but the extensive practice and the diamond glasscutter finally allowed us to control the soaking conditions.

### 2.2.6 Overlap problem

On the way to collect the data set from the Crystal III, the author has been suffered from the serious “overlap” problem (**Figure 2-7**). Although the micro-focus beam in BL32XU in SPring-8 has been enabled the crystallographers to determine a lot of membrane protein structures from micro crystals, enough crystal widths (typically at least ~30 µm) are necessary in order to collect the complete data set from single crystal. In the case of the plate-shaped Crystal III (**Figure 2-5**), the length of the long axis of the unit cell is relatively long ($c$ = 225) and the corresponding side of the crystal is usually quite thin (typically 5-10 µm). These plate-shaped crystals were highly reproducible and diffracted well, but when the crystal was placed on the loop as shown in **Figure 2-7**, the serious overlapping of the diffraction spots was observed. It might be possible to avoid the problem by moving the detector closer to the crystal, but this is not the best solution in this case because the camera distance and the maximum resolution are tradeoff. This “overlap” problem was overcome by following two strategies; (1) The oscillation angle was reduced from typical 1.0˚ to 0.3-0.6˚ in order to limit the diffraction spots in one image. (2) The orientation of the crystals was controlled to balance the widths to be used for data collection and the “overlap” problem. One possible way is to mount the crystal not parallel to the loop (left panel), but it is technically difficult to control the orientation of the crystals strictly. Another way is to use the angled-tip versions of Micromounts loop (right panel). This is easy and highly reproducible way to control the orientation of the crystals on the loop, and was actually critical to avoid the “overlap” problem.

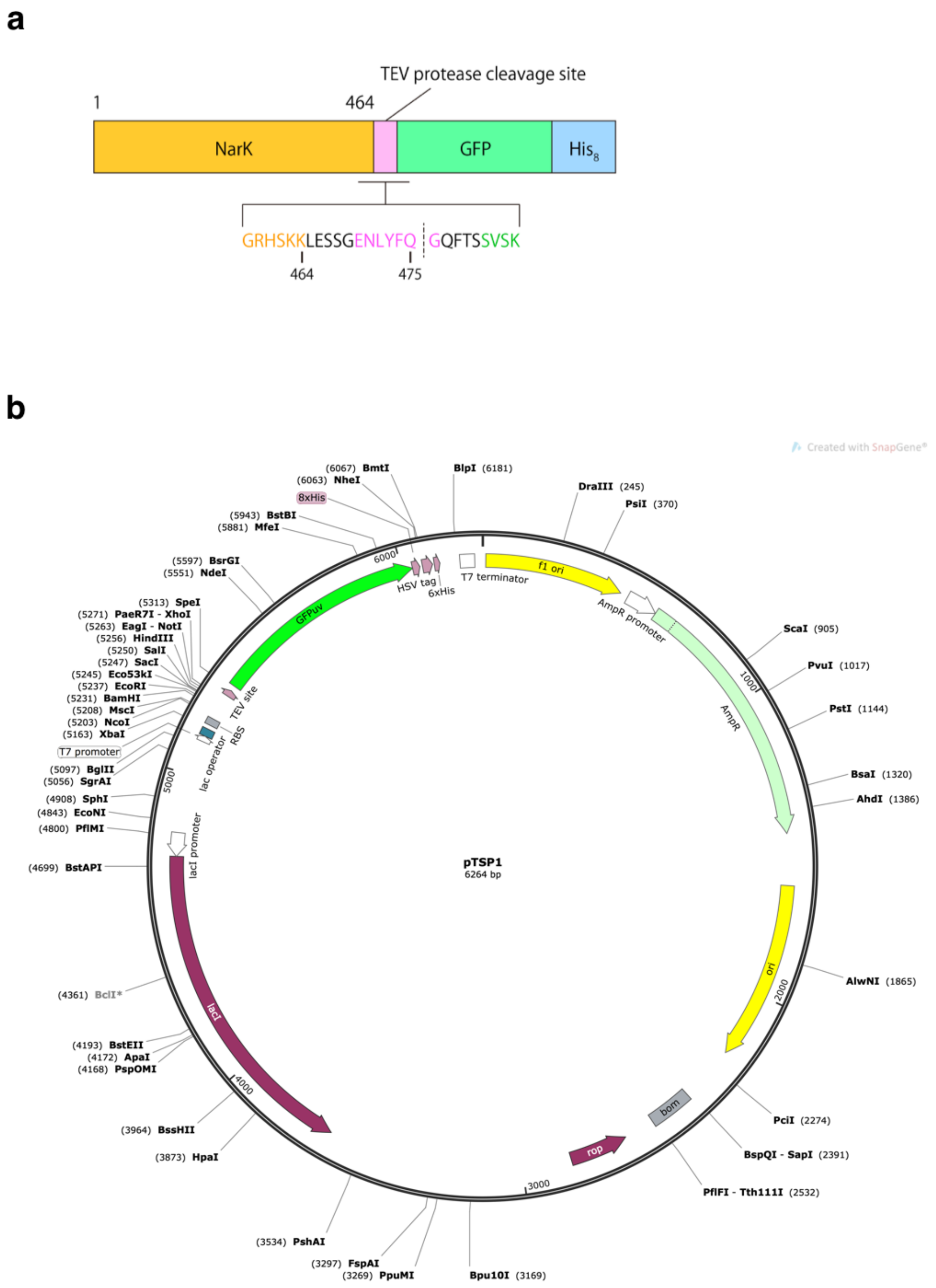
### 2.2.7 Phase determination, model building, and structure refinement

The structure was determined by the single isomorphous replacement with anomalous scattering (SIRAS) method, using the native and $CH_3HgCl$-soaked crystals (Crystal III) (**Figure 2-11**). It was important to set some of the parameters, for example the number of the heavy atoms to search and the resolution cut off. NCS averaging and solvent flattering has greatly improved the quality of the density map. The final models of the Crystal I structure (2.35 Å resolution, $R_{\mathrm{work}} / R_{\mathrm{free}}$ = 0.2037/0.2318), the Crystal II structure (2.40 Å resolution, $R_{\mathrm{work}} / R_{\mathrm{free}}$ = 0.2170/0.2475), and the Crystal III structure (2.40 Å resolution, $R_{\mathrm{work}} / R_{\mathrm{free}}$ = 0.2191/0.2368) were refined using data sets from crystals I, II, and III, respectively. The crystal packing of NarK in crystals I (or II) and III are shown in **Figure 2-12**. The stereochemical quality of the models was validated by Ramachandran plot analysis (**Figure 2-13**). Data collection and refinement statistics are shown in **Table 2-5**. The coordinates and the structure factors have been deposited in the Protein Data Bank (PDB), under the accession codes 4U4V (*apo* inward-open state), 4U4T (nitrate-bound inward-open state) and 4U4W (nitrate-bound occluded state)

## 2.3 Summary of this chapter and perspectives

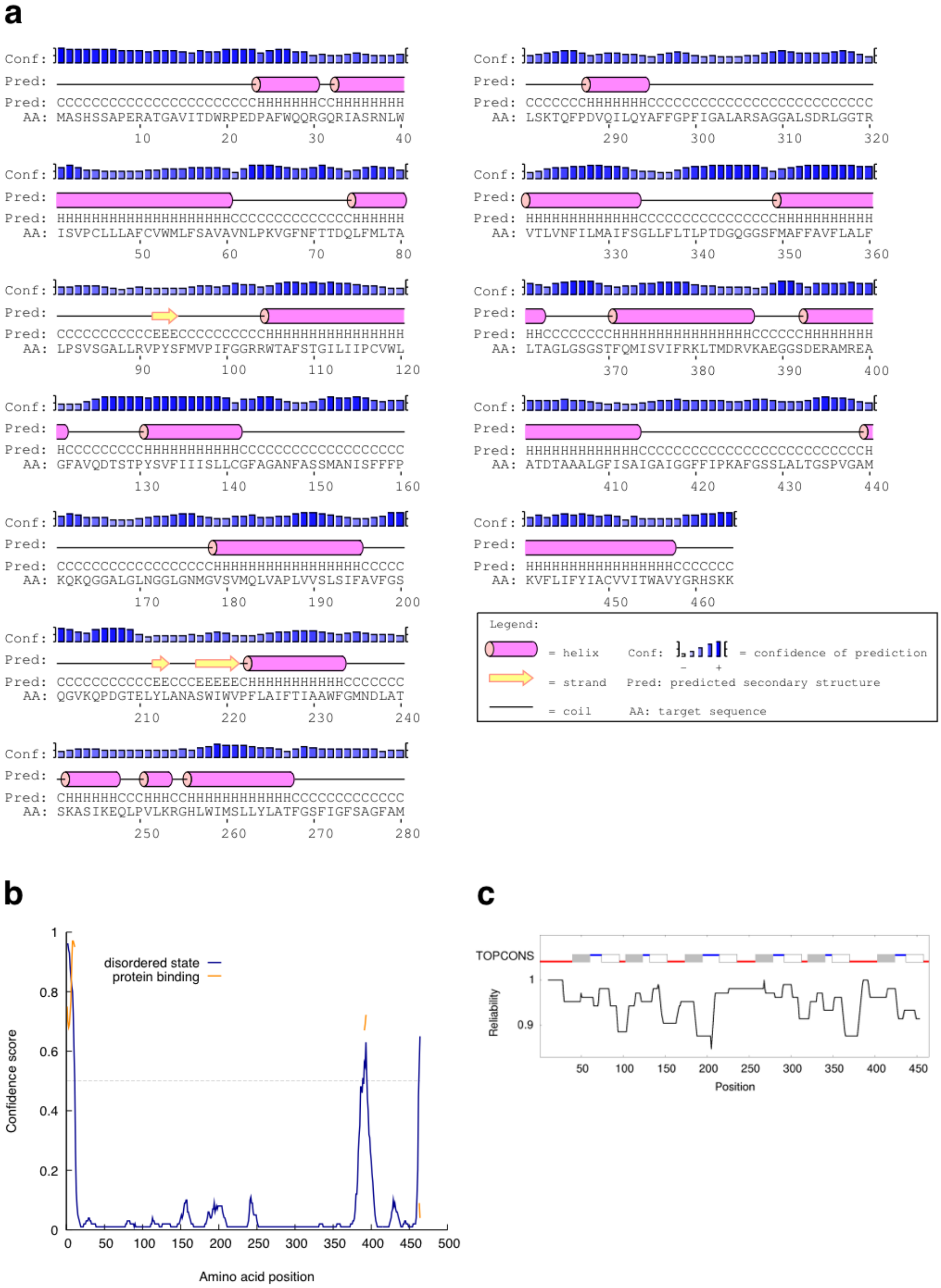
The author cloned two well-known NNP family genes, *narK* and *narU* from the genomic DNA of *Escherichia coli*, and identified that NarK is a more promising candidate for structural analysis by using the GFP-FSEC-based high throughput screening method (**Figure 2-3**). NarK was purified by Ni-NTA affinity chromatography and size exclusion chromatography (**Figure 2-4**), and crystallized by LCP method. Three different forms of crystals (Crystal I, II, III) were obtained and they all diffracted to 2.3-2.4 Å (**Figure 2-5**). For obtaining the high-quality crystals, it was a key point to add nitrate in all the purification step and also in the crystallization conditions. The other important points for crystallization were to concentrate the protein to ~25 mg/mL, to use the glass plates for crystallization, and carefully optimize the buffer pH in the crystallization conditions. In the process of data collection, the “overlap” problem was occurred (**Figure 2-7**), but the author overcame this problem by using the uncommon bending loos for fishing the crystals and by optimizing the setting for data collection in the SPring-8 BL32XU synchrotron beamline. After collecting the native data set of NarK, the author tried several ways to solve the structure; *i.e.* molecular replacement (**Figure 2-6**), crystallization of SeMet derivatives (**Figure 2-8**), and co-crystallization with various heavy atom reagents (**Figure 2-9**), however either these approaches did not work. It took more than one year to solve the phase problem but finally, by applying the heavy atom soaking method on the LCP crystals (**Figure 2-10**), the author solved the phase problem (**Figure 2-11**), and determined the high-resolution structures of NarK in three different states in the transport cycle (**Table 2-4**). In the following chapter, the author will describe the structural feature of NarK precisely and discuss its transport mechanism in molecular level.

## Figures and Tables of Chapter 2



## Figure 2-1 Schematic diagram of the NarK crystallization construct.

(**a**) Expression constructs of full-length NarK. The TEV protease cleavage site and a GFP-His$_8$-tag were fused into the C-terminus of inserted gene. (**b**) The expression vector map.



```
Pred: CCCCCCCCCCCCCCCCCCCCCCCHHHHHHHCCHHHHHHHH
  AA: MASHSSAPERATGAVITDWRPEDPAFWQQRGQRIASRNLW
              10        20        30        40

Pred: HHHHHHHHHHHHHHHHHHHHCCCCCCCCCCCCCCHHHHHH
  AA: ISVPCLLLAFCVWMLFSAVAVNLPKVGFNFTTDQLFMLTA
              50        60        70        80

Pred: CCCCCCCCCCCEEECCCCCCCCCCHHHHHHHHHHHHHHHH
  AA: LPSVSGALLRVPYSFMVPIFGGRRWTAFSTGILIIPCVWL
              90       100       110       120

Pred: HCCCCCCCCCHHHHHHHHHHHCCCCCCCCCCCCCCCCCCC
  AA: GFAVQDTSTPYSVFIIISLLCGFAGANFASSMANISFFFP
             130       140       150       160

Pred: CCCCCCCCCCCCCCCCCHHHHHHHHHHHHHHHHHHCCCCC
  AA: KQKQGGALGLNGGLGNMGVSVMQLVAPLVVSLSIFAVFGS
             170       180       190       200

Pred: CCCCCCCCCCCEECCCEEEEECHHHHHHHHHHHCCCCCCC
  AA: QGVKQPDGTELYLANASWIWVPFLAIFTIAAWFGMNDLAT
             210       220       230       240

Pred: CHHHHHHCCCHHHCCHHHHHHHHHHHHHCCCCCCCCCCCC
  AA: SKASIKEQLPVLKRGHLWIMSLLYLATFGSFIGFSAGFAM
             250       260       270       280

Pred: CCCCCCCHHHHHHHCCCCCCCCCCCCCCCCCCCCCCCCCC
  AA: LSKTQFPDVQILQYAFFGPFIGALARSAGGALSDRLGGTR
             290       300       310       320

Pred: HHHHHHHHHHHHHCCCCCCCCCCCCCCCHHHHHHHHHHHH
  AA: VTLVNFILMAIFSGLLFLTLPTDGQGGSFMAFFAVFLALF
             330       340       350       360

Pred: HHCCCCCCCHHHHHHHHHHHHHHHHHCCCCCCHHHHHHHH
  AA: LTAGLGSGSTFQMISVIFRKLTMDRVKAEGGSDERAMREA
             370       380       390       400

Pred: HHHHHHHHHHHHHCCCCCCCCCCCCCCCCCCCCCCCCCCH
  AA: ATDTAAALGFISAIGAIGGFFIPKAFGSSLALTGSPVGAM
             410       420       430       440

Pred: HHHHHHHHHHHHHHHHHCCCCCCC
  AA: KVFLIFYIACVVITWAVYGRHSKK
             450       460
```

## Figure 2-2 Secondary structure and topology prediction.

(**a**) Secondary structure of full-length NarK was predicted using the program Psipred (http://bioinf.cs.ucl.ac.uk/psipred/). (**b**) Disordered probability of amino acids of NarK predicted using the program DISOPRED3[59]. (**c**) The membrane topology prediction predicted by using the program TOPCONS (http://topcons.cbr.su.se/).

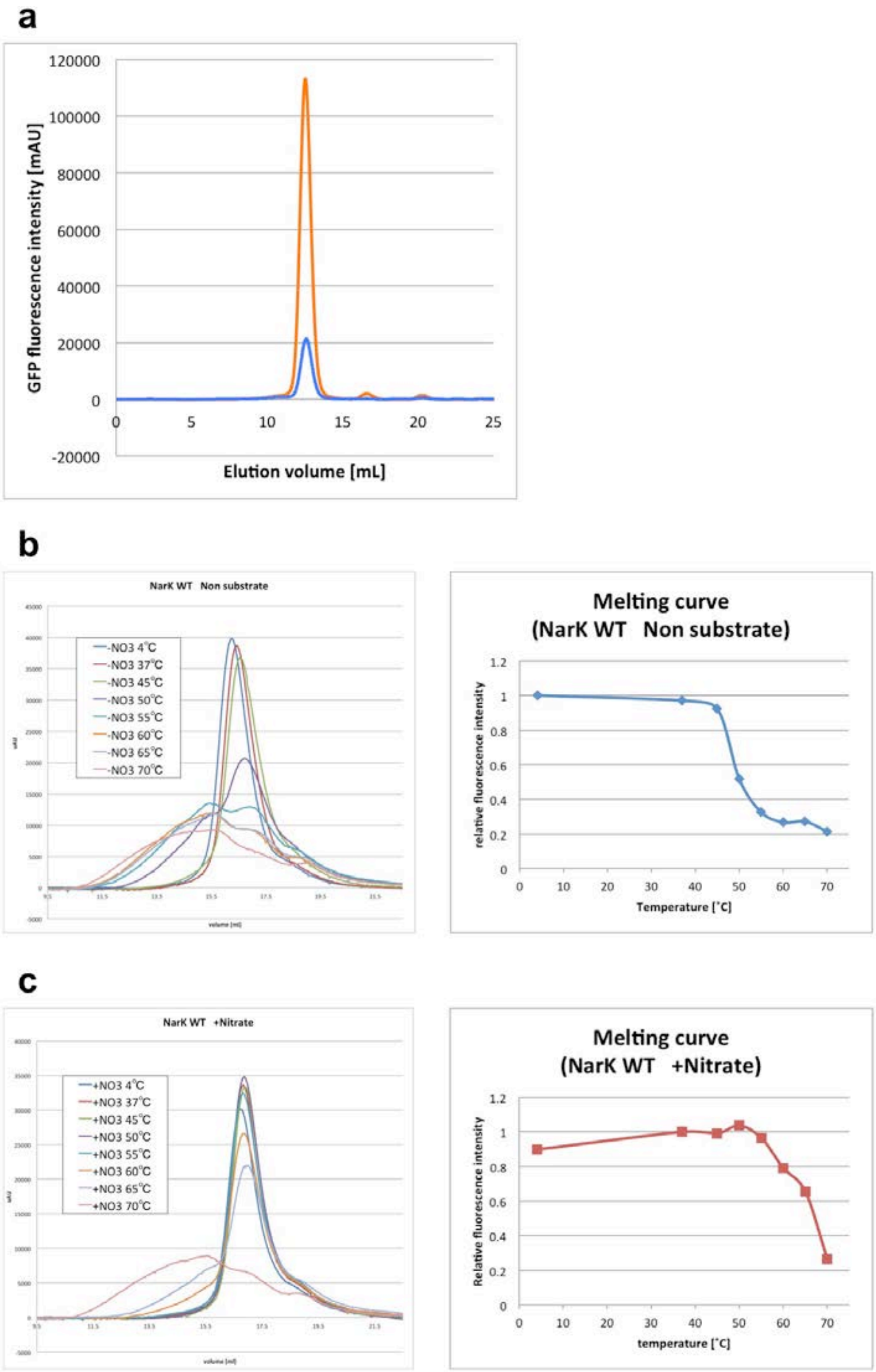


**Figure 2-3 FSEC-based expression screening and thermo-stability assay.**

(**a**) Chromatograms of full-length NarK (orange) and NarU (blue) in the C-terminus GFP-fusion construct (Superdex 200 10/300 GL). (**b**) and (**c**) FSEC-TS analysis of NarK WT in the presence or absence of nitrate. The estimated melting temperature (Tm) values were (b) ~50˚C without nitrate and (c) ~65˚C with nitrate.

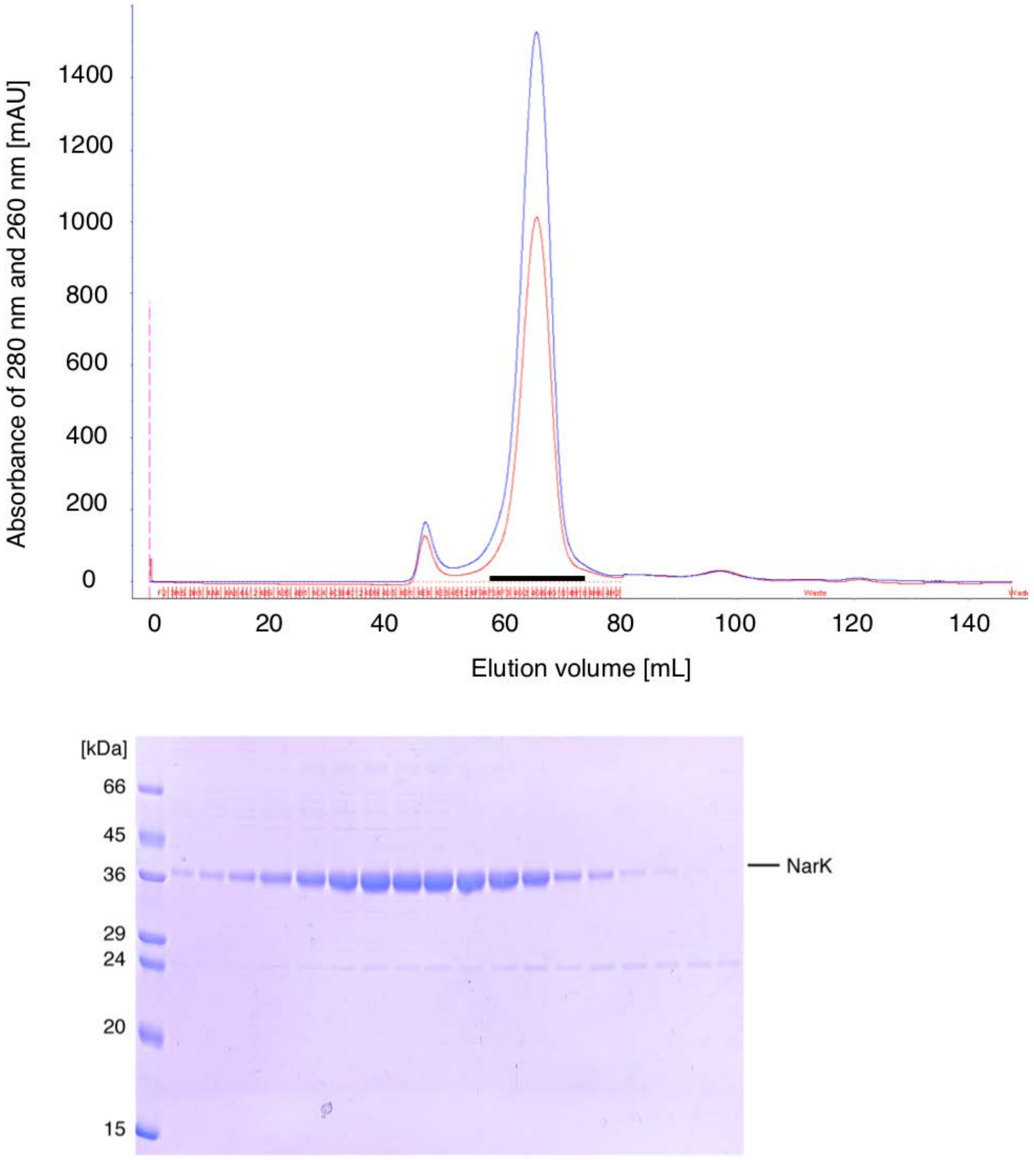
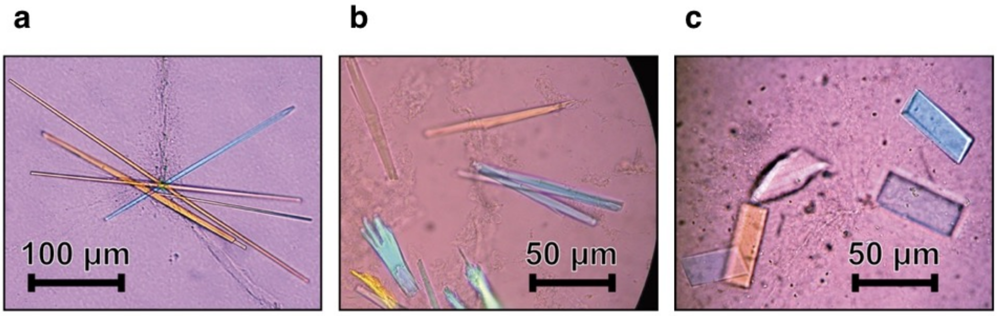


**Figure 2-4 Size-exclusion chromatograms and SDS-PAGE analysis of native NarK.**

Chromatograms of full-length NarK analyzed by HiLoad 16/600 Superdex 200 pg are shown in upper panel. The absorbance of 280 nm and 260 nm are drawn in blue and red lines, respectively. The fractions indicated in black solid line in the upper panel was analyzed by SDS-PAGE with coomassie brilliant blue staining shown in lower panel. The molecular-weight markers were loaded in the left lane and labeled in kDa.



31% PEG500DME
100 mM MES pH6.5
100 mM $NH_4NO_3$

36% PEG400
100 mM MES pH6.0
10 mM $Zn(OAc)_2$
3% D-sorbitol

27% PEG500DME
100 mM MES p5.8
100 mM $NH_4NO_3$

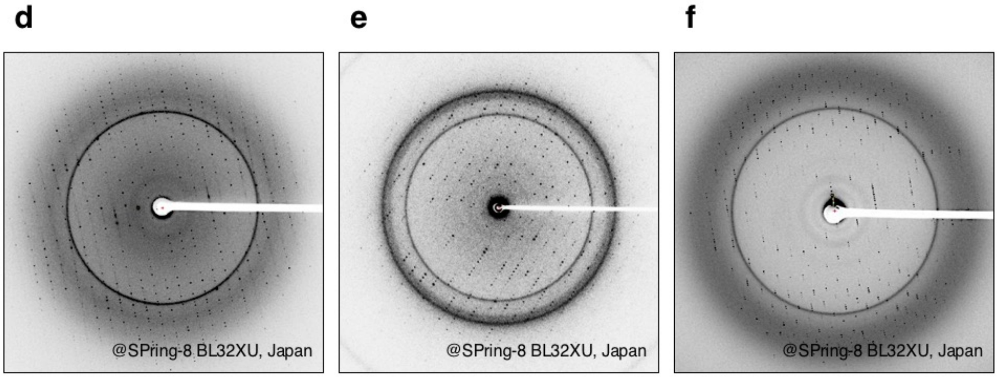
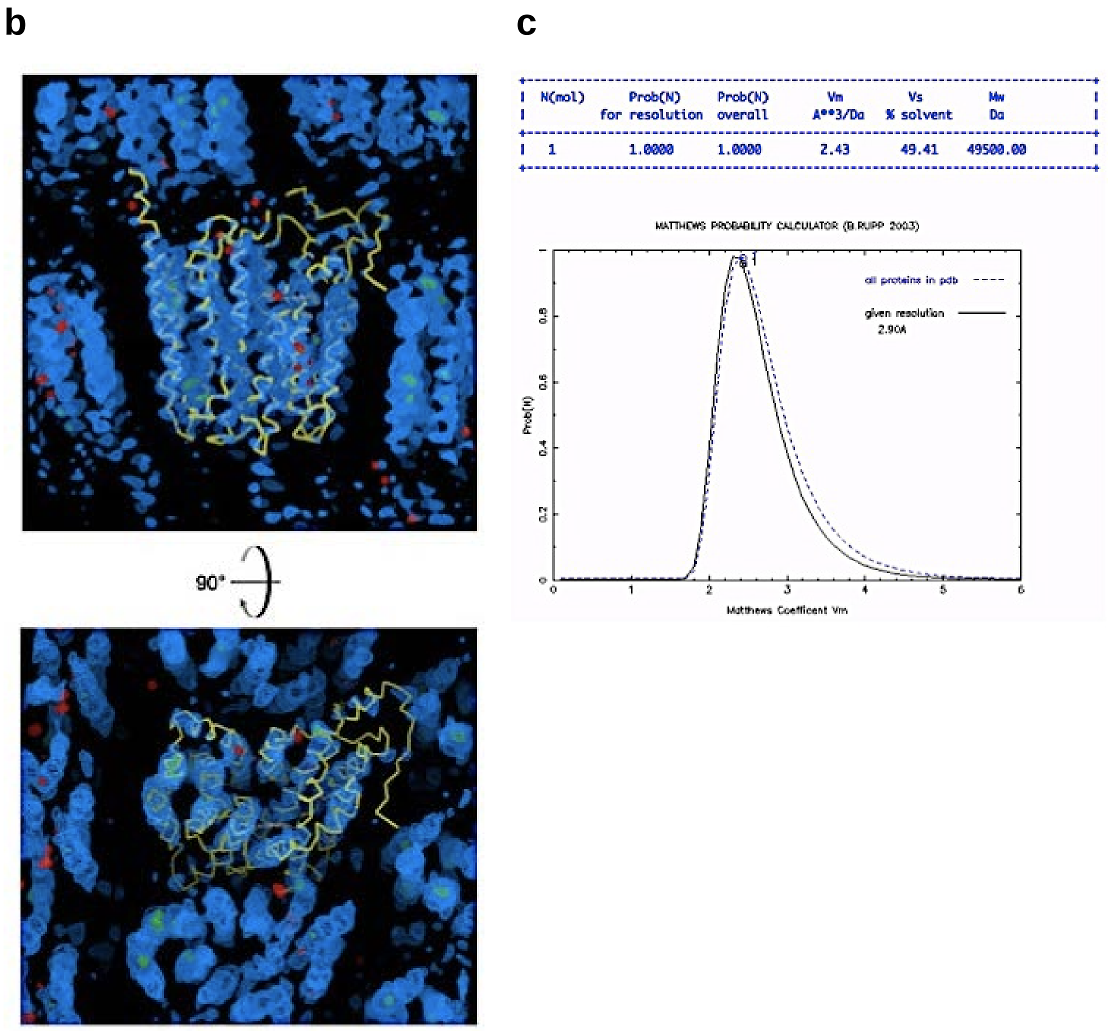


**Figure 2-5 Crystals and X-ray diffraction images of native NarK.**

(**a**), (**b**). and (**c**) Representative photographs of (a) Crystal I, (b) Crystal II, and (c) Crystal III. The crystallization conditions are also shown. (**c**), (**d**). and (**e**) X-ray diffraction images of (d) Crystals I, (e) Crystals II, and (f) Crystals III. All data were collected at SPring-8 BL32XU (Hyogo, Japan).

