



University of  
Massachusetts  
Amherst

## THE KEY QUESTION IN SYMBIOTIC NITROGEN FIXATION: HOW DOES HOST MAINTAIN A BACTERIAL SYMBIONT?

Item Type	Dissertation (Open Access)
Authors	Oztas, Onur
DOI	<a href="https://doi.org/10.7275/9588108.0">10.7275/9588108.0</a>
Download date	2026-06-12 23:27:55
Link to Item	<a href="https://hdl.handle.net/20.500.14394/20214">https://hdl.handle.net/20.500.14394/20214</a>

**THE KEY QUESTION IN SYMBIOTIC NITROGEN FIXATION: HOW DOES HOST MAINTAIN A  
BACTERIAL SYMBIONT?**

A Dissertation Presented

by

ONUR OZTAS

Submitted to the Graduate School of the  
University of Massachusetts Amherst in partial fulfillment  
of the requirements for the degree of

DOCTOR OF PHILOSOPHY

MAY 2017

Molecular and Cellular Biology Program

© Copyright by Onur Oztas 2017

All Rights Reserved

**THE KEY QUESTION IN SYMBIOTIC NITROGEN FIXATION: HOW DOES HOST MAINTAIN A BACTERIAL SYMBIONT?**

A Dissertation Presented

by

ONUR OZTAS

Approved as to style and content by:

---

Dong Wang, Chair

---

Graham Walker, Member

---

Alice Cheung, Member

---

Elizabeth Vierling, Member

---

Dominique Alfandari, director  
Molecular and Cellular Biology

DEDICATION

*To my wife, my mother, my father, and  
my sister*

*Dedicated to the memories of my grandparents*

*Mehmet Emin Gul*

*Naciye Gul*

*Ibrahim Oztas*

*Hacer Oztas*

## ACKNOWLEDGMENTS

I would like to thank my advisor Dr. Dong Wang for his mentoring and support during my PhD education; my dissertation committee: Dr. Graham Walker, Dr. Alice Cheung and Dr. Elizabeth Vierling for their guidance, advice and support; Dr. Anne Gershenson, Dr. Tao He, and Dr. Mary Roberts for their contributions in the DNF2 project; Dr. Huairong Pan and Christina Stonoha for their contributions in the Syntaxin project; Derrick Deming, Bryan Teefy and Ben Roberts for their contributions in the purification of the DNF2 protein; Johanna L'Heureux for her contribution in the ENOD project. I also would like to thank the Molecular and Cellular Biology Program especially to Dr. Barbara Osborne and Sarah Czerwonka, and the Department of Biochemistry and Molecular Biology.

I have special thanks to my wife Esra. She is my love, my best friend, my mentor... She is everything in my life. I felt her love and support in each step of my PhD education. She is responsible for my success. Thank you for making my life so wonderful.

I have many thanks to my family: my mother Zeliha, my father Bayram and my sister Ozge. I could not achieve this without them. It has been very hard to be far away from them, but I felt their unconditional support in every step of my education. And, I have many thanks to my new family Ebru and Ozge Esenlik for their love and support.

I would like to thank Mine Canakci for her friendship, help and support every time I needed, and Furkan Ayaz for his companionship, support and encouragement.

## ABSTRACT

THE KEY QUESTION IN SYMBIOTIC NITROGEN FIXATION: HOW DOES HOST MAINTAIN A BACTERIAL SYMBIONT?

MAY 2017

ONUR OZTAS, B.S., ISTANBUL TECHNICAL UNIVERSITY

M.S., KOC UNIVERSITY

Ph.D., UNIVERSITY OF MASSACHUSETTS AMHERST

Directed by: Professor Dong Wang

The fact that plants cannot use nitrogen in the gaseous form makes them dependent on the levels of usable nitrogen forms in the soil. Legumes overcome nitrogen limitation by entering a symbiotic association with rhizobia, soil bacteria that convert atmospheric nitrogen into usable ammonia. In root nodules, bacteria are internalized by host plant cells inside an intracellular compartment called the symbiosome where they morphologically differentiate into nitrogen-fixing forms by symbiosome-secreted host proteins.

In this project, I explained the host proteins required to maintain bacterial symbionts and described their delivery to the symbiosome. I showed that the SYNTAXIN 132 (SYP132) gene in the model legume *Medicago truncatula* undergoes alternative cleavage and polyadenylation during transcription, giving rise to two t-SNARE protein isoforms. One of the isoforms, SYP132A, is a component of a nodule-specific secretory pathway required for the delivery of host proteins to the symbiosome.

Among the host proteins targeted to the symbiosome is DNF2. I discovered that DNF2 belongs to a novel class of phosphatidylinositol-specific phospholipase C (PI-PLC) enzymes that cleaves proteins containing glycosylphosphatidylinositol (GPI), a glycolipid that is attached to the C-termini of proteins as membrane anchors. I demonstrated that among DNF2 substrates are GPI-

anchored lysin motif (LysM) domain proteins, whose LysM domains detect microbial surface molecules (such as bacterial peptidoglycan) and initiate defense responses. The timely action of DNF2 cleaves LysM domain proteins (LYMs) off the membrane of nascent symbiosomes, suppressing defense responses and ensuring the long-term survival of intracellular rhizobia. I further discovered that the GPI-anchored arabinogalactan-like proteins: ENOD16 and ENOD20 are other DNF2 substrates. These symbiosome-localized proteins play significant roles in nodule development. My study revealed that, once bacteria are internalized, the host cell delivers the DNF2 enzyme via a SYP132A-dependent nodule-specific protein secretory pathway into symbiosomes to cleave LYMs and ENOD16/20, thereby nitrogen-fixing bacteria can survive inside host cell. My study explains the molecular mechanism of how host plant maintains intracellular bacterial symbionts, which is one of the critical steps of symbiotic nitrogen fixation.

## TABLE OF CONTENTS

	Page
ACKNOWLEDGMENTS.....	v
ABSTRACT.....	vii
LIST OF FIGURES.....	xiii
CHAPTER	
1. INTRODUCTION.....	1
1.1 Symbiotic Nitrogen Fixation.....	1
1.1.1 Nodulation Signaling.....	4
1.1.2 Bacterial Invasion.....	7
1.1.3 Nodule Development.....	9
1.1.4 Nitrogen-fixing Symbiosomes.....	10
1.2 Dissertation Statement.....	13
2. NODULE-SPECIFIC PROTEIN DELIVERY BY A SYMBIOTIC SNARE PROTEIN.....	15
2.1 Introduction.....	15
2.1.1 Secretory Pathway.....	15
2.1.2 Alternative Cleavage and Polyadenylation (APA).....	19
2.1.3 Research Question.....	22
2.2 Experimental Strategy.....	23
2.3 Results.....	23
2.3.1 <i>SYP132</i> Gene Generates Two Isoforms by APA.....	23
2.3.2 <i>SYP132</i> Isoforms Have Different Expression Pattern.....	25

2.3.3	Syp132 Isoforms Localize Differently.....	25
2.3.4	SYP132 Isoforms Have Different Functions .....	26
2.4	Discussion .....	32
2.5	Materials and Methods.....	34
2.5.1	Molecular Cloning.....	34
2.5.2	Plant Growth Conditions and Inoculation .....	34
2.5.3	Hairy Root Transformation .....	35
2.5.4	Protein Analysis .....	35
2.5.5	Microscopy.....	36
3.	DNF2: A NOVEL PI-PLC SECRETED TO THE SYMBIOSOME .....	37
3.1	Introduction .....	37
3.1.1	Phosphatidylinositol-Specific Phospholipase C .....	37
3.1.2	Glycosylphosphatidylinositol (GPI)-Anchored Proteins .....	40
3.1.3	Research Question.....	43
3.2	Experimental Strategy.....	43
3.3	Results.....	44
3.3.1	DNF2 is Specifically Expressed in Nodules .....	44
3.3.2	DNF2 is Essential for Nodule Development.....	46
3.3.3	DNF2 Localizes to the Symbiosome .....	49
3.3.4	DNF2 is a Phosphatidylinositol-Specific Phospholipase C.....	51
3.4	Discussion .....	57
3.5	Materials and Methods.....	60
3.5.1	Plant Growth Conditions and Inoculation .....	60

3.5.2	Molecular Cloning .....	60
3.5.3	Hairy Root Transformation .....	60
3.5.4	Microscopy.....	61
3.5.5	Live/Dead Staining .....	61
3.5.6	Protein Analysis .....	61
3.5.7	Microsome and Symbiosome Isolation.....	61
3.5.8	Triton X-114 Partitioning .....	62
3.5.9	GUS Reporter Assay .....	62
4.	DNF2 SUBSTRATES: LYM1 AND LYM2.....	63
4.1	Introduction .....	63
4.1.1	LysM Domain Containing Proteins.....	63
4.1.2	Research Question .....	66
4.2	Experimental Strategy.....	67
4.3	Results.....	67
4.3.1	LYM1 and LYM2 are Found in Nodules .....	67
4.3.2	LYM1 and LYM2 Localize to the Infection Thread .....	74
4.3.3	LYM1 and LYM2 are cleaved by DNF2 .....	74
4.3.4	Cleavage of LYM1 and LYM2 are Required for Nodule Development .....	77
4.4	Discussion .....	82
4.5	Materials and Methods.....	84
4.5.1	Plant Growth Conditions and Inoculation .....	84
4.5.2	Molecular Cloning.....	84

4.5.3 Hairy Root Transformation .....	85
4.5.4 Microscopy.....	85
4.5.5 Live/Dead Staining .....	86
4.5.6 Protein Analysis .....	86
4.5.7 Microsome and Symbiosome Isolation.....	86
4.5.8 Triton X-114 Partitioning .....	86
4.5.9 PCR Reverse-Screening for Tnt1 Insertion Mutants .....	87
5. DNF2 SUBSTRATES: ENOD16 AND ENOD20 .....	88
5.1 Introduction .....	88
5.1.1 Arabinogalactan Proteins.....	88
5.1.2 ENOD16 and ENOD20 Proteins .....	90
5.1.3 Research Question .....	91
5.2 Experimental Strategy.....	94
5.3 Results.....	94
5.3.1 ENOD16 and ENOD20 are Exclusively Expressed in Nodules .....	94
5.3.2 ENOD16 and ENOD20 Localize to the Symbiosomes.....	94
5.3.3 ENOD16 and ENOD20 are Crucial for Nodule Development .....	99
5.4 Discussion .....	107
5.5 Materials and Methods.....	111
5.5.1 Plant Growth Conditions and Inoculation .....	111
5.5.2 Molecular Cloning.....	111
5.5.3 Hairy Root Transformation .....	111
5.5.4 Microscopy.....	112

5.5.5 SYTO9 Staining .....	112
5.5.6 PCR Reverse-Screening for Tnt1 Insertion Mutants .....	112
5.5.7 Phenolics Staining .....	112
6. SUMMARY AND FUTURE DIRECTIONS .....	113
6.1 Summary .....	113
6.2 Future Directions .....	115
APPENDICES	
A. A SYMBIOTIC SNARE PROTEIN GENERATED BY ALTERNATIVE TERMINATION OF TRANSCRIPTION .....	117
B. THE PRIMER TABLE .....	118
C. DOES SUCCINOGLYCAN SHIELD SYMBIONT BACTERIA FROM HOST IMMUNITY?.....	119
BIBLIOGRAPHY .....	125

## LIST OF FIGURES

Figure	Page
1.1. Symbiotic association between <i>Medicago truncatula</i> and <i>Sinorhizobium meliloti</i> .....	3
1.2. The early nodulation signaling.....	6
1.3. The genetic regulation of nitrogen-fixation in symbiosomes.....	12
1.4. How does host cell maintain bacteria inside symbiosome?.....	14
2.1. SNARE-mediated vesicle fusion .....	18
2.2. The alternative terminal exon APA.....	21
2.3. Two SYP132 isoforms are generated through alternative cleavage and polyadenylation.....	24
2.4. SYP132 isoforms show different expression profiles.....	27
2.5. Western Blot analysis of SYP132 isoforms .....	28
2.6. Differential subcellular localization of SYP132 protein isoforms.....	29
2.7. SYP132C is essential for root growth.....	30
2.8. SYP132A is required for bacteroid maturation.....	31
2.9. Regulation of protein trafficking in nodule cells through APA of <i>t-SNARE</i> gene.....	33
3.1. Phosphatidylinositol-specific Phospholipase C.....	39
3.2. Structure of the GPI-anchor.....	42
3.3. DNF2 expression pattern .....	45
3.4. DNF2 is required for the survival of rhizobial bacteria inside host cells.....	47
3.5. Complementation of <i>dnf2</i> .....	48
3.6. DNF2 is targeted to the symbiosome through the nodule-specific protein secretory pathway .....	50
3.7. Protein modelling of DNF2 .....	53
3.8. DNF2 is a phosphatidylinositol-specific phospholipase C cleaving GPI-anchored proteins .....	54

3.9.	Localization of GPI-anchored GFP protein.....	55
3.10.	DNF2 cleaves GPI-anchored GFP protein.....	56
3.11.	DNF2 is required to maintain intracellular bacteria in nitrogen-fixing symbiosis .....	59
4.1.	LysM containing defense receptors.....	65
4.2.	Lym1 and Lym2 are downregulated in nodules.....	69
4.3.	LYM1 and LYM2 are detected in microsomes and symbiosomes .....	70
4.4.	Genotyping of <i>lym1</i> and <i>lym2</i> mutants .....	71
4.5.	Testing LYM1 and LYM2 antibodies.....	72
4.6.	Detection of LYM1 and Lym2 in symbiosomes.....	73
4.7.	LYM1 and LYM2 localize to the infection thread.....	75
4.8.	LYM1 and LYM2 are cleaved by DNF2 .....	76
4.9.	The overexpression of LYM1 and LYM2 using ENOD16 promoter .....	78
4.10.	Detection of LYM overexpression level .....	79
4.11.	Hyperexpression of LYM1 and LYM2 caused <i>dnf2</i> phenotype .....	80
4.12.	Silencing of LYM1 and LYM2 partially rescued <i>dnf2</i> phenotype .....	81
4.13.	DNF2-mediated inactivation of LYM1 and LYM2 blocks host defense mechanism .....	83
5.1.	ENOD16 and ENOD20 proteins.....	92
5.2.	ENOD16 and ENOD20 contain glycosylation motifs .....	93
5.3.	ENOD16 and ENOD20 have nodulation-specific expression .....	95
5.4.	ENOD16 and ENOD20 are expressed in all nodule zones.....	96
5.5.	ENOD16 localizes to the symbiosome .....	97
5.6.	ENOD20 localizes to the symbiosome .....	98
5.7.	Genotyping of <i>enod16</i> mutant lines .....	101
5.8.	Genotyping of <i>enod20</i> mutant lines .....	102
5.9.	The growth phenotypes of rhizobia-inoculated <i>enod16</i> and <i>enod20</i> plants .....	103

5.10	The <i>enod16</i> and <i>enod20</i> plants produce defective nodules.....	104
5.11.	Phenolics accumulation is observed in <i>enod16</i> and <i>enod20</i> nodules .....	105
5.12.	The <i>dnf2</i> bacteroid phenotype is rescued by the expression of GFP-ENOD16 and GFP-ENOD20 in <i>dnf2</i> nodules.....	106
5.13.	DNF2 cleaves ENOD16 and ENOD20 inside symbiosomes .....	110
6.1.	Host cell maintains intracellular bacteria by secreting DNF2 enzyme into the symbiosomes.....	114
C.1.	<i>ExoY</i> is only expressed by infection thread-localized bacteria inside nodules .....	121
C.2.	Bacteria constitutively expressing <i>ExoY</i> cause defective nodules .....	122
C.3.	<i>ExoY</i> mutant bacteria did not produce nodules in <i>lym1</i> and <i>lym2</i> plants .....	123
C.4.	<i>ExoY</i> overexpression in bacteria partially rescues the <i>dnf2</i> phenotype.....	124