**a**

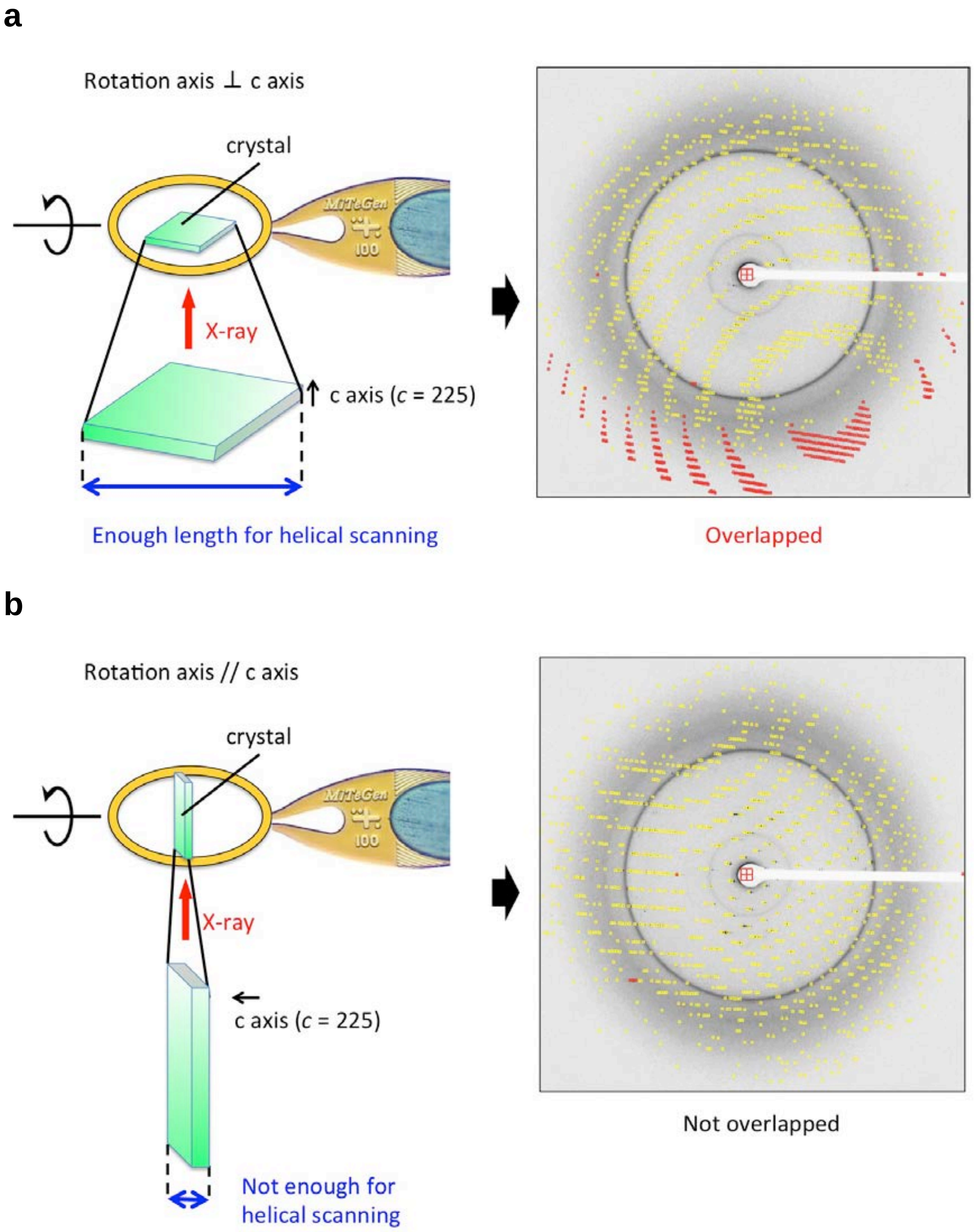
| | | Program | | | |
|---|---|---|---|---|---|
| | | ClustalO | LAIGN | ClustalW2 | FFAS |
| Protein name | PDB ID | Identity [%] | Identity [%] | Identity [%] | Identity [%] |
| LacY | 2CFG | 13.7 | 16.0 | 10.6 | 12.0 |
| GlpT | 1PW4 | 17.3 | 14.6 | 9.1 | 10.9 |
| EmrD | 2GFP | 13.4 | 18.0 | 8.0 | 9.6 |
| FucP | 3O7Q | 15.1 | 16.4 | 8.9 | 8.2 |
| PepTso | 2XUT | 13.8 | 17.0 | 8.9 | 10.5 |



```
+----------------------------------------------------------------------------+
I  N(mol)      Prob(N)       Prob(N)        Vm          Vs          Mw       I
I          for resolution  overall       A**3/Da   % solvent      Da        I
+----------------------------------------------------------------------------+
I   1          1.0000        1.0000        2.43        49.41    49500.00    I
+----------------------------------------------------------------------------+
```

**Figure 2-6 Molecular replacement.**

(**a**) Amino acids sequence identities with other MFS transporters calculated by Clustal Omega (ClustalO), LALIGN, ClustalW2, Fold & Function Assignment System (FFAS). The MFS transporters that have been reported before the molecular replacement trials are shown. (**b**) The 2*Fo*-*Fc* electron density maps as a result of the molecular replacement (contoured at 2.0 σ). Poly-alanine model of GlpT (PDB ID: 1PW4) was used as a search model. $R_{\mathrm{free}}/R_{\mathrm{work}}$ = 46.2/51.9. (**c**) Matthew's coefficients were predicted from the space group, cell dimensions and molecular weight of NarK, and are shown on the probability distribution of Matthew's coefficients from deposited to Protein Data Bank (PDB). The probabilities are maximum, assuming the presence of one molecule in the asymmetric unit of Crystal II.


a
Rotation axis ⊥ c axis
crystal
MiTeGen
100
X-ray
c axis (c = 225)
Enough length for helical scanning
Overlapped
b
Rotation axis // c axis
crystal
MiTeGen
100
X-ray
c axis (c = 225)
Not enough for helical scanning
Not overlapped

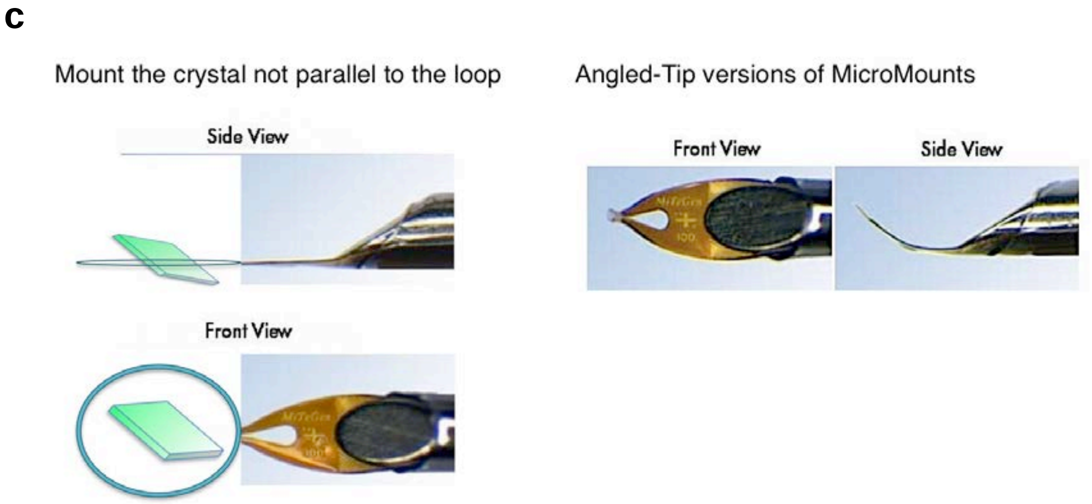


**Figure 2-7 Overlap problem.**

(**a**), (**b**) The schematic diagrams of the plate-shaped crystal on the loop and the representative resulted diffraction images. The diffraction spots of overlapped and not-overlapped are colored in red and yellow, respectively. When the rotation axis of the goniometer and the *c* axis of the crystal are nearly perpendicular, the “overlap” problem is likely to be occurred especially when the *c* axis of the crystal is relatively long. (**c**) The possible ways to avoid the overlap problem.

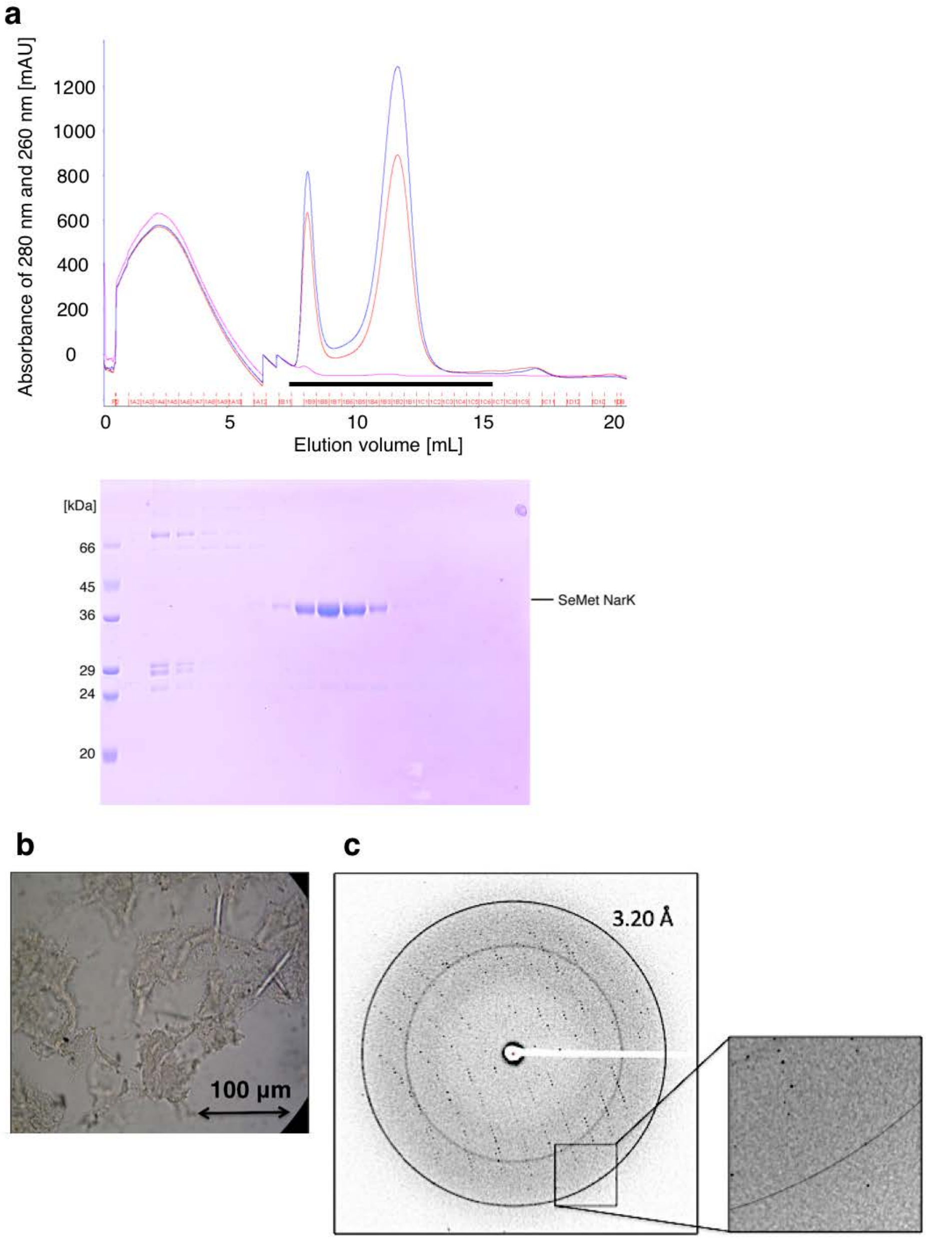
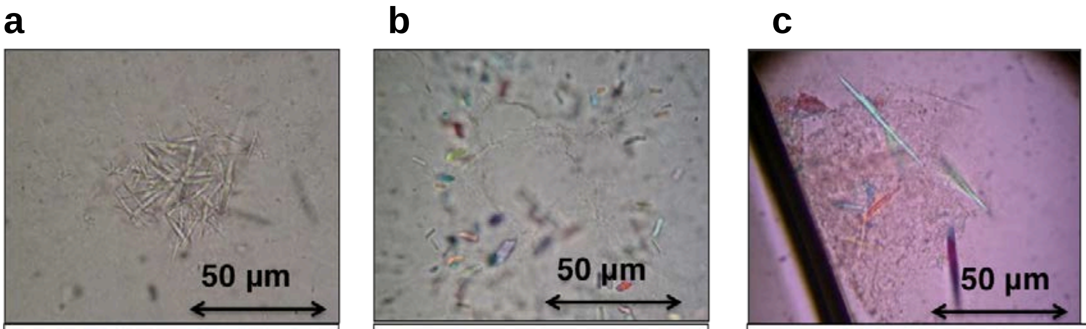


**Figure 2-8 Purification, crystallization, and X-ray diffraction images of SeMet NarK.**

(**a**) Chromatogram of full-length SeMet NarK analyzed by Superdex 200 10/300 GL column are shown in upper panel. The absorbance of 280 nm and 260 nm are drawn in blue and red lines, respectively. The fractions indicated in black solid line in the upper panel was analyzed by SDS-PAGE with coomassie brilliant blue staining shown in lower panel. The molecular-weight markers were loaded in the left lane and labeled in kDa. (**b**) Crystals of SeMet NarK. (**c**) X-ray diffraction image of SeMet NarK collected at SPring-8 BL32XU (Hyogo, Japan).



【Condition】
2 mM/1 mM MeHgCl
35% PEG 550MME
100 mM MES pH5.90
100 mM NaCl
100 mM Succinic acid

【Condition】
2 mM PheHg(OAc)
35% PEG 600
100 mM MES pH6.00
100 mM NaCl
100 mM $NaNO_3$
4% γ-butyrolactone

【Condition】
2 mMPCMB
35% PEG 500DME
100 mM HEPES-NaOH pH7.0
100 mM $Zn(OAc)_2$
100 mM $Mg(CHOO)_2$

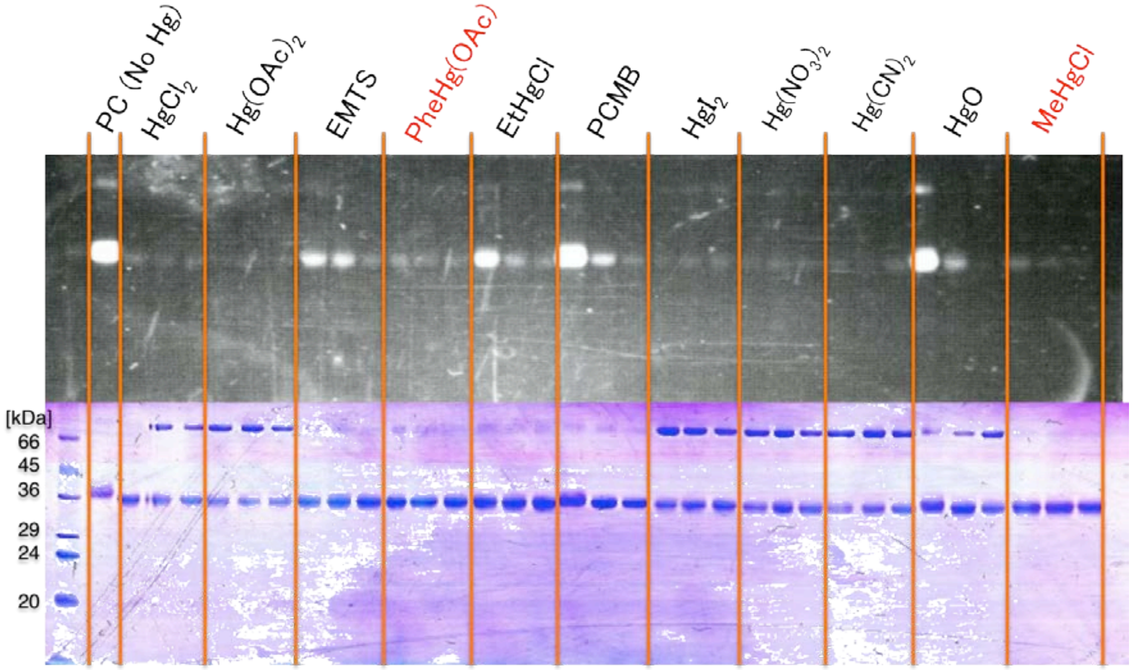
**d**

| Hg-derivative | Conc. | XAFS | Resolution | Data set |
|---|---|---|---|---|
| MeHgCl | 2 mM | ○ | ~ 3.8 Å | × |
| 〃 | 1 mM | - | ~ 20 Å | × |
| PheHg(OAc) | 2 mM | - | UN | - |
| PCMB | 2 mM | - | ~ 3.5 Å | × |

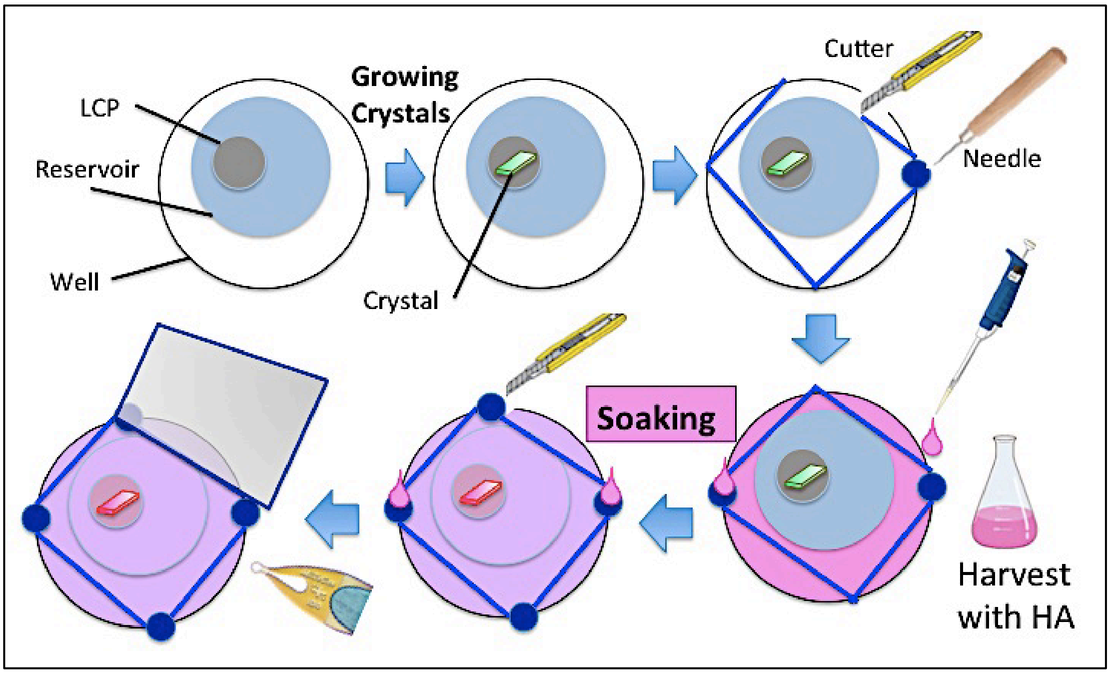
**Figure 2-9** **Co-crystallization with heavy atom reagents.**

(**a**), (**b**), and (**c**) Co-crystals with (a) MeHgCl, (b) PheHg(OAc), and (c) PCMB. (**d**) The result of x-ray diffraction experiments using the co-crystals with heavy atom-containing reagents shown in (a), (b), and (c).

**a**


PC (No Hg)
HgCl$_2$
Hg(OAc)$_2$
EMTS
PheHg(OAc)
EtHgCl
PCMB
HgI$_2$
Hg(NO$_3$)$_2$
Hg(CN)$_2$
HgO
MeHgCl
[kDa]
66
45
36
29
24
20

**b**


LCP
Reservoir
Well
Growing Crystals
Crystal
Cutter
Needle
Soaking
Harvest with HA

**c**

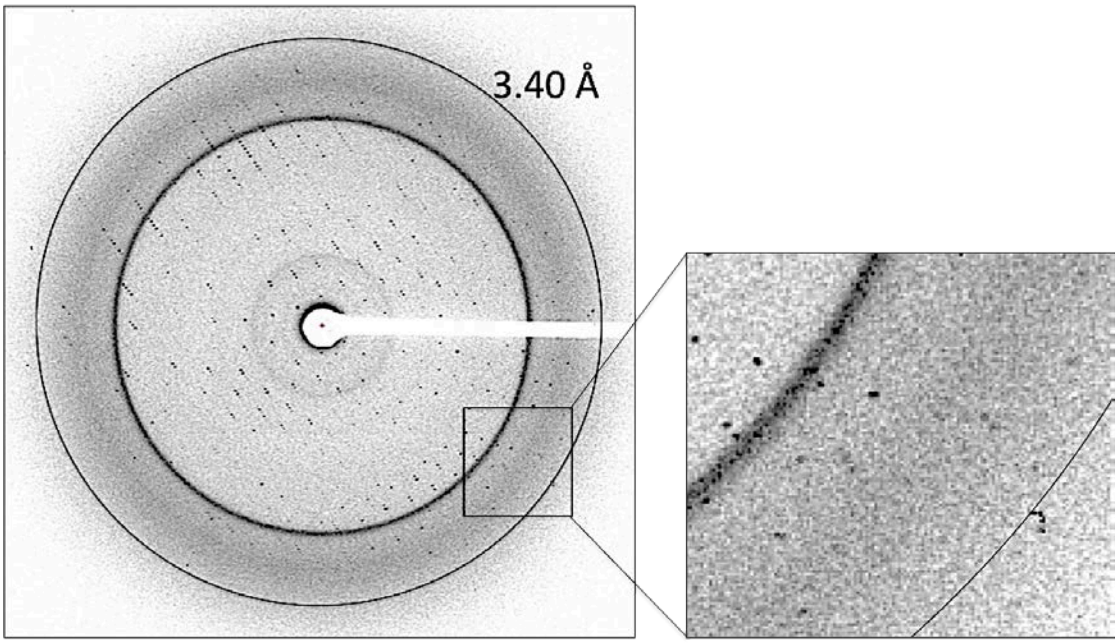
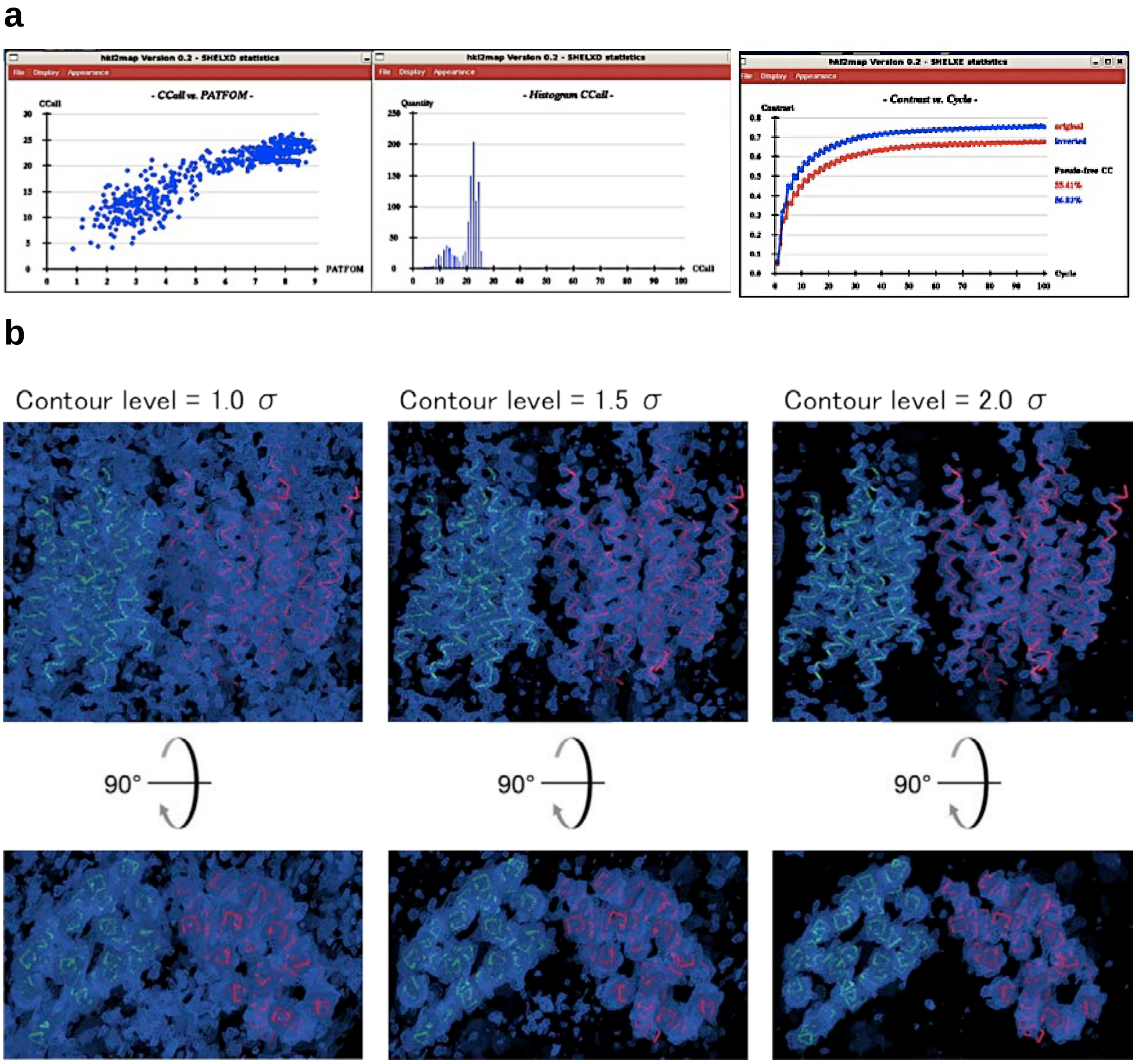


Figure 2-10 Soaking with heavy atom reagents.

(**a**) The Fluorescence Detection of Heavy Atom Labeling (FD-HAL) competition assay. The molar ratio of heavy atom reagents : purified NarK are 2:1, 4:1, and 8:1. (**b**) Schematic explanation for heavy atom soaking into LCP crystals. HA: Heavy atom. (**c**) X-ray diffraction image of MeHgCl-soaked NarK collected at SPring-8 BL32XU (Hyogo, Japan).



## Figure 2-11 Experimentally phasing by SIRAS method.

(**a**) Experimental phasing using the program SHELX C/D/E. The diffraction data sets of native and MeHgCl-soaked Crystal III were used for the single isomorphous replacement with anomalous scattering (SIRAS) method. (**b**) The initial electron density maps of NarK (Crystal III) after the phase calculation, heavy atom refinement, and density modification using the autoSHARP pipeline.

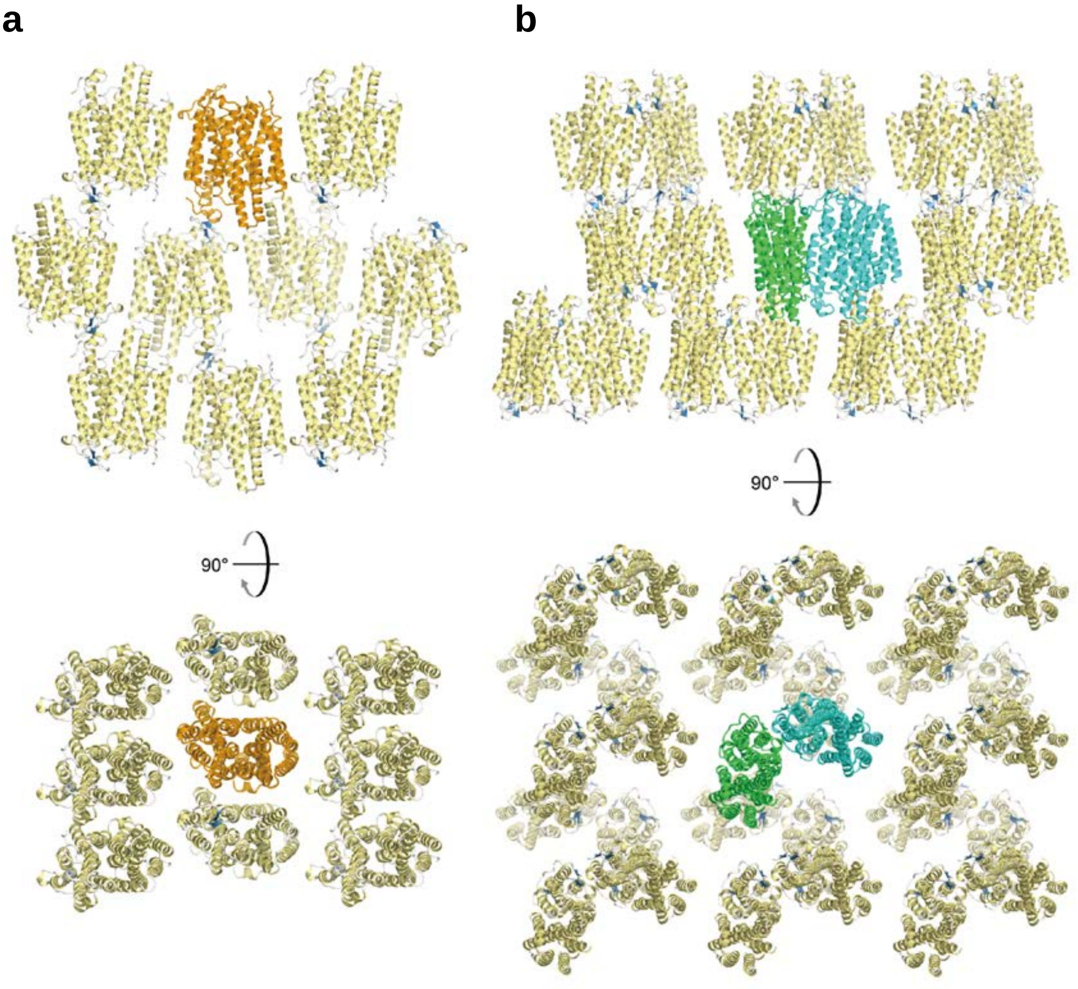
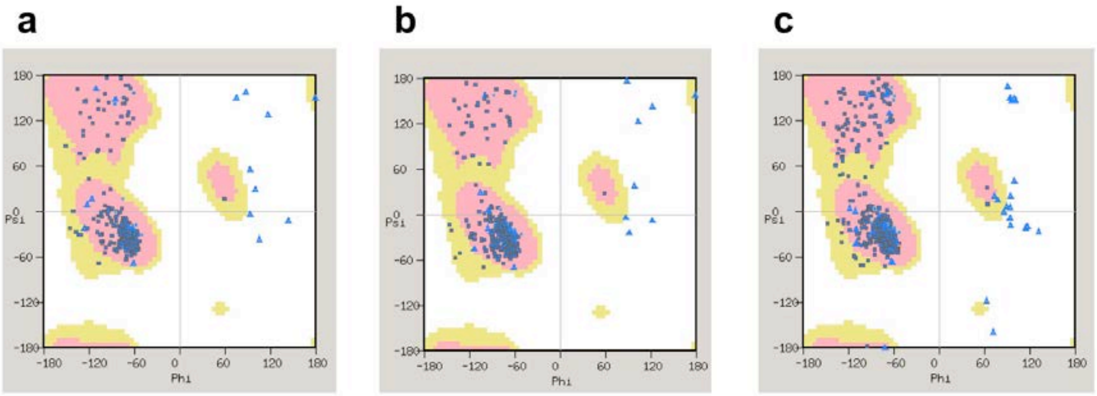


**Figure 2-12 Crystal packing of NarK in LCP crystals.**

(**a**) One molecule in the asymmetric unit of Crystals I is colored in orange. (**b**) Two molecules in the asymmetric unit of Crystals III are colored in light blue and green. The crystal packing, viewed from the membrane (upper panel) and the intracellular side (lower panel).



## Figure 2-13 Ramachandran plots.

(**a**) Ramachandran plots of Crystal I. In preferred regions: 421 amino acids (96.56%). In allowed regions: 15 amino acids (3.44%). Outliers: 0 amino acid (0.00%). (**b**) Ramachandran plots of Crystal II. In preferred regions: 425 amino acids (97.70%). In allowed regions: 10 amino acids (2.30%). Outliers: 0 amino acid (0.00%). (**c**) Ramachandran plots of Crystal III. In preferred regions: 836 amino acids (97.44%). In allowed regions: 22 amino acids (2.56%). Outliers: 0 amino acid (0.00%).

**Table 2-1 Macromolecule-production information.**

| | |
|---|---|
| Source organism | Escherichia coli |
| DNA source | UniProtKB: P10903 |
| Expression vector | Modified pET-28a |
| Expression host | *E. coli* strain C41(DE3) Δ*acrB* |
| Amino acid sequence of the construct | MASHSSAPERATGAVITDWRPEDPAFWQQRGQRIASRNLWISVPCLLLAFCVWMLFSAVAVNLPKVGFNFTTDQLFMLTALPSVSGALLRVPYSFMVPIFGGRRWTAFSTGILIIPCVWLGFAVQDTSTPYSVFIIISLLCGFAGANFASSMANISFFFPKQKQGGALGLNGGLGNMGVSVMQLVAPLVVSLSIFAVFGSQGVKQPDGTELYLANASWIWVPFLAIFTIAAWFGMNDLATSKASIKEQLPVLKRGHLWIMSLLYLATFGSFIGFSAGFAMLSKTQFPDVQILQYAFFGPFIGALARSAGGALSDRLGGTRVTLVNFILMAIFSGLLFLTLPTDGQGGSFMAFFAVFLALFLTAGLGSGSTFQMISVIFRKLTMDRVKAEGGSDERAMREAATDTAAALGFISAIGAIGGFFIPKAFGSSLALTGSPVGAMKVFLIFYIACVVITWAVYGRHSKKLESSGENLYFQ(GQFTSSVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKFICTTGKLPVPWPTLVTTFSYGVQCFSRYPDHMKRHDFFKSAMPEGYVQERTISFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYITADKQKNGIKANFKIRHNIEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSKLSKDPNEKRDHMVLLEFVTAAGITHGMDELYKHHHHHHHH) |

The residues cleaved by TEV protease are in parentheses.

The cloning artifacts are underlined.

Colorings of amino acid sequence of the construct are as follows; amino acid sequence of NarK from *Escherichia coli* in orange, TEV protease recognition sequences in pink, GFPuv A206K mutant in green, His$_8$-tag in blue, and other cloning artifacts in black.