## CHAPTER 1

### INTRODUCTION

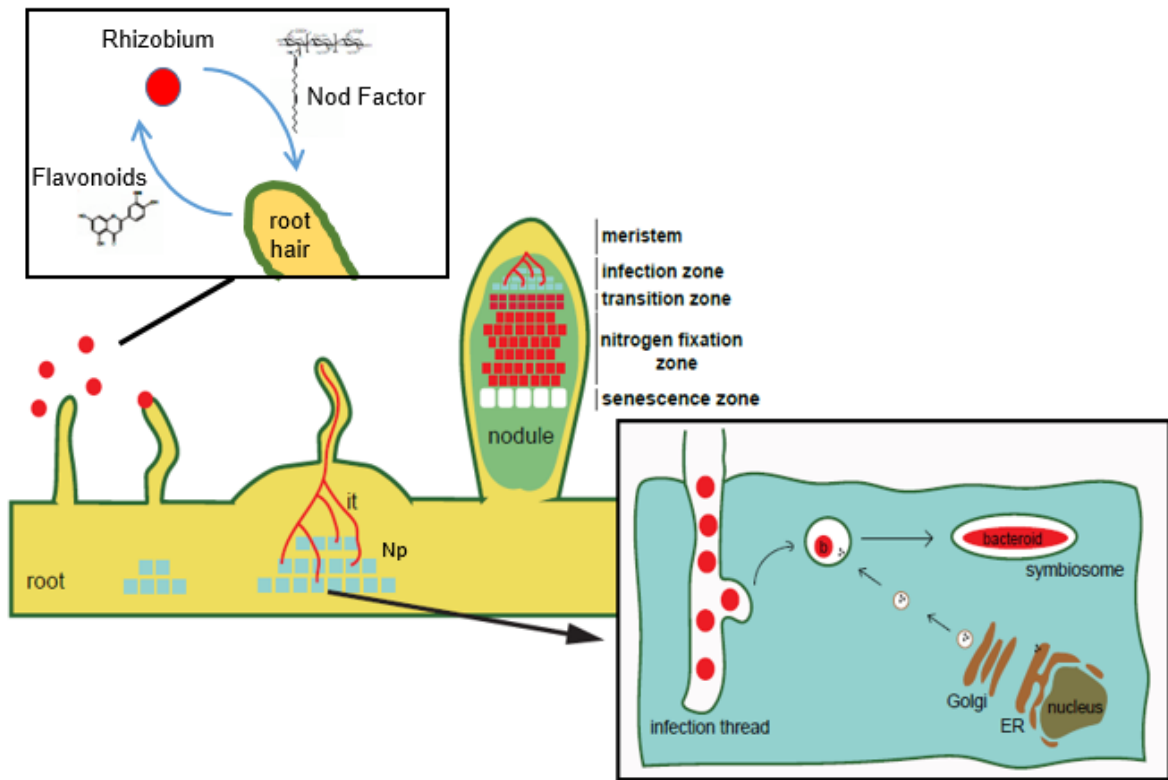
#### 1.1 Symbiotic Nitrogen Fixation

The presence of nitrogen (N) in the structures of DNA, RNA and protein makes it an essential component for life on Earth. The most abundant gas in the atmosphere is molecular nitrogen ( $N_2$ ), approximately with the concentration of 78%; however, its triply bonded diatomic form ( $N\equiv N$ ) prevents molecular nitrogen from reacting with other chemicals, or being used in metabolic reactions in the cell. Most organisms including plants, therefore, can utilize nitrogen in compound forms, not in its chemically inert gaseous form (UNEP Year Book, 2014).

Plants obtain metabolically reactive nitrogen forms such as ammonium ( $NH_4^+$ ) or nitrate ( $NO_3^-$ ) from the soil. That soil nitrogen depletes over time limits plant growth. Hence, high yields in agriculture requires synthetic nitrogen fertilizers, manufactured using the Haber-Bosch process, to enrich soil in utilizable nitrogen compounds. Inefficient application of nitrogen fertilizers, however, results in detrimental environmental impacts: air and water pollution. Excess nitrogen contributes to global warming and ozone depletion. It also runs off and leaches into streams and rivers causing dead zones, hypoxic areas in oceans and lakes caused by algae overgrowth, contaminating groundwater, and leading to biodiversity loss (Oldroyd, 2013).

Leguminous plants, known as the family Fabaceae, overcome usable nitrogen limitation by entering a symbiotic association with nitrogen-fixing rhizobial bacteria that reduce inert nitrogen to ammonia via the nitrogenase enzyme complex. This interaction occurs in a special root organ, called the nodule, in which bacteria produce ammonia and trade it with the host plant for a carbon source (Gibson et al., 2008).

The communication between *Medicago truncatula* and *Sinorhizobium meliloti* triggers root hair curling and the formation of a nodule primordium from which the nodule meristem arises. The bacteria penetrate the nodule through root hairs in a structure called the infection thread and invade host plant cells when they reach the nodule primordium. The intracellular bacteria differentiate into bacteroids inside an intracellular compartment, called the symbiosome, by the secreted host proteins. The process ends with a fully developed nodule, consisting of different zones harboring different types of cells: the meristem (Zone I), the infection or invasion zone (Zone II), the transition zone or the interzone (Zone II/III), the nitrogen fixation zone (Zone III) and the senescence zone (Zone IV) (Gibson et al., 2008) (Fig. 1.1).



**Figure 1.1 Symbiotic interaction between *Medicago truncatula* and *Sinorhizobium meliloti***

Host flavonoids exuded into the soil induces Nod factor production in the bacteria. Nod Factor captured by host receptors triggers various host responses: root hair curling, nodule primordium (Np) formation and penetration of bacteria into the nodule through root hairs in a structure called the infection thread (it). Bacteria (b) that are endocytosed by host plant cells differentiate into bacteroids, the nitrogen-fixing form, inside an intracellular compartment called the symbiosome by secreted proteins. The process results in formation of a fully developed nodule, including different zones as shown (Gibson et al., 2008).

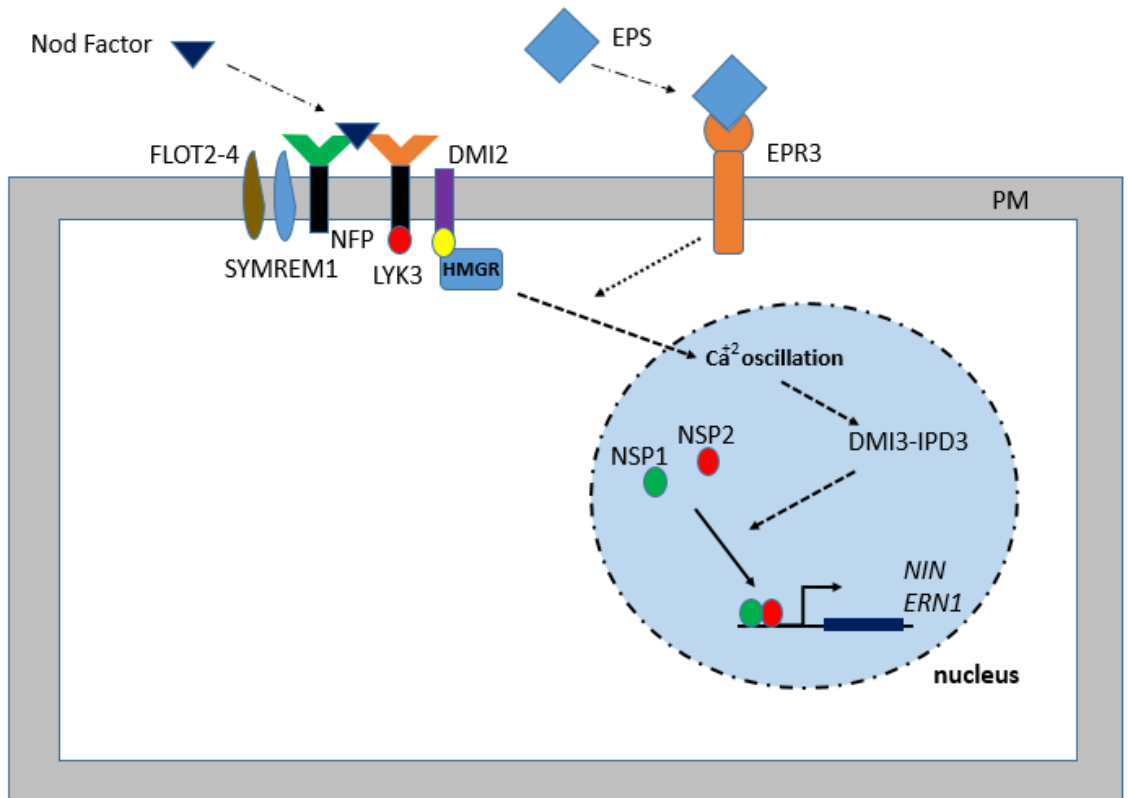
### 1.1.1 Nodulation Signaling

The symbiotic interaction relies on a molecular dialogue between rhizobia and the host legume: the host plant root exudes a mixture of flavonoids into the soil (Oldroyd, 2013). As a response, rhizobia produce nodulation factors (Nod factors) that elicit various host responses, such as nodule development and transcription, to promote rhizobia colonization (Dénarié et al., 1996) (Fig 1.1). Nod factors consist of an acylated chitooligosaccharide backbone decorated with functional group substitutions at the terminal or non-terminal residues (Miller and Oldroyd, 2012). These decorations as well as length and saturation degree of the N-acyl group vary between rhizobial species and determine the specificity of interactions with different legume species (Roche et al., 1991). Nod factors contain N-acetylglucosamine (GlcNac) that interact with LysM domains on host proteins (Gust et al., 2012). To stimulate symbiotic signaling, Nod factor binds two forms of LysM receptor-like kinases of *Medicago truncatula*: LysM Receptor Kinase 3 (LYK3) and Nod Factor Perception (NFP) (Oldroyd, 2013). In addition, the exopolysaccharide (EPS) molecules surrounding the symbiotic bacteria are essential for bacterial invasion and its specificity. EPS was shown to be perceived by a LysM domain containing receptor, EPR3, in *Lotus japonicus* (Kawaharada et al., 2015).

With Nod factor recognition, LYK3 and NFP form a receptor complex with DMI2, a leucine-rich-repeat (LRR)-containing receptor-like kinase. This hetero-complex associate with Symbiotic Remorin 1 (SYMREM1), which is predicted to play a role in localizing the Nod factor receptors in microdomains, together with Flotillin-like 2 (FLOT2) and FLOT4. The cytoplasmic domains of LYK3 and DMI2 have kinase activities, whereas the kinase domain of NFP is non-functional. That LYK3 promotes autophosphorylation and trans-phosphorylates NFP upon Nod factor binding is essential for Nod factor signaling. DMI2-interacting 3-Hydroxy-3-Methylglutaryl-CoA Reductase (HMGR) is required to produce a second messenger that links Nod Factor signaling to nucleus-

associated oscillations in calcium (Oldroyd, 2013). DMI1, an inner nuclear membrane-localized channel of *Medicago truncatula*, generates Nod factor-induced calcium spiking in the nucleus (Ané et al., 2004), activating the nuclear proteins called Calcium- and Calmodulin-dependent Serine/ Threonine Protein Kinase (CCaMK), also known as DMI3 (Lévy et al., 2004). This protein binds and phosphorylates IPD3, which is an essential component of the symbiosis signaling pathway.

Downstream of CCaMK-IPD3 complex formation, the transcription factors Nodulation Signaling Pathway 1 (NSP1) and NSP2 form a hetero-complex and induce nodulation-specific genes, such as Nodule Inception (NIN) and ERN1 transcription factors (Oldroyd, 2013) (Fig 1.2).



**Figure 1.2 The early nodulation signaling pathway in a host cell**

To trigger symbiotic signaling, Nod factor binds LYK3 and NFP resulting in receptor complex formation with DMI2. This hetero-complex associate with SYMREM1 and FLOT2-4. DMI2 and its binding protein, HMGR, lead to calcium oscillation in the nucleus by an unknown mechanism. Nod factor-induced calcium spiking in the nucleus promotes DMI3-IPD3 interaction. This leads to NSP1-NSP2 heterodimer formation, which induces the expressions of *NIN* and *ERN1* transcription factors. The exopolysaccharide (EPS) molecules binding to EPR3 are also crucial for the activation of nodulation signaling.

### 1.1.2 Bacterial Invasion

Following this signal exchange, free-living rhizobia infect root hairs and travel through a hollow tube, called the infection thread, to the nodule primordium, the region where they are internalized by cortical cells inside an organelle, known as the symbiosome. The intracellular bacteria terminally differentiate into the nitrogen-fixing bacteroid in a host-controlled manner (Gibson et al., 2008; Oldroyd et al., 2011) (Fig 1.1).

The association of Nod factor with the symbiotic receptors elicits various host physiological responses to enable rhizobial invasion of legume cells. Nod factor activates the symbiotic pathway that leads to  $\text{Ca}^{+2}$  influx in root epidermal cells, resulting in the alkalinization of the cytosol, followed by depolarization of the plasma membrane. By altering cytoskeletal dynamics, root hair tip growth is reoriented toward the site on the plasma membrane with the greatest Nod factor concentration (Esseling et al., 2003), ending up with a deformed and curled root hair and forming an infection pocket (Geurts et al., 2005). Inhibition of reactive oxygen species (ROS:  $\text{O}^{\cdot-2}$ ,  $\text{H}_2\text{O}_2$ , and  $\text{HO}\cdot$ ) formation by Nod factor is also critical for root hair deformation and curling (Lohar et al., 2007). Inside the infection pockets bacteria become entrapped and divide to establish infection foci from where infection threads arise (Oldroyd et al., 2011).

Rhizobia traverse the root epidermis and cortex inside the infection threads to reach the nodule primordium. Infection thread formation presumably requires host cell wall degradation in the region that bacteria localize. The host cell wall further invaginates along with the extended plasma membrane, and forms a hollow tube, topologically outside of the host cell. The bacteria reside inside the glycoprotein matrix and can only divide at the tip of the growing infection thread, which is regulated by host cell wall dynamics (Brewin, 2004). In contrast to the initial stages of the bacterial invasion, Nod factor induces ROS efflux, which presumably cross-links cell wall components in the lumen of the infection thread (Ramu et al., 2002). This is also supported

by the Nod Factor-dependent upregulation of the *rip1* gene of *Medicago truncatula*, which encodes a putative peroxidase playing a role in hydrogen peroxide-dependent cross-linking of cell wall proteins (Prell and Poole, 2006). In addition, the activation of the Nod factor-dependent symbiotic pathway induces the expression of the *early nodulin (ENOD)* genes involved in different steps of nodule development (D’Haeze and Holsters, 2002; Oldroyd and Downie, 2008).

Within the infection thread, invading rhizobia use a variety of polysaccharides. One of the possible functions of them is protection from stresses (Lohar et al., 2006). The lack of succinoglycan, a type of exopolysaccharide, in *Sinorhizobium meliloti* leads to the accumulation of antimicrobial phenolics and phytoalexins and the upregulation of host defense genes, showing the role of succinoglycan in dampening host immunity (Gibson et al., 2008). Cyclic  $\beta$ -glucan is an additional bacterial molecule that is required to modulate host defense response by suppressing the formation of antimicrobial phytoalexins (D’Antuono et al., 2008). In addition, bacterial outer-membrane localized lipopolysaccharides inhibit the release of ROS into the infection thread (Albus et al., 2001).

Once the branched infection thread network reaches the nodule primordium, the bacteria are internalized by the host cell and surrounded by a plant-derived membrane. An infected cell contains thousands of symbiosomes and each symbiosome contains only one bacterium. The rod-shaped bacteria inside the symbiosomes elongate and terminally differentiate into branch-shaped bacteroids. The bacteroid completely fills the symbiosome resulting in a very thin space (peribacteroid space) between the bacteroid and the host-derived membrane (peribacteroid membrane) (Gibson et al., 2008). The morphological differentiation is accompanied by several rounds of genomic endoreduplication, the replication of the genome without cell division (Mergaert et al., 2006). Bacteroid formation is under the control of the host and mediated by

various nodule-specific cysteine-rich (NCR) peptides. These signal-sequence containing peptides are delivered to the symbiosome via a DNF1-dependent nodule-specific secretory pathway (Van de Velde et al., 2010). The endoplasmic reticulum (ER)- localized DNF1 encodes a subunit of signal peptidase complex that cleaves signal peptides of cargo proteins from their N-terminals (Wang et al., 2010). The NCR peptides are 30-50 amino acid residues long and are structurally similar to defensins, antimicrobial peptides in plants. *Medicago truncatula* contains about 600 NCR peptide genes, about more than 5% of the nodule transcriptome. The NCR genes, mostly, are specifically expressed in the nodules and induced by the infection of rhizobium (Durgo et al., 2015). Among the well-characterized NCR peptides are NCR247 and NCR211. While NCR247 has been shown to change the bacterial transcriptome and translation, and inhibit bacterial division inside symbiosome, NCR211 is required to prevent the bacteroid from degeneration (Kim et al., 2015). An inner membrane protein of bacteria, BacA, plays an essential role in importing these NCR peptides into the cytosol of the bacteria (Mergaert et al., 2006).

### **1.1.3 Nodule Development**

Nod factor recognition elicits the differentiation of quiescent root cortical cells into an actively dividing meristem, leading to the formation of a nodule primordium at the root cortex. These cells undergo endoreduplication and become enlarged, then, are invaded by the bacteria carried within infection thread from the root surface (Gibson et al., 2008). The spatio-temporal auxin/cytokinin balance is associated with the nodule development. The fact that Nod Factor elevates the localized cytokinin levels in the root cortex and pericycle, at the site of inoculation, dampens the localized polar auxin transport, resulting in the initiation of nodule formation. Once initiated, auxin activity is increased in the nodule meristem, suggesting that the auxin has a role in the later stages of nodule organogenesis (Frugier et al., 2008; Mathesius et al., 1998).