## Table 2-2 Primers used in this study.

| Primer name | Fw/ Rv | Sequence |
|---|---|---|
| NarK (WT) *cloning* | Fw | 5'- GCGCCCATGGCCAGTCACTCATCCGCCCCCGAAAG -3' |
|  | Rv | 5'- GCGCCTCGAGTTTTTTAGAATGCCGACCATATAC -3' |
| NarU (WT) *cloning* | Fw | 5'- GCGCCCATGGCCCTGCAAAATGAGAAAAATAGTC -3' |
|  | Rv | 5'- GCGCCTCGAGTTTTTGGCTGAACTTCCGCCGAC -3' |
| NarK (G268A) | Fw | 5'- GTATCTGGCAACCTTCGCCTCCTTCATCGGCTTC -3' |
|  | Rv | 5'- GAAGCCGATGAAGGAGGCGAAGGTTGCCAGATAC -3' |
| NarK (ΔG3) | Fw | 5'- CTAGTGCTTCCACTTTCCAGATGATTTCCGTG -3' |
|  | Rv | 5'- CTTCCTGGCGCTGTTCCTGACCGCTGCGCTGG -3' |
| NarK (G408A) | Fw | 5'- GACACGGCGGCGGCGCTGCCTTTCATCTCTG -3' |
|  | Rv | 5'- CAGAGATGAAAGGCAGCGCCGCCGCCGTGTC -3' |
| NarK (G418A) | Fw | 5'- GATTGGCGCGATTGGTGCCTTCTTTATCCCGAAAG -3' |
|  | Rv | 5'- CTTTCGGGATAAAGAAGGCACCAATCGCGCCAATC -3' |
| NarK (R89K) | Fw | 5'- GGTTTCTGGCGCGTTATTAAAGGTTCCATACTCCTTTATGC -3' |
|  | Rv | 5'- GCATAAAGGAGTATGGAACCTTTAATAACGCGCCAGAAACC -3' |
| NarK (R305K) | Fw | 5'- GTTTATTGGTGCGCTGGCGGCTTCTGCAGGTGGTG -3' |
|  | Rv | 5'- CACCACCTGCAGAAGCCGCCAGCGCACCAATAAAC -3' |
| NarK (Y263F) | Fw | 5'- GTGGATTATGAGCCTGCTGTTTCTGGCAACCTTCGG -3' |
|  | Rv | 5'- CCGAAGGTTGCCAGAAACAGCAGGCTCATAATCCAC -3' |
| NarK (F147A) | Fw | 5'- CTTTGCTGGCGCGAACGCCGCATCCAGTATGG -3' |
|  | Rv | 5'- CCATACTGGATGCGGCGTTCGCGCCAGCAAAG -3' |
| NarK (F267A) | Fw | 5'- GCTGTATCTGGCAACCGCCGGCTCCTTCATCGG -3' |
|  | Rv | 5'- CCGATGAAGGAGCCGGCGGTTGCCAGATACAGC -3' |
| NarK (G85A) | Fw | 5'- CTGCCTTCGGTTTCTGCCGCGTTATTACGTG -3' |
|  | Rv | 5'- CACGTAATAACGCGGCAGAAACCGAAGGCAG -3' |
| NarK (G141A) | Fw | 5'- CATTTCTCTGCTGTGCGCCTTTGCTGGCGCGAAC -3' |
|  | Rv | 5'- GTTCGCGCCAGCAAAGGCGCACAGCAGAGAAATG -3' |
| NarK (G144A) | Fw | 5'- CTGTGCGGCTTTGCTGCCGCGAACTTCGCATCCAG -3' |
|  | Rv | 5'- CTGGATGCGAAGTTCGCGGCAGCAAAGCCGCACAG -3' |
| NarK (G164A) | Fw | 5'- CCGAAACAGAAGCAGGCTGGCGCGCTGGGTCTG -3' |
|  | Rv | 5'- CAGACCCAGCGCGCCAGCCTGCTTCTGTTTCGG -3' |

| | | |
|---|---|---|
| NarK (G165A) | Fw | 5'- GAAACAGAAGCAGGGTGCCGCGCTGGGTCTGAATG -3' |
| | Rv | 5'- CATTCAGACCCAGCGCGGCACCCTGCTTCTGTTTC -3' |
| NarK (G168A) | Fw | 5'- GCAGGGTGGCGCGCTGGCTCTGAATGGTGGTCTG -3' |
| | Rv | 5'- CAGACCACCATTCAGAGCCAGCGCGCCACCCTGC -3' |
| NarK (G171A) | Fw | 5'- CGCGCTGGGTCTGAATGCTGGTCTGGGCAACATGG -3' |
| | Rv | 5'- CCATGTTGCCCAGACCAGCATTCAGACCCAGCGCG -3' |
| NarK (G172A) | Fw | 5'- CGCTGGGTCTGAATGGTGCTCTGGGCAACATGGGC -3' |
| | Rv | 5'- GCCCATGTTGCCCAGAGCACCATTCAGACCCAGCG -3' |
| NarK (G174A) | Fw | 5'- CTGAATGGTGGTCTGGCCAACATGGGCGTCAGC -3' |
| | Rv | 5'- GCTGACGCCCATGTTGGCCAGACCACCATTCAG -3' |
| NarK (G177A) | Fw | 5'- GGTCTGGGCAACATGGCCGTCAGCGTCATGCAG -3' |
| | Rv | 5'- CTGCATGACGCTGACGGCCATGTTGCCCAGACC -3' |
| NarK (G308A) | Fw | 5'- CTGGCGCGTTCTGCAGCTGGTGCATTATCTGAC -3' |
| | Rv | 5'- GTCAGATAATGCACCAGCTGCAGAACGCGCCAG -3' |
| NarK (G309A) | Fw | 5'- GGCGCGTTCTGCAGGTGCTGCATTATCTGACCG -3' |
| | Rv | 5'- CGGTCAGATAATGCAGCACCTGCAGAACGCGCC -3' |

NcoI and XhoI restriction sites are colored blue and red, respectively.

**Table 2-3 Buffer composition.**

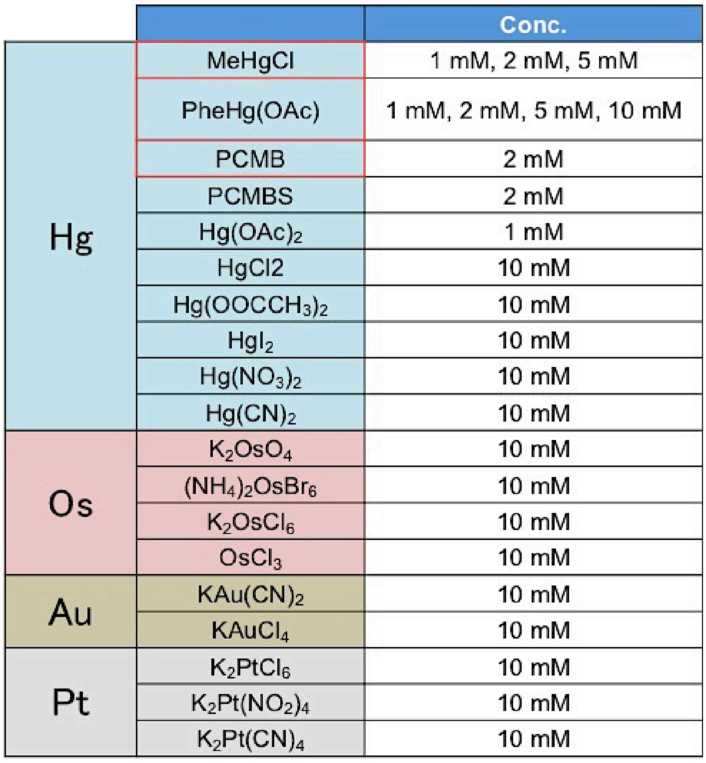
| Name | The purification step | Composition |
|---|---|---|
| Buffer A’ | Resuspension of cells | 50 mM Tris-HCl (pH 8.0), 300 mM NaCl,<br>0.1 mM phenylmethylsulfonyl fluoride (PMSF),<br>4 mM β-mercaptoethanol (β-ME) |
| Buffer A | Resuspension of cells | 50 mM Tris-HCl (pH 8.0), 300 mM NaCl, 20 mM $NaNO_3$,<br>0.1 mM PMSF, 4 mM β-ME |
| Buffer B | Solubilization | 50 mM Tris-HCl (pH 8.0), 300 mM NaCl, 25 mM $NaNO_3$,<br>0.1 mM PMSF, 4 mM β-ME, 2% n-dodecyl-β-D-maltoside (DDM), 0.4% cholesteryl hemisuccinate (CHS) |
| Buffer C | Equilibration of Ni-NTA column | 50 mM Tris-HCl (pH 8.0), 300 mM NaCl, 25 mM $NaNO_3$,<br>0.1 mM PMSF, 4 mM β-ME, 0.05% DDM, 0.01% CHS,<br>10 mM imidazole |
| Buffer D | Wash in Ni-NTA affinity chromatography | 50 mM Tris-HCl (pH 8.0), 300 mM NaCl, 25 mM $NaNO_3$,<br>4 mM β-ME, 0.05% DDM, 0.01% CHS,<br>20 mM imidazole |
| Buffer E | Elution in Ni-NTA affinity chromatography | 50 mM Tris-HCl (pH 8.0), 300 mM NaCl, 25 mM $NaNO_3$,<br>4 mM β-ME, 0.05% DDM, 0.01% CHS,<br>300 mM imidazole |
| Buffer F | Dialysis | 50 mM Tris-HCl (pH 8.0), 300 mM NaCl, 25 mM $NaNO_3$,<br>4 mM β-ME, 0.05% DDM, 0.01% CHS |
| Buffer G | SEC | 20 mM Tris-HCl (pH 8.0), 150 mM NaCl, 25 mM $NaNO_3$,<br>4 mM β-ME, 0.05% DDM, 0.01% CHS |

## **Table 2-4 Data collection and refinement statistics.**

| | Apo, inward-open | Nitrate-bound, inward-open | Nitrate-bound, occluded | CH$_3$HgCl derivative | SeMet derivative |
|---|---|---|---|---|---|
| Data collection | | | | | |
| X-ray source | SPring-8 BL32XU | | | | |
| Wavelength (Å) | 1.0000 | 1.0000 | 1.0000 | 1.0000 | 0.9794 |
| Space group | $P2_12_12_1$ | $P2_12_12_1$ | $P2_12_12_1$ | $P2_12_12_1$ | $P2_12_12_1$ |
| Cell dimensions | | | | | |
| $a$, $b$, $c$ (Å) | 35.5, 106.5, 124.6 | 35.4, 109.4, 126.0 | 68.8, 86.1, 221.6 | 68.8, 86.3, 222.4 | 35.4, 105.7, 125.4 |
| $\alpha, \beta, \gamma$ (°) | 90, 90, 90 | 90, 90, 90 | 90, 90, 90 | 90, 90, 90 | 90, 90, 90 |
| Resolution (Å) | 50.0–2.35 | 50.0–2.4 | 50.0–2.4 | 50.0–3.4 | 50.0–3.3 |
| | (2.49–2.35) | (2.54–2.4) | (2.54–2.4) | (3.52–3.4) | (3.39–3.33) |
| $R_{\mathrm{sym}}$ | 0.147 (1.391) | 0.243 (1.600) | 0.165 (0.1111) | 0.445 (1.516) | 0.258 (0.649) |
| $I/\sigma(I)$ | 7.33 (1.12) | 6.79 (1.08) | 7.00 (1.10) | 10.0 (2.3) | 12.7 (4.6) |
| Completeness (%) | 88.3 (87.7) | 99.8 (99.0) | 97.2 (98.6) | 95.7 (100.0) | 100.0 (100.0) |
| Redundancy | 5.2 (4.6) | 6.4 (6.3) | 4.2 (4.1) | 9.0 (9.3) | 14.2 (14.6) |
| CC$_{1/2}$ | 0.996 (0.598) | 0.994 (0.477) | 0.993 (0.601) | 0.989 (0.515) | - (-) |
| | | | | | |
| Refinement | | | | | |
| Resolution (Å) | 50.0–2.35 | 50.0–2.4 | 50.0–2.4 | | |
| No. reflections | 18,217 | 19,921 | 50,895 | | |
| $R_{\mathrm{work}}/R_{\mathrm{free}}$ | 0.2037/0.2318 | 0.2170/0.2475 | 0.2191/0.2368 | | |
| No. atoms | | | | | |
| Protein | 3,311 | 3292 | 6,486 | | |
| Ligand/ion/lipid | 40 | 77 | 224 | | |
| Water | 26 | 31 | 104 | | |
| *B*-factors | | | | | |
| Protein | 63.1 | 45.1 | 52.7 | | |
| Ligand/ion/lipid | 63.2 | 45.1 | 52.8 | | |
| Water | 49.1 | 37.7 | 47.1 | | |
| R.m.s. deviations | | | | | |
| Bond length (Å) | 0.003 | 0.009 | 0.002 | | |
| Bond angle (°) | 0.736 | 1.169 | 0.671 | | |
| Ramachandran | | | | | |
| Favored (%) | 98.17 | 97.95 | 97.9 | | |
| Allowed (%) | 1.83 | 2.05 | 2.1 | | |
| Outliers (%) | 0 | 0 | 0 | | |

Values in parentheses are for highest-resolution shell.

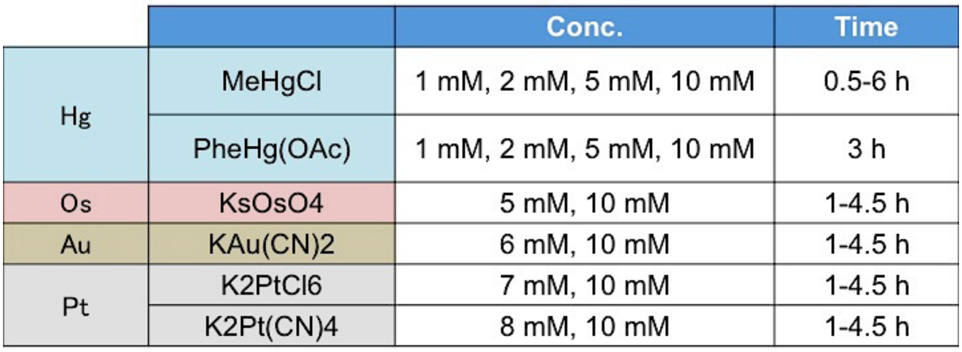
**Table 2-5 The heavy atom reagents used for co-crystallization.**



| | | Conc. |
|---|---|---|
| Hg | MeHgCl | 1 mM, 2 mM, 5 mM |
| | PheHg(OAc) | 1 mM, 2 mM, 5 mM, 10 mM |
| | PCMB | 2 mM |
| | PCMBS | 2 mM |
| | $Hg(OAc)_2$ | 1 mM |
| | HgCl2 | 10 mM |
| | $Hg(OOCCH_3)_2$ | 10 mM |
| | $HgI_2$ | 10 mM |
| | $Hg(NO_3)_2$ | 10 mM |
| | $Hg(CN)_2$ | 10 mM |
| Os | $K_2OsO_4$ | 10 mM |
| | $(NH_4)_2OsBr_6$ | 10 mM |
| | $K_2OsCl_6$ | 10 mM |
| | $OsCl_3$ | 10 mM |
| Au | $KAu(CN)_2$ | 10 mM |
| | $KAuCl_4$ | 10 mM |
| Pt | $K_2PtCl_6$ | 10 mM |
| | $K_2Pt(NO_2)_4$ | 10 mM |
| | $K_2Pt(CN)_4$ | 10 mM |

The conditions that crystals were obtained are in red frame.

**Table 2-6 The heavy atom reagents used for soaking.**



| | | Conc. | Time |
|---|---|---|---|
| Hg | MeHgCl | 1 mM, 2 mM, 5 mM, 10 mM | 0.5-6 h |
| | PheHg(OAc) | 1 mM, 2 mM, 5 mM, 10 mM | 3 h |
| Os | KsOsO4 | 5 mM, 10 mM | 1-4.5 h |
| Au | KAu(CN)2 | 6 mM, 10 mM | 1-4.5 h |
| Pt | K2PtCl6 | 7 mM, 10 mM | 1-4.5 h |
| | K2Pt(CN)4 | 8 mM, 10 mM | 1-4.5 h |

# Chapter 3 Crystal structures of NarK

## 3.1 Introduction

In **Chapter 2**, the author obtained three forms of the crystals (Crystal I, II, and III) from different crystallization conditions. Three different NarK structures (forms I, II, and III) were determined from Crystal I, II, and III, at the resolutions of 2.35, 2.40, and 2.40 Å, respectively. In this chapter, I describe three distinct crystal structures of NarK.

## 3.2 Materials and Methods

All molecular graphics were rendered using the program CueMol (http://www.cuemol.org/). The electrostatic potential of the molecular surface was calculated using the programs PDB2PQR[59] and APBS[60] within the program CueMol (http://www.cuemol.org/ja/). The multiple amino acid sequence alignment was generated using the programs CLUSTALW[61] and ESPript[62].

## 3.3 Results

### 3.3.1 Overall structure of NarK

To elucidate the structural basis for the nitrate/nitrite antiport mechanism of NarK, the author determined the crystal structures of NarK from *Escherichia coli*. Three distinct types of crystals (Crystal I, II, and III) were obtained, using the lipidic cubic phase (LCP) method[63]. This method provides a three-dimensional lipid-bilayer-like environment for membrane proteins and the structures of various membrane proteins has been determined by this powerful and efficient crystallization technique[10,64–73]. The phases were determined by the single isomorphous replacement with anomalous scattering (SIRAS) method, using the mercury-derivative crystal. The crystal structures of NarK were finally determined in three different states, the nitrate-bound occluded, nitrate-bound inward-open, and *apo* inward-open states, at resolutions of 2.4, 2.4, and 2.35 Å, respectively (**Table 2-4**). Each asymmetric unit of Crystals I and II contained one inward-open molecule, while in the asymmetric unit of Crystal III, there were two occluded molecules (Mol A and B). Since Mol A and B have almost the same structure (r.m.s.d. of 0.38 Å for all Cα atoms), Mol A is used in the following discussion.

The overall structures of NarK in both the occluded and inward-open states adopt a canonical MFS-fold with 12 TM helices, with TM1–TM6 and TM7–TM12 forming the N and C bundles, respectively (**Figure 3-1**). These N and C bundles are connected with a linker loop between TM6 and TM7, which is disordered in the crystal structures (**Figure 3-1**). These overall structural features are consistent with the previously reported structures of NarK[27] and NarU[26]. The substrate-binding site is located at the interface between the N and C bundles, in the middle of the TM segment (**Figure 3-1**). While the substrate-binding site is accessible from the cytoplasmic side in the inward-open structures, it is completely sequestered in the occluded state structure (**Figure 3-2**).

### 3.3.2 Structural change of the TM helices in the C bundle

While no structural difference between the occluded and inward-open structures was observed in the N bundle, large structural changes were observed in the TM helices in the C bundle. The TM10 and TM11 helices are straight in the occluded structure, whereas they are bent around the conserved Gly residues (G363, G365, G367, G408, G414, G417 and G418) in the inward-open structures (**Figure 3-3** and **Figure 3-4**). Among them, G408, G414, G417 and G418 are included in the conserved Nitrate Signature motif 2. The cytosolic halves of TM10 and TM11 are rotated by 31 and 26 degrees, respectively, away from the center of the transporter (**Figure 3-3**). Similar bending of TM10 and TM11 was also observed in the inward-facing NarK structures (PDB accession code: 4JR9)[27]; however, the TM10 and TM11 helices are further tilted toward the N bundle side in the present crystal structure (**Figure 3-5**). The present crystal structures further reveal the structural change in TM7: as compared to the occluded state structure, the cytoplasmic half of TM7 in the inward-open state structure is rotated by 18 degrees toward the periphery, at the conserved G268 residue in the inward-open structure (**Figure 3-3**). As TM7 is located adjacent to TM10 and TM11 and forms tight interactions (**Figure 3-1**), the structural change of TM7 may be coupled with those of TM10 and TM11. G268 is strictly conserved among the NNP family transporters, including NarU from *Escherichia coli* and NRT2.1 from *Arabidopsis thaliana* (**Figure 3-4**), suggesting that a similar bending of TM7 occurs in other NNP transporters.

### 3.3.3 Structural change in the transport path

In the crystal structure of the nitrate-bound occluded state, both the cytosolic and periplasmic sides are tightly closed by hydrophilic and hydrophobic interactions between the N and C bundles (**Figure 3-6** and **Figure 3-7**), and many of these interactions were not observed in the NarU structure previously reported as the occluded state (PDB accession code: 4IU8, Mol A)[26] (**Figure 3-5**). Furthermore, the present high-resolution crystal structures reveal the atomic details of these interactions, including water-mediated interactions (**Figure 3-5** and **Figure 3-7**). In the cytoplasmic pathway, F370 and L407 in the C bundle hydrophobically interact with F147, M151, and L167 in the N bundle, forming the interaction layer C1 (**Figure 3-6**). Furthermore, the main-chain carbonyl groups of A404 in TM11 and A148 in TM4 form a water-mediated hydrogen bond (**Figure 3-6**). These interactions in layer C1 occlude the pathway from the substrate-binding site toward the cytosolic side. In addition to these hydrophobic interactions, TM4 and TM5 in the N bundle form a number of hydrophobic and hydrophilic interactions with TM10 and TM11 in the C bundle, thereby closing the transport pathway (layers C2 and C3; **Figure 3-6**). In layer C2, the side chain of F156 in TM4 and the main chain of P159, K160, and Q161, in the loop between TM4 and TM5, form hydrophobic interactions with the side chains of M382 in TM10 and M396 and A400 in TM11 (**Figure 3-6**). The guanidinium group of R378 in TM10 hydrogen bonds with the main-chain carbonyl groups of F158 and K160, and the main-chain carbonyl group of A400 hydrogen bonds with the hydroxyl group of S155 in TM4 (**Figure 3-6** and **Figure 3-8**). In layer C3, part of the N-terminal loop consisting of P20 and E21 forms van der Waals contacts with the cytosolic terminus of TM11, including D392, E393, and M396 (**Figure 3-6**). These interactions between the N and C bundles tightly close the pathway to the cytoplasmic side, to prevent the intrusion of water or other solvent molecules into the substrate-binding site. In the structure of NarU[26], the cytoplasmic pathway is reportedly closed by a thin gate consisting of only two residues, F145 and F367 (corresponding to F147 and F370 in NarK, respectively), and the extensive interactions involving TM4 and TM5 in the N bundle and TM10 and TM11 in the C bundle are missing (**Figure 3-5**). However, given that several residues of NarK involved in the above interactions, including R378, are conserved in NarU (**Figure 3-4**), the cytoplasmic side of NarU would be further closed by the interactions involving TM4, TM5, TM10, and TM11 in the occluded state. Therefore, the previously reported occluded state structure of NarU[26] may actually represent the partially-occluded inward-facing state (**Figure 3-5**).

In contrast to their conformations in the present occluded structure of NarK, in the inward-open structure, TM7, TM10, and TM11 in the C bundle are bent at the conserved Gly residues (**Figure 3-3**). As a result, none of the interactions in layers C1, C2 and C3 is formed (**Figure 3-6**). The transport pathway is thereby formed between the N and C bundles, and the substrate-binding site is accessible from the cytosolic side (**Figure 3-6**). The previous structure of NarK bound to the antibody Fab fragment is reportedly in the inward-facing conformation (PDB accession code: 4JR9)[27]. However, the hydrophobic interactions between F147, M151 and L407 are still formed, and thus the substrate-binding site is not accessible in the previous structure (**Figure 3-5**). Therefore, the previously reported structure of NarK is still in the intermediate, partially-occluded inward-facing state. In contrast, TM10 and TM11 in the present inward-open structure are further rotated toward the periphery, thereby completely opening the cytoplasmic pathway (**Figure 3-6** and **Figure 3-5**).

### 3.3.4 Nitrate recognition in the substrate-binding site

In the nitrate-bound occluded structure, a strong electron density peak was observed in the substrate-binding site (**Figure 3-9**). This flat, triangular density peak was present in a location consistent with those of the bound nitrate and nitrite ions in the NarU[26] and NarK[27] structures, respectively. The crystallization conditions for the occluded-state crystals (Crystal III) contained ammonium nitrate. Therefore, the author assigned this peak as a nitrate ion. The present crystal structures have clarified the interactions between the substrate and the protein side chains (**Figure 3-9**), as compared to the previous crystal structures of NarU[26] and NarK[27]. The substrate-binding site is mainly formed by the side chains of F49, R89, F147, N175, Y263, F267, and R305 (**Figure 3-9** and **Figure 3-10**), which are highly conserved among the NNP family transporters[74,75] (**Figure 3-4**). The nitrate ion is sandwiched between F147 and F267, and forms a π-π stacking interaction with the π electrons of their phenyl groups (**Figure 3-9** and **Figure 3-10**). The oxygen atoms of the nitrate ion are further recognized by three sites: A, B, and C. In Site A, two oxygen atoms of the nitrate ion form a bidentate salt bridge with the NH1 and NH2 atoms of the R89 guanidinium group. This guanidinium group of R89 is anchored by the main-chain carbonyl groups of G144 in TM4 and S411 in TM11, and together they form a rigid substrate-binding site (**Figure 3-9** and **Figure 3-10**). In Site B, two oxygen atoms of the nitrate ion form a bifurcated hydrogen bond with the phenol hydroxyl group of Y263 in TM7. This hydroxyl group hydrogen bonds with the R305 guanidinium group, which in turn hydrogen bonds with the S366 hydroxyl group (**Figure 3-9** and **Figure 3-10**). As a result, the guanidinium group of R305 is oriented perpendicularly to the substrate nitrate ion. In Site C, the side-chain amide group of N175 and the phenyl group of F49 form an interaction site for the nitrate ion (**Figure 3-9** and **Figure 3-10**). Overall, these conserved amino acid residues create a suitable substrate-binding site for a flat anion with delocalized π electrons.

Similar nitrate/nitrite recognition by Arg side chain and/or aromatic groups is also observed in several proteins. For example, in molybdenum-containing NAD(P)H:nitrate reductase, the nitrate ion is reportedly recognized by two Arg residues[76] (**Figure 3-11**). In contrast, the manner of nitrate recognition by the other nitrate transporter, NRT1.1, is quite different from that of NarK, as NRT1.1 recognizes nitrate by Thr and His residues[12,13] (**Figure 3-12**).

In the nitrate-bound inward-open structure, a triangular electron density peak of a nitrate ion was also observed in the substrate-binding site (**Figure 3-7**). However, several differences were observed in the recognition manner of the substrate nitrate, as compared with that in the occluded state. In the inward-open structure, the bending of TM7, TM10, and TM11 leads to a slight movement of the side chains of Y263 in TM7 and R305 in TM8, away from the center of the transporter (**Figure 3-9**). As a result, the following changes were observed in the interactions in Sites A and B. In Site A, the distance between one oxygen atom of the nitrate ion and the R89 guanidinium group is increased from 2.7 Å to 3.1 Å (**Figure 3-9** and **Figure 3-10**). In Site B, the distance between the nitrate ion and the Y263 phenol group is increased from 2.4 Å to 3.9 Å, and only one oxygen atom is within hydrogen bonding distance with the Y263 phenol group (**Figure 3-9** and **Figure 3-10**). These changes in Sites A and B increase the volume of the substrate-binding site in the inward-open state structure (37.6 and 50.5 Å$^3$ for the occluded and inward-open structures, respectively; **Figure 3-13**). Taken together, the overall structural change in the transition from the nitrate-bound occluded to the inward-open state, including the bending of TM7, TM10, and TM11, induces the movement of Y263 and R305 in Site B, which enlarges the volume of the substrate-binding site. This change in the substrate-binding site may result in decreased substrate affinity in the inward-open state, thereby facilitating substrate release toward the cytoplasmic side.

In contrast, the *apo* inward-open structure was similar to the nitrate-bound inward-open structure (r.m.s.d. of 0.51 Å for 435 Cα atoms; **Figure 3-10** and **Figure 3-14**). While no electron density peak is observed in the substrate-binding site in the *apo* structure of the inward-open state (**Figure 3-7**), the positions of the side chain atoms in the substrate-binding site are superimposable on those of the nitrate-bound structure (**Figure 3-14**). This suggested that the inward-open state is also stable without the substrate, and that the binding and release of the substrate can occur without any large structural changes.

## 3.4 Discussion and Perspectives

The author determined the crystal structures of NarK from *Escherichia coli* in three different states in the transport cycle (nitrate-bound occluded, nitrate-bound inward-open and apo inward-open states) at 2.35–2.4 Å. Although the crystal structures of NarU and NarK were reported from the separate groups just two months after the author have determined the structure of NarK, the transport mechanism including the structural changes in the transport cycle and the driving force for substrate transport remained elusive because of the medium resolutions and the partially disordered structures. The higher resolution structures determined in this study represent the completely occluded and inward-open conformations, in contrast to the previously reported partially occluded and partially inward-open conformations. Structural comparison between inward-open and occluded structures suggested that the conserved glycine residues in TM10, TM11, and TM7 in C-bundles are important for conformational change and substrate transport.

The two nitrate-bound structures clarified the nitrate recognition mechanism of NarK in atomic resolution, and further revealed that the local structural change of Tyr263 in TM7 and Arg305 in TM8 in the substrate binding site occurred in the conformational change between inward-open and occluded states. These structural findings suggested that the coupling between the substrate recognition and the opening/closing of the transport pathway is accomplished by the orchestrated movement of the TM helices in the C-bundle and the key residues Tyr263 and Arg305 in the substrate-binding site. Moreover, considering that nitrate and nitrite can be bound to the same biding site competitively, and that the putative substrate-translocation pathway is positively charged and lacking protonatable residues, it seems more likely that NarK functions as a nitrate/nitrite antiporter, rather than co-transporter using proton motive force. However, these hypotheses are only based on the structural information and therefore are too speculative without any evidence from the functional analyses. Therefore, in the next chapter, the author will discuss the molecular mechanism of NarK in combination of structural information and the results of various functional analyses.

## Figures and Tables of Chapter 3

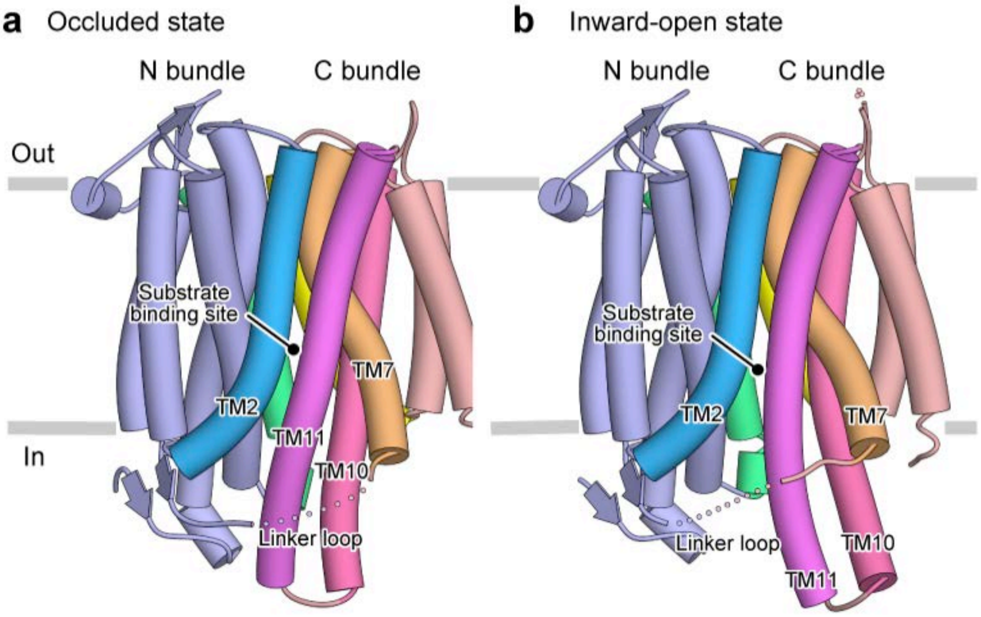


**Figure 3-1 Overall structure of NarK.**

(**a**) and (**b**) Overall structures of NarK in the occluded (a) and inward-open (b) states. Each molecule contains 12 TM helices, forming the N bundle (TM1–6) and the C bundle (TM7–12). TM2, TM5, TM7, TM8, TM10 and TM11 are colored light blue, green, amber, yellow, pink and purple, respectively. Other TM helices in the N and C bundles are colored pale blue and light pink, respectively. Gray bars indicate the approximate location of the lipid bilayer. The figures depicting the molecular structures were prepared using CueMol (http://www.cuemol.org/).

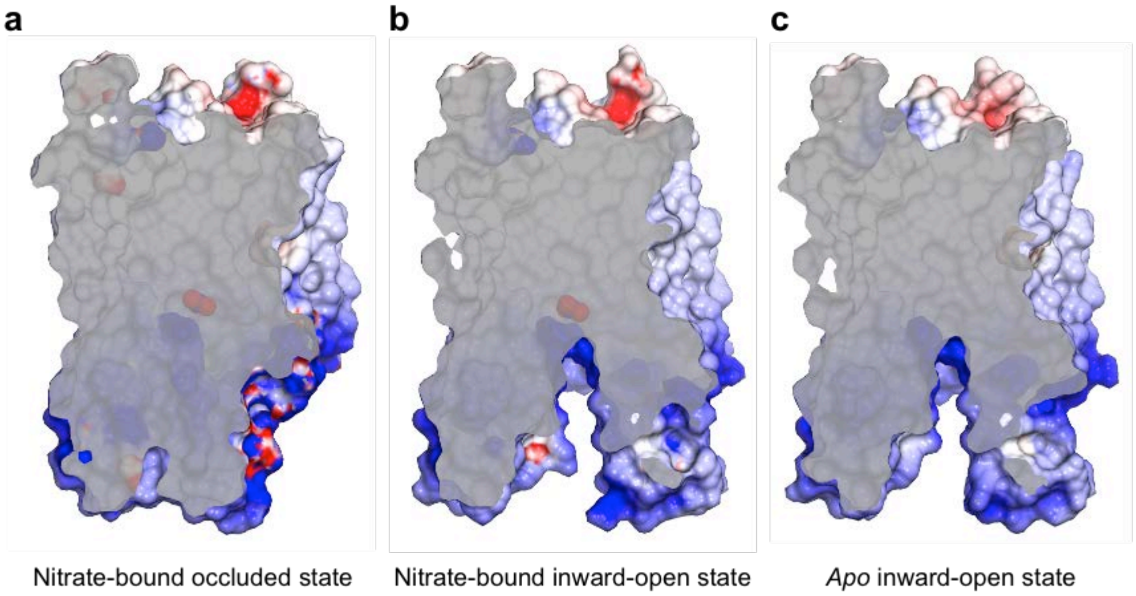


**Figure 3-2 Cross-sections of the crystal structures of NarK.**

(**a**), (**b**), and (**c**) Cross-sections of the crystal structures of NarK, colored semi-transparent gray, showing the solvent accessible surface calculated with a 1.4 Å radius sphere (assuming a water molecule). The molecular surfaces are colored according to the electrostatic potential, ranging from blue (positive) to red (negative). Bound nitrate molecules are shown as CPK models. (a) Nitrate-bound occluded state (Mol A). (b) Nitrate-bound inward-open state. (c) *Apo* inward-open state.

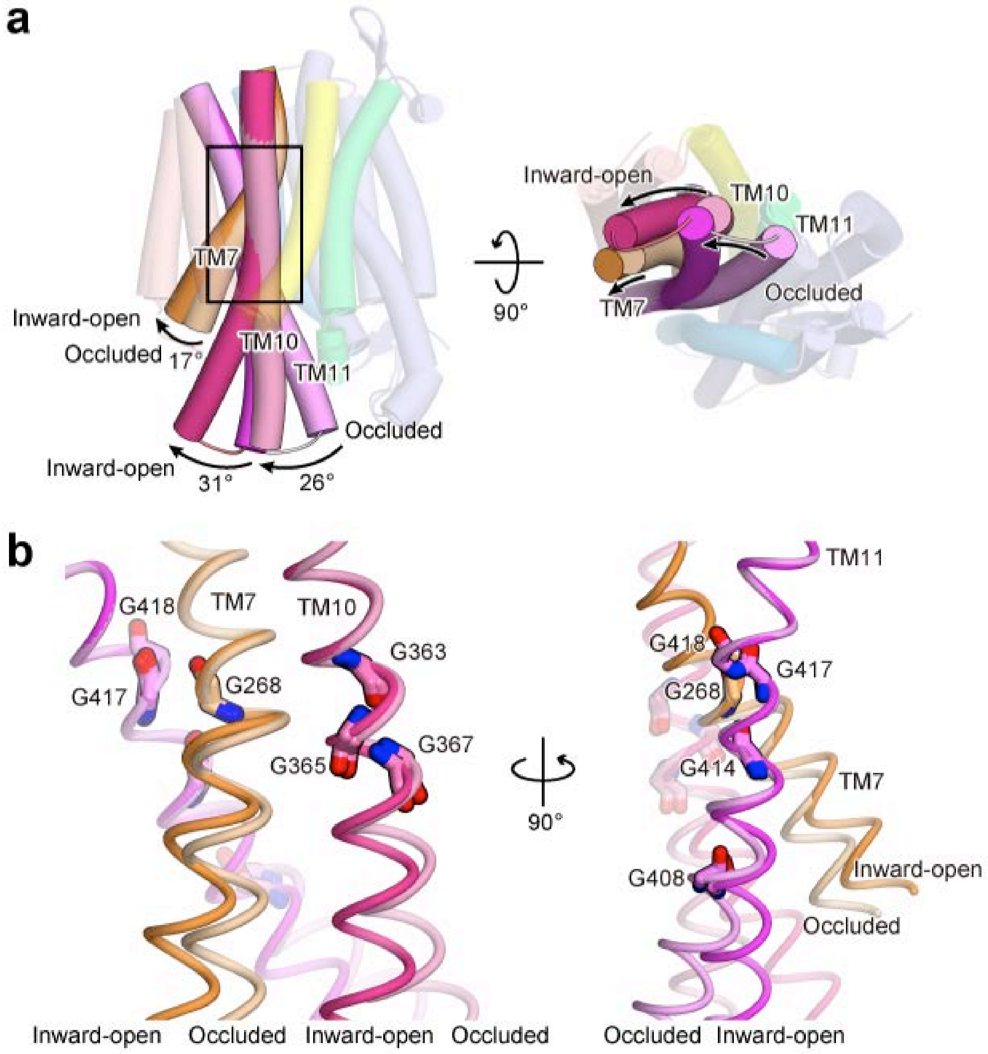
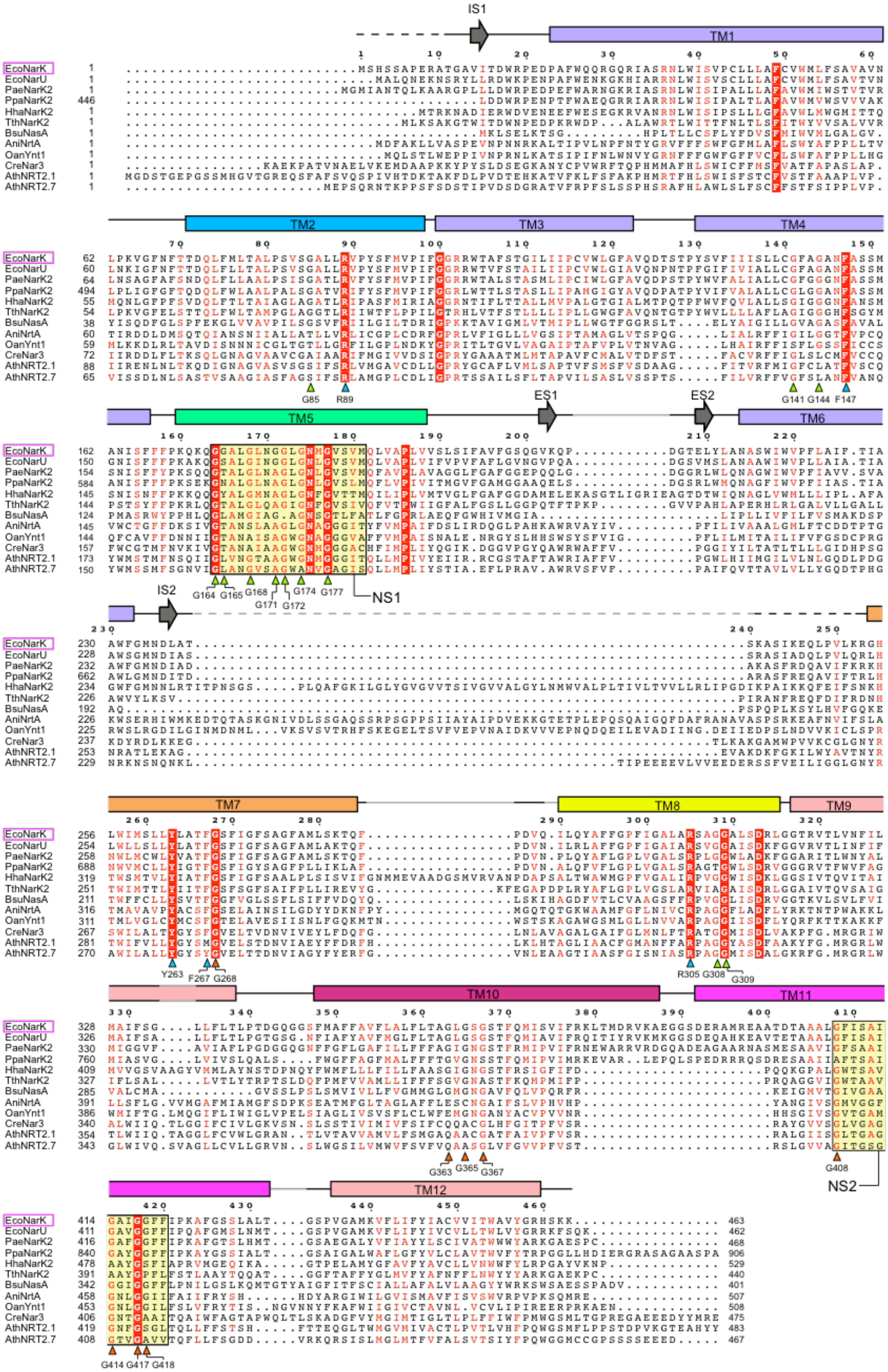


**Figure 3-3 Structural comparison of the occluded and inward-open states.**

(**a**) Overall structures of NarK in the occluded and inward-open states, viewed from the plane of the membrane (left panel) and the cytoplasmic side (right panel). The two structures were superimposed based on the Cα atoms of the N bundle residues (1–233). The colors of the helices in the inward-open state correspond to those in Figure 3-1 (a) and (b). TM7, TM10 and TM11 in the occluded state are depicted with pale colors. The rectangle indicates the region highlighted in (b). (**b**) Close-up views around the conserved Gly residues in TM7, TM10 and TM11. The occluded and inward-open structures were superimposed as in (a). The Gly residues are shown in stick models.


IS1
TM1
TM2
TM3
TM4
ES1
ES2
TM5
TM6
IS2
NS1
TM7
TM8
TM9
TM10
TM11
TM12
NS2
EcoNarK
EcoNarU
PaeNarK2
PpaNarK2
HhaNarK2
TthNarK2
BsuNasA
AniNrtA
OanYnt1
CreNar3
AthNRT2.1
AthNRT2.7
G85
R89
G141
G144
F147
G164
G165
G168
G171
G172
G174
G177
Y263
F267
G268
R305
G308
G309
G363
G365
G367
G408
G414
G417
G418

## Figure 3-4 Multiple amino acid sequence alignment of NNP family proteins.

Sequence alignment of *Escherichia coli* NarK (EcoNarK, GI:16129186), *Escherichia coli* NarU (EcoNarU, GI:16129428), *Pseudomonas aeruginosa* PAO1 NarK2 (PaeNarK2, GI:15599071), *Paracoccus pantotrophus* NarK2 (PpaNarK2, GI:14348603), *Halomonas halodenitrificans* NarK2 (HhaNarK2, GI:18413617), *Thermus thermophilus* HB8 NarK2 (TthNarK2, GI:6689413), *Bacillus subtilis* subsp. subtilis str. 168 NasA (BsuNasA, GI:255767099), *Aspergillus nidulans* FGSC A4 NrtA (AniNrtA, GI:67517614), *Ogataea angusta* Ynt1 (OanYnt1, GI:3849986), *Chlamydomonas reinhardtii* Nar3 (CreNar3, GI:437765), *Arabidopsis thaliana* NRT2.1 (AthNRT2.1, GI:15223123) and *Arabidopsis thaliana* NRT2.7 (AthNRT2.7, GI: 15241440). The secondary structure of EcoNarK, determined by the present crystal structures, is indicated above the sequences. The colors of the TM helices correspond to those in Figure 3-1 (a) and (b). The α-helices (TM1-12, as described in the main text) and β-strands (IS1-2 and ES1-2, representing the intracellular and extracellular strands, respectively) are indicated by cylinders and arrows, respectively. Strictly conserved residues among the twelve molecules are highlighted in red boxes, and highly conserved residues are indicated by red letters. The conserved Gly residues in TM7, 10 and 11 are indicated by orange triangles. The conserved Gly residues in TM2, 4, 5 and 8 that were mutated for the *in vivo* nitrate uptake assay are indicated by light green triangles. The residues involved in substrate recognition that were mutated for the *in vivo* nitrate uptake assay are indicated by light blue triangles. NS1 and NS2 are representing the Nitrate Signature motif 1 and 2, respectively.

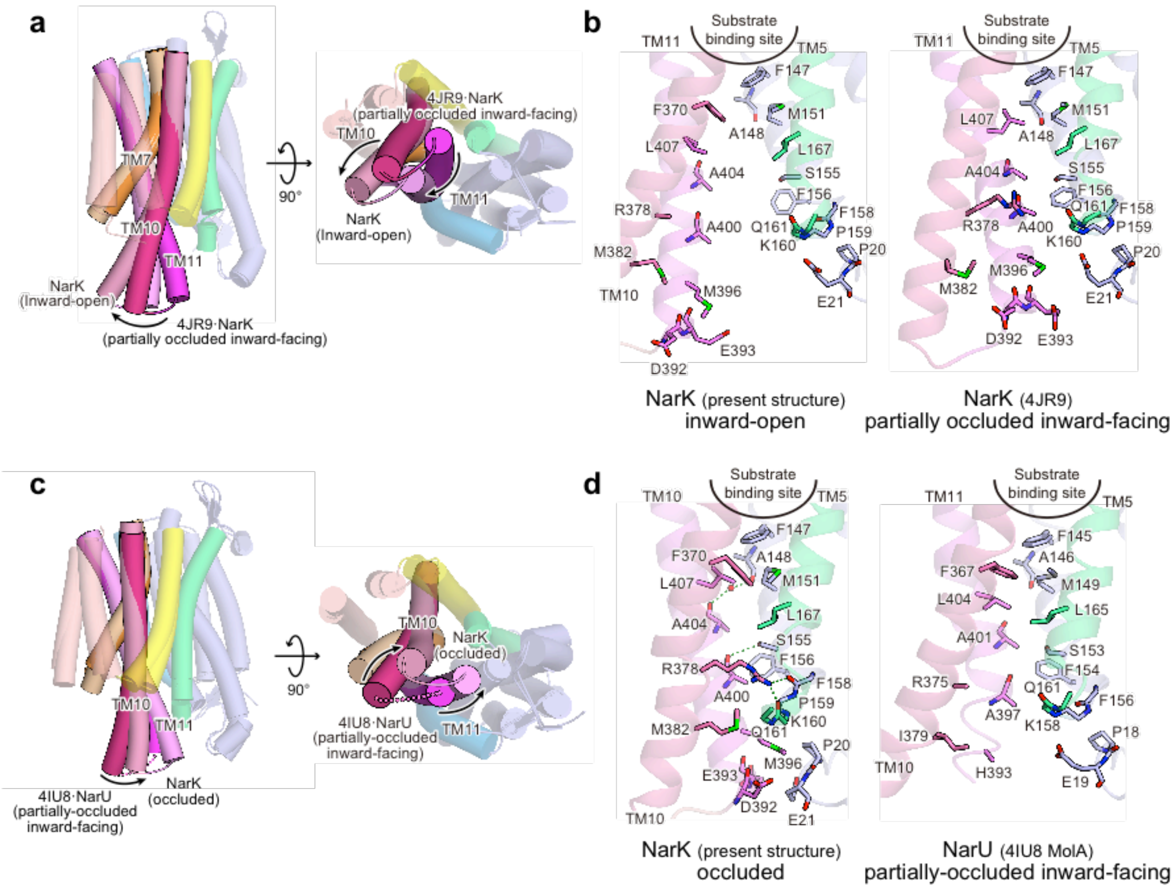


**Figure 3-5 Comparison of the present and previously reported structures of nitrate transporters.**

(**a**) and (**b**) Structural comparison of the inward-open states of NarK in the present crystal structure and the previously reported structure (PDB accession code: 4JR9). (a) Superimposed structures, viewed from the plane of the membrane (left panel) and the cytoplasmic side (right panel). (b) The cytosolic transport pathway of NarK in the inward-open states in the present crystal structure, left panel, and the previously reported structure (4JR9, right panel). (**c**) and (**d**) Structural comparison between NarK in the nitrate-bound occluded state (Mol A, present crystal structure) and NarU in the nitrate-bound partially-occluded inward-facing state (PDB accession code: 4IU8, Mol A). (c) Superimposed structures, viewed from the plane of the membrane (left panel) and the cytoplasmic side (right panel). (d) The cytosolic transport pathway of NarK in the nitrate-bound occluded state (present crystal structure, left panel) and NarU in the nitrate-bound partially-occluded inward-facing state (4IU8, right panel)

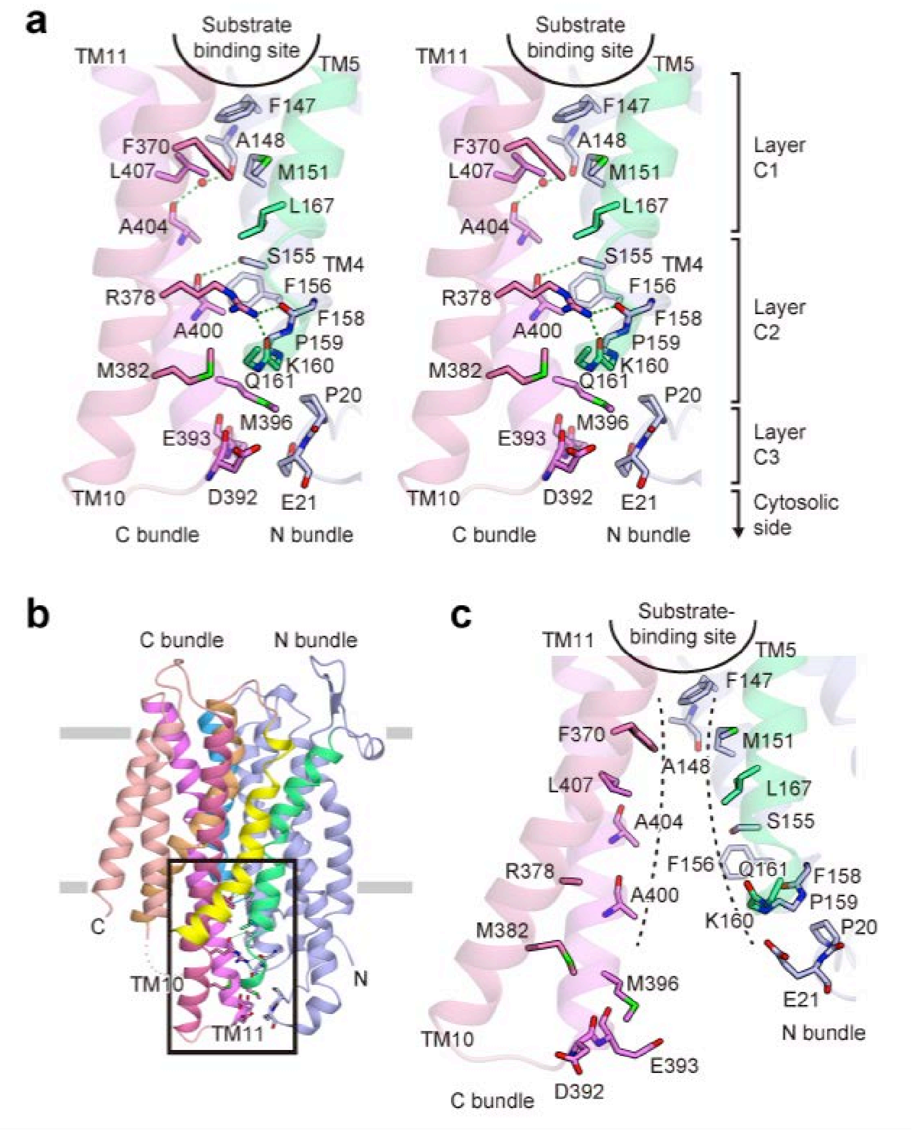


**Figure 3-6 The structural change in the cytosolic transport pathway of NarK.**

(**a**) The cytoplasmic interactions between the N and C bundles of the occluded state (stereo view). Residues related to the interactions are depicted with stick models. The side chains from F158 to Q161 are omitted for clarification. Water molecules are represented by red spheres. The green dashed lines indicate hydrogen bonds. The area highlighted in this panel is indicated by the rectangle in the overall structure shown in (b). (**b**) The overall structure of NarK in the nitrate-bound occluded state. The side chains involved in the cytoplasmic interactions are depicted by stick models. (**c**) The cytoplasmic area between the N and C bundles of the inward-open state. The black dashed lines indicate the transport pathway observed in the inward-open state, which is closed in the occluded state.

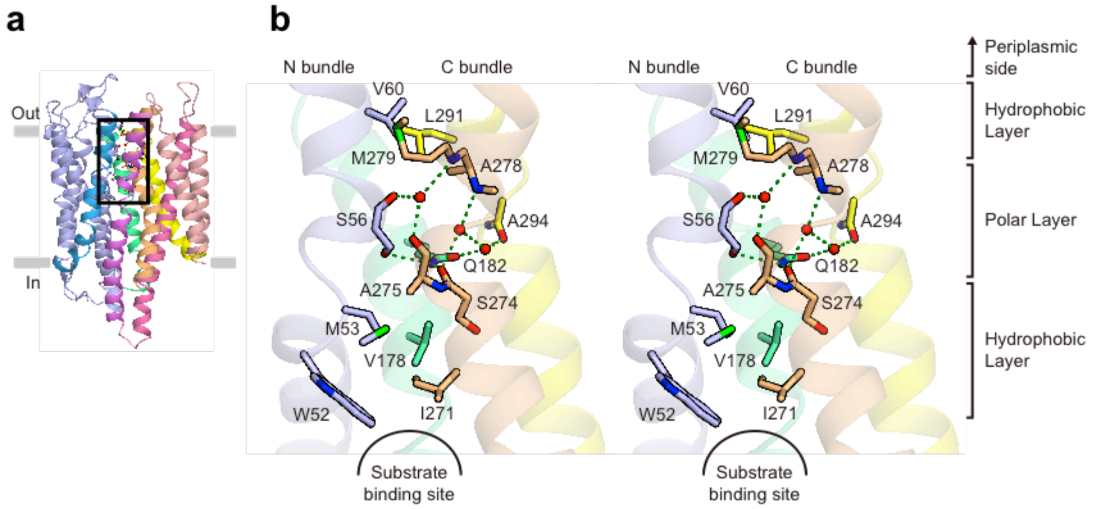


**Figure 3-7** **Periplasmic interactions of NarK.**

(**a**) Overview of the structure of NarK in the nitrate-bound occluded state (Mol A). The colors of the helices in the inward-open state correspond to those in Figure 3-1 (a) and (b). A close-up view of the area indicated by the rectangle is provided in panel (b). (**b**) Stereo view of the interactions between the N and C bundles that occlude the substrate-binding site from the periplasmic side. TM1, TM5, TM7, and TM8 are shown by ribbon models in pale colors. The residues composing the two hydrophobic layers and the polar layers are shown in stick models. Water molecules are depicted by red spheres. Hydrogen bonds are represented by green dashed lines.

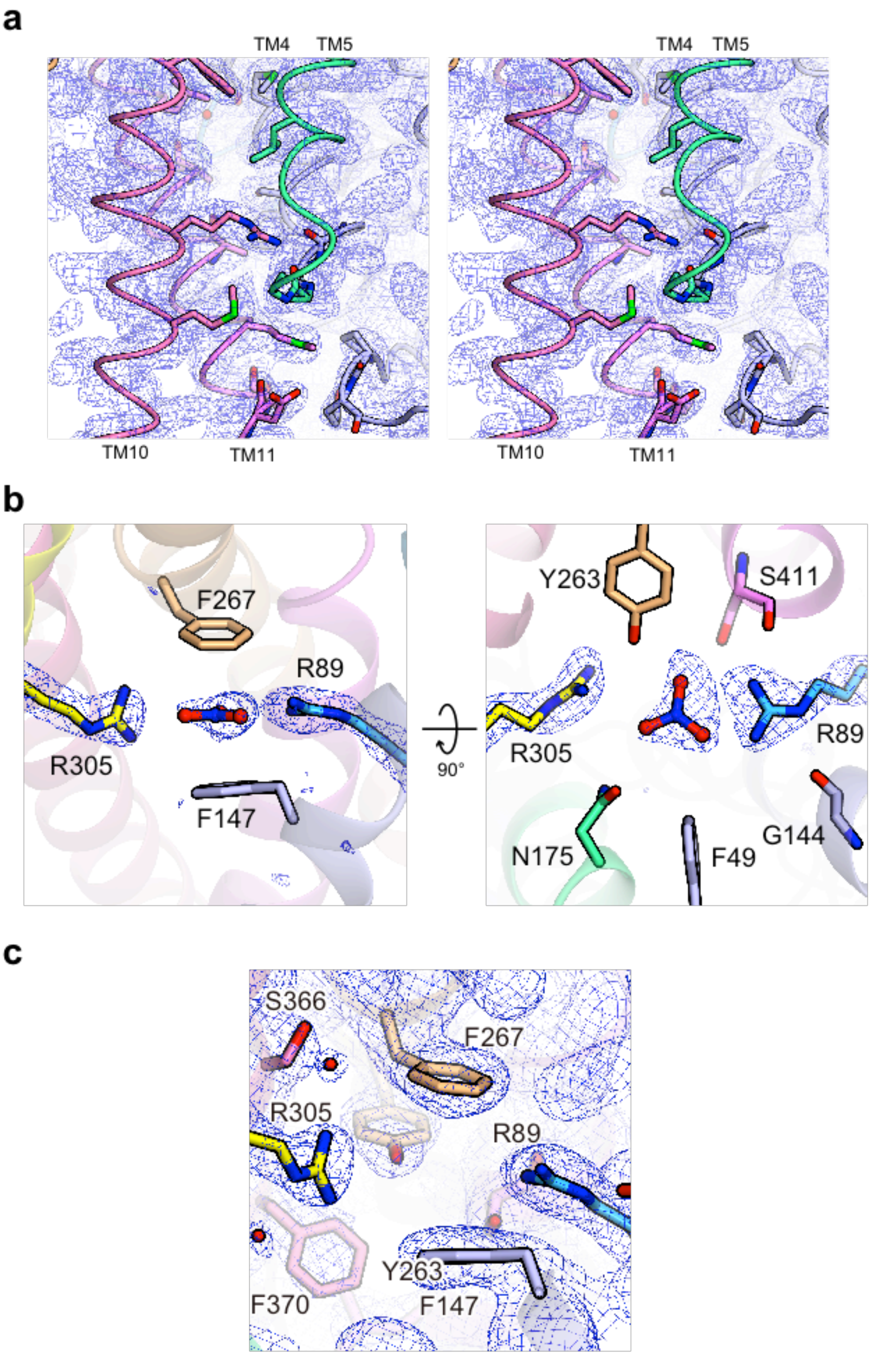


**Figure 3-8 Electron density maps of NarK.**

(**a**) The 2m$F_o$–D$F_c$ electron density map contoured at 1.1 σ around the cytosolic transport path of NarK, in the nitrate-bound occluded state (stereo view). (**b**) The m$F_o$–D$F_c$ omit maps contoured at 4 σ around the substrate-binding site of the nitrate-bound inward-open state. The nitrate ion and the side chains of R89 and R305 were omitted in the map calculation. (**c**) The 2m$F_o$–D$F_c$ electron map contoured at 1.1 σ around the substrate-binding site of NarK, in the *apo* inward-open state.

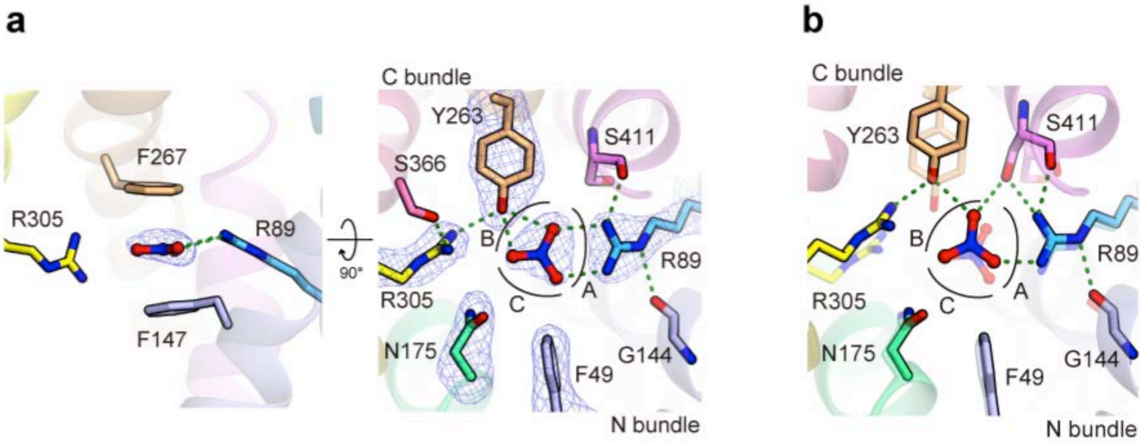
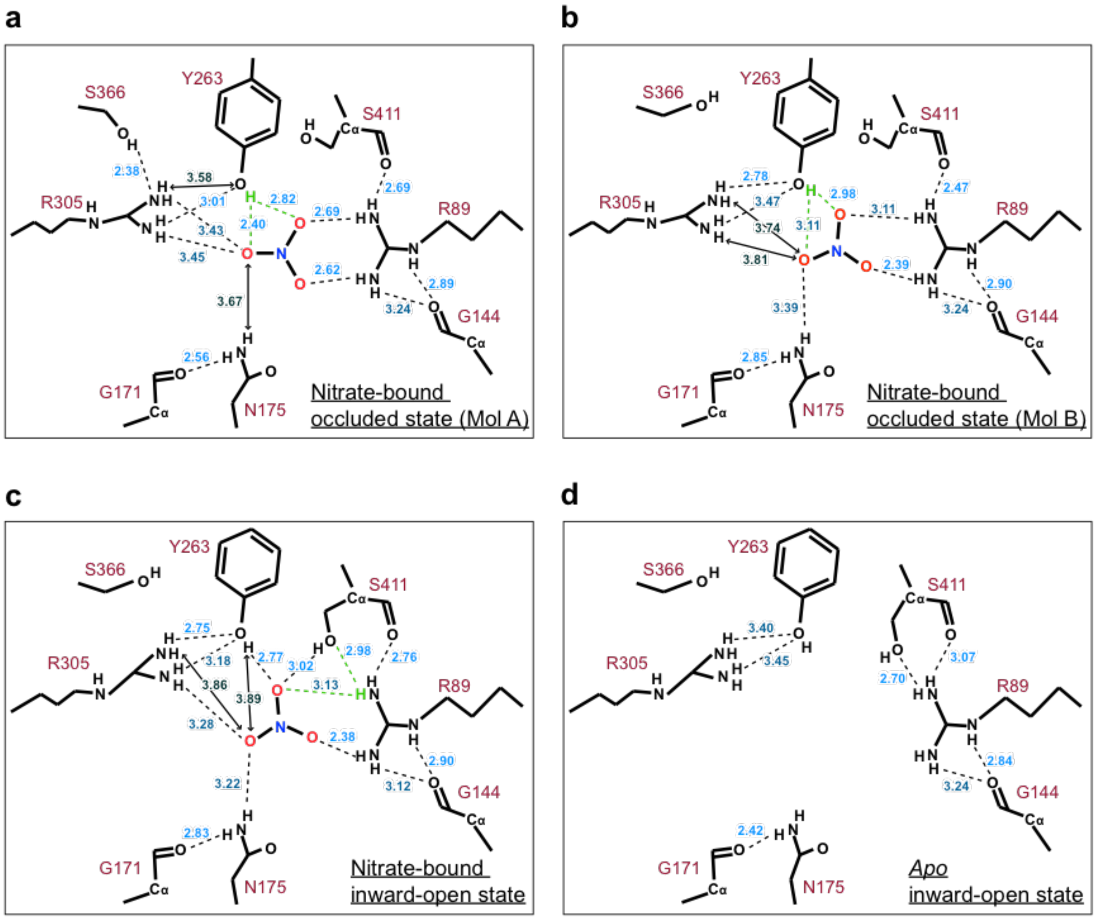


**Figure 3-9 The nitrate recognition mechanism and the structural change in the substrate-binding site of NarK.**

(**a**) Close-up views of the substrate-binding site in the nitrate-bound occluded state (Mol A), viewed from two different directions. The electron density map of the m$F_o$–D$F_c$ omit maps, contoured at 4 σ for nitrate ion (left panel) and nitrate ion, F49, R89, N175, Y263 and R305 (right panel), respectively. The colors of the carbon atoms represent the TM helices with the same color scheme as in Figure 3-1 (a) and (b). The green dashed lines indicate the hydrogen bonds up to 3.1 Å. (**b**) Close-up view of the substrate-binding site in the nitrate-bound inward-open state. The residues in the occluded state are semi-transparent, for comparison.