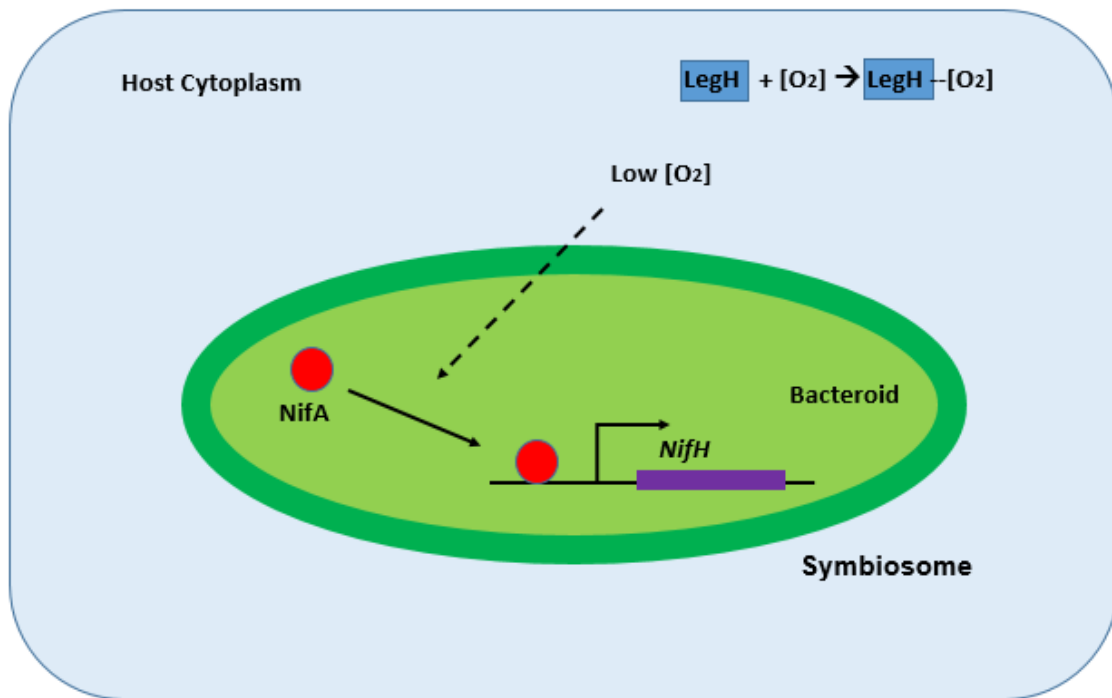
Legumes form a symbiotic relationship with rhizobia inside a root nodule. Root nodules are classified as determinate or indeterminate. While a determinate nodule has a transient meristem, derived from outer cortical cells, an indeterminate nodule has a tip-growing, persistent meristem, originated from inner cortical cells. *Medicago truncatula* forms indeterminate nodules. A fully developed indeterminate nodule possess the meristem (Zone I), the invasion zone (Zone II) where the bacteria carried by infection thread invade the host cells, the interzone (Zone II/III) where the bacteria differentiate into bacteroids, the nitrogen-fixing zone (Zone III) including the nodule cells with nitrogen-fixing mature bacteroids, and the senescent zone (Zone IV) in which both host cells and bacteroids degenerate (Gibson et al., 2008; Oldroyd et al., 2011).

#### **1.1.4 Nitrogen-fixing Symbiosomes**

Inside symbiosomes, rhizobia fix atmospheric nitrogen into ammonia by the nitrogenase enzyme complex, which consists of two protein subunits: a molybdenum-iron (MoFe) protein and Fe-containing protein. While *NifK* and *NifD* encode the MoFe protein subunits, *NifH* encodes the Fe-containing protein. The MoFe protein activity depends on a FeMo cofactor (FeMo-Co), whose synthesis requires the *NifB*, *NifV*, *NifN* and *NifE* genes (Rosenberg and Gophna, 2011; Udvardi and Poole, 2013). Nitrogenase activity is irreversibly inhibited by oxygen (O<sub>2</sub>), thereby requires a microoxic environment provided by host leghemoglobin proteins (Oldroyd, 2013).

Leghemoglobins are the most abundant nodule proteins. They bind oxygen in the host cytoplasm with high affinity and deliver it to bacteroids of infected cells. Nitrogenase is not made under aerobic conditions. The absence of oxygen is sensed by the membrane-anchored FixL protein in *Sinorhizobium meliloti*, which phosphorylates the transcriptional regulator FixJ (Gilles-Gonzalez et al., 1991). Fix J induces the expression of the transcription factor, NifA, which regulates the nifHDK operon encoding the components of nitrogenase enzyme complex

(Foussard et al., 1997) (Fig 1.3). The expressions of leghemoglobin and NifH are good indicators of proper nodule development and nitrogen fixation.



**Figure 1.3 The genetic regulation of nitrogen-fixation in the symbiosome**

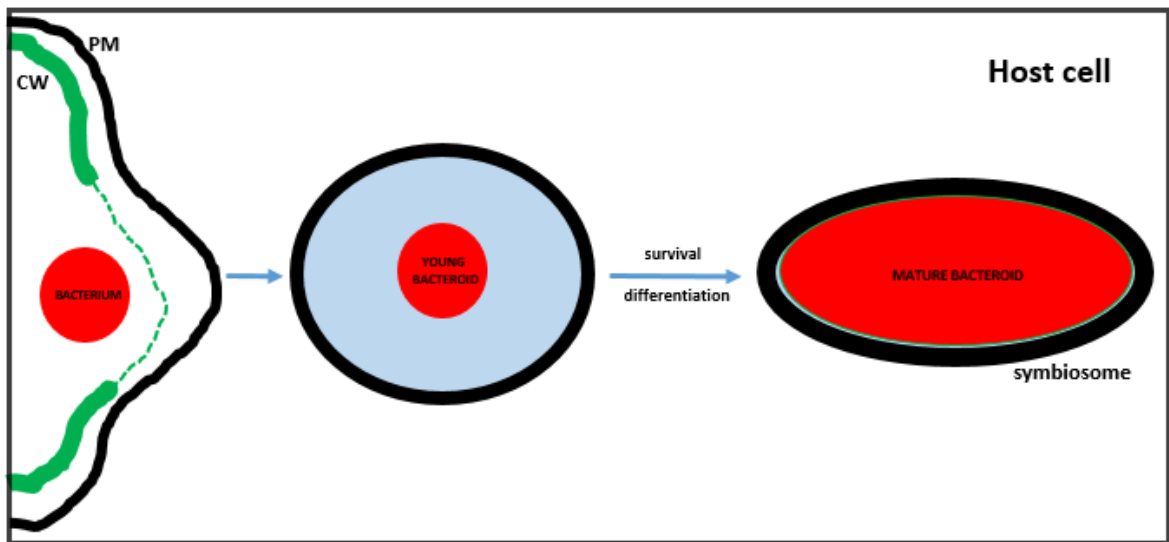
Leghemoglobin proteins provide the microoxic environment by binding to oxygen in the host cytoplasm. The decrease in oxygen level leads to the activation of the transcriptional factor, NifA, which regulates the nifHDK operon encoding the components of nitrogenase enzyme complex, such as NifH.

## 1.2 Thesis Statement

The fact that world population is increasing dramatically and is expected to count 9 billion by 2045, elevates the demand for food, and therefore high-yield agriculture. Agricultural productivity requires the use of chemical nitrogen fertilizers to enrich the soil in metabolically usable nitrogen forms, whose application, however, detrimentally affects the environment, especially leading to the pollution of water resources. Therefore, there is need to find an alternative way to feed the crops starving for nitrogen.

Legumes overcome nitrogen shortage by developing root nodules in which they form a symbiotic association with nitrogen-fixing bacteria. The transfer of this ability to non-legume crops, such as maize and wheat, might be a method to reduce or eliminate nitrogen fertilizer use. Hence, understanding the molecular mechanism of symbiotic nitrogen fixation is essential, requiring identification of the essential plant and bacterial molecules for the symbiotic association between legumes and rhizobia.

In my thesis, I described how a host plant maintains the nitrogen-fixing bacteria during the nodule development (Fig 1.4). In Chapter 2, I explain a nodule-specific regulation of protein trafficking to the symbiosomes by alternative cleavage and polyadenylation of the *Medicago truncatula* *SYNTAXIN 132* gene, encoding two target-membrane Soluble NSF Attachment Protein Receptor (t-SNARE) isoforms. In Chapter 3, I explain the function of a secreted host enzyme, named DNF2, in the survival of bacteria within symbiosomes. In Chapter 4, I depict how the action of DNF2 blocks host defense response, inactivating two LysM domain-containing defense receptors, LYM1 and LYM2. Finally, I outline the roles of two early nodulin proteins: ENOD16 and ENOD20, that are additional substrates of DNF2, in nodulation.



**Figure 1.4 How does the host cell maintain bacteria inside symbiosome?**

Symbiotic nitrogen fixation occurs in an intracellular compartment of plant host cell. Bacteria are endocytosed and differentiated into a nitrogen-fixing form inside symbiosomes. However, the host mechanisms that maintain endosymbiotic bacteria are largely unknown.

## CHAPTER 2

### NODULE-SPECIFIC PROTEIN DELIVERY BY A SYMBIOTIC SNARE PROTEIN

The work presented in this chapter was performed in a collaboration with Huairong Pan. The figures in this chapter are derived from the original research article: Huairong Pan, Onur Oztas, Xiaowei Zhang, Xiaoyi Wu, Christina Stonoha, Ertao Wang, Bin Wang and Dong Wang. A symbiotic SNARE protein generated by alternative termination of transcription. 2016, 2 (2): 15197, *Nature Plants*.

#### 2.1 Introduction

##### 2.1.1 Secretory Pathway

Proteins are delivered to the subcellular compartments and extracellular space by the secretory pathway (Sigurbjörnsdóttir et al., 2014). These proteins destined for secretion possess a signal peptide at their N-terminus that is cleaved by the signal peptidase complex within the ER. DNF1 protein of *Medicago truncatula* is a subunit of the signal peptidase complex required to direct the legume proteins to a bacterial symbiont-containing symbiosome (Auclair et al., 2012; Wang et al., 2010). The processed proteins are labeled as cargoes of the secretory pathway and packaged into carrier vesicles to be transported from the ER to the Golgi, where they undergo sorting and further modifications to acquire their functional forms. Proteins further exit the ER in secretory vesicles to be carried to their final destinations. The vesicles fuse with the target membranes and release their contents once they reach the destination (Sigurbjörnsdóttir et al., 2014).

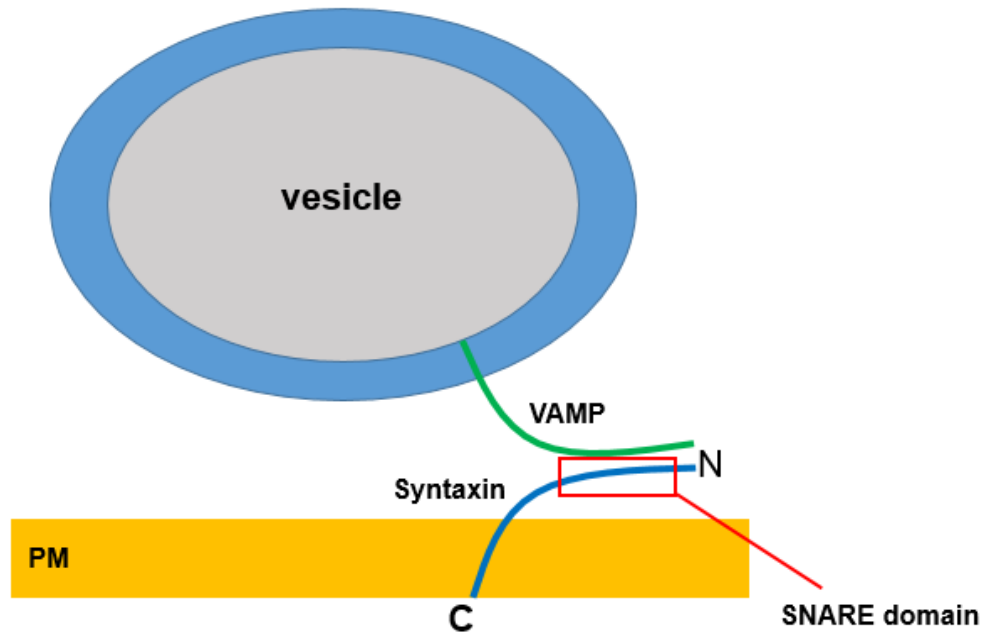
Vesicle fusion is mediated by SNARE (soluble NSF attachment protein receptor where NSF stands for N-ethyl-maleimide-sensitive fusion protein) proteins, that are classified as v-SNAREs and t-SNAREs on the basis of whether they localize on vesicle or target membrane, respectively.

Among the well-known SNARE families are syntaxin, VAMP and SNAP-25. While SNAP-25 is anchored to the target membrane via palmitoylation, syntaxin on the target membrane and VAMP on the vesicle contain transmembrane domains in their C-termini. SNAREs possess heptad repeats that have the ability to form coiled coil structures. One coiled-coil from syntaxin and VAMP and two coiled-coil of SNAP-25 intertwine to assemble a SNARE core complex that brings secretory vesicle and target membrane close enough for fusion. This process is regulated by a chaperone protein, named Sec1, and Rab proteins. By binding to a chaperone protein, named nSec-1, with high affinity, syntaxin adopts a closed conformation that prevents forming a SNARE core complex with SNAP-25 and VAMP. In contrast, Rab proteins facilitate the transition from a closed to open state, thereby SNAREs can assemble a coiled-coil complex enabling vesicle fusion and release of the cargo in a  $\text{Ca}^{2+}$  dependent manner. The SNARE complex on the target membrane dissociates after the completion of membrane fusion via the recruitment of  $\alpha$ -SNAP and NSF from the cytoplasm, followed by ATP hydrolysis. Hence, SNAREs are recycled for subsequent fusion reactions (Bonifacino and Glick, 2004; Chen and Scheller, 2001).

Plants contain higher numbers of SNAREs compared to other eukaryotes. Arabidopsis possess approximately 65 SNARE proteins that play roles in growth and development, as well as in biotic and abiotic stress responses (Kim and Brandizzi, 2012).

SNAREs show distinct subcellular localizations. Each v-SNARE on a different type of secretory vesicles couples with a unique cognate t-SNARE, facilitating the specificity for the fusion of vesicles with appropriate target membranes (Dun et al., 2010) (Fig. 2-1). This SNARE type-dependent specificity leads to the regulation of protein trafficking to different subcellular compartments simultaneously. *Medicago truncatula* possess two v-SNARE proteins with nodulation-specific roles, called MtVAMP721d and MtVAMP721e. These proteins localize to the region where the host cells internalize the invading bacteria. Silencing of *MtVAMP721d* and

*MtVAMP721e* inhibits the bacterial release from the infection thread, and hence symbiosome formation (Ivanov et al., 2012).



**Figure 2.1 SNARE-mediated vesicle fusion**

Vesicle fusion is mediated by SNARE proteins. To bring secretory vesicle and target membrane close enough for fusion, the Syntaxin protein on the target membrane and the VAMP protein on the vesicle interact by their SNARE domains, resulting in the formation of a heterocomplex. This SNARE complex also contains SNAP-25, which is not shown in this diagram for simplicity. The Syntaxin proteins contain a transmembrane domain in their C-termini.