## **Figure 3-10** Schematic drawings of the hydrogen bond geometry in the substrate-binding site of each NarK molecule.

(**a**), (**b**), (**c**), and (**d**) Alternative hydrogen-bonding configurations possible for (a) the nitrate-bound occluded state (Mol A), (b) the nitrate-bound occluded state (Mol B), (c) the nitrate-bound inward-open state, and (d) the *apo* inward-open state are shown. The potential hydrogen bonds and electrostatic interactions within 3.5 Å are depicted as black dashed lines. Distances over 3.5 Å are represented by black arrows. The bifurcated hydrogen bonds and the hydrogen atoms participating in them are colored light green.

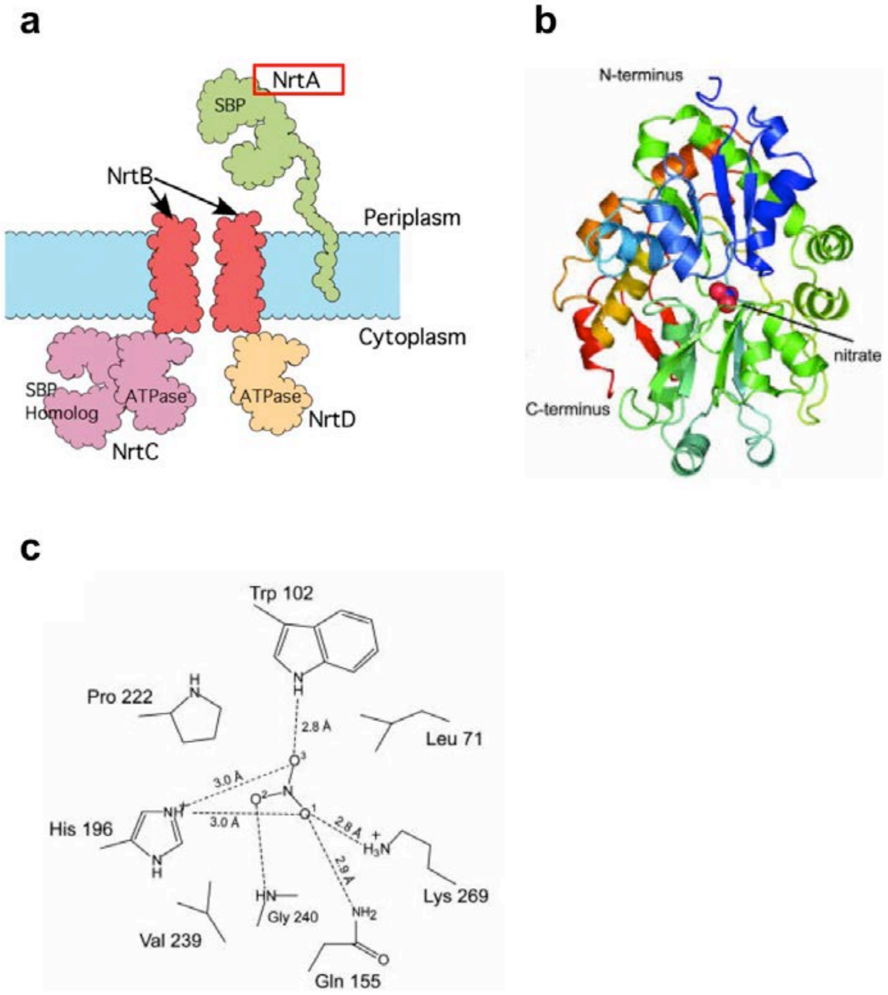


**Figure 3-11 Structure and nitrate recognition mechanism of NrtA, a component of ABC-type nitrate transporter.**

(**a**) The schematic diagram of ABC-type NrtABCD transporter. (**b**) Overall structure of nitrate-bound NrtA from *Synechocystis* sp. PCC 6803. (**c**) Nitrate recognition mechanism of NrtA. The nitrate ion is recognized by Trp, His, Lys, Gly, and Gln residues. The figures are cited from Koropatkin, Nicole M., Himadri B. Pakrasi, and Thomas J. Smith. "Atomic structure of a nitrate-binding protein crucial for photosynthetic productivity." *Proceedings of the National Academy of Sciences* 103.26 (2006): 9820-9825.

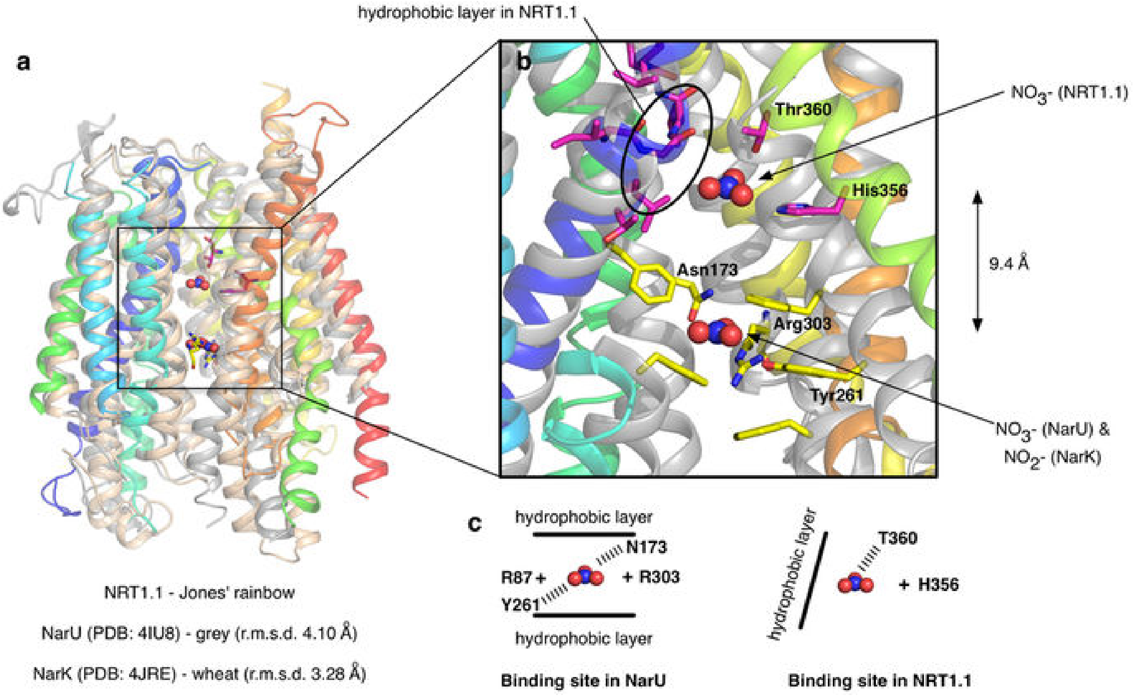


**Figure 3-12 Structure and nitrate recognition mechanism of AtNRT1.1.**

NRT1.1 from *Arabidopsis thaliana* (AtNRT1.1) recognizes nitrate by Thr and His residues. The Figures are cited from Parker, Joanne L., and Simon Newstead. "Molecular basis of nitrate uptake by the plant nitrate transporter NRT1. 1." *Nature*507.7490 (2014): 68-72.

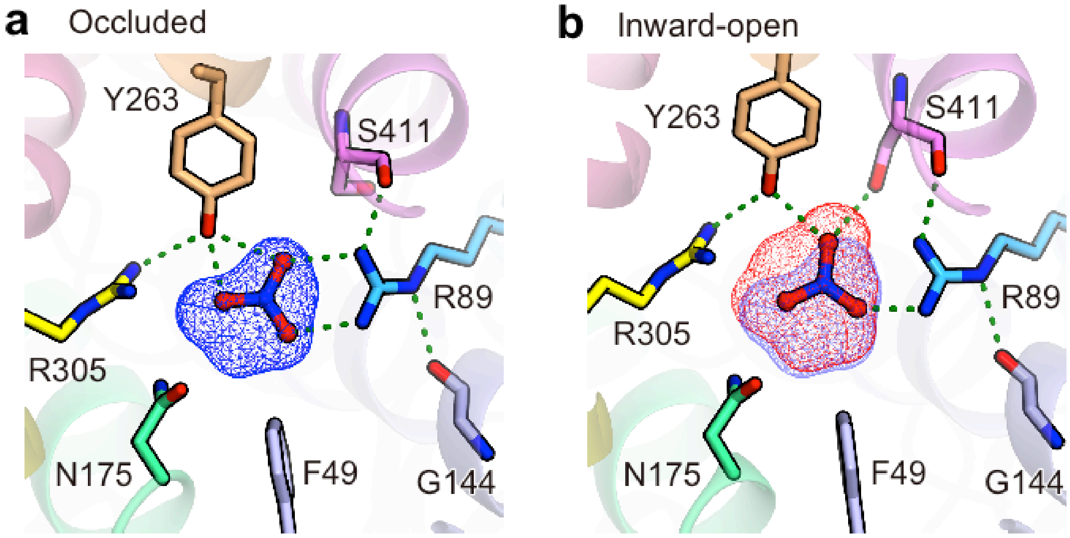
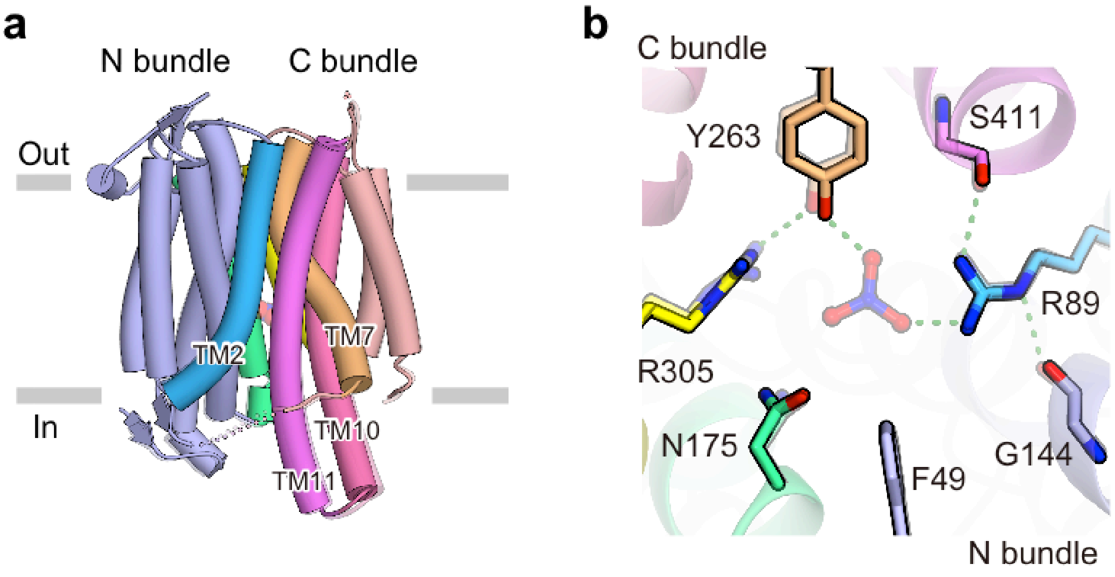


**Figure 3-13 Volumes of the cavity in the substrate-binding site.**

(**a**) The cavity in the substrate-binding site of the occluded state. The blue mesh represents the surface of the cavity calculated by the program MSMS, using a 1.2 Å radius sphere as a probe. The green dashed lines represent the hydrogen bonds up to 3.1 Å. (**b**) The cavity in the substrate-binding site of the inward-open state. The red mesh represents the surface of the cavity, calculated as in panel (a). The cavity in the occluded state is also shown in a blue transparent mesh, for comparison.



## Figure 3-14 Structure of NarK in the apo inward-open state.

(**a**) Overall structure of the *apo* inward-open state, shown as a cartoon model. (**b**) Close-up view around the substrate-binding site. The structure of the nitrate-bound inward-open structure, superimposed over all of the Cα atoms, is shown for comparison in both panels. The colors of the helices and residues correspond to those in Figure 3-1 (a) and (b). The semi-transparent green dashed lines represent the hydrogen bonds found in the nitrate-bound inward-open state.

# Chapter 4 Functional analysis of NarK

## 4.1 Introduction

In **Chapter 3**, the author determined the high-resolution crystal structures of NarK in three different states in the transport cycle. Structural comparison between inward-open and occluded structures suggested the important role of conserved glycine residues in C-bundle. Furthermore, it is suggested from the structures that the substrate recognition and the opening/closing of the transport pathway is coupled by the orchestrated movement of the TM helices in the C-bundle and Tyr263 and Arg305 in the substrate-binding site. Moreover, the molecular function of NarK is suggested to be a nitrate/nitrite antiporter. However, these hypotheses are only based on the structural information. In order to confirm these hypotheses suggested in **Chapter 3**, various kinds of functional analyses were performed. In this chapter, the detail information of these functional analyses is described, and based on the structural information and the results of functional analyses, the transport mechanism of NarK is discussed.

## 4.2 *in vitro* functional analysis using proteo-liposome

### 4.2.1 Materials and Methods

#### 4.2.1.1 Preparation of mutants

All mutants were created by site-directed mutagenesis. The expression and purification were performed in the same manner as described in **Chapter 2**. The DNA primers for introducing mutations into the NarK gene are shown in **Table 2-2**.

### 4.2.1.2 Preparation of proteo-liposome

For the preparation of NarK-reconstituted proteoliposomes, soybean lipid (L-α-phosphatidylcholine type IV-S; Sigma) was dissolved in chloroform and evaporated at room temperature. The lipid film deposited on the interior of a plastic tube was resuspended to a final concentration of 10 mg $mL^{-1}$ in the extra-liposome buffer, containing 25 mM HEPES-NaOH (pH 7.0), and 50 mM of either $NaNO_3$, NaCl, $NaHCO_3$, HCOONa, $CH_3COONa$ or $CH_3CH_2COONa$. The suspension was sonicated for 30 sec. to form unilamellar vesicles. Purified proteins were reconstituted into liposomes, at a lipid/protein ratio of 100:1 (w/w), by three rounds of freezing and thawing.

### 4.2.1.3 Nitrite influx assay using the proteo-liposome

The proteo-liposomes prepared in **4.2.1.2** were centrifuged at 210,000×g for 30 min. The nitrite uptake reaction was initiated by resuspending the pellet in intra-liposome buffer, containing 25 mM HEPES-NaOH (pH 7.0) and 50 mM $NaNO_2$. After an incubation at room temperature for 1 hour, the reaction mixture was loaded on a Sephadex G-50 column (0.5×4.0 cm) equilibrated with 25 mM HEPES-NaOH (pH 7.0). The eluted fraction was mixed with sodium dodecyl sulfate (SDS), to disrupt the liposomes, and then with HCl, in order to lower the pH, to final concentrations of 0.5% and 0.1 M, respectively. The solution was mixed with DAF-FM[77] and incubated at room temperature for 2 min. The solution was diluted 10-fold with 25 mM HEPES-NaOH (pH 7.0), and incubated at room temperature for 2 min. The fluorescence of triazolofluorescein was monitored with excitation at 500 nm and emission detected at 515 nm, using an F-7000 fluorescence spectrophotometer (Hitachi) (**Figure 4-1**). To generate the membrane potential in proteoliposomes, the author added 10 mM KCl to the intra-liposome buffer, and 100 mM KCl with 10 µM valinomycin to the extra-liposome buffer. To change the pH of the intra- and extra-liposome buffers, MES-NaOH (pH 6.0) and Tris-HCl (pH 8.0) were used.

### 4.2.2 Results

#### 4.2.2.1 NarK is a nitrate/nitrite antiporter

To understand the exact function of NarK, an *in vitro* assay system was developed. In this assay, NarK-reconstituted proteoliposomes and the fluorescent probe, DAF-FM[77] were used. Although DAF-FM was originally developed as a fluorescent probe for nitric oxide, it also reacts with nitrite under acidic conditions to form a fluorescent compound, triazolofluorescein, but not with other ions, including nitrate[77] (**Figure 4-1**). In this system, the amount of nitrite accumulated in the liposomes were quantified by measuring the fluorescence derived from the formation of triazolofluorescein. The uptake of nitrite was observed when nitrate was present inside the liposomes (**Figure 4-2**). In contrast, almost no nitrite uptake was observed when either acetate or propionate, which are larger than nitrate, was inside the liposomes. Anions with similar or smaller sizes than nitrate; *i.e.*, chloride, carbonate and formate, exhibited weaker nitrite uptake activities than nitrate. While the lower pH resulted in increased transport activity in spite of the presence or absence of the membrane potential, the pH gradient across the membrane had no effect on the activity, suggesting that the $H^+$ concentration gradient across the membrane does not drive the substrate transport. Taken together, the flux of nitrite occurs in the presence of nitrate (and its near-cognate anions) as a counteranion, which is transported toward the outside of the liposomes. Although our results do not preclude the possibility that NarK is driven by $H^+$, they clearly demonstrate that NarK is a nitrate/nitrite antiporter.

#### 4.2.2.2 Mutational analysis using nitrite influx assay

By using the nitrite influx assay system, mutational analyses were conducted.

The Y263F and R305K mutants were purified as described in **2.1.3** and reconstituted into the proteo-liposomes. The Y263F and R305K mutants showed no transport activity in our reconstituted system (**Figure 4-2**).

## 4.3 Genetic analysis (Mutational analysis of the substrate recognition residues)

### 4.3.1 Materials and Methods

#### 4.3.1.1 Preparation of mutants

All mutants were created by site-directed mutagenesis as described in **4.2.1.1**.

#### 4.3.1.2 Construction of *E. coli* strain

To construct the assay strain with the sextuplet gene knock-out (Δ*napA*-*napB*, Δ*narZ*-*narU*, Δ*nirC*, Δ*narK*), each gene's knockout allele was constructed and transferred step-by-step to the destination strain BW25113, principally by the method of Datsenko and Wanner[78], using appropriate PCR primers (NAP1 and NAP2 for Δ*napA*-*napB*, NAR1 and NAR2 for Δ*narZ*-*narU*; **Table 4-1**) or P1 phage transduction from a systematic *E. coli* knockout library strain (Δ*nirC*, Δ*narK*)[79]. The λ DE3 lysogenic strains, which carries the inducible T7 polymerase system for the gene expression[80], were isolated by spotting λ DE3 lysate together with appropriate helper phage lysate onto the host strain.

#### 4.3.1.3 Genetic analysis of the transport activity of NarK mutants

The genetic analysis of the NarK mutants was performed essentially as described previously[24]. In brief, in the presence of trimethylamine N-oxide (TMAO) under anaerobic conditions, *E. coli* assimilates TMAO, which is reduced to the strong base trimethylamine (TMA) by endogenous TMAO reductases, resulting in the formation of pale yellow spots on MacConley-glucose-TMAO agar plates. In contrast, if nitrate is assimilated by *E. coli* cells during anaerobic growth, it prevents TMAO reduction by out-competing TMAO reductase for electrons from the quinol pool[24], resulting in the formation of dark red spots (**Figure 4-3**). The *E. coli* strain constructed for this assay lacks all of the endogenous nitrate/nitrite transporters and channels (NarK, NarU and NirC). Therefore, the colors of the cells grown anaerobically in the presence of both nitrate and TMAO completely depend on the nitrate uptake activity of the plasmid-encoded NarK (**Figure 4-3**).

Under the non-induced conditions in the absence of lactose or IPTG, the assay strain expresses a stable amount of T7 polymerase, which is sufficient for the assays of NarK mutants with severe phenotypes (“no-induction system”; **Figure 4-4**). In contrast, to evaluate the nitrate transport activity of NarK and its mutants depending on their expression levels, another expression system was constructed. In the presence of pLysS, the T7 polymerase activity is reduced, thus, allowing more stringent control of expression from the T7 promoter, depending on the IPTG concentration (“inducible system”). This inducible system was used for the assays of mutants involved in substrate recognition under three different conditions (IPTG = 0 μM, 23 μM, and 67 μM) (**Figure 4-4**).

Transformants harboring the wild type (WT) or mutant *narK* gene derived from the exogenous plasmid were cultured overnight at 37°C in liquid LB medium, supplemented with appropriate antibiotics, diluted 50-fold in the same medium, and then grown to an $OD_{600}$ of 0.6 at 37°C. The refreshed log-phase cultures were spotted on modified MacConkey-glucose-nitrate-TMAO plates[24], containing 0.5 g $KNO_3$, 0.8 g D-glucose and 2.0 g trimethylamine N-oxide (TMAO) in 200 mL MacConkey agar medium, and grown anaerobically at 30°C for 12 hours using AnaeroPack™ jar systems (AnaeroPack Kenki 5% A-07, Mitsubishi Gas Chemical Company, Inc., Tokyo, Japan).

### 4.3.2 Results

#### 4.3.2.1 Mutational analysis of glycine residues in TM helices

To confirm the importance of the conserved Gly residues in TM7, TM10, and TM11, a genetic analysis of NarK mutants of these Gly residues was performed. In this analysis, we examined the complementation activity of NarK mutants introduced by a plasmid, using a mutant *E. coli* strain lacking the operons of all nitrate and nitrite transporters, as well as two of the three nitrate reductases (BW25113 Δ*nirC* Δ*narZ-narU* Δ*narK* Δ*napA-napG-napH-napB* (DE3))[81]. Using the "no-induction system", the complementation activities of NarK mutants were visualized by the color of *E. coli* spots on MacConkey-glucose-nitrate-TMAO agar plates under anaerobic conditions[24]. The results showed that the G408A and G418A mutants of TM11 and the triple Ala mutant of G363, G365, and G367 in TM10 (ΔG3 mutant) completely abolished the nitrate uptake activity (**Figure 4-4**). These results suggested the importance of the flexibility of TM10 and TM11 conferred by these Gly residues. Furthermore, the G268A mutant of TM7 also completely abolished the nitrate uptake activity, suggesting that the bending of TM7 around the G268 residue is critical for the nitrate transport activity by NarK.

#### 4.3.2.2 Mutational analysis of the substrate recognition residues

To examine the importance of the residues involved in the substrate recognition, the structure-based genetic analysis of NarK mutants was performed using the "inducible system". The results showed that the Ala mutants of both F147 and F267, which form π-π interactions with nitrate (**Figure 3-9**), did not restore the nitrate-transport activity to the same level as that of wild-type NarK, while complementation was observed with higher IPTG concentrations (23 μM and 67 μM) (**Figure 4-4**). This result suggested that the π-π stacking interaction between the substrate nitrate and the phenyl groups of F147 and F267 is important, but not critical for the transport activity. The R89K mutant also abolished the complementation activity, but rescued the nitrate-uptake activity under the conditions with the highest IPTG concentration (67 μM). This result suggested that the positive charge at the R89 position is important for the substrate recognition. In contrast, the Y263F and R305K mutants completely abolished the complementation activity, even under the conditions with the higher IPTG concentrations (23 μM and 67 μM). The results obtained with the Y263F mutant suggested that the phenol hydroxyl group of Y263, which hydrogen bonds with both the substrate nitrate and R305 guanidinium group, is critical for the transport activity. In addition, the results from the R305K mutant suggested that the guanidinium group of R305 itself, rather than simply a positive charge, is essential for the transport activity.

Taken together, the results from genetic and *in vitro* analyses support our proposal that Y263 and R305 in the substrate-binding site are the key residues for coupling the bending of the TM helices with the substrate recognition.

## 4.4 Molecular dynamics simulation

### 4.4.1 Materials and Methods

The MD simulations starting from the *apo* occluded, nitrate-bound occluded, and *apo* inward-open states were performed. The present crystal structures of the nitrate-bound occluded and *apo* inward-open states were used as initial models. All of the lipid molecules observed in the crystal structures were removed, while all of the water molecules observed in the crystal structures were kept. The disordered loop connecting the N and C bundles was modeled by the program MODELLER[82]. For the *apo*-occluded simulation, the bound nitrate ion was removed from the crystal structure. The missing side chains and hydrogen atoms were built with the program VMD[83]. The prepared structures were then embedded in a fully hydrated POPE bilayer[84]. In the crystal structure, several monoolein molecules were observed around the TM segments, and these were used as indicators for the locations of the membrane. The lipid molecules overlapping the protein were removed, resulting in 320 POPE molecules included overall. The lipid-protein complex was then hydrated to form the 96 Å × 96 Å × 96 Å simulation box. Sodium and chloride ions were then added, to neutralize the system with a salt concentration of 150 mM. The molecular topologies and parameters from the CHARMM36 force-field parameters, with $\varphi$, $\psi$ cross-term map correction (CMAP)[84], were used.

Molecular dynamics simulations were performed with the program NAMD 2.8[85]. The systems were first energy minimized for 1,000 steps with fixed positions of the non-hydrogen protein atoms, and then for another 1,000 steps with 10 kcal/mol restraints for the non-hydrogen protein atoms. For the equilibration, the system was first subjected to a 100-ps simulation run with harmonic restraints on the non-hydrogen protein atoms in the NPAT ensemble, to relax the solute molecules, and then was further equilibrated by a 1-ns simulation run with harmonic restraints on the non-hydrogen protein atoms in the NPT ensemble. Finally, 250 ns MD simulations without any restraints were performed in the NPT ensemble for the production runs. In these simulations, constant pressure (1 atm) and temperature (300K) were maintained using Langevin dynamics and a Langevin piston, respectively. The particle mesh Ewald (PME) method was employed for the calculation of the electrostatic interactions[86]. The equation of motion was integrated with a time step of 2 fs.

### 4.4.2 Results

To clarify the dynamic coupling mechanism between the state transition and the substrate recognition, molecular dynamics (MD) simulations were performed based on the present crystal structure. In the simulation starting from the occluded state with the bound nitrate (crystal structure), no large structural change in the overall conformation was observed during the 250 ns simulation (**Figure 4-5**), suggesting that the overall structure of the occluded state is stable in the lipid bilayer environment. The structure of the substrate-binding site, including the nitrate recognition manner, remained unchanged during the simulation, showing that the substrate-binding site of the occluded state structure is stable with the bound nitrate.

Next, to examine the effect of the nitrate binding and release on the occluded state structure, the three independent simulations starting from the occluded state without the bound nitrate were performed. Interestingly, a large structural change was observed in the first 10 ns, mainly in the TM10 and TM11 helices in the C bundle (**Figure 4-5**). The interaction between TM4-TM5 and TM10-TM11 was lost in the first 10 ns, including the hydrogen bonds between R378 and the loop connecting TM4 and TM5. The transport path from the cytosolic side to the substrate-binding site opened, and water molecules entered the substrate-binding site. After 250 ns in the simulation, the distance between TM4-TM5 and TM10-TM11 was increased to a similar level as that in the inward-open structure (**Figure 4-5**). In addition, the MD simulation starting from the *apo* inward-open structure was also performed. The result showed that the *apo* inward-open structure is stable, and no structural change toward the occluded state was observed without any substrates during the 250-ns simulation (**Figure 4-6**).

Taken together, these results suggested that the negatively-charged nitrate ion, bound between the positively-charged R89 and R305 residues, relaxes the electrostatic repulsion between them, thereby facilitating the formation of the occluded state. In contrast, without the substrate nitrate ion, the electrostatic repulsion between R89 and R305 allows the N and C bundles to segregate from each other, thereby facilitating the formation of the inward-open state. The similar regulation mechanism by the electric repulsion was proposed in the well-studied anion exchanger, the triose-phosphate/phosphate translocator (TPT)[69].

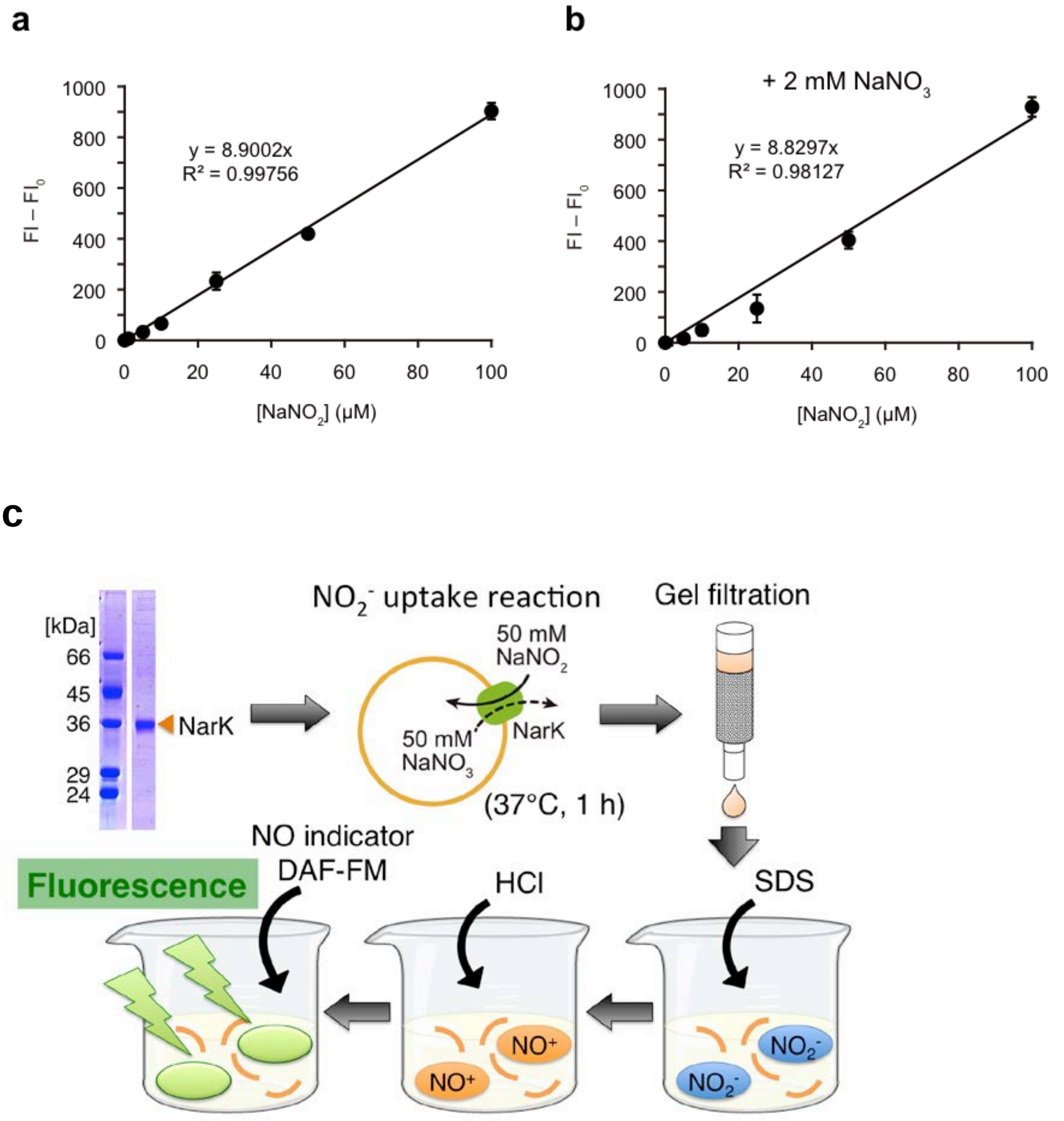
## 4.5 Discussion and Perspectives

The results of *in vitro* liposome-based functional analysis clearly showed that NarK is a nitrate/nitrite antiporter. This conclusion is also consistent with the results of molecular dynamics simulations with and without nitrate in the substrate-binding site. The present crystal structures and structure-based genetic analyses have elegantly explained the coupling mechanism between the substrate recognition and the structural changes that underlie the nitrate/nitrite antiport mechanism. This is also supported by the liposome-based mutational analyses using the purified Y263F and R305K mutants.

Based on the structural and functional analyses, the author proposed the working model of nitrate/nitrite antiport by NarK which explains the half of the transport cycle involving the inward-open state (**Figure 4-7**). In the nitrate-bound occluded state, the negative charge of the nitrate in the substrate-binding site relaxes the electrostatic repulsion between the two positively charged arginine residues (R89 and R305), thereby enabling these two arginine residues to be in close proximity in the low-dielectric intramembrane environment. The following structural change to the inward-open state induces the bending of the helices, TM7, TM10, and TM11, around the conserved Gly residues, which opens up the transport pathway from the substrate-binding site to the cytoplasmic side between the N and C bundles. These bending motions of the TM helices are coupled with the movement of the side chains of Y263 in TM7 and R305 in TM8 away from the N-bundle side, which enlarges the volume of the substrate-binding site and decreases the affinity of NarK for nitrate, thereby facilitating nitrate release to the cytoplasm. The key residues for this coupling mechanism (Y263, R305, G268) are strictly conserved in the NNP family (**Figure 3-4**), suggesting that other NNP family members, including NRT2.1 of *Arabidopsis thaliana*, recognize nitrate ions in a similar manner. The electrostatic repulsion between the positive charges of R89 and R305 prohibits the movement of Y263 and R305 toward the N-bundle side, thus preventing the closure of the transport pathway by the movement of TM7, TM10 and TM11 of the C bundle toward the N-bundle side. In the next step, a nitrite ion in the cytoplasm is bound to the substrate-binding site in a similar manner to that of nitrate, which may facilitate the conformational change to the outward-open state via the nitrite-bound occluded state. In this structural change between the occluded and outward-open states, the other conserved Gly residues in TM2, TM4, TM5, and TM8 may play important roles in the conformational flexibility

(**Figure 3-4**). Further understanding of the remaining half of the transport cycle awaits the determination of the outward open structure at atomic resolutions.