### 2.1.2 Alternative Cleavage and Polyadenylation

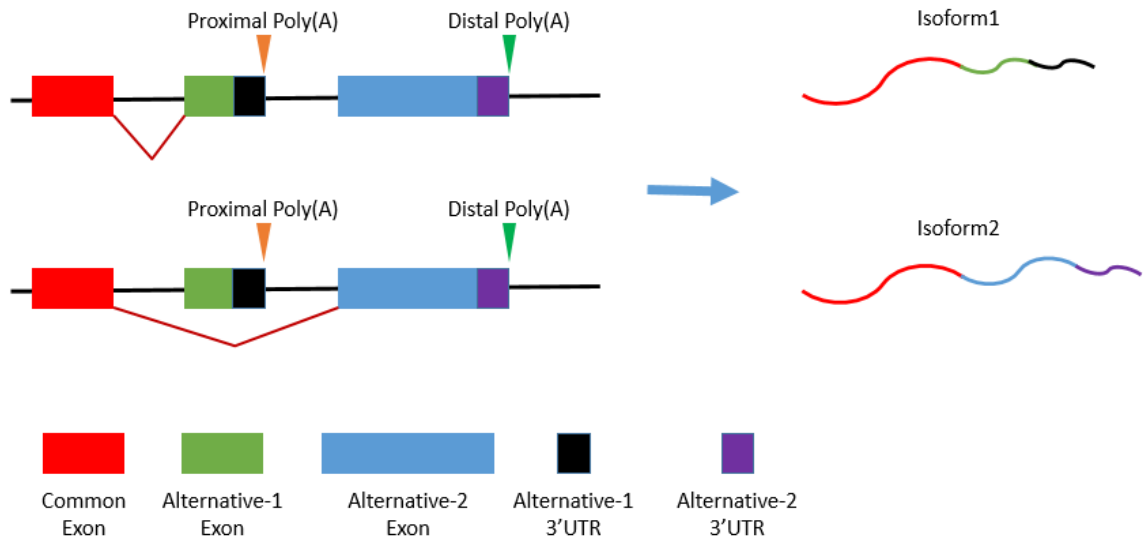
Most messenger RNAs (mRNAs) possess a poly(A) tail, consisting of multiple adenosine monophosphates, at their 3' ends, which is required for their nuclear export, stability, and translation by ribosomes (Sachs, 1990). During RNA polymerase II-mediated transcription, the poly(A) tail is added to transcripts via two-coupled reactions. An endonuclease-containing protein complex first cleaves the nascent RNA. Then, a poly(A) polymerase (PAP) synthesizes a poly(A) tail on the 3' terminus of the cleaved RNA (Richard and Manley, 2009). This process, called polyadenylation, requires cis-acting RNA elements and core and auxiliary proteins. The poly(A) signal (PAS) is the essential cis-element that drives cleavage of the nascent RNA, while U- or GU-rich downstream sequence elements (DSEs) and upstream sequence elements (USEs) are required to increase the cleavage efficiency. PAS is a six-nucleotide motif mainly in the form of AAUAAA and located 15-30 nucleotides upstream of the cleavage site. Once PAS and DSE are recognized by the cleavage and polyadenylation specificity factor (CPSF) and the cleavage stimulating factor (CSTF), respectively, the nascent RNA is cleaved between these elements (Mandel et al., 2008). CPSF and CSTF are multi-protein complexes. CPSF is composed of CPSF4, CPSF2, CPSF3, FIP1L1, WDR33 and CPSF1, which is the subunit recognizing PAS (82), whereas CSTF consists of CSTF1, CSTF2 and CSTF3 proteins. Cleavage and polyadenylation also involve the scaffold protein simplekin, cleavage factors Im (CFIm) and CFIIIm, and poly(A) binding protein.

Most protein coding genes contain at least two poly(A) sites, termed proximal and distal based on their positions relative to the coding region or 3'UTR (Elkon et al., 2013). For instance, 60% of Arabidopsis genes and between 47 and 82% of rice genes contain more than one poly(A) site (Hunt, 2011). By differential usage of these sites one gene can encode distinct transcripts with different 3' ends, a phenomenon named alternative cleavage and polyadenylation (APA). All

eukaryotes, from yeast to human, have alternative poly(A) sites that generate mRNA isoforms differing either in their coding sequences or in their 3' untranslated regions (UTRs). This RNA-processing mechanism regulates the function, stability, localization, and translation efficiency of target RNAs. Many plant genes have multiple PASs and are processed by APA (Tian and Manley, 2016).

Four different types of APA have been described: tandem 3' untranslated region (UTR), alternative terminal exon, intronic, and internal exon. In the tandem 3' untranslated region (UTR) APA, the alternative poly(A) sites are present within the same terminal exon, thereby producing isoforms with different 3'UTR length. The additional three types of APA events result in isoforms that differ in their coding sequences in addition to their 3'UTRs. Distinct mRNA isoforms, that differ in their last exon, are produced by alternative splicing via the alternative terminal exon APA (Fig. 2.2). In the intronic APA, the cleavage at the cryptic intronic poly(A) signal leads to a longer internal exon, becoming the terminal one. Premature polyadenylation occurs within the coding region in the internal exon APA (Elkon et al., 2013).

Multiple factors play roles in the selection of poly(A) signal (PAS), whether it is proximal or distal, during APA. Polyadenylation is associated with transcription, thereby chromatin structure. Hence, the RNA polymerase II elongation rate, the recruitment of APA regulatory factors by transcription machinery and nucleosome density in the region around the PAS determines the choice of PAS. Among the other factors influencing PAS selection is the expression and recruitment of core polyadenylation machinery components such as CSTF64 (Takagaki et al., 1996). Also, the presence of RNA binding proteins interacting the region near PASs, N6-methyladenosine (m6A), and inhibition of polyadenylation by the splicing factor U1 small nuclear ribonucleoprotein (U1 snRNP) also influence the PAS usage (Tian and Manley, 2016).



**Figure 2.2 The alternative terminal exon APA**

Alternative splicing, combined with alternative cleavage and polyadenylation, generates two distinct mRNA isoforms using different poly(A) sites (proximal or distal) in the 3' UTRs. These isoforms differ in their coding sequences and their 3' UTRs. This phenomenon is termed as alternative terminal exon APA (Elkon et al., 2013).

### 2.1.3 Research Question

Rhizobia are endocytosed into host cells where they individually become enclosed in a peribacteroid membrane, resulting in the formation of intracellular symbiosome compartment. This unique compartment is a novel destination for host protein secretion, which is essential for the bacterial symbiont to acquire a nitrogen-fixing phenotype. The host cell therefore needs to create an alternate path to deliver proteins specifically to the symbiosome without disrupting the existing path to the extracellular space.

The secretory pathway consists of two significant steps: accurate cargo packing into transport vesicles in the ER and their targeting to appropriate compartments, associated with an interaction between v-SNARE and t-SNARE proteins. Whether a reprogramming occurs in these steps of the secretory pathway in an infected host cell is a crucial question. Discovery of DNF1 suggested a nodule-specific cargo selection in the ER (Wang et al., 2010). Identification of two v-SNAREs induced during nitrogen fixing symbiosis in *Medicago truncatula*, suggests the presence of transport vesicles specifically targeted towards the symbiosome. Both findings support the idea of a nodule-specific protein secretory pathway. However, a peribacteroid membrane-localized t-SNARE protein required for a symbiosome-specific SNARE complex has not been identified.

Here, I identified a *t-SNARE* gene, *SYP132*, encoding two protein isoforms through alternative cleavage and polyadenylation. One of the isoforms, SYP132C, encodes a plasma membrane-localized t-SNARE that is required for the protein trafficking to extracellular space, whereas SYP132A encodes a nodule-specific t-SNARE that tags symbiosomes for protein secretion. This study explains how a host cell marks the symbiosomes as a new destination for protein secretion and distinguishes plasma membrane and peribacteroid membrane during nitrogen-fixing symbiosis.

## 2.2 Experimental Strategy

In this study, I discovered that the alternative cleavage and polyadenylation of *M. truncatula* *SYNTAXIN132* (*SYP132*) gene results in encoding two different isoforms: Canonical *SYNTAXIN132* (*SYP132C*) and Alternative *SYNTAXIN132* (*SYP132A*), that differ in their transmembrane domains and SNARE domains. First, I analyzed the expressions of *SYP132C* and *SYP132A* using the *M. truncatula* gene expression database. I further checked the subcellular localizations of these protein isoforms under confocal microscopy. To understand whether these isoforms possess a functional difference, I silenced *SYP132C* and *SYP132A* genes, and investigated their roles in root development and nodule formation.

## 2.3 Results

### 2.3.1 *SYP132* Gene Generates Two Isoforms by APA

*Syntaxin132* is encoded by a single gene the chromosome 2 of *Medicago truncatula*. I noticed that the *SYP132* gene is represented on the *Medicago truncatula* Gene Expression Atlas by two probe sets, corresponding to two distinct transcripts (Benedito et al., 2008). Each transcript has 12 exons in common and a different terminal (13<sup>th</sup>) exon, named the canonical terminal exon (exon 13C) and the alternative terminal exon (exon 13A). The terminal exon encodes a transmembrane domain and a part of the SNARE domain. I found that these two mature transcripts are generated by alternative cleavage and polyadenylation mechanism, where transcription terminates either at exon 13C or exon 13A (Fig. 2.3). When exon 13A is produced, it replaces exon 13C through alternative splicing.



### **2.3.2 SYP132 Isoforms Have a Different Expression Pattern**

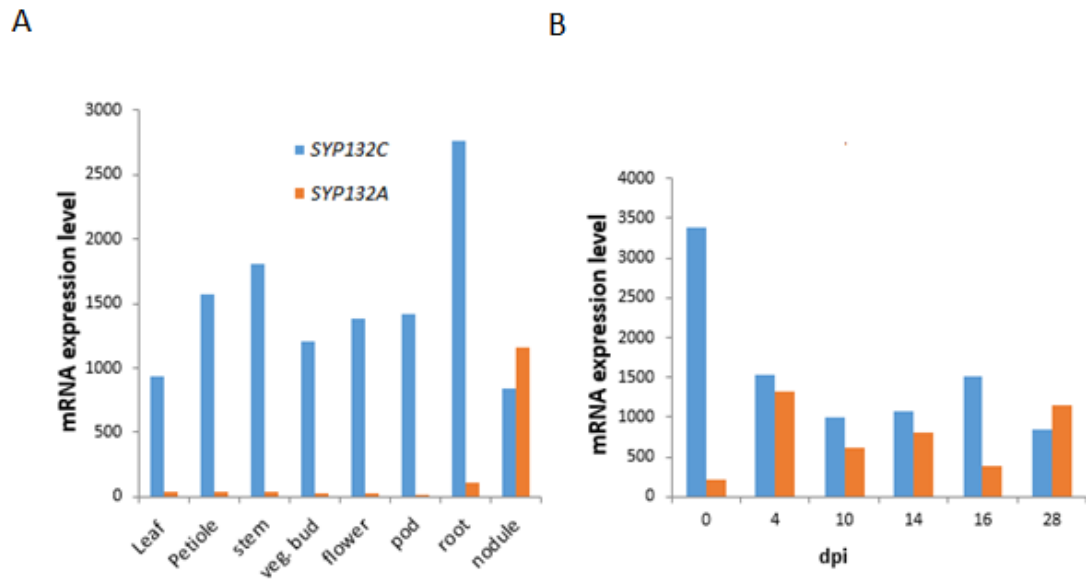
*SYP132A* and *SYP132C* have different coding sequences at their 3' ends. I analyzed the expression of *SYP132A* and *SYP132C* by blasting the distinct sequences at the 3' end on the *Medicago truncatula* Gene Expression Atlas. I found that each isoform has a unique expression pattern. While *SYP132C* is constitutively expressed in every tissue, the *SYP132A* has a nodule-specific expression pattern (Fig. 2.4A), also confirmed by Real-Time PCR analysis (Pan and Oztas et al., 2016). In addition, I noticed that the treatment of legumes with rhizobia upregulates *SYP132A*, while downregulating *SYP132C* (Fig. 2.4B). The expression data indicate that *SYP132C* is the predominant form functioning in all tissues and *SYP132A* is the nodule-specific form.

### **2.3.3 SYP132 Isoforms Localize Differently**

*SYP132A* and *SYP132C* isoforms differ in their SNARE domains and transmembrane domains at the carboxy-termini, suggesting differential localization. To determine the subcellular localization, I fused each isoform to green fluorescent protein (GFP) and produce in nodules under the control of a *DNF1* promoter, that is co-expressed with the *SYP132* gene. When I tested the expression of each recombinant construct by the Western Blotting using an anti-GFP antibody, I detected both isoforms at expected sizes (Fig. 2.5). I further investigated the subcellular localization of the GFP-fused isoforms in rhizobia-containing nodules under a confocal microscope. I found that the *SYP132C* localizes to the plasma membrane in the infected nodule cells, whereas the *SYP132A* both localizes to the peribacteroid membrane surrounding a bacteroid and the plasma membrane (Fig. 2.6A). That *SYP132A* is found to localize on the plasma membrane is most probably in spite of ectopic expression. To understand better the symbiosome localization of the isoforms, I examined isolated symbiosomes under a confocal microscope. I detected *SYP132A* protein on the peribacteroid membrane, but not *SYP132C* (Fig. 2.6B).

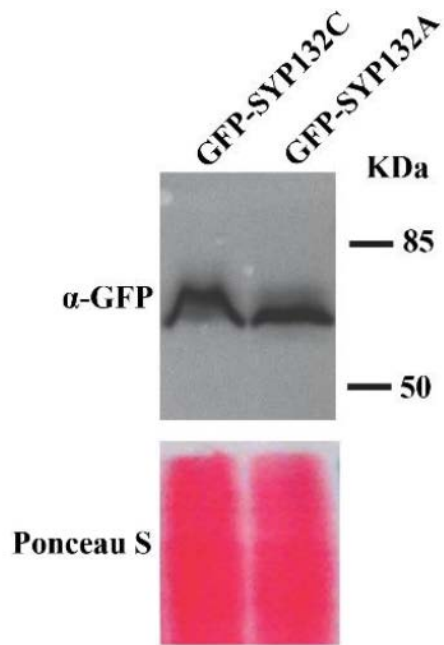
#### 2.3.4 SYP132 Isoforms Have Different Functions

To investigate their roles, I silenced *SYP132C* and *SYP132A* isoforms by generating RNAi constructs targeting the terminal exon of *SYP132A* or *SYP132C*, and one targeting both transcripts (*SYP132 RNAi*). Huairong Pan confirmed the reduction in transcript levels, as well as RNAi silencing specificity by real-time PCR (Pan and Oztas et al., 2016). I showed that knockdown of *both transcripts or SYP132C* alone prevented root growth, implying that *SYP132C* isoform is an essential t-SNARE protein for root development (Fig. 2.7). Huairong Pan analyzed the phenotype of *SYP132A* silencing in WT nodules. Unlike empty vector control, *SYP132A*-silenced plants developed white and small nodules (Fig. 2.8A&B) and contained small and undifferentiated intracellular bacteria (Fig. 2.8C). Consistent with their morphology, *SYP132A*-silenced nodules produced less NifH protein than control nodules (Fig. 2.8D). That knockdown of *SYP132A* caused *dnf1*-like phenotype (Wang et al., 2010) indicates that *SYP132A* is a peribacteroid membrane-localized t-SNARE required for secreting host secretory proteins important in bacterial differentiation (Van de Velde et al., 2010) to the symbiosome.



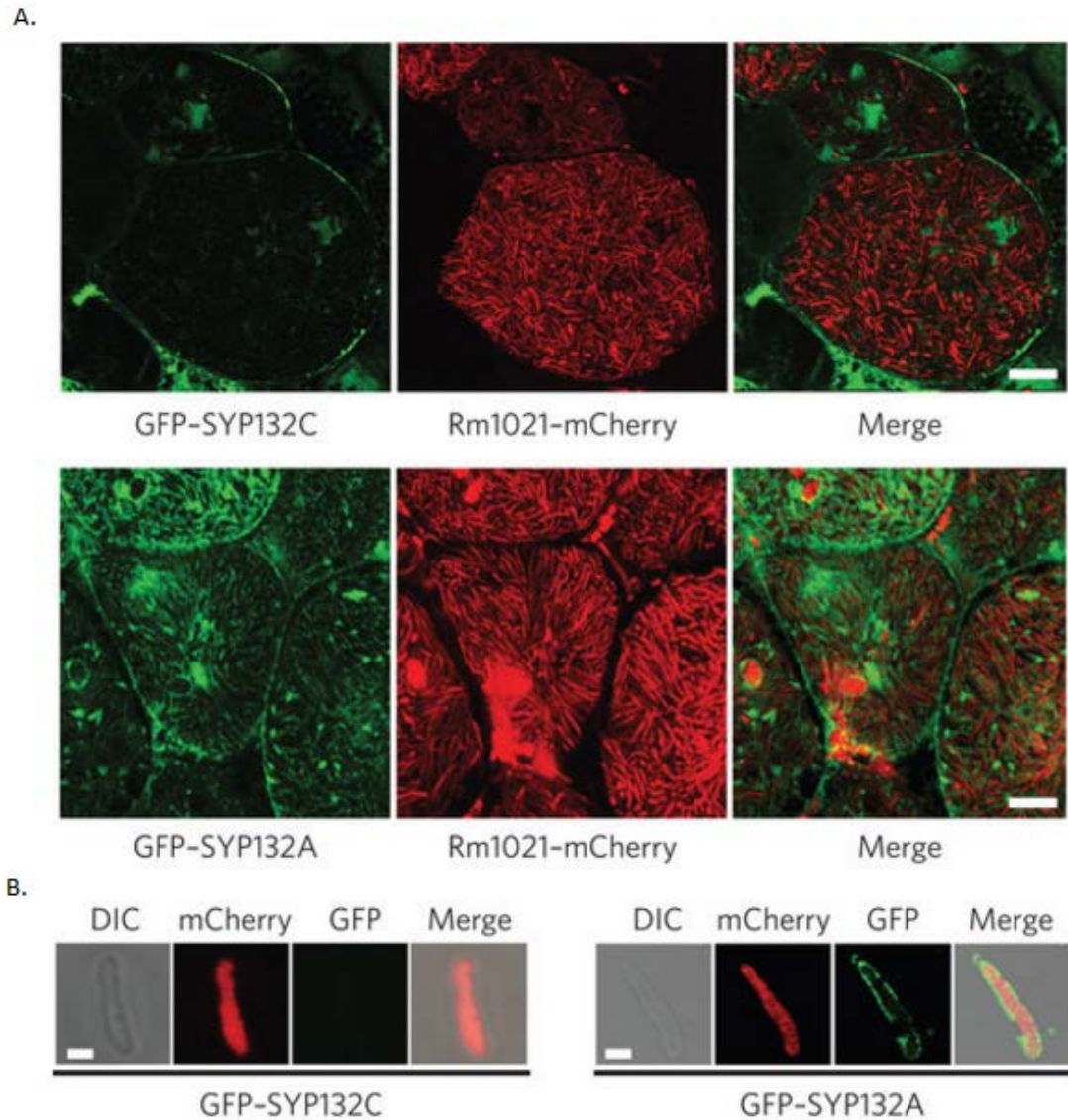
**Figure 2.4 SYP132 isoforms show different expression profiles**

**(A)** *SYP132C* and *SYP132A* expression levels in different tissues. **(B)** Temporal profile of *SYP132C* and *SYP132A* expression in nodule. Dpi stands for days post infection. Values are levels of Affymetrix probe signal based on microarray data from the MtGEA database (<http://mtgea.noble.org>).