## Figures and Tables of Chapter 4



### Figure 4-1 Standard curve of DAF-FM as a nitrite-specific probe.

(**a**) Relationship between the nitrite ion concentration and the fluorescence intensity of triazolofluorescein, a nitrite-adduct of DAF-FM. (**b**) The presence of nitrate does not affect the triazolofluorescein formation by nitrite and DAF-FM. FI is the fluorescence intensity and $FI_0$ is the FI with 0 μM $NaNO_2$. All error bars represent the standard deviation (s.d.) of three independent trials. (**c**) Liposome-based nitrite-flux assay system.

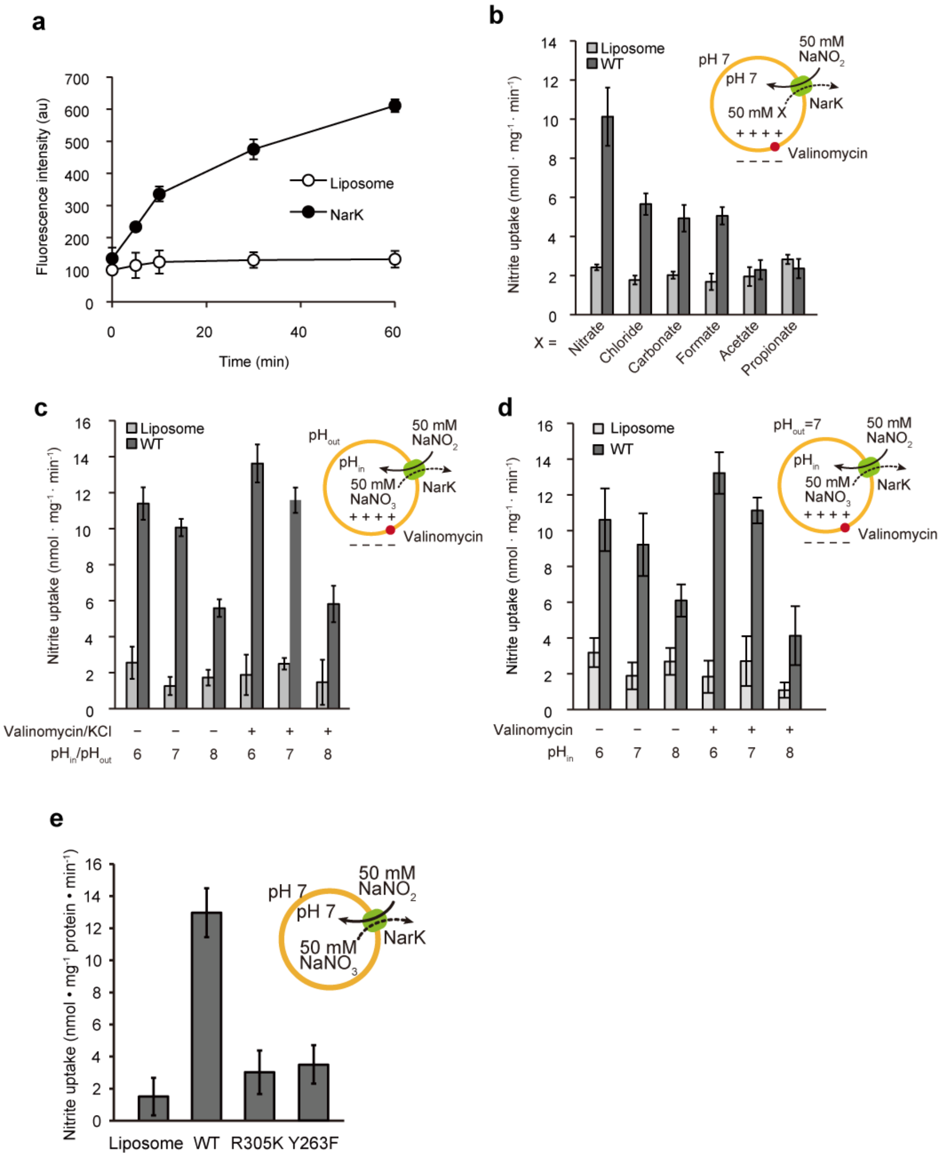
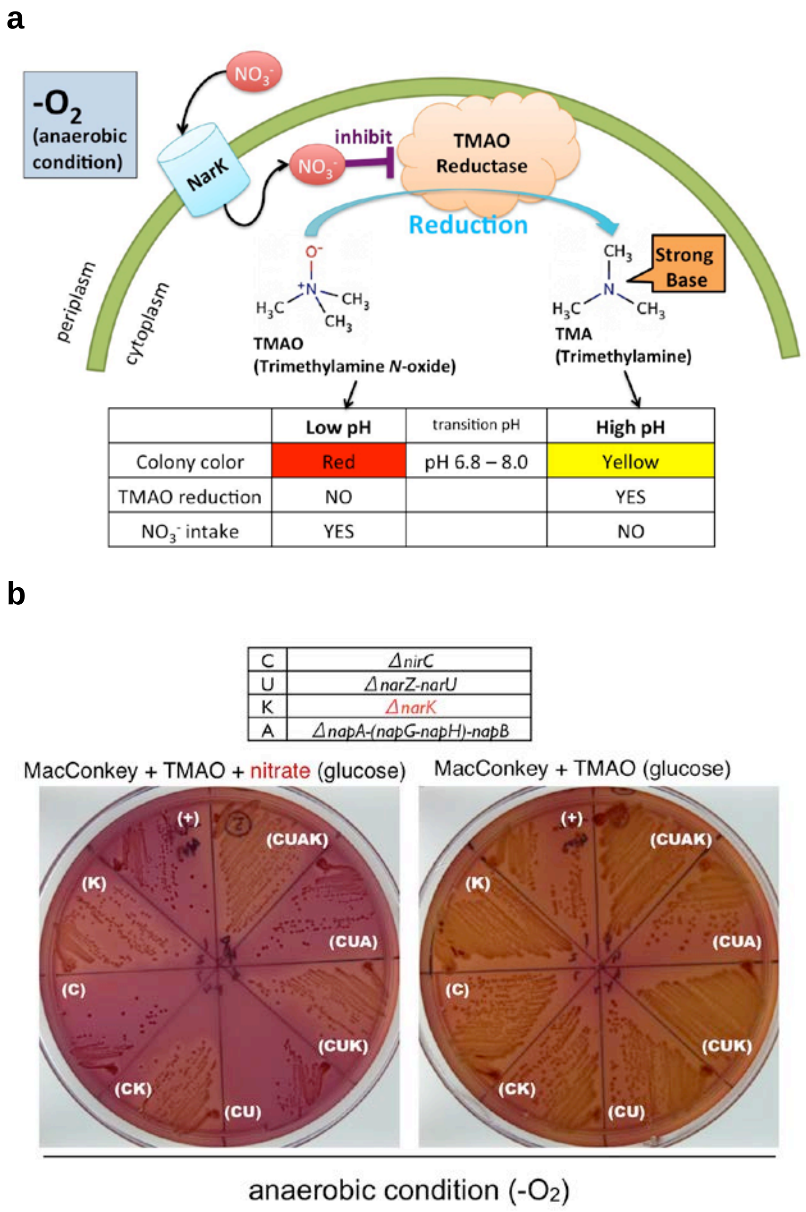


**Figure 4-2 Liposome-based nitrite-flux assay of NarK.**

(**a**) Time-dependent nitrite influx driven by the nitrate concentration gradient across the membrane at pH 7, in the presence of a membrane potential, measured at 37˚C. The intra- and extra-liposomal solutions contained 50 mM $NaNO_3$ and $NaNO_2$, respectively. (**b**) Substrate specificity of NarK. The intra-liposomal solution contained 50 mM $NaNO_3$, NaCl, $Na_2CO_3$, HCOONa, $CH_3COONa$ or $C_2H_5COONa$. The extra-liposomal solution contained 50 mM $NaNO_2$, and was the same in all measurements. (**c**) pH and membrane-potential dependence of NarK activity. (**d**) Liposome-based nitrite uptake assay in the presence of pH concentration gradients across the membrane. All error bars represent the standard deviation (s.d.) of three independent trials. (**e**) Liposome-based nitrite-flux assay of NarK mutants (Y263F and R305K). All error bars represent the s.d. of three independent trials.

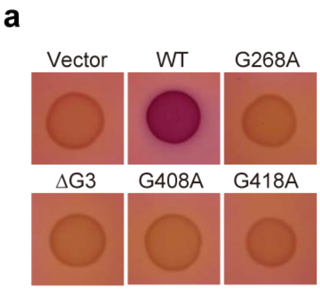


| | Low pH | transition pH | High pH |
|---|---|---|---|
| Colony color | Red | pH 6.8 – 8.0 | Yellow |
| TMAO reduction | NO | | YES |
| $NO_3^-$ intake | YES | | NO |

| | |
|---|---|
| C | *ΔnirC* |
| U | *ΔnarZ-narU* |
| K | *ΔnarK* |
| A | *ΔnapA-(napG-napH)-napB* |

## Figure 4-3 The principle of color spot assay.

(**a**) The principle of color spot assay system. (**b**) Preliminary experiment using gene knock-out *E. coli* strains.



| | Nitrate uptake |
|---|---|
| Vector | − |
| WT | ++ |
| G268A | − |
| ΔG3<br>(G363A, G365A, G367A) | − |
| G408A | − |
| G418A | − |

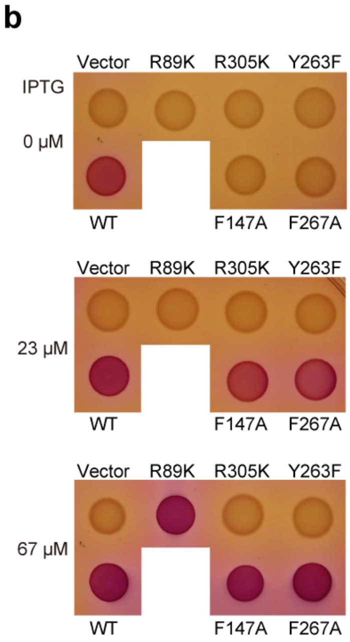


Figure 4-4 **Genetic analysis of the nitrate transport activity of NarK.**

(**a**) Genetic analysis of the nitrate transport activity of NarK mutants of the conserved Gly residues involved in the bending of TM7, TM10 and TM11, using the no-induction system. *E. coli* cultures were spotted on MacConley-glucose-TMAO agar plates under anaerobic conditions, and their colors were monitored. *E. coli* cells expressing intact NarK accumulate nitrate, generating dark-red spots, while those expressing inactive NarK generate pale-yellow spots. The results of this genetic analysis are summarized in the table on the right. (**b**) Genetic analysis of the nitrate transport activity of NarK mutants directly involved in substrate recognition, using the inducible system. The results are summarized in the table on the right.

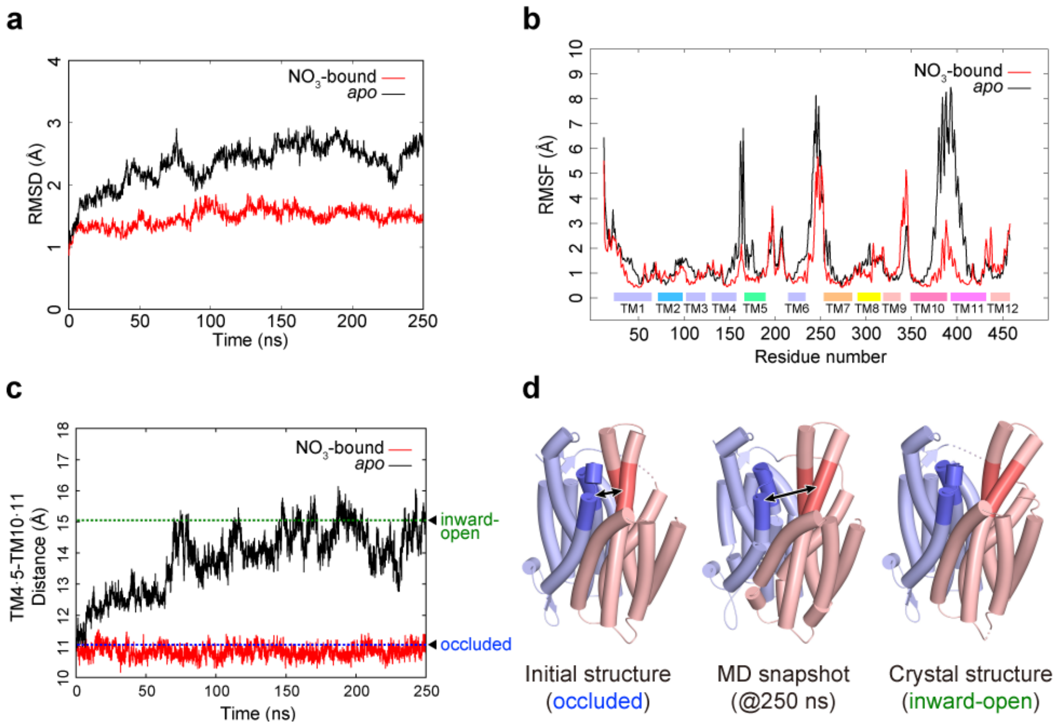


**Figure 4-5 Molecular dynamics simulations of NarK.**

(**a**) and (**b**) Structural deviations and fluctuations observed in the molecular dynamics simulations. (a) Plots of the RMS deviations of the overall Cα atoms from the initial crystal structures as a function of time, during the nitrate-bound and *apo* occluded simulations. (b) Plots of the time-averaged RMS fluctuations of each residue from the initial crystal structures, during the nitrate-bound and *apo* occluded simulations. The locations of the TM segments are indicated by the boxes, which are colored as in Figure 3-1 (a) and (b). (**c**) Plots of the distances between the centroids of the Cα atoms of the cytosolic parts of TM4-TM5 (A148 to G171) and TM10-TM11 (T369 to L380, and A399 to F409), as a function of time, during the nitrate-bound and *apo* occluded simulations. The distances between the centroids of the Cα atoms of the cytosolic parts of TM4-TM5 and TM10-TM11 of two crystal structures (occluded and inward-open state) are shown by blue and green dashed lines, respectively, and indicated by black rectangles. (**d**) The initial structure, the snapshot structure of the *apo* simulation at 250 ns, and the crystal structure in the inward-open state are shown. The N and C bundles are colored pale blue and pink, respectively. The regions in TM4, TM5, TM10, and TM11 used for the distance measurement are highlighted in deep blue and red colors and indicated by black arrow

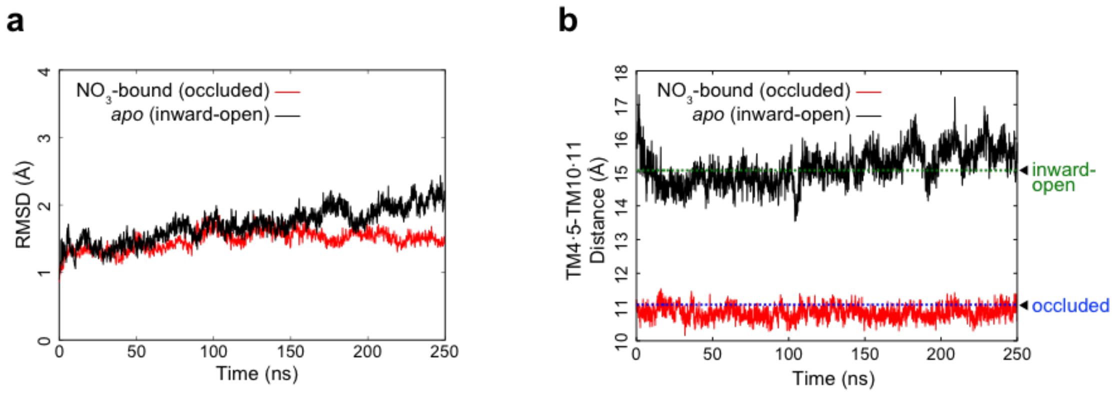


**Figure 4-6 MD simulation starting from the apo inward-open structure.**

(**a**) Plots of the RMS deviations of the overall Cα atoms from the initial crystal structures (nitrate-bound occluded and *apo* inward-open states), as a function of time, during the simulations. (**b**) Plots of the TM4•5-TM10•11 distance as a function of time, during the nitrate-bound occluded and *apo* inward-open simulations. The TM4•5-TM10•11 distances of two crystal structures (occluded and inward-open state) are shown by blue and green dashed lines, respectively, and indicated by black rectangles.

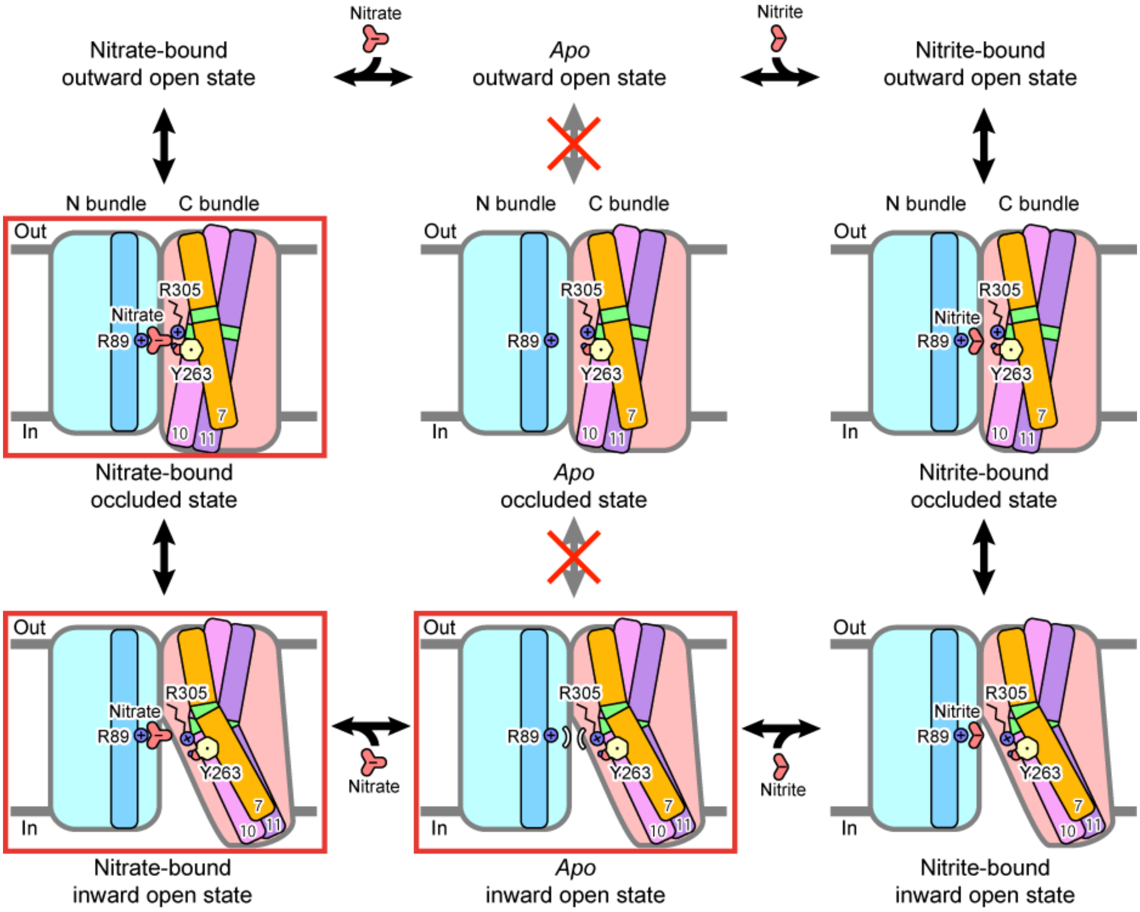


**Figure 4-7 Working model of nitrate/nitrite antiport by NarK.**

The key residues in the coupling between the structural change and the substrate recognition (R89, Y263, and R305) and substrates (nitrate and nitrite) are shown. The TM helices that become bent during the transport cycle (TM7, TM10, and TM11) in the C bundle are illustrated as amber, pink and purple sticks, respectively, while those in the N bundle are represented by light blue sticks. The approximate locations of the important Gly residues for the TM bending are colored green. The states corresponding to the present crystal structures are highlighted by red rectangles.

**Table 4-1 Primers for constructing *E. coli* strain used in genetic analyses.**

| Name | Knock-out gene | Sequence |
|---|---|---|
| NAP1 | ΔnapA-napB | 5'- TATGAAAGCTAACGCCGTTGCGGCCGCTGCGGCGGCTGCCGTGTAGGCTGGAGCTGCTTC -3' |
| NAP2 | ΔnapA-napB | 5'- GGTAAAGGTATTCCCCACGATTGGCGCGGTATCGGCCTGCCATATGAATATCCTCCTTA -3' |
| NAR1 | ΔnarZ-narU | 5'- AATAGTCGTTATCTTTTGCGCGACTGGAAACCAGAAAATCGTGTAGGCTGGAGCTGCTTC -3' |
| NAR2 | ΔnarZ-narU | 5'- GATCGCGACCTTCATCATCCAGCCAGTTAACGTTCTTCATCATATGAATATCCTCCTTA -3' |

# Chapter 5　５年以内に雑誌等で刊行予定のため、非公開。

# Chapter 6 General Discussion

## 6.1 Comparison with previously reported structures of NNP family

While the author was on the way of the structural and functional analyses of NarK, two other groups reported the structures of prokaryotic NNP family transporters[26,27] (**Table 1-1**). These structures revealed the key amino acid residues that are involved in substrate gating and transport, and suggested that they transport substrates via an alternating-access mechanism[28] (**Figure 1-4**). Here the author will discuss again the differences between the previously reported and the present crystal structures.

### 6.1.1 Comparison with structures of NarU

Just two months after the author determined the structure of NarK, Yan *et al.* (in Yigong Shi's laboratory) reported the crystal structures of NarU from *Escherichia coli* at 3.0–3.1 Å resolutions in occluded and partially inward-open states[26]. These two structures are well-superimposed to the present crystal structures of occluded and inward-open states, respectively (**Figure 6-1 a**, **c**), however, some structural changes were observed as described in **3.3.2** and **3.3.3** (**Figure 3-5 c**, **d**).

The r.m.s.d. between the structure of NarU reported in the occluded state (PDB

ID: 4IU8 chain A) and the present crystal structure of NarK in occluded state (PDB ID: 4U4W chain A) is 1.2 Å for all Cα atoms (**Figure 6-1 a**). From the structural comparison between these two occluded structures, the differences were observed mainly in TM10 and slightly in TM11 and TM7 (**Figure 3-5 c**, **d** and **Figure 6-1 a**). This slight movement of TM7 resulted in the slight movement of Try263 side chain (**Figure 6-1 b**). The r.m.s.d. between the structure of NarU reported in the partially inward-facing state (PDB ID: 4IU8 chain B) and the present crystal structure of NarK in inward-open state (PDB ID: 4U4T) is 0.79 Å for all Cα atoms. These two inward-conformation structures are quite similar but there are differences in TM10 and TM11 (**Figure 6-1 c**). The slight movement of Try263 side chain was also observed (**Figure 6-1 d**). These structural comparison suggested that these two previously reported structures are both the intermediate states between the present occluded and inward-open structures (**Figure 3-5 c**, **d**), and therefore the author renamed the conformations of these NarU structures as partially occluded inward-facing state.

Yan *et al*. suggested that the structural changes in TM10 and TM11 are important for substrate transport. The present crystal structures further revealed that the kink movement in TM7 around the conserved Gly268 residue plays the critical roles for conformational changes. As TM7 is located adjacent to TM10 and TM11 and is tightly interacted by the hydrophobic interactions (**Figure 3-1**), the structural change of TM7 may be coupled with those of TM10 and TM11. The structural changes of TM10, TM11, and TM7 are transmitted to the substrate binding site via Tyr263 in TM7. Based on the structural comparison, Yan *et al*. proposed the transport mechanism in which bending of transmembrane helices within the N- and C-domains, rather than rigid-body movement of these two domains, might be occurred with substrate transport (**Figure 6-3 a**). It remains unclear whether both the N- and C-domains become open to the cytoplasm in the inward-open conformation or not.

Although the architecture of substrate binding site and the putative substrate translocation pathway was revealed by the NarU structures, the substrate recognition mechanism was not completely reliable because the quality of the density maps of NarU was quite poor (**Figure 6-1 e**) compared to the present crystal structures (**Figure 3-8** and **Figure 3-9**) and no density peak correspond to nitrate was observed (**Figure 6-1 f**). This is possibly due to the medium resolution (3.0-3.1 Å) or the incompleteness of structure refinement (The Ramachandran outliers in the PDB file of 4IU8 is 4.42%). The detailed nitrate recognition mechanism including the hydrogen bonding networks was revealed by the present high-resolution nitrate-bound structures (**Figure 3-9**).

The structural comparison revealed that the structures of NarK and NarU from

*Escherichia coli* are structurally similar. This is consistent with the facts that their sequence identity is ~76% and they both belong to the NarK2-type clade (discussed in **6.5**). Yan *et al*. concluded that NarU functions most likely as a symporter rather than an antiporter, using both the sodium and potassium ion as the co-transported cation based on the results of liposomal stopped-flow assay. In this assay, Yan *et al*. evaluated the activity of the transporter reconstituted in the liposome indirectly by tracing the changes of light scattering signals caused by the deflating and reswelling of proteo-liposomes. Although it seems to be true that NarU can transport both of nitrate and nitrite, it is unlikely that NarU transports $Na^+$ and/or $K^+$ as counter ions, because the putative substrate translocation pathway is highly positively charged and lacks neither the protonatable residues nor the potential binding site for either $Na^+$ nor $K^+$. Although there is a possibility that NarU transport the substrate in the different manner from NarK, it would be more likely that both NarK and NarU transport the substrates in the same manner: *i.e.* nitrate/nitrite antiport mechanism.

### 6.1.2 Comparison with previously reported NarK structures

Two more months after the publish of the paper of NarU, Zheng *et al.* (in Tamir Gonen's laboratory) reported the structures of NarK from *Escherichia coli* in *apo* and nitrite-bound form at the resolutions of 2.6 and 2.8 Å, respectively[27] (**Table 1-1**). These structures were essentially the same conformation. The structural comparisons to the present crystal structures of occluded and inward-open conformations revealed the structural differences in TM10, TM11, and TM7 (**Figure 6-2 a**, **c**) and also slightly in the substrate binding site (**Figure 6-2 b**, **d**). These previously reported structures of NarK bound to the antibody Fab fragment is reported as the inward-facing conformation (PDB ID: 4JR9)[27]. However, the hydrophobic interactions in the cytosolic transport pathway are still formed, and thus the substrate-binding site is not accessible in the previous structure (**Figure 3-5 a**, **b**). Therefore, the previously reported structure of NarK is still in the intermediate, partially-occluded inward-facing state.

In the previously reported NarK structure, a strong electron-density peak was observed in the substrate binding site. The m$F_o$–D$F_c$ omit map contoured at 3.0 $\sigma$ is shown in **Figure 6-2 f**. The shape of the density is well-fit to that of nitrite which was soaked to the crystals, and thus Zheng *et al.* assigned the nitrite to this density. In the 2m$F_o$-D$F_c$ electron density map, however, the density of nitrite is not clear (**Figure 6-2 e**), possibly due to the lower affinity of NarK for nitrite, as compared to that for nitrate (**Figure 6-5**), or flexibility in the binding site. The author determined the structure of NarK from *Escherichia coli* in nitrate-bound form. These structures revealed that both nitrate and nitrite can be bound to the same binding site competitively (**Figure 6-2 b**, **d**,

**g**), which support the nitrate/nitrite antiport mechanism proposed in this study. Zheng *et al*. also concluded that NarK is a nitrate/nitrite antiporter, but this conclusion was mainly based on the structural evidence that the positively charged substrate-translocation pathway of NarK lacks any protonatable residues, and they did not perform any functional analysis; they only cited and summarized the results of the previous mutagenesis and functional studies. It should be noted that the inconsistence of the previous studies led to the proposals of various controversial transport mechanisms of prokaryotic NNP family transporters. Although the structural information is greatly supportive for the nitrate/nitrite antiport model, the possibility of other transport mechanisms (e.g. nitrite uniporter, nitrate uniporter etc…) could not be excluded. In this study, the author demonstrated that NarK functions as a tightly-coupled nitrate/nitrite antiporter by the nitrite influx assay system, and revealed the molecular mechanism of substrate transport, in which the substrate recognition and the opening/closing of the transport pathway is tightly coupled by the orchestrated movement of the three TM helices and the key residues Tyr263 and Arg305 in the substrate-binding site (**Figure 4-7**). This transport model is completely different from the canonical rocker switch model proposed by Zheng *et al*., in which the conformational changes are mediated by the rigid-body movement of N- and C-bundles (**Figure 6-3 b**).

## 6.2 Conformational change to the outward open state

Structural and functional analyses suggested that NarK transports substrates via an alternating-access mechanism[28]. During the transport cycle, the conformational changes from the outward-open and inward-open states via the intermediate occluded state should be occurred (**Figure 1-4**). In this study, we determined the occluded and inward-open structures of NarK, and the other groups reported the intermediate conformations between the occluded and inward-open structures presented in this study. These structural and functional analyses revealed the detailed molecular mechanism of the half of the transport cycle of prokaryotic NNP family transporters. In contrast, any structure of outward-open conformation has not been reported yet.

The structural comparison between the occluded and inward-open structures, the structural changes were observed mainly in the transmembrane helices in C-bundles which contain the Nitrate Signature motif 2 (NS2). Although no structural difference is observed in the other TM helices (i.e., TM2, TM4, TM5, and TM8), they include many conserved Gly residues (**Figure 3-4**). Especially, G164, G165, G168, G171, G172, G174, and G177 in TM5 are included in the Nitrate Signature motif 1 (NS1)[74,75]. These glycine-rich motifs are the unique feature of NNP family and it is suggested that they are involved in the transport activity of eukaryotic nitrate transporter NrtA[87]. For investigation of the importance of these Gly residues, the genetic analysis of their mutants was performed. The results also indicated the importance of the Gly residues, G85 in TM2, G141 and G144 in TM4, G164, G168, G172 and G177 in TM5, and G308 and G309 in TM8, for the nitrate uptake activity (**Figure 6-6**). This result suggests that these conserved Gly residues in TM2, TM4, TM5, and TM8 also play important roles in the nitrate uptake activity. These Gly residues may increase the conformational flexibility and contribute to the structural change between the occluded and outward-open states. Another possibility is that NarK adopts the "full-open" inward-open conformation in which transmembrane helices in both the C- and N-bundles are kinked by the conserved Gly residues. It should be noted that the glycine recent study revealed that in the other MFS transporters such as GLUT5 and XylE in Sugar Porter (SP) family, the conformational changes are caused by structural changes in TM7 and TM10, in which the hinge points contain the conserved glycine residues[36,39,88].

## 6.3 Recognition mechanism of nitrite and other anions by bacterial nitrate transporter NarK

The locations of the nitrate ions found in the present NarK structures are consistent with that of the nitrite ion in the previous crystal structure of NarK[27], suggesting that both nitrate and nitrite bind to the same substrate-binding site competitively. The present atomic-resolution structures of the nitrate-bound NarK provided further structural insights into the nitrite ion recognition mechanism (**Figure 6-4**). Given that nitrite also has a negative charge and delocalized π electrons, it can bind to the substrate-binding site between the phenyl groups of F147 and F267, as observed in the nitrate-bound structures. The interactions in Sites A and B are also possible, in the case of nitrite. The nitrite oxygen atoms can form a bidentate salt bridge with the NH1 and NH2 atoms of R89, and hydrogen bond with the hydroxyl group of Y263 (**Figure 6-4**), as in the nitrate-bound structures (**Figure 3-9**). The nitrogen atom of nitrite has lone-pair electrons, and thus it could further hydrogen bond with the hydroxyl group of Y263 (**Figure 6-4**). In contrast, nitrite has only two oxygen atoms, and is smaller than nitrate. Therefore, the interactions involving one of the three oxygen atoms of nitrate are impossible in the case of nitrite. In our nitrite-bound model, the side-chain amide group of N175 (*i.e.*, Site C) is not involved in substrate recognition (**Figure 6-4**). This model is consistent with the previously reported NarK structure in complex with nitrite[27]. This notion is also consistent with the results of ITC measurement, showing that the affinity of NarK for nitrite is much lower than that for nitrate (**Figure 6-5**). The estimated $K_d$ values of NarK were approximately 160 µM for nitrate and 2-3 mM for nitrite.