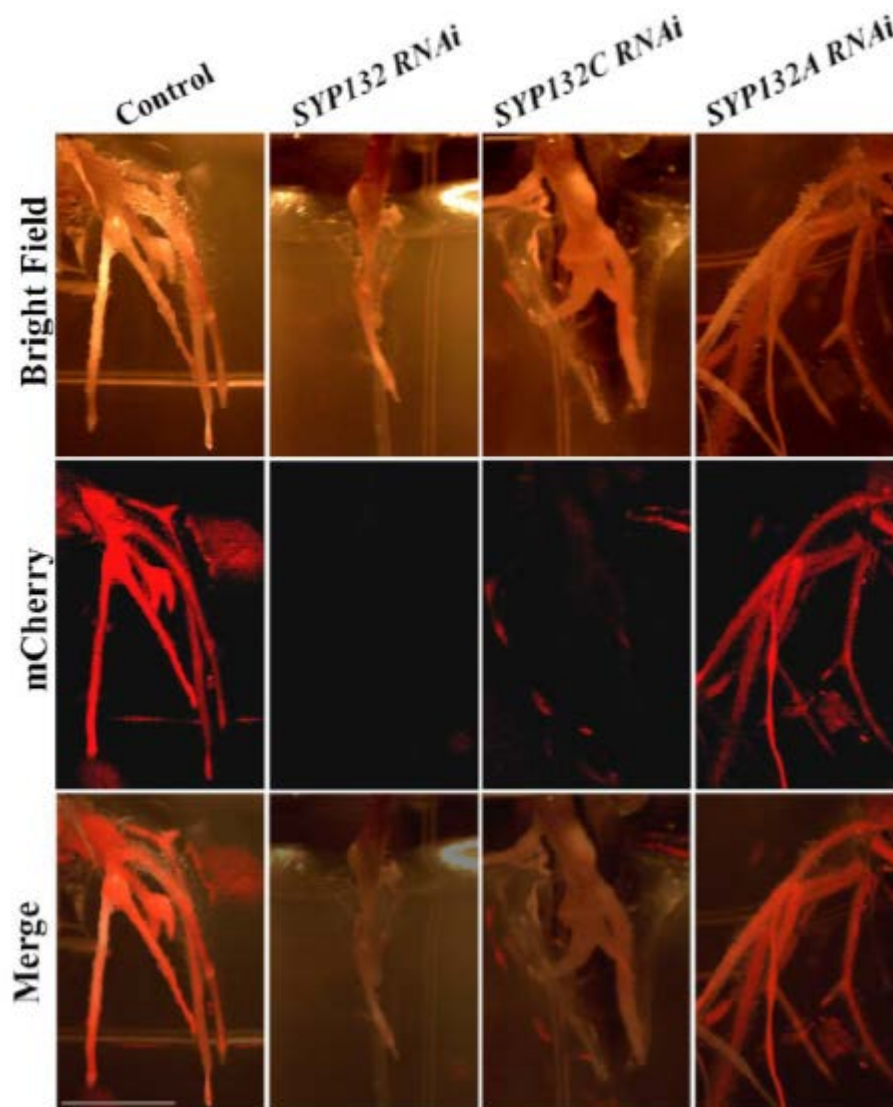
**Figure 2.5 Western blot analysis of SYP132 isoforms**

Nodule extracts of WT plants expressing *DNF1-promoter::GFP-SYP132* isoforms were separated on SDS-PAGE and probed with an anti-GFP antibody. The expected size of GFP-SYP132 isoforms is 62 kDa. Ponceau S staining confirmed equal protein loading.



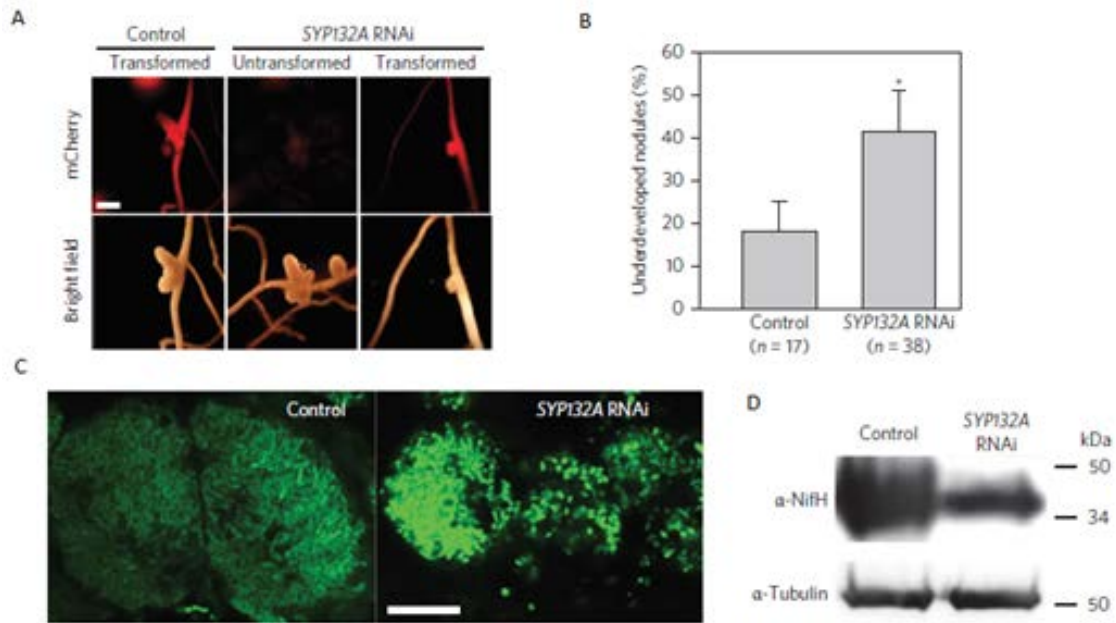
**Figure 2.6 Differential subcellular localization of SYP132 protein isoforms**

**(A)** Subcellular localization, by confocal microscopy, of full-length SYP132C and SYP132A, each fused with GFP, in nodule cells infected with *Rm1021* expressing an mCherry marker. Both SYP132C and SYP132A were driven from the *DNF1* promoter. Scale bars, 20  $\mu$ m. **(B)** GFP-SYP132C and GFP-SYP132A signal on individual symbiosomes isolated from transgenic nodules. Symbiosomes were isolated from mCherry-*Rm1021* infected nodules and checked under confocal microscope immediately. Scale bars, 2  $\mu$ m.



**Figure 2.7 SYP132C is essential for root growth**

*SYP132 RNAi* refers to a construct aimed at a region common to both *SYP132C* and *SYP132A*. The *pHellsGate8* vector was modified with a mCherry marker, allowing transformed roots to be visualized under a fluorescence stereoscope. Only non-fluorescent roots lacking mCherry signals emerged from *SYP132 RNAi* and *SYP132C RNAi* plants, suggesting that these roots did not silence their intended targets. These roots similarly lacked kanamycin resistance, and grew poorly in Fahraeus medium with the antibiotic. Bar = 1 cm. Images are representatives from more than four independent experiments, each containing around 20 independently transformed roots per construct.



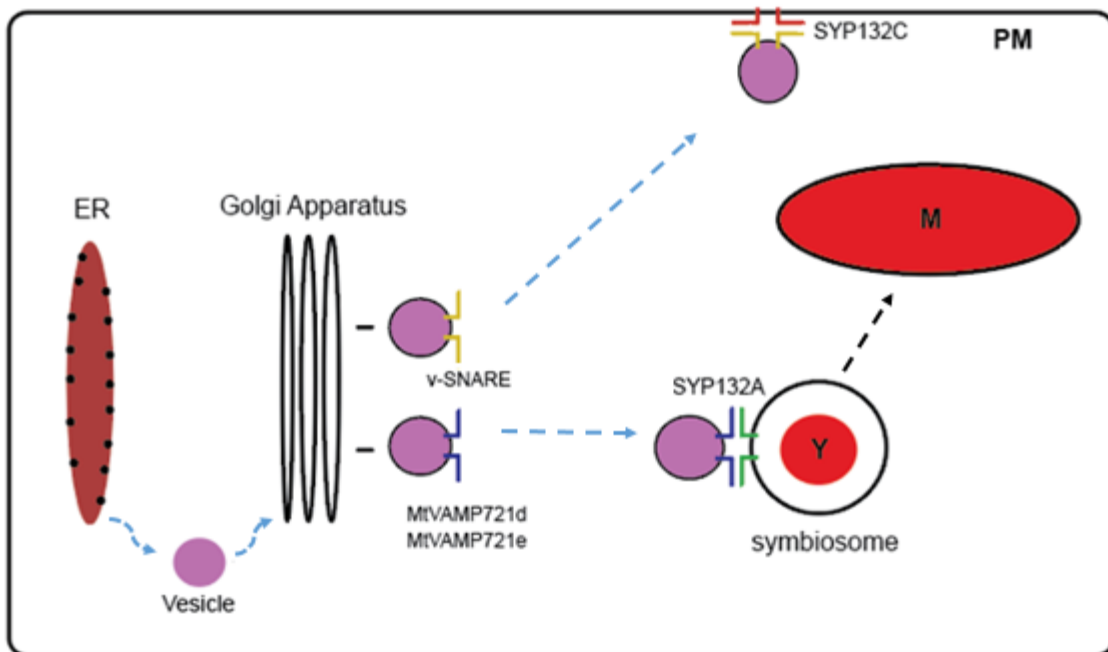
**Figure 2.8 Syp132A is required for bacteroid maturation**

**(A)** Morphology of *SYP132A*-silenced nodules and the control. The upper panel shows the fluorescence of the mCherry reporter encoded in the *pHellsgate8* RNAi vector. The lower panel shows the morphology of *SYP132A*-silenced nodules and control nodules. Scale bar, 2 mm. **(B)** Statistical analysis to show that *SYP132 RNAi* plants have more white nodules comparing to the empty vector control. \*\*\* $P < 0.001$ , Student's t-test. n = the number of transformed plants for each vector. **(C)** SYTO9 staining of bacteroids in control and *SYP132A*-silenced nodules. Scale bar, 40  $\mu$ m. **(D)** NifH protein levels in *SYP132A*-silenced nodules and control nodules by Western blot. Tubulin was used as a loading control.

## 2.4 Discussion

In this study, I described a novel mechanism to label a target membrane for protein delivery through APA of a t-SNARE gene. I found that, by the differential regulation of transcription termination, *SYP132* gene of *Medicago truncatula* encodes two distinct isoforms that differ in their expression patterns, localization, and biological roles. While SYP132C is produced in every tissue, SYP132A is only expressed in nodules. SYP132C only localizes to the plasma membrane (PM), whereas SYP132A localizes to the peribacteroid membrane (PBM), as well as to the PM. SYP132C is essential for root development, while SYP132A is required for bacteroid differentiation. Taking all these data into consideration, I conclude that SYP132C labels the PM for protein secretion to the extracellular space, and SYP132A specifies the PBM for delivery of host proteins to symbiosomes.

How a host cell reprograms protein trafficking once symbiotic bacteria are internalized into the symbiosomes can be explained by *SYP132* gene regulation by APA. The uninfected nodule cells only produce the SYP132C isoform. Once cells are infected, they induce the expression of SYP132A through APA to label the symbiosomes for secretion of specialized host proteins. To mediate the fusion of secretory vesicles with the PBM, SYP132A likely forms a SNARE complex with VAMP721d/e, v-SNAREs upregulated in infected nodule cells (Ivanov et al., 2012) (Fig. 2.9). The symbiotic bacteria are enclosed by a PBM, which is derived from the plant PM. The PBM differentiates in phase with bacteroid development and acquires vacuole membrane properties (Verma et al., 1995), becoming different than the original membrane. That PBM and PM vary in their molecular composition presumably leads to the distinct localization of SYP132C and SYP132A isoforms differing in their transmembrane domains. Furthermore, SYP132C and SYP132A vary in their SNARE domains, implying that they have the ability to pair with different families of v-SNAREs, and therefore different types of secretory vesicles.



**Figure 2.9 Regulation of protein trafficking in infected nodule cells through APA of *t-SNARE* gene**

SYP132 gene encodes two different isoforms: Syp132C and Syp132A. In infected nodule cells, Syp132C and SYP132A localizes to the PM and the PBM, respectively. While SYP132C is required for protein secretion to extracellular space, SYP132A is involved in protein secretion to the symbiosome. SYP132A likely forms a complex with VAMP721d/e to mediate the fusion of symbiosome-targeting secretory vesicles with the PBM.

## 2.5 Materials and Methods

### 2.5.1 Molecular Cloning

For the GFP–SYP132 fusions, the protein-coding sequences of *SYP132A* and *SYP132C* were first cloned into the *pCR8/GW/TOPO* vector (Life Technologies), and introduced into *pDNF1–pMDC43* using Gateway recombination facilitated by the LR clonase (Life Technologies). The original *pMDC43* vector has a *35S* promoter, which was replaced by a 3 kb *DNF1* promoter (Curtis and Grossniklaus, 2003).

For RNA interference, the regions between 3049–3474 bp and 5410–5659 bp of the *SYP132* genomic sequence (downstream of the start codon), as well as the region between 345–594 bp of the *SYP132* coding sequence (for *SYP132C*, *SYP132A* and *SYP132*, respectively) were first cloned into the *pCR8/GW/TOPO* vector and recombined into the *pHellsGate8* vector modified with a *35S* promoter-driven mCherry fluorescent marker for plant transformation (Helliwell and Waterhouse, 2005).

All constructs were transformed into the ARqua1 strain of *Agrobacterium rhizogenes* through electroporation. The presence of the target plasmids was checked by PCR and sequencing.

### 2.5.2 Plant Growth Conditions and Inoculation

All experiments were performed on *M. truncatula* ecotype A17. Plants were grown under a 16 h/22 °C–8 h/18 °C light/dark cycle.

For nodulation, *Sinorhizobium meliloti* strain *Rm1021 pHemA::mCherry* was used (Leong et al., 1985). The bacteria cells used for inoculation were suspended in 1/2 BNM liquid medium (2.0 mM Ca(SO<sub>4</sub>)<sub>2</sub>, 0.5 mM KH<sub>2</sub>PO<sub>4</sub>, 0.5 mM Mg(SO<sub>4</sub>)<sub>2</sub>, 50 μM Na<sub>2</sub>EDTA, 50 μM FeSo<sub>4</sub>·7H<sub>2</sub>O, 16 μM ZnSo<sub>4</sub>·7H<sub>2</sub>O, 50 μM H<sub>3</sub>BO<sub>3</sub>, 50 μM MnSO<sub>4</sub>, 1 μM Na<sub>2</sub>MoO<sub>4</sub>·H<sub>2</sub>O, 0.1 μM CuSO<sub>4</sub>, 0.1 μM CoCl<sub>2</sub>·6H<sub>2</sub>O, and 2.0 mM MES·KOH (pH = 6.5)) to an optical density (600 nm) of 0.05. Ten

milliliters of the liquid suspension was used for each plant. Phenotypes were checked 21 days post inoculation (dpi).

### **2.5.3 Hairy Root Transformation**

The transformation of plasmids into plants was carried out by hairy root transformation (Haney and Long, 2010). The seeds of *A17* plants were sterilely germinated to obtain seedlings. The root tips of 1 day old seedlings were cut with a razor blade and incubated with *A. rhizogenes* strain ARqua1 containing the desired plant expression vector on Fahraeus medium (FÅHRAEUS, 1957) for one week. After all emerging roots were removed; plants were transferred to selective media with either 25 mg/L kanamycin (remained on selective media for 15 days) or 10 mg/L hygromycin (for 10 days). Then, the plants were transferred to a mixture of turface and sand (1:1) with 1/2× Gamborg's B5 Basal Salt medium (Sigma-Aldrich) for one week. Transgenic roots were selected based on antibiotic resistance or mCherry fluorescence.

### **2.5.4 Protein Analysis**

Nodules from transformed roots were harvested into a precooled extract solution containing 0.5 M sucrose, 10 mM DTT, 1% V/V protease inhibitor cocktail (P8340, Sigma-Aldrich), and 50 mM Tris-HCl, pH 7.4. Tissues were ground by hand using pestles, filtered through two layers of Miracloth (Calbiochem), and then centrifuged to obtain protein extracts. These extracts were boiled in 1× Laemmli buffer and proteins were separated by SDS–PAGE. Proteins were then transferred to nitrocellulose membranes (Adventec) for immunoblotting analysis. The membrane was incubated with an anti-tubulin antibody (at 1:4,000 dilution, Sigma), anti-NifH antibody (1: 4,000, AgriSera) or anti-GFP antibody (1: 2,000, Life Technologies). A secondary antibody (anti-mouse or anti-rabbit) conjugated with horseradish peroxidase was incubated

with the membrane and signals were visualized by enhanced chemiluminescence (Thermo Scientific) on a G-box (New England Biogroup).

### **2.5.5 Microscopy**

In confocal microscopy, 21 dpi transgenic nodules were hand sectioned using double-edged razor blades and mounted on microscope slides. For SYTO9, nodule slices were stained in 5  $\mu$ M SYTO9 (Life Technologies) with 50 mM Tris-HCl (pH 7.0) and 25 mg ml<sup>-1</sup> sucrose for 15 min. For single symbiosome confocal microscopy, young nodules were ground in 10% freshly depolymerized paraformaldehyde in 1  $\times$  TBS (pH 7.5) and put on ice for 10 min. The suspension containing symbiosomes was mounted onto a microscope slide and observed immediately. Samples were observed under an Olympus FLUOVIEW FV1000 confocal laser scanning microscope. GFP/SYTO9 signal was detected using excitation with a 485-nm laser and emission with a 490–540 nm band pass filter. The mCherry signal was detected using excitation with a 587-nm laser and emission with a 575–675 nm band pass filter.

## CHAPTER 3

### DNF2: A NOVEL PI-PLC SECRETED TO THE SYMBIOSOME

The work presented in this chapter was performed in a collaboration with Tao He, Erik Limpens, Elena Federova, Ton Bisseling, Mary Roberts, and Anne Gershenson.

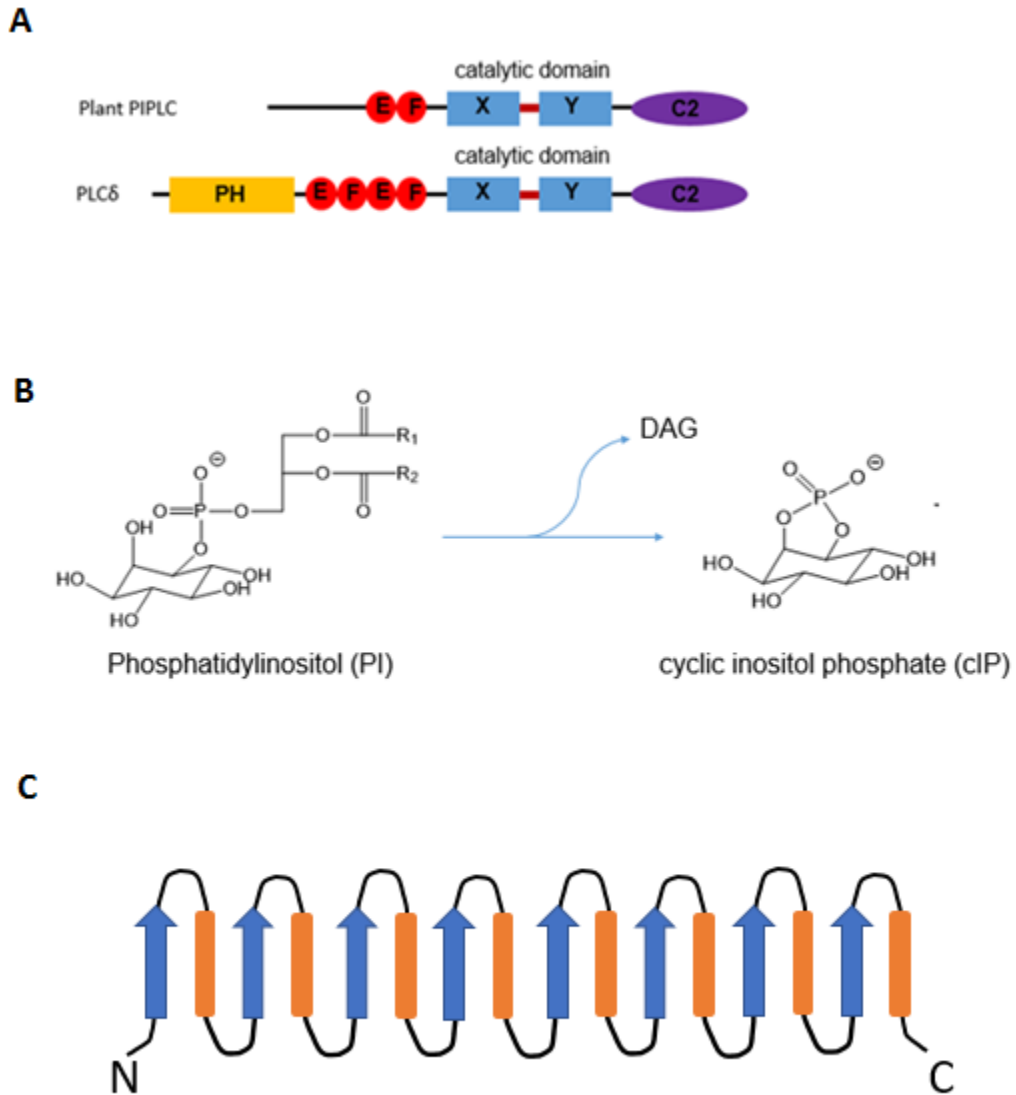
#### 3.1 Introduction

##### 3.1.1 Phosphatidylinositol-Specific Phospholipase C

Phospholipase C (PLC) constitutes a class of enzymes that cleave phospholipids, such as phosphatidylcholine (PC) and phosphatidylinositol (PI). Here, PI-specific PLCs will be discussed. Eukaryotic PI-PLCs hydrolyze phosphatidylinositol-4,5-bisphosphate (PIP<sub>2</sub>) to inositol-1,4,5-trisphosphate (IP<sub>3</sub>) and DAG, which initiate intracellular calcium release and protein kinase C activation, respectively. Eukaryotic PI-PLCs are classified into six families (PLC $\beta$ ,  $\gamma$ ,  $\delta$ ,  $\epsilon$ ,  $\zeta$  and  $\eta$ ) based on their domain structure. Most of eukaryotic PI-PLCs contain a PH (pleckstrin homology) domain, an EF hand motif for Ca<sup>2+</sup> binding, a C2 domain mediating interactions between Ca<sup>2+</sup> and lipids (or proteins) and a catalytic (XY) domain, which consists of a X-domain containing catalytic residues and a Y-domain involved in substrate recognition (Goñi et al., 2012) (Fig. 3.1A). PI-PLCs are also found in bacteria. The ones from *Bacillus cereus*, *Bacillus thuringiensis* and *Listeria monocytogenes* have been characterized (Griffith and Ryan, 1999). Bacterial PI-PLCs catalyze the hydrolysis of free phosphatidylinositol (PI) to DAG and inositol-1,2-(cyclic)-phosphate (cIP) and 1-inositol phosphate (IP) (Fig. 3.1B). In addition, bacterial PI-PLCs, but not eukaryotic PLCs, can process glycosylphosphatidylinositol (GPI) anchors, a glycolipid structure added to the C-terminus of proteins that stably anchors proteins with the appropriate signal (GPI anchor attachment site) to the outer leaflet of the cell membrane (Goñi et al., 2012). Bacterial and eukaryotic PI-PLCs both possess a distorted ( $\beta\alpha$ )<sub>8</sub>-barrel, or TIM barrel, (Fig. 3.1C) as a structural motif with a surprisingly large structural similarity in their first half of the TIM

barrel, including the catalytic residues especially two essential histidine residues, and a much weaker similarity in the second half. Unlike eukaryotic PI-PLCs, bacterial PI-PLCs do not require  $\text{Ca}^{2+}$  for their activity, explained by the presence of an arginine residue on the  $\text{Ca}^{2+}$  binding site of eukaryotic ones (Heinz et al., 1998).

All plant PI-PLCs identified so far are members of the  $\zeta$  class and possess all the domains except the PH domain. The interaction of plants with both pathogens and symbionts, as well as hyperosmotic stress or drought have been reported to induce PI-PLC activation in plants (Chen et al., 2011).



**Figure 3.1 Phosphatidylinositol-specific Phospholipase C**

**(A)** Most eukaryotic PI-PLCs contain a PH (pleckstrin homology) domain, an EF hand motif for  $\text{Ca}^{2+}$  binding, a C2 domain mediating interactions between  $\text{Ca}^{2+}$  and lipids (or proteins) and a catalytic (XY) domain. Plant PI-PLCs do not possess a PH domain and their EF hand motifs are smaller. **(B)** Bacterial PI-PLCs catalyze the hydrolysis of free phosphatidylinositol (PI) to DAG and inositol-1,2-(cyclic)-phosphate (cIP). **(C)** Schematic representation of the topology of the TIM barrel. While blue arrow indicates  $\beta$ -strands, orange rectangular indicates  $\alpha$ -helices.

### 3.1.2 Glycosylphosphatidylinositol (GPI)- Anchored Proteins

About 0.5% of cellular proteins in eukaryotes are glycosylphosphatidylinositol (GPI)-anchored.

GPI is a glycolipid structure post-translationally added to the C-terminus of proteins that serves to anchor proteins to the outer leaflet of the cell membrane (Paulick and Bertozzi, 2008).

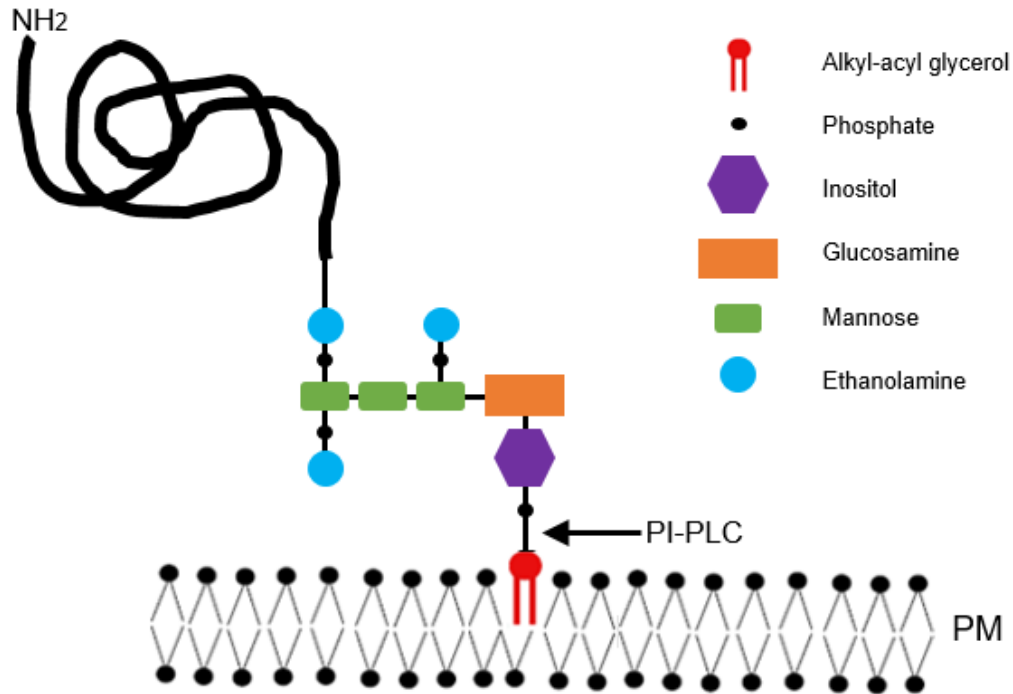
The GPI anchor possess a common conserved core. A phosphoethanolamine linker connects the  $\alpha$ -carboxyl group at the C-terminus of a protein to a trimannosyl-non-acetylated glucosamine (Man3-GlcN). This glycan is linked to phosphatidylinositol (PI), attaching the GPI to the cell membrane (Fig. 3.2). Based on the protein and species of origin, the conserved core undergoes extensive modifications during secretion from the cell: the mannose residues can have additional sugars or ethanolamine; the inositol ring can be acetylated; the length, saturation, and hydroxylation of the fatty acids and the types of linkage to the glycerol backbone can differ; diacylglycerol can be replaced with ceramide (Mayor and Riezman, 2004).

The assembly of GPI anchor occurs in the ER (Paulick and Bertozzi, 2008). The synthesis begins on the cytoplasmic surface with the transfer of N-acetylglucosamine (GlcNAc) to PI by a multi-enzyme complex, which is composed of PIG-A, PIG-C, PIG-H, PIG-P, GPI1, DPM2 and PIG-Y in mammalian cells. The de-N-acetylation of by PIG-L follow this reaction. The completed GPI intermediate is translocated to the luminal face of the ER membrane by flippases. The acylation of the inositol moiety by PIG-W generates GlcN-acyl-PI, to which three mannose and three ethanolamine phosphates are added by sequential enzymatic reactions. The synthesized moiety is transferred to the C-terminus of a protein with the appropriate signal -GPI anchor attachment site- by the GPI-transamidase, which is a multiprotein complex composed of PIG-K, PIG-T, GPAA1, PIG-S and PIG-U subunits (Taylor and Hooper, 2011).

GPI-anchored proteins play roles in diverse biological processes such as immune responses, cellular communication, membrane protein transport, cell adhesion, and signal transduction

(Paulick and Bertozzi, 2008). Deficiency in GPI-anchor biosynthesis in mice causes embryonic death (Nozaki et al., 1999). The parasitic protozoa, Trypanosomas, possess GPI-anchored variant surface glycoproteins (VSGs) on their surfaces that are required to evade host surveillance (Ferguson, 1999). In yeast, GPI-anchored proteins are involved in targeting mannoproteins to incorporate into the cell wall, which is essential for cell wall biosynthesis, and thereby survival. In addition, GPI biosynthesis is required for cell-wall synthesis, morphogenesis, and pollen tube development in Arabidopsis (Ferguson et al., 2009).

(GPI)-PLCs can cleave the GPI-anchor and release proteins from the membrane. GPI-PLC hydrolysis of the phosphodiester between inositol and lipid part of the anchor produces a free 1,2-diacylglycerol and glycopeptide-bound inositol cyclic-1,2-phosphate. GPI-anchor variants with ceramide in their structures instead of diacylglycerol show resistance to PI-PLC-mediated cleavage (McConville and Ferguson, 1993). GPI-anchor- cleaving PI-PLCs have not been identified in plants and animals although many GPI-anchored proteins exist (Eisenhaber et al., 2001).



**Figure 3.2 Structure of the GPI-anchor**

The GPI anchors possess a common conserved core. A phosphoethanolamine linker connects the  $\alpha$ -carboxyl group at the C-terminus of protein to a trimannosyl-non-acetylated glucosamine (Man<sub>3</sub>-GlcN). This glycan is linked to phosphatidylinositol (PI) attaching the GPI to the cell membrane. GPI-anchored proteins can be cleaved by PI-PLC. The arrow indicates the site of cleavage.

### 3.1.3 Research Question

Nitrogen-fixing symbiosis is established and maintained by host proteins destined to the symbiosomes by a nodule-specific protein secretory pathway. Targeting towards the symbiosome is associated with an N-terminal signal sequence cleaved by a nodule-specific signal peptidase complex in the ER. Several proteins with such signals have been identified, implying the existence of various symbiosome-localized host proteins involved in symbiotic nitrogen fixation. Some are well-characterized, such as NCR peptides mediating bacteroid differentiation. Nevertheless, many other symbiosome-targeted proteins and their roles in nitrogen-fixing symbiosis have yet to be determined.

Here, I identified a secreted host protein, DNF2, required for the survival of intracellular symbiotic bacteria. I found that DNF2 is a PI-PLC enzyme that cleaves GPI-anchored proteins inside the symbiosomes.