The present *in vitro* functional analysis revealed that chloride ions weakly facilitate the nitrate-uptake activity (**Figure 4-2**), suggesting that the chloride ion binds to the substrate-binding site to be transported as a counteranion. Although the chloride ion lacks π electrons, it can bind between the R89 and R305 residues to induce the formation of the occluded-state. This is consistent with the results of the genetic analysis of F147 and F267, showing that the π-π stacking interaction with the substrate is important, but not essential for the transport activity. The results of the *in vitro* functional analysis also showed that carbonate and formate ions weakly facilitate the nitrate-uptake activity (**Figure 4-2**). Given that carbonate ion is solely protonated to form bicarbonate ion under the assay conditions, the proton of the bicarbonate cannot fit well within the substrate-binding site (**Figure 6-4**). This may reduce the affinity of the bicarbonate ion to the substrate-binding site, as compared to that of the nitrate ion, despite their similar molecular shapes. The formate ion has a hydrogen atom bound to the carbon atom, while the nitrite ion has lone pair electrons at the corresponding nitrogen atom. Thus, the formate ion also cannot fit well within the substrate-binding site, although it has a similar molecular shape as the nitrite ion (**Figure 6-4**). In the absence of sufficient amounts of nitrite in cytoplasm, however, there is a possibility that NarK uses small anions such as chloride, formate, and carbonate as the “second-best” driving force for uptake of nitrate.

In contrast, the present *in vitro* analysis showed that acetate and propionate ions do not facilitate the nitrate-uptake activity (**Figure 4-2**). The bulky methyl and ethyl groups may completely prevent the interaction of these anions with the substrate-binding site, and thus they are not transported as counteranions. These results with the acetate and propionate ions also suggested that NarK is a tightly-coupled antiporter, which does not transport substrates in the absence of counteranions (**Figure 4-7**).

## 6.4 Nitrate metabolic system in bacteria

Nitrate is used as a terminal electron acceptor for nitrate respiration in anaerobic conditions as described in **1.1**. Given that prokaryotic NarK-type transporters function as nitrate/nitrite antiporter under the physiological conditions, sufficient amounts of nitrite are necessary in cytoplasm in order to repeat the transport cycle continuously. The membrane-embedded nitrate reductase catalyzes the reduction reaction from nitrate to nitrite in cytoplasm. The imbalance of extracellular nitrate and intracellular nitrite leads to the lack of intracellular nitrate for reduction, or the accumulation of excess amount of cytotoxic nitrite in intracellular area, which are severe disadvantages for biological system using the nitrate reduction pathway. Although there is a possibility that NarK can use the other small anions including chloride, formate, and carbonate as the driving force to uptake nitrate suggested by the liposomal analyses (**Figure 4-2**) and these anions can be fit to the substrate binding site (**Figure 6-4**), the highly controlled system for keeping the balance of extracellular nitrate and intracellular nitrite is needed for maintaining the nitrogen homeostasis.

### 6.4.1 Nitrate metabolism in *Escherichia coli*

*Escherichia coli* is a Gram-negative, facultatively anaerobic bacteria that is commonly found in various environments including the mammalian gut. In the presence of oxygen, *E. coli* synthesizes ATP by aerobic respiration and acquires energy. In the absence of oxygen, however, they switch from aerobic respiration to anaerobic respiration. Nitrate is important as a major nitrogen source but also is utilized as a substrate for nitrate respiration, one of the most efficient anaerobic respiration.

Many transporters/channels and reductases are involved in the complex nitrogen metabolic system in *E. coli* (**Figure 6-7**). The nitrate taken up by NarK or NarU in NNP family is reduced in cytoplasm by nitrate reductases. *E. coli* possesses two homologous cytoplasmic membrane-bound quinol-nitrate oxidoreductases, NarGHI and NarZYW, both of which composed of three subunits: a molybdenum cofactor-containing Mo-bisMGD catalytic subunit (NarG or NarZ; approx. 1250 residues), an [Fe-S] cluster-containing electron transfer subunit (NarH or NarY; approx. 500 residues) and a hemecontaining membrane anchor subunit (NarI or NarW; approx. 220 residues)[89]. NarK and NarU are coded closely to the *narGHJI* and *narZYVW* genes, respectively, and form the operons. These transporters and reductases are expressed in different conditions. The expression of NarK and NarGHI are highly induced during anaerobic growth in the presence of nitrate, whereas during severe nitrogen source (nitrate) starvation or slow growth, the expression of NarK and NarGHI are repressed, and instead the contributions of constitutive housekeeping enzymes, NarU and NarZYW, become much more dominant[20]. This system is similar to the nitrate-inducible and constitutive nitrate transport systems (iLATS/cLATS and iHATS/cHATS) in plants[90].

The toxic product nitrite is exported again by NNP transporters in a manner of nitrate/nitrite antiport as discussed in **Chapter 4**, or further reduced by the cytoplasmic nitrite reductase NirBD to ammonium ion, which is required for nitrogen assimilation. This nitrite reductase is coded closely to another transport protein, NirC, and composing the *nirBCD* operon. NirC functions as a bi-directional proton-coupled nitrite channel[91]. Nitrite can be re-taken up or excreted by NirC, which supplies a driving force for NNP family transporters or facilitates the cytoplasmic detoxification.

Another nitrite reduction system is existing in periplasmic side. The periplasmic nitrate and nitrite reductases NapAB and NrfA reduce the nitrate and nitrite in periplasm, respectively. The resulted ammonium ion is imported by ammonium transporter AmtB and immediately used for biosynthesis of biomolecules. This pathway is mainly used in the absence of abundant nitrate in the environment.

In this way, nitrogen metabolism in *E. coli* has been achieved by various factors working coordinately.

### 6.4.2 Nitrate metabolism in *Pseudomonas aeruginosa* PAO1

During the oxygen respiration, electrons are transported from donor molecules such as NADH via membrane-localized multiprotein complexes along a redox cascade to the electron acceptors. The free energy generated is recovered by the formation of $H^+$ or $Na^+$ gradient at the membrane, ultimately driving the activity of ATP synthases. In the absence of oxygen, many microorganisms can replace oxygen by an alternative electron acceptor such as nitrate. Denitrifying bacteria accomplish the whole process of denitrification. One of the most studied denitrifying bacteria *Pseudomonas aeruginosa* form biofilms in the environment and during infection, and their growth are dependent on anaerobic respiration via denitrification. During the typical denitrification process, nitrate (Nar), nitrite (Nir), nitric oxide (Nor), and nitrous oxide (Nos) reductases catalyze the reduction reactions from $NO_3^-$, via $NO_2^-$, NO, and $N_2O$, finally to $N_2$. After nitrate is taken up, nitrate reductase catalyzes the following reaction: $NO_3^- + 2e^- \rightarrow 2H^+ + NO_2^-$

+ $H_2O$. This reaction is occurred in cytoplasm but all the following reduction reactions from $NO_2^-$ to $N_2$ are conducted in periplasmic side. Given that the NNP family transporter in denitrifying bacteria functions as a nitrate/nitrite antiporter, it is a reasonable system for both the uptake of nitrate and the extrusion of cytotoxic nitrite produced by nitrate reductase. The variations of NNP family in prokaryotes are discussed in **6.5**.

### 6.4.3 Denitrification supercomplex

The genes in operons are transcribed onto the same polycistronic mRNA and regulated strictly such that they are all turned on or off together. Recent studies indicated that the protein assembly is efficiently accomplished by gene organization into operons[92]. Recently, the interactions between proteins involving the denitrification pathway in *P. aeruginosa* PAO1 were reported and the model of denitrification supercomplex was proposed[93]. Nitrate transporter-reductase operons are also seen in *P. aeruginosa* PAO1, in which two *nark*-like genes *narK1* and *narK2* are tandemly coded located upstream of the structural genes of nitrate reductase and chaperon protein (*narGHJI*), forming the *narK1K2GHJI* operon. Given that the prokaryotic NNP transporter functions as a nitrate/nitrite antiporter as well as NarK in *Escherichia coli*, the assembly of nitrate reductase and nitrate transporter coded in a same operon seems to be a reasonable system for preventing the accumulation of excess amount of cytotoxic nitrite in cytoplasm, and also for generating the driving force for nitrate/nitrite antiporters to repeat the transport cycle continuously. Although the structure of membrane-bound nitrate reductase NarGHI has already been reported[89], there is no structural evidence for such a transporter-reductase complex. Further functional and structural analyses including the single particle analysis using the cryo-electron microscopicy[94–96] should be needed.

## 6.5 Variation of NNP family in prokaryotes

NNP family proteins are widely conserved in prokaryotes. There are two types of *narK*-like genes (*narK1* and *narK2*). All the prokaryotic species possess the NarK2 subfamily proteins, while the NarK1 subfamily is conserved only in part of prokaryotes including *Pseudomonas aeruginosa*, *Paracoccus denitrificans*, *Mycobacterium tuberculosis*, *Thermus thermophilus*. The sequence similarity between NarK1 and NarK2 from *P. aeruginosa* PAO1 is only 28% (**Figure 6-9**). The sequence similarity of two NNP transporters of *Escherichia coli* (NarK and NarU) to NarK1 and NarK2 from *P. aeruginosa* PAO1 are approximately 25% and 75%, respectively (**Figure 6-9**), suggesting that both NarK and NarU belong to NarK2 subfamily.

In **Chapter 3** and **Chapter 4**, structural and functional analyses revealed that the transport mechanism of NarK from *Escherichia coli* in NarK2 subfamily is nitrate/nitrite antiport. The substrate-recognition residues are almost completely conserved between NarK1 and NarK2 subfamily (**Figure 6-9**), however the function of NarK1 was suggested to be a $NO_3^-/H^+$ symporter[97,98]. The homology model of NarK1 and NarK2 from *P. aeruginosa* PAO1 based on the structure of NarK suggested that the nitrate recognition mechanism is quite similar to that of NarK[27] and NarU[26]. The putative substrate-translocation pathways of the models of both NarK1 and NarK2 lack protonatable residues as seen in NarK[27] and NarU[26], suggesting that protons are unlikely to be co-transported. The remaining possible transport mechanisms for NarK1 and NarK2 are (I) $NO_3^-$ uniport, (II) $NO_2^-$ uniport, and (III) $NO_3^-/NO_2^-$ antiport. Whether $NO_3^-/NO_2^-$ antiport mechanism shown in **Chapter 3** and **Chapter 4** can be applied to all the NarK-like transporters in prokaryotes or not is still unclear and further studies are necessary.

## 6.6 Summary of this study

In this study, to elucidate the dynamic mechanism of the substrate transport by NarK, the crystal structures of NarK from *Escherichia coli* were determined in three different states in the transport cycle (nitrate-bound occluded, nitrate-bound inward-open and apo inward-open states) at 2.35–2.4 Å. These high-resolution structures represent the occluded and inward-open conformations, in contrast to the previously reported partially occluded and partially inward-open conformations. Furthermore, a novel nitrite influx assay system to detect the NarK activity *in vitro* using NarK-reconstituted proteo-liposomes was developed and the results demonstrated that NarK is a tightly-coupled nitrate/nitrite antiporter. The combination of X-ray crystallography, *in vitro* nitrite influx assay, genetic analyses and an *in silico* molecular dynamics simulation enabled us to reveal the novel mechanism of substrate transport by NarK, which clearly explains how the transport of nitrate is tightly coupled with that of nitrite: the coupling between the substrate recognition and the opening/closing of the transport pathway is accomplished by the orchestrated movement of the three TM helices and the key residues Tyr263 and Arg305 in the substrate-binding site. These results will facilitate the precise understanding of the general molecular mechanisms of MFS secondary transporters, in addition to the nitrate/nitrite exchange mechanism of the NNP family.

これ以降の内容は、５年以内に雑誌等で刊行予定のため、非公開。

## Figures and Tables of Chapter 6

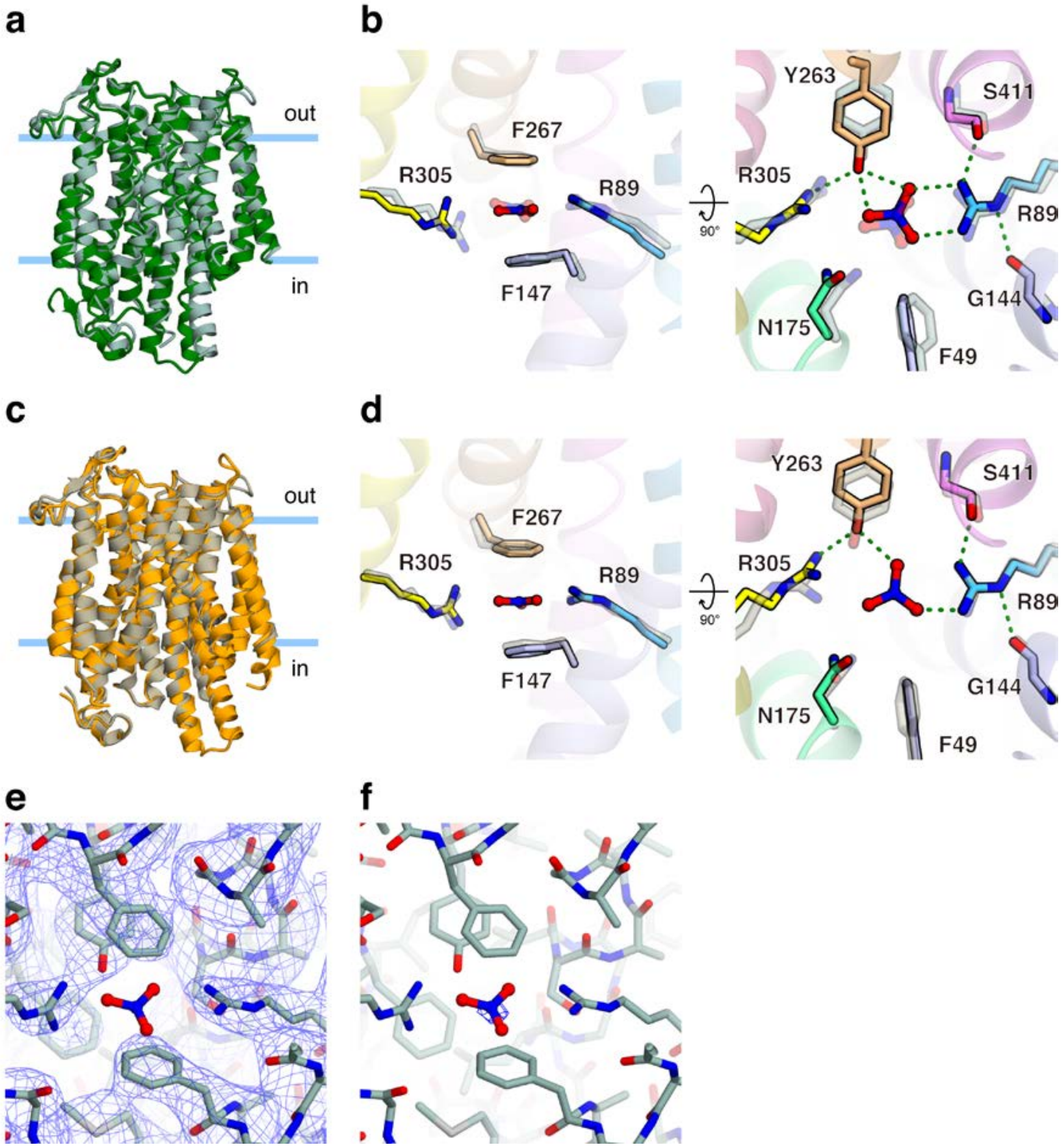


**Figure 6-1 Structural comparison with NarU.**

(**a**), (**b**) Structural comparison of NarK in the nitrate-bound occluded state (PDB ID: 4U4W, MolA, present crystal structure) colored in green with NarU in the nitrate-bound occluded state (PDB ID: 4IU8, MolA) colored in greenish grey. The two structures were superimposed based on the Cα atoms of the N bundle residues. The superposed overall structures are shown in (a). Close-up views of the substrate-binding site from two different directions are shown in (b). The residues of substrate binding site of NarU are depicted in semi-transparent color. (**c**), (**d**) Structural comparison of NarK in the nitrate-bound inward-open state (PDB ID: 4U4T, present crystal structure) colored in orange with NarU in the *apo* partially-inward-facing state (PDB ID: 4IU8, MolB) colored in yellowish grey. The superposed overall structures are shown in (c). Close-up views of the substrate-binding site from two different directions are shown in (d). The residues of substrate binding site of NarU are depicted in semi-transparent color. (**e**), (**f**) Electron density maps of NarU. The 2m$F_o$–D$F_c$ electron map contoured at 1.1 σ around the substrate binding site of NarU in the nitrate-bound occluded state is shown in (e). The m$F_o$–D$F_c$ omit map contoured at 3.0 σ around the substrate-binding site of NarU in the nitrate-bound inward-open state is shown in (f). The nitrate ion was omitted in the map calculation.

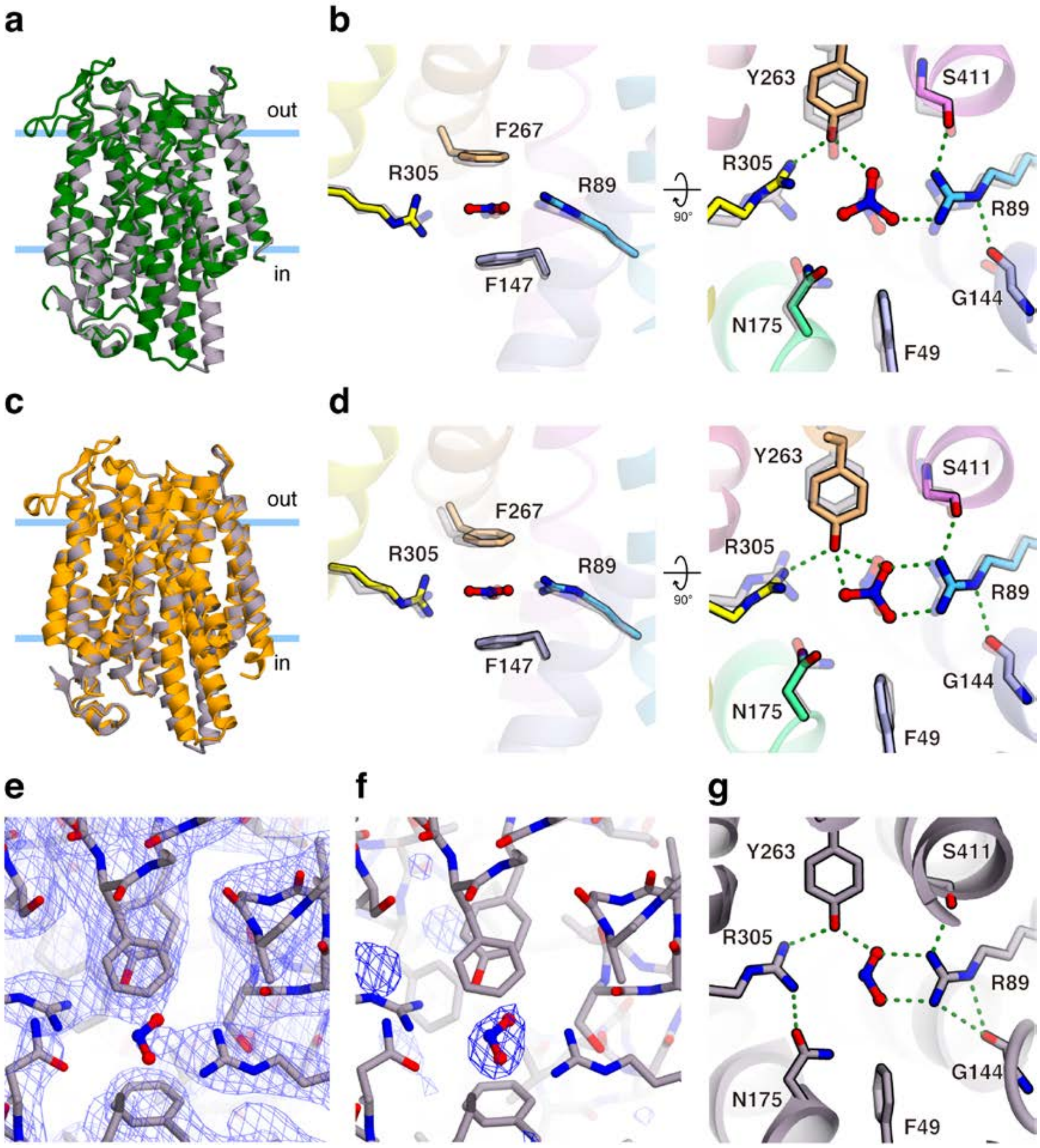


**Figure 6-2** **Structural comparison with previously reported NarK.**

(**a**), (**b**) Structural comparison of NarK in the nitrate-bound occluded state (PDB ID: 4U4W, MolA, present crystal structure) colored in green with NarK in the nitrite-bound partially inward-facing state (PDB ID: 4JRE) colored in reddish grey. The two structures were superimposed based on the Cα atoms of the N bundle residues. The superposed overall structures are shown in (a). Close-up views of the substrate-binding site from two different directions are shown in (b). The residues of substrate binding site of nitrite-bound NarK (PDB ID: 4JRE) are depicted in semi-transparent color. (**c**), (**d**) Structural comparison of NarK in the nitrate-bound inward-open state (PDB ID: 4U4T, present crystal structure) colored in orange with NarU in the *apo* partiall-inward-facing state (PDB ID: 4IU8, MolB) colored in reddish grey. The superposed overall structures are shown in (c). Close-up views of the substrate-binding site from two different directions are shown in (d). The residues of substrate binding site of nitrite-bound NarK (PDB ID: 4JRE) are depicted in semi-transparent color. (**e**), (**f**) Electron density maps of NarU. The 2m$F_o$–D$F_c$ electron map contoured at 1.1 σ around the substrate binding site of nitrite-bound NarK (PDB ID: 4JRE) is shown in (e). The m$F_o$–D$F_c$ omit map contoured at 3.0 σ around the substrate-binding site of nitrite-bound NarK (PDB ID: 4JRE) is shown in (f). The nitrite ion was omitted in the map calculation. (**g**) Close-up view of the nitrite-binding site in previously reported NarK (PDB ID: 4JRE).

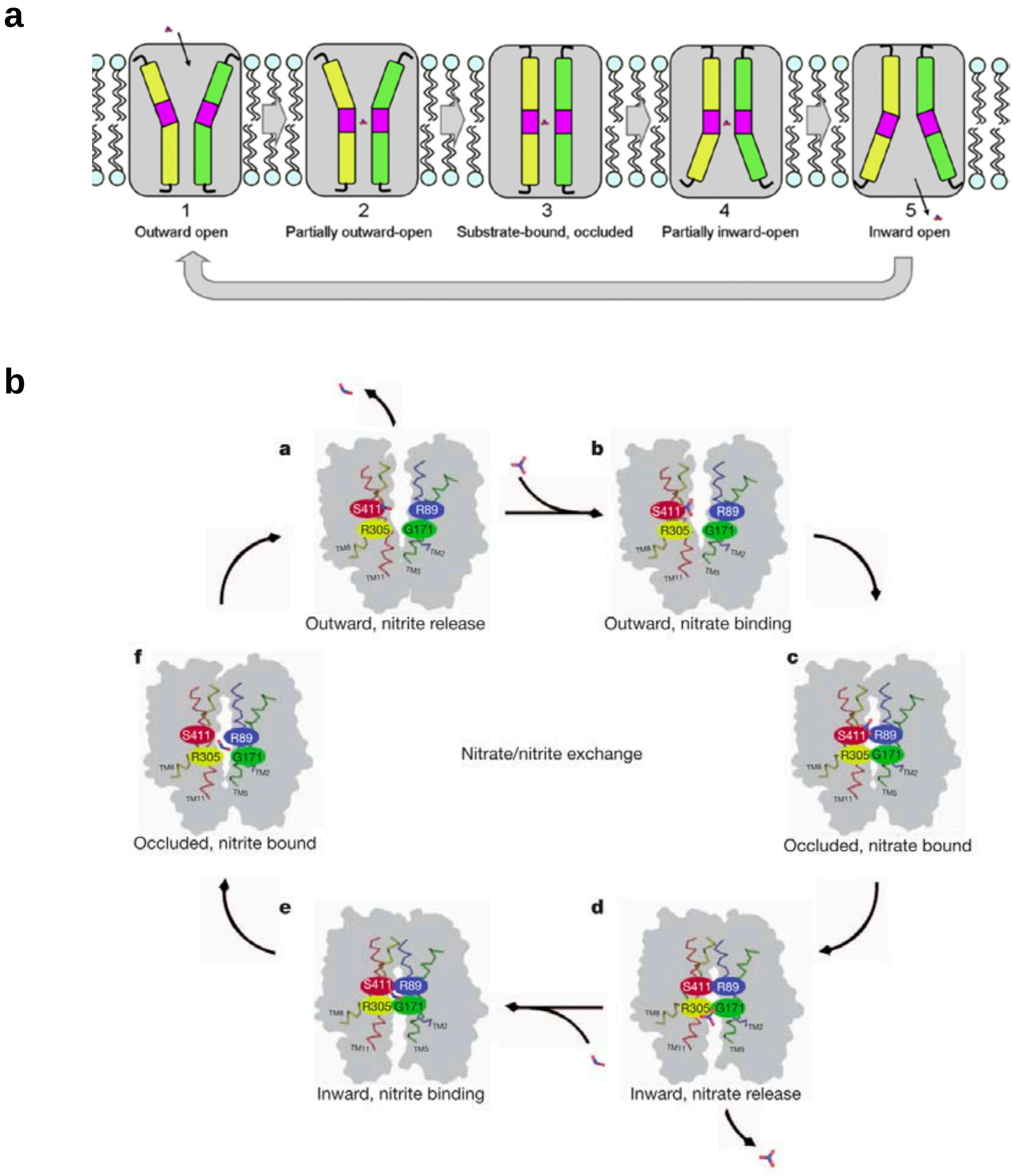


**Figure 6-3 Previously proposed transport models of prokaryotic NNP family transporters.**

(**a**) The transport model of NarU proposed by Yan *et al.* (***Cell Rep.***, 2013). The figure was cited from "Yan, H. *et al.* (2013) 'Structure and mechanism of a nitrate transporter.', ***Cell reports***. **3**(3), pp. 716–23. doi: 10.1016/j.celrep.2013.03.007." (**b**) The rocker switch transport model of NarK proposed by Zheng *et al.* (***Nature***, 2013) The figure was cited from "Zheng, H., Wisedchaisri, G. and Gonen, T. (2013) 'Crystal structure of a nitrate/nitrite exchanger', ***Nature***, **497**(7451), pp. 647–651. doi: 10.1038/nature12139."

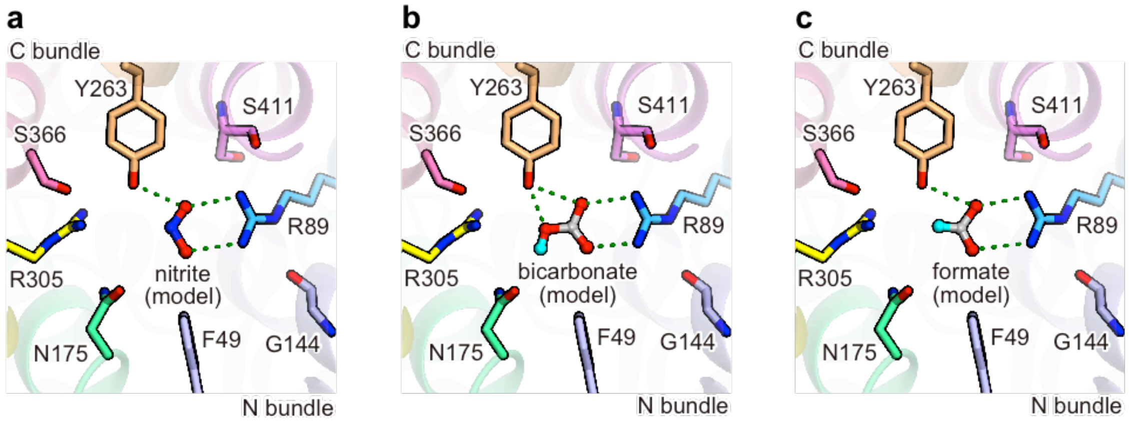


**Figure 6-4** **Recognition models of nitrite and other anions by NarK.**

(**a**), (**b**) and (**c**) The recognition models of (a) nitrite, (b) bicarbonate, and (c) formate by NarK, based on the present high-resolution nitrite-bound structures. The possible hydrogen bonds between nitrite and the R89 and Y263 side-chains are indicated by green dashed lines.

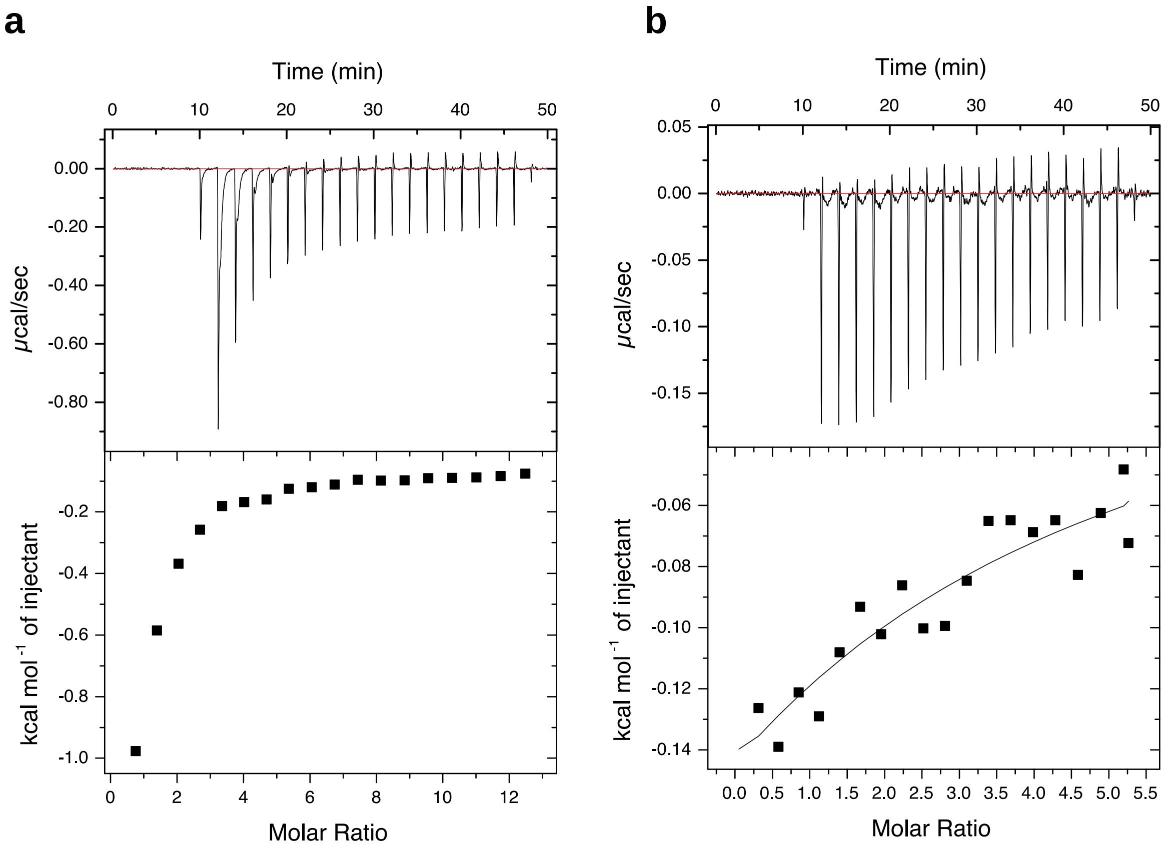
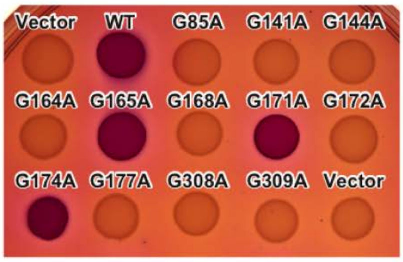


**Figure 6-5 Isothermal titration calorimetry analysis of NarK.**

(**a**), (**b**) The results of Isothermal Titration Calorimetry (ITC) on NarK titrated with (a) 10 mM $NaNO_3$ and (b) 10 mM $NaNO_2$. The estimated $K_d$ values were 162 μM for nitrate and 2-3 mM for nitrite. The protein concentration was 0.045 mM (2.25 mg/mL). The measurements were conducted at 4˚C and the buffer composition of the protein sample was 150 mM NaCl, 20 mM Tris-HCl pH 8.0, 0.05% DDM, 0.01% CHS, and 4 mM β-ME.



| | TM | Nitrate uptake |
|---|---|---|
| Vector | – | – |
| WT | – | ++ |
| G85A | TM2 | – |
| G141A | TM4 | – |
| G144A | TM4 | – |
| G164A | TM5 | – |
| G165A | TM5 | ++ |
| G168A | TM5 | – |
| G171A | TM5 | ++ |
| G172A | TM5 | – |
| G174A | TM5 | ++ |
| G177A | TM5 | – |
| G308A | TM8 | – |
| G309A | TM8 | – |

**Figure 6-6 Genetic analysis of Glycine mutants in TM2, TM4, TM5, and TM8.**

Genetic analysis of the nitrate transport activities of NarK mutants of the conserved Gly residues in TM2, TM4, TM5, and TM8 by the "no-induction system". The results are summarized in the right panel.

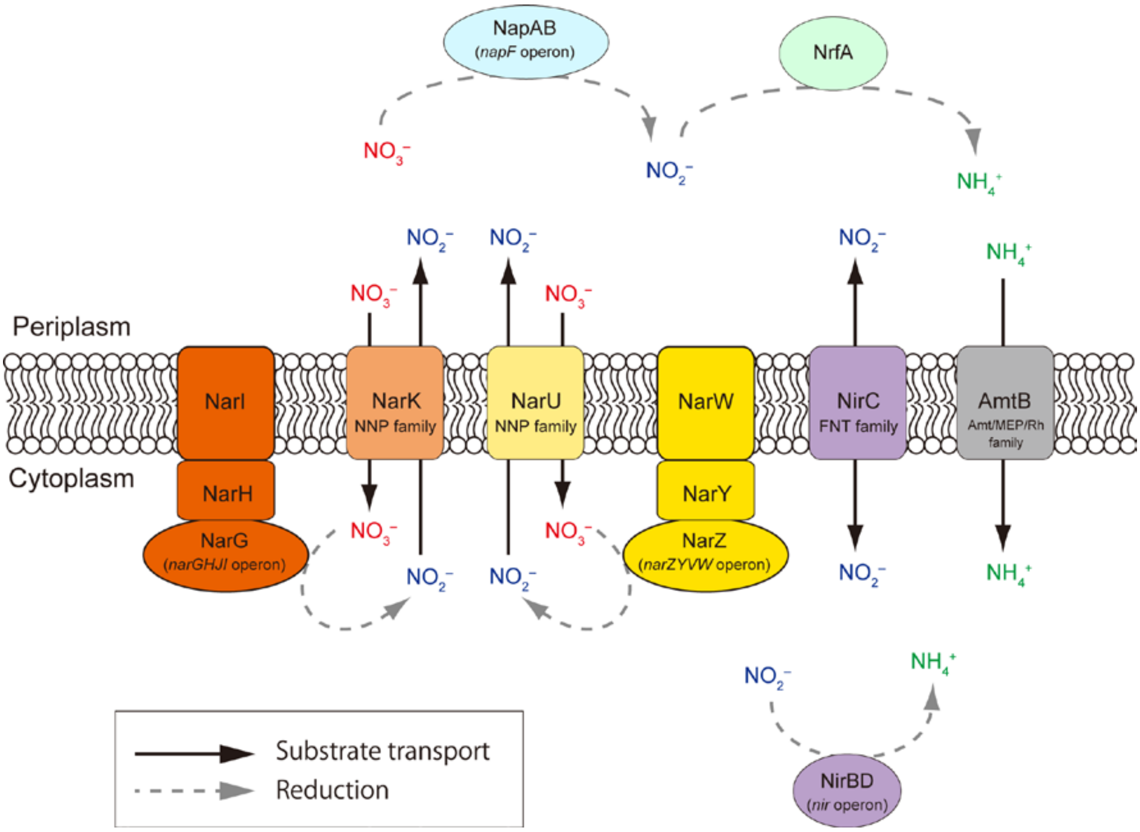


**Figure 6-7** **Nitrogen metabolic pathway in *Escherichia coli.***

Nitrate transporters NarK and NarU in NNP family, membrane-bound nitrate reductases NarGHI and NarZYW, a nitrite channel NirC in FNT family, a NADH-dependent cytoplasmic nitrite reductase NirBD, an ammonium transporter AmtB in Amt/MEP/Rh family, a periplasmic nitrate reductase NapAB, and a periplasmic nitrite reductase NrfA are shown. The proteins coded in a same operon are colored in the same colors. The black solid arrows indicate the substrate transport. The gray dashed arrows indicate the reduction reaction.

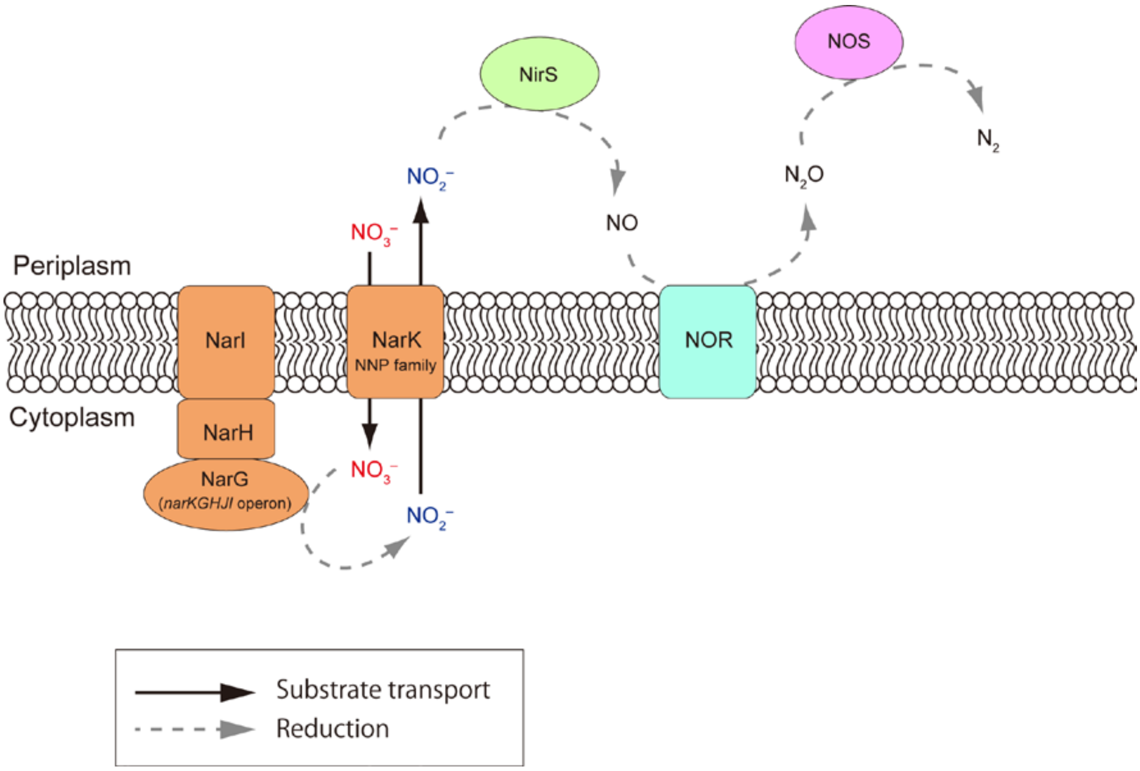


**Figure 6-8 Denitrification pathway in *Pseudomonas aeruginosa* PAO1.**

A nitrate transporter NarK in NNP family, a membrane-bound nitrate reductase NarGHI, periplasmic nitrite reductase NirS, a membrane-bound nitric oxide reductase NOR, a periplasmic nitrous oxide reductase NOS are shown. The proteins coded in a same operon are colored in the same colors. The black solid arrows indicate the substrate transport. The gray dashed arrows indicate the reduction reaction.

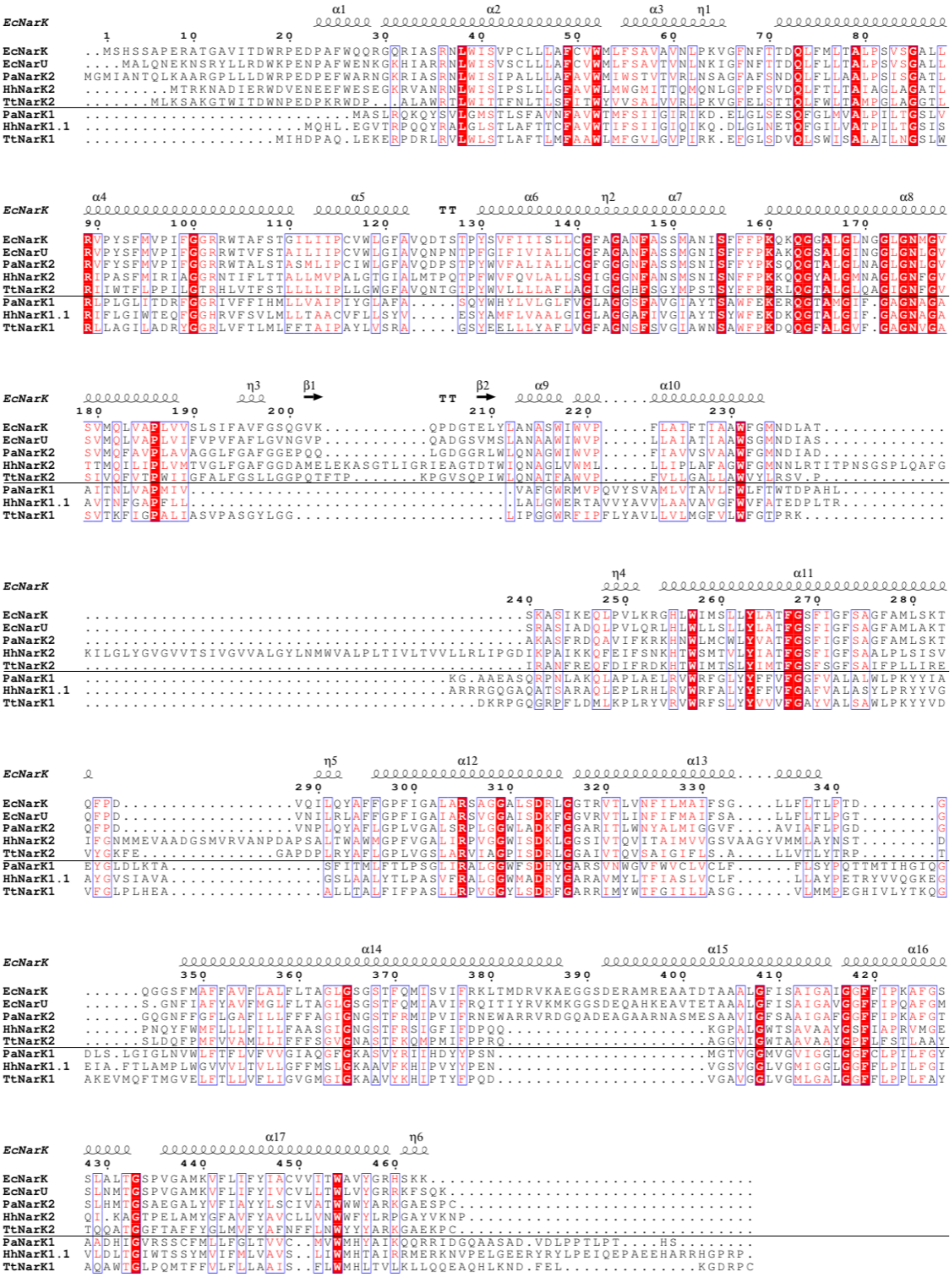


**Figure 6-9 Amino acid sequence alignments between NarK1 and NarK2 subfamily in prokaryotes.**

Sequence alignments of NarK1 and NarK2 proteins from prokaryotes calculated by using the program ClustalO. The secondary structure and sequence of NarK (PDB accession code: 4U4T) are also shown.

# References

1. Stewart, V. Nitrate respiration in relation to facultative metabolism in enterobacteria. *Microbiol. Rev.* **52,** 190–232 (1988).
2. Bonete, M., Martínez-Espinosa, R., Pire, C., Zafrilla, B. & Richardson, D. J. Nitrogen metabolism in haloarchaea. *Saline Systems* (2008). doi:10.1186/1746-1448-4-9
3. Lundberg, J. O., Weitzberg, E., Cole, J. a & Benjamin, N. Nitrate, bacteria and human health. *Nat. Rev. Microbiol.* **2,** 593–602 (2004).
4. Martínez-Espinosa, R. M., Cole, J. a, Richardson, D. J. & Watmough, N. J. Enzymology and ecology of the nitrogen cycle. *Biochem. Soc. Trans.* **39,** 175–8 (2011).
5. Dixon, R. & Kahn, D. Genetic regulation of biological nitrogen fixation. *Nat. Rev. Microbiol.* **2,** 621–31 (2004).
6. Nacry, P., Bouguyon, E. & Gojon, A. Nitrogen acquisition by roots: physiological and developmental mechanisms ensuring plant adaptation to a fluctuating resource. *Plant Soil* **370,** 1–29 (2013).
7. Omata, T. Structure, function and regulation of the nitrate transport system of the cyanobacterium Synechococcus sp. PCC7942. *Plant Cell Physiol.* **36,** 207–13 (1995).
8. Tsay, Y.-F., Chiu, C.-C., Tsai, C.-B., Ho, C.-H. & Hsu, P.-K. Nitrate transporters and peptide transporters. *FEBS Lett.* **581,** 2290–300 (2007).
9. Léran, S. *et al.* A unified nomenclature of NITRATE TRANSPORTER 1/PEPTIDE TRANSPORTER family members in plants. *Trends Plant Sci.* **19,** 5–9 (2014).
10. Doki, S. *et al.* Structural basis for dynamic mechanism of proton-coupled symport by the peptide transporter POT. *Proc. Natl. Acad. Sci. U. S. A.* **110,** 11343–8 (2013).
11. Newstead, S. *et al.* Crystal structure of a prokaryotic homologue of the mammalian oligopeptide-proton symporters, PepT1 and PepT2. *EMBO J.* **30,** 417–26 (2011).
12. Parker, J. L. & Newstead, S. Molecular basis of nitrate uptake by the plant nitrate transporter NRT1.1. *Nature* **507,** 68–72 (2014).
13. Sun, J. *et al.* Crystal structure of the plant dual-affinity nitrate transporter NRT1.1. *Nature* **507,** 73–7 (2014).
14. Guettou, F. *et al.* Structural insights into substrate recognition in proton-dependent oligopeptide transporters. *EMBO Rep.* **14,** 804–10 (2013).

15. Lyons, J. a *et al.* Structural basis for polyspecificity in the POT family of proton-coupled oligopeptide transporters. *EMBO Rep.* **15,** 886–93 (2014).

16. Guettou, F. *et al.* Selectivity mechanism of a bacterial homolog of the human drug-peptide transporters PepT1 and PepT2. *Nat. Struct. Mol. Biol.* **21,** 728–31 (2014).

17. MacGregor, B. J., Biddle, J. F., Harbort, C., Matthysse, A. G. & Teske, A. Sulfide oxidation, nitrate respiration, carbon acquisition, and electron transport pathways suggested by the draft genome of a single orange Guaymas Basin Beggiatoa (Cand. Maribeggiatoa) sp. filament. *Mar. Genomics* **11,** 53–65 (2013).

18. Pao, S. S., Paulsen, I. T. & Saier, M. H. Major facilitator superfamily. *Microbiol. Mol. Biol. Rev.* **62,** 1–34 (1998).

19. Rowe, J. J., Ubbink-Kok, T., Molenaar, D., Konings, W. N. & Driessen, a. J. M. NarK is a nitrite-extrusion system involved in anaerobic nitrate respiration by Escherichia coli. *Mol. Microbiol.* **12,** 579–586 (1994).

20. Clegg, S. J., Jia, W. & Cole, J. a. Role of the Escherichia coli nitrate transport protein, NarU, in survival during severe nutrient starvation and slow growth. *Microbiology* **152,** 2091–100 (2006).

21. Moir, J. W. & Wood, N. J. Nitrate and nitrite transport in bacteria. *Cell. Mol. Life Sci.* **58,** 215–224 (2001).

22. Noji, S., Nohno, T., Saito, T. & Taniguchi, S. The narK gene product participates in nitrate transport induced in Escherichia coli nitrate-respiring cells. *FEBS Lett.* **252,** 139–143 (1989).

23. DeMoss, J. A. & Hsu, P. Y. NarK enhances nitrate uptake and nitrite excretion in Escherichia coli. *J. Bacteriol.* **173,** 3303–10 (1991).

24. Clegg, S., Yu, F., Griffiths, L. & Cole, J. a. The roles of the polytopic membrane proteins NarK, NarU and NirC in Escherichia coli K-12: Two nitrate and three nitrite transporters. *Mol. Microbiol.* **44,** 143–155 (2002).

25. Jia, W., Tovell, N., Clegg, S., Trimmer, M. & Cole, J. A single channel for nitrate uptake, nitrite export and nitrite uptake by Escherichia coli NarU and a role for NirC in nitrite export and uptake. *Biochem. J.* **417,** 297–304 (2009).

26. Yan, H. *et al.* Structure and mechanism of a nitrate transporter. *Cell Rep.* **3,** 716–23 (2013).

27. Zheng, H., Wisedchaisri, G. & Gonen, T. Crystal structure of a nitrate/nitrite exchanger. *Nature* **497,** 647–51 (2013).

28. Radestock, S. & Forrest, L. R. The alternating-access mechanism of MFS transporters arises from inverted-topology repeats. *J. Mol. Biol.* **407,** 698–715 (2011).

29. Abramson, J. *et al.* Structure and mechanism of the lactose permease of Escherichia coli. *Science* **301,** 610–5 (2003).

30. Huang, Y., Lemieux, M. J., Song, J., Auer, M. & Wang, D.-N. Structure and mechanism of the glycerol-3-phosphate transporter from Escherichia coli. *Science* **301,** 616–20 (2003).

31. Yin, Y. Structure of the Multidrug Transporter EmrD from Escherichia coli. *Science (80-. ).* **312,** 741–744 (2006).

32. Dang, S. *et al.* Structure of a fucose transporter in an outward-open conformation. *Nature* **467,** 734–8 (2010).

33. Pedersen, B. P. *et al.* Crystal structure of a eukaryotic phosphate transporter. *Nature* **496,** 533–6 (2013).

34. Sun, L. *et al.* Crystal structure of a bacterial homologue of glucose transporters GLUT1-4. *Nature* **490,** 361–6 (2012).

35. Deng, D. *et al.* Crystal structure of the human glucose transporter GLUT1. *Nature* **510,** 121–5 (2014).

36. Nomura, N. *et al.* Structure and mechanism of the mammalian fructose transporter GLUT5. *Nature* 1–7 (2015). doi:10.1038/nature14909

37. Deng, D. *et al.* Molecular basis of ligand recognition and transport by glucose transporters. *Nature* (2015). doi:10.1038/nature14655

38. Quistgaard, E. M., Löw, C., Moberg, P., Trésaugues, L. & Nordlund, P. Structural basis for substrate transport in the GLUT-homology family of monosaccharide transporters. *Nat. Struct. Mol. Biol.* **20,** 766–8 (2013).

39. Wisedchaisri, G., Park, M.-S., Iadanza, M. G., Zheng, H. & Gonen, T. Proton-coupled sugar transport in the prototypical major facilitator superfamily protein XylE. *Nat. Commun.* **5,** 4521 (2014).

40. Taniguchi, R. *et al.* Outward- and inward-facing structures of a putative bacterial transition-metal transporter with homology to ferroportin. *Nat. Commun.* **6,** 8545 (2015).

41. Goehring, A. *et al.* Screening and large-scale expression of membrane proteins in mammalian cells for structural studies. *Nat. Protoc.* **9,** 2574–85 (2014).

42. Ward, J. J., Sodhi, J. S., McGuffin, L. J., Buxton, B. F. & Jones, D. T. Prediction and Functional Analysis of Native Disorder in Proteins from the Three Kingdoms of Life. *J. Mol. Biol.* **337,** 635–645 (2004).

43. Hirata, K. *et al.* Achievement of protein micro-crystallography at SPring-8 beamline BL32XU. *J.*

*Phys. Conf. Ser.* **425,** 12002 (2013).

44. Kabsch, W. XDS. *Acta Crystallogr. D. Biol. Crystallogr.* **66,** 125–132 (2010).

45. Waterman, D. G. *et al.* The DIALS framework for integration software. *Ccp4 Newsl. Protein Crystallogr.* 16–19 (2013).

46. Evans, P. R. & Murshudov, G. N. How good are my data and what is the resolution? *Acta Crystallogr. Sect. D Biol. Crystallogr.* **69,** 1204–1214 (2013).

47. Schneider, T. R. & Sheldrick, G. M. Substructure solution with SHELXD. *Acta Crystallogr. D. Biol. Crystallogr.* **58,** 1772–1779 (2002).

48. Sheldrick, G. M. A short history of SHELX. *Acta Crystallogr. A.* **64,** 112–122 (2008).

49. Fortelle, E. D. La, Irwin, J. & de la Fortelle, E. SHARP : A Maximum-Likelihood Heavy-Atom Parameter Refinement Program for the MIR and MAD Methods. *Methods Enzym.* **276,** 472–494 (1997).

50. Abrahams, J. P. & Leslie, A. G. Methods used in the structure determination of bovine mitochondrial F1 ATPase. *Acta Crystallogr. D. Biol. Crystallogr.* **52,** 30–42 (1996).

51. Cowtan, K. D. & Main, P. Improvement of macromolecular electron-density maps by the simultaneous application of real and reciprocal space constraints. *Acta Crystallogr. D. Biol. Crystallogr.* **49,** 148–57 (1993).

52. Kelley, L. a & Sternberg, M. J. E. Protein structure prediction on the Web: a case study using the Phyre server. *Nat. Protoc.* **4,** 363–71 (2009).

53. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D. Biol. Crystallogr.* **60,** 2126–2132 (2004).

54. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D. Biol. Crystallogr.* **66,** 486–501 (2010).

55. Adams, P. D. *et al.* PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallogr. D. Biol. Crystallogr.* **58,** 1948–1954 (2002).

56. Adams, P. D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D. Biol. Crystallogr.* **66,** 213–221 (2010).

57. McCoy, A. J. *et al.* Phaser crystallographic software. *J. Appl. Crystallogr.* **40,** 658–674 (2007).

58. Kawate, T. & Gouaux, E. Fluorescence-Detection Size-Exclusion Chromatography for Precrystallization Screening of Integral Membrane Proteins. *Structure* **14,** 673–681 (2006).

59. Dolinsky, T. J. *et al.* PDB2PQR: Expanding and upgrading automated preparation of biomolecular structures for molecular simulations. *Nucleic Acids Res.* **35,** 522–525 (2007).

60. Unni, S. *et al.* Web Servers and Services for Electrostatics Calculations with APBS and PDB2PQR. *J. Comput. Chem.* **32,** 1488–1491 (2011).

61. Thompson, J. D., Higgins, D. G. & Gibson, T. J. CLUSTAL W: Improving the Sensitivity of Progressive Multiple Sequence Alignment Through Sequence Weighting, Position-Specific Gap Penalties and Weight Matrix Choice. *Nucleic Acids Res.* **22,** 4673–4680 (1994).

62. Robert, X. & Gouet, P. Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res.* **42,** 320–324 (2014).

63. Landau, E. M. & Rosenbusch, J. P. Lipidic cubic phases: a novel concept for the crystallization of membrane proteins. *Proc. Natl. Acad. Sci. U. S. A.* **93,** 14532–5 (1996).

64. Kumazaki, K. *et al.* Structural basis of Sec-independent membrane protein insertion by YidC. *Nature* **509,** 516–20 (2014).

65. Kumazaki, K. *et al.* Crystal structure of Escherichia coli YidC, a membrane protein chaperone and insertase. *Sci. Rep.* **4,** 7299 (2014).

66. Nishizawa, T. *et al.* Structural basis for the counter-transport mechanism of a H+/Ca2+ exchanger. *Science* **341,** 168–72 (2013).

67. Kato, H. E. *et al.* Crystal structure of the channelrhodopsin light-gated cation channel. *Nature* **482,** 369–74 (2012).

68. Taniguchi, R. *et al.* Structural insights into ligand recognition by the lysophosphatidic acid receptor LPA6. *Nature* **548,** 356–360 (2017).

69. Lee, Y. *et al.* Structure of the triose-phosphate/phosphate translocator reveals the basis of substrate specificity. *Nat. Plants* **3,** 825–832 (2017).

70. Shihoya, W. *et al.* Activation mechanism of endothelin ETB receptor by endothelin-1. *Nature* 1–21 (2016). doi:10.1038/nature19319

71. Takeda, H. *et al.* Structural basis for ion selectivity revealed by high-resolution crystal structure of Mg2+ channel MgtE. *Nat. Commun.* **5,** (2015).

72. Miyauchi, H. *et al.* Structural basis for xenobiotic extrusion by eukaryotic MATE transporter. *Nat. Commun.* **8,** (2017).

73. Tsuchiya, H. *et al.* Structural basis for amino acid export by DMT superfamily transporter YddG. *Nature* **534,** 417–420 (2016).

74. Trueman, L. J., Richardson, A. & Forde, B. G. Molecular cloning of higher plant homologues of the high-affinity nitrate transporters of Chlamydomonas reinhardtii and Aspergillus nidulans. *Gene* **175,** 223–31 (1996).

75. Unkles, S. E. *et al.* Two perfectly conserved arginine residues are required for substrate binding in a high-affinity nitrate transporter. *Proc. Natl. Acad. Sci. U. S. A.* **101,** 17549–17554 (2004).

76. Fischer, K. *et al.* Structural basis of eukaryotic nitrate reduction: crystal structures of the nitrate reductase active site. *Plant Cell* **17,** 1167–79 (2005).

77. Kojima, H. *et al.* Detection and imaging of nitric oxide with novel fluorescent indicators: diaminofluoresceins. *Anal. Chem.* **70,** 2446–53 (1998).

78. Datsenko, K. a & Wanner, B. L. One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. *Proc. Natl. Acad. Sci. U. S. A.* **97,** 6640–5 (2000).

79. Baba, T. *et al.* Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. *Mol. Syst. Biol.* **2,** 2006.0008 (2006).

80. Studier, F. W., Rosenberg, a H., Dunn, J. J. & Dubendorff, J. W. Use of T7 RNA polymerase to direct expression of cloned genes. *Methods Enzymol.* **185,** 60–89 (1990).

81. Stewart, V. & MacGregor, C. H. Nitrate reductase in Escherichia coli K-12: involvement of chlC, chlE, and chlG loci. *J. Bacteriol.* **151,** 788–99 (1982).

82. Sali, A., Potterton, L., Yuan, F., van Vlijmen, H. & Karplus, M. Evaluation of comparative protein modeling by MODELLER. *Proteins* **23,** 318–26 (1995).

83. Humphrey, W., Dalke, A. & Schulten, K. VMD: visual molecular dynamics. *J. Mol. Graph.* **14,** 33–8, 27–8 (1996).

84. Klauda, J. B. *et al.* Update of the CHARMM all-atom additive force field for lipids: validation on six lipid types. *J. Phys. Chem. B* **114,** 7830–43 (2010).

85. Phillips, J. C. *et al.* Scalable molecular dynamics with NAMD. *J. Comput. Chem.* **26,** 1781–802 (2005).

86. Darden, T., York, D. & Pedersen, L. Particle mesh Ewald: An N⋅ log(N) method for Ewald sums in large systems. *J. Chem. Phys.* **98,** 10089 (1993).

87. Unkles, S. E. *et al.* Alanine scanning mutagenesis of a high-affinity nitrate transporter highlights the requirement for glycine and asparagine residues in the two nitrate signature motifs. *Biochem. J.* **447,** 35–42 (2012).

88. Sun, L. *et al.* Crystal structure of a bacterial homologue of glucose transporters GLUT1-4. *Nature* **490,** 361–6 (2012).

89. Bertero, M. G. *et al.* Insights into the respiratory electron transfer pathway from the structure of nitrate reductase A. *Nat. Struct. Biol.* **10,** 681–7 (2003).

90. Wang, Y.-Y., Hsu, P.-K. & Tsay, Y.-F. Uptake, allocation and signaling of nitrate. *Trends Plant*

*Sci.* **17,** 458–67 (2012).

91. Lu, W. *et al.* Structural and functional characterization of the nitrite channel NirC from Salmonella typhimurium. *Proc. Natl. Acad. Sci.* **109,** 18395–18400 (2012).

92. Shieh, Y. W. *et al.* Operon structure and cotranslational subunit association direct protein assembly in bacteria. *Science (80-. ).* **350,** 678–680 (2015).

93. Rohde, M. *et al.* Protein Network of the Pseudomonas aeruginosa Denitrification. *J. Bacteriol.* **198,** 1401–1413 (2016).

94. Cheng, Y. Single-Particle Cryo-EM at Crystallographic Resolution. *Cell* **161,** 450–457 (2015).

95. Henderson, R. Overview and future of single particle electron cryomicroscopy. *Arch. Biochem. Biophys.* **581,** 19–24 (2015).

96. Frank, J. Improving the resolution of asymmetric structures Advances in the field of single-particle cryo-electron microscopy over the last decade. *Nat. Protoc.* **12,** 209–212 (2017).

97. Sharma, V., Noriega, C. E. & Rowe, J. J. Involvement of NarK1 and NarK2 proteins in transport of nitrate and nitrite in the denitrifying bacterium Pseudomonas aeruginosa PAO1. *Appl. Environ. Microbiol.* **72,** 695–701 (2006).

98. Goddard, A. D., Moir, J. W. B., Richardson, D. J. & Ferguson, S. J. Interdependence of two NarK domains in a fused nitrate/nitrite transporter. *Mol. Microbiol.* **70,** 667–81 (2008).

# Original papers

1) Fukuda, M., Takeda, H., Kato, H. E., Doki, S., Ito, K., Maturana, A. D., Ishitani, R., Nureki, O. (2015). Structural basis for dynamic mechanism of nitrate/nitrite antiport by NarK. *Nature Communications*, *6*, 7097. Retrieved from http://dx.doi.org/10.1038/ncomms8097

# Acknowledgements

First, I would like to thank Professor Osamu Nureki for his supervision and encouragement during the bachelor's, master's and Ph.D. courses. Besides Prof. Nureki, I have been assisted by many members of the Nureki laboratory. In particular, I would like to thank Dr. Ryuichiro Ishitani for guidance and assistance with performing experiments and writing papers. I also would like to thank Dr. Tomohiro Nishizawa, Dr. Hiroshi Nishimasu, Dr. Takanori Nakane, Dr. Tomoya Tsukazaki, Dr. Motoyuki Hattori, and Dr. Ryohei Ishii for technical advice and many useful discussion, Ms. Rieko Yamazaki for secretarial assistance, and Ms. Arisa Kurabayashi, Ms. Hitomi Kawamura, Ms. Keiko Ogomori, Ms. Masae Miyazaki, and Ms. Sanae Okazaki for the powerful and lots of technical supports.

This study was performed in collaboration with many scientists. I would like to thank Dr. Koichi Ito (The University of Tokyo) for the collaboration with construction of the gene knockout *E. coli* strains and genetic analysis, Dr. Andrés D. Maturana (Nagoya University) for the experiments using fluorescent probes before the liposome-based nitrite-flux assay, Dr. Ryuichiro Ishitani (The University of Tokyo) for the molecular dynamics simulations, Dr. Hideaki E. Kato (Stanford University) for technical advice and experimental support as a supervisor, Dr. Hironori Takeda (Kyoto Sangyo University) for the *in vitro* assays using proteoliposomes, Dr. Shintaro Doki (The University of Tokyo) for purification of proteins for *in vitro* assays,

During the bachelor's, master's and Ph.D. courses, I enjoyed many exciting experiences with the members of the Nureki laboratory. I thank Dr. Akihisa Tsutsumi, Dr. Kayo Nozawa, Dr. Yoshiki Tanaka, Dr. Takashi Higuchi, Dr. Michio Koyama, Dr. Hideaki E. Kato, Dr. Kan Kobayashi, Dr. Megumi Kato, Dr. Hironori Takeda, Dr. Shintaro Doki, Dr. Kazuki Kato, Dr. Kaoru Kumazaki, Dr. Naoki Matsumoto, and Mr. Toshiki Kishimoto, Dr. Wataru Shihoya, Mr. Reiya Taniguchi, Mr. Mizuki Takemoto, Ms. Junko Morita, Mr. Hiroki Omura, Mr. Go Kasuya, Mr. Yongchan Lee, Ms. Mari Yamada, and Mr. Takashi Yamano for stimulating discussions. In particular, I would like to say "Thank you" for Dr. Tsukasa Kusakizako as an honorable and supportive labmate since the undergraduate course.

Finally, I would like to thank my family for continuous support throughout my primary education, the bachelor's, master's and Ph.D. courses, and my life.

Masahiro Fukuda