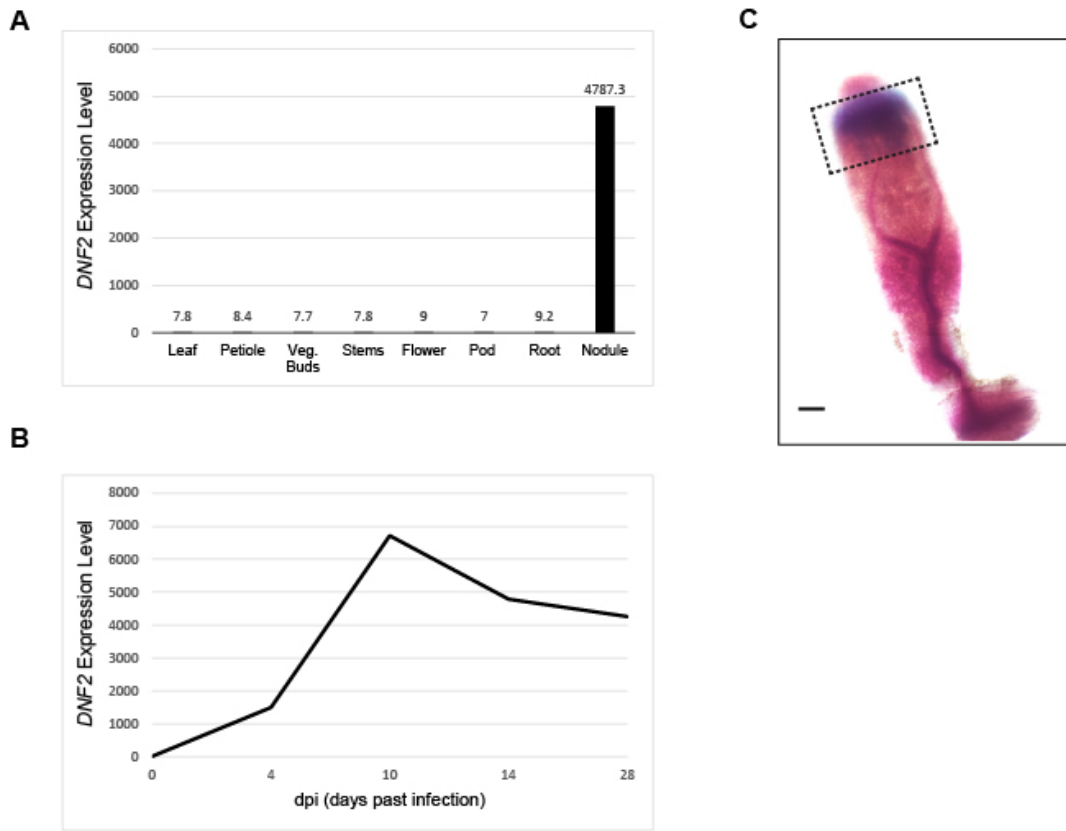
### 3.2 Experimental Strategy

In this study, I found that a symbiosome-localized host secretory protein, named DNF2, is required for the survival of intracellular bacteria inside *M. truncatula* nodule cells. I also discovered that DNF2 is a PI-PLC enzyme cleaving GPI-anchored proteins on the peribacteroid membrane. First, I studied the *dnf2* phenotype to ascertain the role of this protein in nodule development. I further analyzed the *DNF2* expression using the *M. truncatula* gene expression databases and GUS-staining. Then, I investigated the subcellular localization of DNF2 under a confocal microscope and whether DNF2 is a substrate of the nodule-specific secretory pathway. Finally, I examined the function of DNF2 using biochemical assays.

### 3.3 Results

#### 3.3.1 DNF2 is Specifically Expressed in Nodules

I investigated the expression of *DNF2* using the *Medicago truncatula* gene expression atlas (Benedito et al., 2008). *DNF2* is specifically expressed in nodules (Fig. 3.3A) and is induced with rhizobial infection (Fig. 3.3B), indicating that *DNF2* has a nodule-specific function. To determine the spatial pattern of *DNF2* expression, I fused the *DNF2* promoter to a GUS reporter and expressed this construct in nodules. *DNF2* was expressed in the region including infection and transition zone (Fig. 3.3C). Expression of *DNF2* in the area where the bacteria infect host cells and differentiate into bacteroid supported our claim that *DNF2* is essential to maintain intracellular bacteria inside symbiosomes.

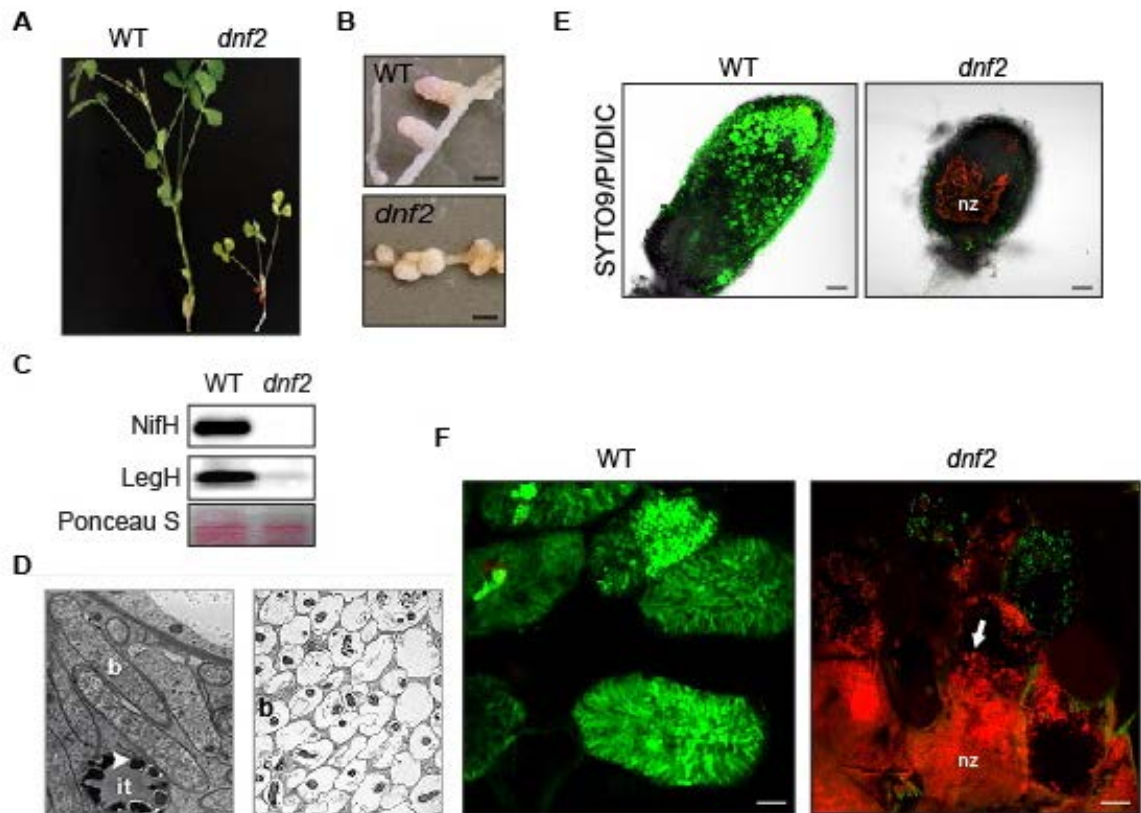


**Figure 3.3** *DNF2* expression pattern.

**(A)** *DNF2* expression level in different tissues. **(B)** Temporal profile of *DNF2* expression in nodules. Values in (A) and (B) are levels of the Affymetrix probe signal based on microarray data from the MtGEA database (<http://mtgea.noble.org>). **(C)** *DNF2* promoter activity in 21 dpi nodules inoculated with *Rm1021* carrying *hemA::lacZ*. A *DNF2 promoter::GUS* transgenic nodule was stained for GUS activity, photographed, briefly fixed, sectioned manually, and then stained with Magenta-Gal to reveal bacterial cells. GUS activity is shown in blue. Dashed line box represents the infection zone plus the nitrogen fixation zone of the nodule. Scale bar, 100  $\mu$ m.

### 3.3.2 DNF2 is Essential for Nodule Development

To investigate the role of the DNF2 protein in symbiotic nitrogen fixation, I inoculated *dnf2* mutant plants, as well as WT plants, with rhizobia. Three weeks after the treatment, the *dnf2* plants were smaller than WT and had yellow leaves, meaning they were nitrogen-starved (Fig. 3.4A). While WT plants had large and pink-colored nodules due to the presence of leghemoglobin, the *dnf2* plants developed small and white nodules (Fig. 3.4B), which were rescued by the ectopic expression of DNF2 coding sequence under the control of its own promoter (Fig. 3.5). Consistent with their phenotypes, the *dnf2* nodules did not produce NifH or leghemoglobin (Fig. 3.4C). I further examined WT and *dnf2* symbiosomes with electron microscopy. In WT, I observed fully differentiated bacteria filling the symbiosomes, resulting in a very thin or not visible peribacteroid space. The *dnf2* nodules however contained symbiosomes with undifferentiated, small bacteria and the peribacteroid space appeared very large (Fig. 3.4D). To understand whether the bacteria are viable, I stained WT and *dnf2* nodules with SYTO9 and propidium iodide, which label alive bacteria green and dead bacteria red, respectively (Boulos et al., 1999). WT nodules, as expected, contained cells infected with differentiated alive bacteria. I observed undifferentiated live, as well as dead bacteria in *dnf2* nodules, and a necrotic zone with smeared, red autofluorescence, as a consequence of accumulation of phenolics, implying that the defense response is activated in the absence of DNF2 protein (Fig. 3.4E&F). Overall, I discovered that DNF2 prevents the internalized bacteria from being degraded by host defenses.



**Figure 3.4 DNF2 is required for the survival of rhizobial bacteria inside host cells.**

Growth (A) and nodule (B) phenotypes of WT and *dnf2*, inoculated with *Sinorhizobium medicae* strain ABS7, at 21 days post inoculation (dpi). Scale bar, 2 mm. (C) Detection of leghemoglobin and NifH in the nodule extracts of WT and *dnf2*. (D) Transmission electron microscope (TEM) image of infection thread (it) and bacteroids surrounded by a symbiosome membrane in WT and of non-differentiated bacteria in symbiosomes of *dnf2*. Scale bar, 1  $\mu$ m. Live/Dead staining of nodules (Scale bar, 0.1 mm) (E) and bacteroids (Scale bar, 10  $\mu$ m) of WT and *dnf2* (F). Green color from SYTO 9 indicates live bacteria and red color from propidium iodide (PI) indicates dead bacteria. Nz stands for necrotic zone.



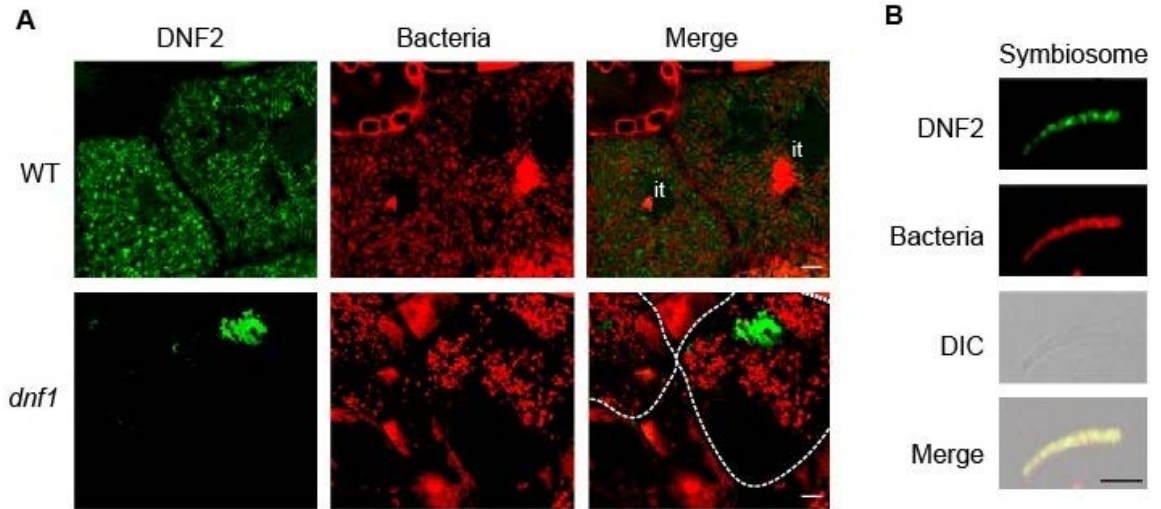
**Figure 3.5 Complementation of *dnf2*.**

Complementation of the *dnf2* phenotype by transformation with DNF2, fused to GFP, under the control of its own promoter. Scale bar, 2 mm.

### 3.3.3 DNF2 Localizes to the Symbiosome

To determine the subcellular localization of DNF2, I fused DNF2 to a GFP protein and introduced it in nodules under the control of its own promoter. I then examined DNF2-expressing transgenic nodules and, isolated symbiosomes under a confocal microscope. I found that DNF2 localized to the symbiosome, but not the infection thread, implying that DNF2 functions inside symbiosomes after the internalization of rhizobia (Fig. 3.6).

Host secretory proteins are delivered to the symbiosome by a nodule-specific protein secretory pathway, which is dependent on a symbiotic t-SNARE protein, SYP132A (Pan and Oztas et al., 2016) and the DNF1 protein, a subunit of the signal peptidase complex located in the ER (Wang et al., 2010). The presence of a signal peptide at the amino terminus suggests that DNF2 is a cargo protein of the nodule-specific protein secretory pathway. To test this, I expressed GFP-fused DNF2 protein in *dnf1* nodules. The absence of DNF1 aborted the delivery of DNF2 to the symbiosome and caused DNF2 to be trapped in ER-like structures (Fig. 3.6). I concluded that DNF2 is localized to the symbiosome through the nodule-specific protein secretory pathway.



**Figure 3.6 DNF2 is targeted to the symbiosome through the nodule-specific protein secretory pathway.**

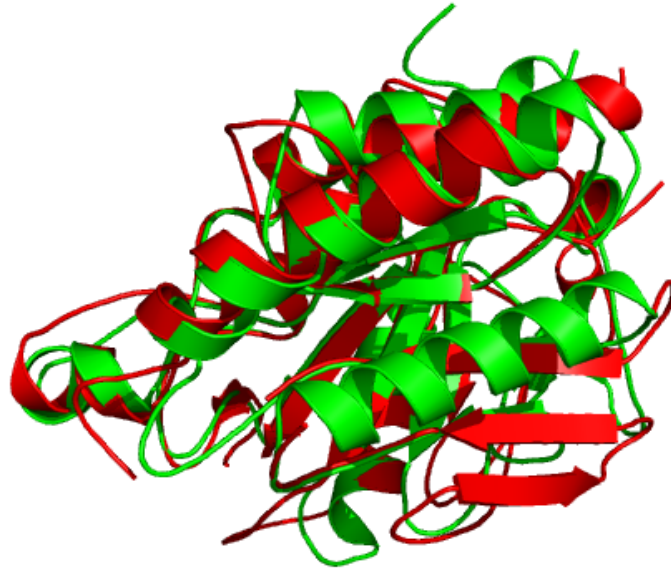
**(A)** Localization of DNF2-GFP translational fusion in nodule cells of WT and *dnf1*, inoculated with *Rhizobium meliloti* 1021 carrying *hemA::mCherry*. The dashed lines indicate membrane borders. Scale bar, 5  $\mu\text{m}$ . It: infection thread. **(B)** The same markers in isolated symbiosomes of WT. Scale bar, 5  $\mu\text{m}$ .

### 3.3.4 DNF2 is a Phosphatidylinositol-Specific Phospholipase C

I predicted that mature DNF2 is structurally similar to bacterial phosphatidylinositol-specific phospholipase C although they possess no sequence similarity (Fig. 3.7). To assess if DNF2 can cleave phospholipids, I produced a recombinant DNF2 protein in *E. coli*. I found that recombinant DNF2 cleaves phosphatidylinositol but no other phospholipids (Fig. 3.8A).

While free PI in eukaryotic cells is predominantly found on the cytosolic side of the lipid bilayer (Yamaji-Hasegawa and Tsujimoto, 2006), DNF2 localizes to the peribacteroid space, which is topologically extracellular. Among PI species, glycosylated phosphatidylinositol (GPI) molecules exist on the outer leaflet of the plasma membrane (Paulick and Bertozzi, 2008). I therefore hypothesized that the biologically relevant substrates of DNF2 are GPI-anchored proteins (GAPs) on the bacteria-facing side of the peribacteroid membrane. To determine whether DNF2 cleaves symbiosome-localized GAPs, I isolated symbiosomes from WT and *dnf2* nodules and detected the existence of cleaved versions of GAPs using a cross-reacting determinant (CRD) antibody that specifically recognizes the exposed cyclic inositol phosphate moiety once a PI-PLC cleaves a GAP (Hooper, 2001). I detected two major bands in WT and, also in *dnf7*, which produces white and small nodules lacking nitrogen fixation, like *dnf2*. This indicates that more than one GAP exist inside WT symbiosomes and that they are cleaved. However, the major bands were absent in *dnf2* symbiosomes (Fig. 3.8B). The CRD antibody detected a purified cleaved GAP from *Trypanosoma brucei*, called sVSG, showing that the antibody works properly. That I detected the same levels of Nodulin 26 protein, a channel protein on the peribacteroid membranes, in all symbiosome fractions proved that the compromised symbiosomes did not cause the lack of the major bands in *dnf2*. Overall, our results support that DNF2 is essential for the cleavage of GAPs inside the symbiosomes.

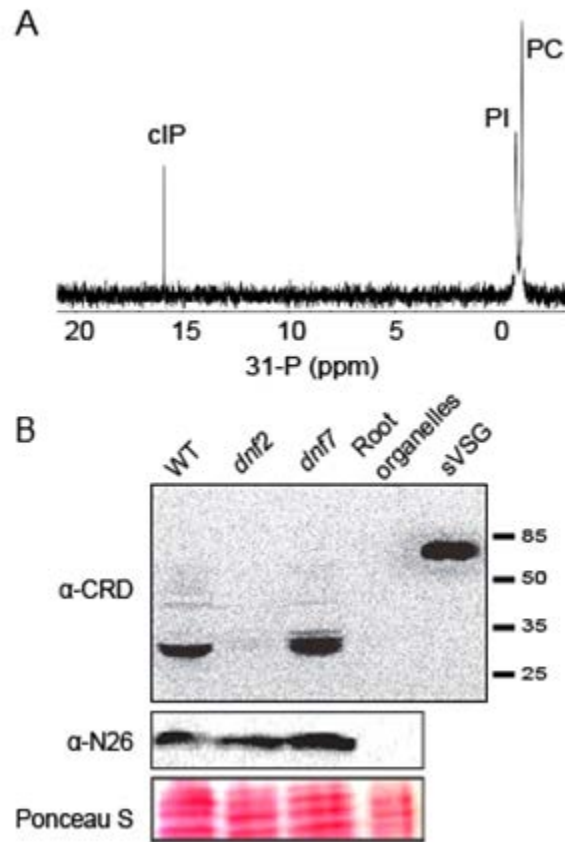
To further test whether GAPs are processed by DNF2, I prepared a GPI-anchored GFP protein under the control of a *DNF2* promoter. I confirmed that this recombinant protein is expressed and secreted to the symbiosome in WT and *dnf2* nodule cells (Fig. 3.9). I further used Triton X-114 extraction to determine the levels of cleaved and intact GAPs in symbiosome extracts. Triton X-114 treated protein extracts are partitioned into two different phases: aqueous and detergent. While aqueous phase contains cleaved GAPs, detergent phase includes intact GAPs, due to the presence of phosphatidylinositol in their structure (Doering et al., 2001). After Triton X-114 extraction, most GFP in WT is enriched in the aqueous phase, indicating that its GPI anchor is cleaved. In contrast, in *dnf2* most GFP remains in the detergent phase, demonstrating that DNF2 is required for a GPI-anchored symbiosome protein to be solubilized (Fig. 3.10). To conclude, DNF2 is a symbiosome-localized PI-PLC that cleaves GAPs.



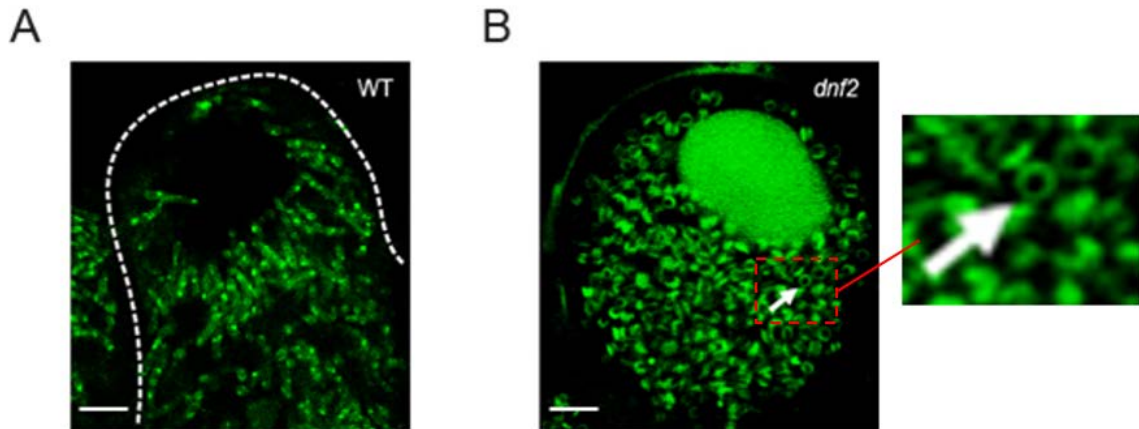
RMSD	1.51
Alignment length	319
Aligned residues	214
Fraction aligned	67.08 %

**Figure 3.7 Protein modeling of DNF2.**

The structure of the DNF2 protein (19-334 aa) without the signal peptide was predicted by using PHYRE (Kelley and Sternberg, 2009), and superimposed with phosphatidylinositol-specific phospholipase C of *Listeria monocytogenes* (PDB: 2plc) by using IPBA webserver ([http://www.dsimb.inserm.fr/dsimb\\_tools/ipba/index.php](http://www.dsimb.inserm.fr/dsimb_tools/ipba/index.php)). Green: phosphatidylinositol-specific phospholipase C of *Listeria monocytogenes*. Red: DNF2 protein. RMSD: root-mean-square deviation.

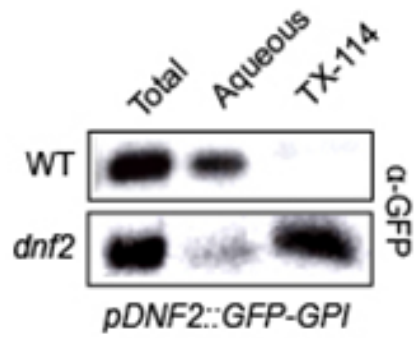


**Figure 3.8 DNF2 is a phosphatidylinositol-specific phospholipase C, cleaving GPI-anchored proteins. (A)** <sup>31</sup>P NMR spectrum (242.8 MHz), obtained at 25°C, showing the cleavage of PI by purified DNF2-intein. The reaction was carried out with 2mM PI/ 2mM PC in 50mM HEPES, pH 7.4, at 25 °C. <sup>31</sup>P resonances are identified as PI at -0.81ppm, PC at -1.03ppm and cIP at 15.93ppm. **(B)** Detection of cleaved GPI-anchored proteins and Nodulin 26, a nodule specific symbiosome membrane protein, in symbiosome extracts of WT, *dnf2* and *dnf7* by CRD antibody and Nodulin 26 antibody, respectively. sVSG, a purified cleaved GPI-anchored protein from *Trypanosoma brucei*. Ponceau S-stained membrane is shown as a loading control



**Figure 3.9 Localization of GPI-anchored GFP protein.**

*GFP* coding sequence, under the control of *DNF2* promoter, with *ENOD16* signal sequence and GPI attachment site was introduced into WT (**A**) and *dnf2* (**B**) roots. The transgenic nodules were viewed under confocal microscope. The dotted line represents the cell boundary. Scale bar, 10  $\mu$ m. The arrow indicates a dark bacteroid surrounded by GFP signal inside symbiosome.



**Figure 3.10 DNF2 cleaves GPI-anchored GFP protein**

The levels of cleaved and uncleaved GPI-anchored GFP protein in WT and *dnf2* by TX-114 partitioning and Western Blot using a GFP antibody. The aqueous phase contains cleaved forms, while TX-114 phase includes uncleaved forms.

### 3.4 Discussion

In this study, I identified a host protein called DNF2 that is secreted to the symbiosome through the nodule-specific protein secretory pathway. DNF2 is only expressed in nodules in the region where bacteroid differentiation occurs, and is induced in response to rhizobial infection. The absence of DNF2 in nodules caused the death of bacteria inside symbiosomes, as well as the accumulation of phenolic compounds, a sign of a defense response, indicating that DNF2 prevents the intracellular bacteria from being degraded by the host cell.

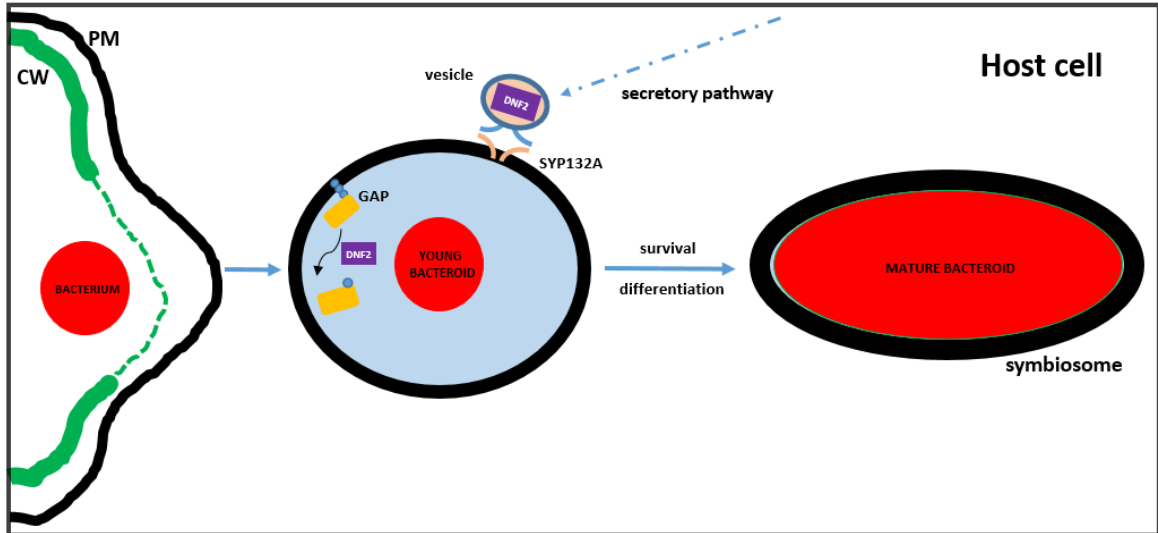
I discovered that DNF2 is a (G)PI-PLC capable of cleaving GAPs on the peribacteroid membrane. Even though plants possess a high number of distinct GAPs, a PI-PLC capable of processing them has not been reported before. Thus, DNF2 is the first (G)PI-PLC identified in plants.

Many homologs of DNF2 are found across the plant kingdom. For instance, Arabidopsis has five DNF2 homologs. The fact that Arabidopsis plants are not able to form a symbiotic interaction with rhizobia, or with fungi, indicates that DNF2 homologs participate in distinct biological processes in plants, not only in symbiotic association. Beyond that, DNF2-like proteins are widespread across several kingdoms. This family of PLCs therefore may represent a general mechanism to regulate the functions of GAPs.

On the basis of our results, two major GAPs with different sizes (Fig 3.8B), presumably with distinct functions, exist inside symbiosomes, showing that DNF2 has more than one substrate. A symbiosome-localized GAP has not been identified before. The induction of host defense response in the absence of DNF2 implies the presence of a GPI-anchored defense receptor inactivated by DNF2. Furthermore, DNF2 likely cleaves various GAPs simultaneously, thereby influencing distinct processes in symbiosomes. This suggests that DNF2 likely act as a checkpoint protein during bacterial development that influences the transition of intracellular bacteria to a bacteroid. By which factors and how DNF2 is regulated should be answered by future studies.

In addition to processing GAPs, DNF2 hydrolyzes the free PI to lipid messengers called DAG and cIP, proposing that free PI on the peribacteroid and bacterial membrane might be a second substrate of DNF2 inside symbiosomes. PI is mostly found on the inner leaflet of the peribacteroid membrane and requires flippase activity to face symbiosome space where DNF2 localizes. In addition, PI is a minor component of the rhizobial membrane (López-Lara et al., 2003). Therefore, it is unlikely that DNF2 processes free PI in the symbiosome, requiring to be proven experimentally.

Here, I introduced a model explaining the DNF2 function in symbiosis. Once bacteria are endocytosed, DNF2 is secreted to the symbiosome via a nodule-specific secretory pathway. From the peribacteroid membrane, DNF2 releases the GAPs participating in the bacterial maintenance and development (Fig. 3.11). DNF2 acts as a host signal activating or inhibiting the functions of symbiosome-localized GAPs.



**Figure 3.11 DNF2 is required to maintain intracellular bacteria in nitrogen-fixing symbiosis**

With the bacterial internalization, DNF2 is secreted to the symbiosome via symbiosome-targeted secretory vesicles interacting with the SYP132A t-SNARE on the peribacteroid membrane. Inside symbiosomes DNF2 cleaves the GAPs involved in the maintenance and development of the intracellular symbiont.

### **3.5 Materials and Methods**

#### **3.5.1 Plant Growth Conditions and Inoculation**

The legume model plant *Medicago truncatula*, ecotypes *Jemalong A17* and *R108*, were used for this study. Plants were inoculated with *Sinorhizobium medicae* strain ABS7 or ABS7-HemA::LacZ (Leong et al., 1985) or with *Sinorhizobium meliloti* strain *Rm1021-mCherry* in ½X BNM medium to obtain nodules (21 dpi) in different experiments. Plants were grown under a 16-hour, 22°C/8-hour, 18°C light/dark cycle.

#### **3.5.2 Molecular Cloning**

For the DNF2-GFP construct, the *DNF2* coding sequence was cloned into the pCR8/GW/TOPO vector (Life Technologies), and recombined with the vector including 1.6 kb promoter of the *DNF2* gene using the Gateway technology (Invitrogen). For the GPI-anchored GFP construct, the last 170 bp of *ENOD16* coding sequence was recombined with pMDC43 (Curtis and Grossniklaus, 2003) including 1.6 kb of *DNF2* promoter and the first 78 bp of the *DNF2* coding sequence.

#### **3.5.3 Hairy Root Transformation**

The introduction of plasmids into plants was carried out by hairy root transformation (Haney and Long, 2010). The seeds of *A17* plants were sterilely germinated to obtain seedlings. The root tips of one day-old seedlings were cut with a razor blade and incubated with *A. rhizogenes* strain ARqua1 containing the desired plant expression vector on Fahraeus medium (FÅHRAEUS, 1957) for one week. After all emerging roots were removed; plants were transferred to selective media with either 25 mg/L kanamycin (remained on selective media for 15 days) or 10 mg/L hygromycin (for 10 days). Then, the plants were transferred to a mixture of turface and sand (1:1) with 1/2× Gamborg's B5 Basal Salt medium (Sigma-Aldrich) for one week. Transgenic roots were selected based on antibiotic resistance or mCherry fluorescence.

### **3.5.4 Microscopy**

For confocal microscopy, nodules were hand-sectioned using double-edged razor blades and mounted on microscope slides. Samples were observed under an Olympus FLUOVIEW FV1000 confocal laser scanning microscope. GFP signal was detected using excitation with a 485 nm laser and emission with a 490-540 nm band pass filter, whereas mCherry signal was detected using excitation with a 587 nm laser and emission with a 575-675 nm band pass filter.

### **3.5.5 Live/Dead Staining**

For live/dead staining, nodule slices were stained in 5  $\mu$ M SYTO9 and propidium iodide (Life Technologies) with 50 mM Tris-HCl (pH 7.0) and 25 mg/ml sucrose for 15 min.

### **3.5.6 Protein Analysis**

Antibodies used in this study were diluted in TBS-Tween (1%) supplemented with 5% powdered milk or 3% Bovine Serum Albumin (BSA) as follows: anti-CRD (Hooper, 2001) 1/5000; anti-GFP (JL-8, Clontech) 1/10000; E6 (Weaver et al., 1994) 1/10000; anti-mouse (Sigma-Aldrich) 1/10000; anti-rabbit (GenScript) 1/10000; anti-NifH (1:2500); anti-Leghemoglobin (1:10000). The signals were detected by using ECL-detection kit (Thermo Scientific).

### **3.5.7 Microsome and Symbiosome Isolation**

The nodules were collected in extraction buffer (0.5 M sucrose, 10 mM DTT, 50 mM Tris-HCl pH 7.5) including protease inhibitor cocktail (Sigma P9599) on ice. The tissue was ground by mortar and pestle and filtered through four layers of Miracloth (Calbiochem), followed by centrifugation for 10 min at 200 x g to remove debris. The flow through was centrifuged for 15

min at 1600 x g to collect the rough symbiosomes (Rosendahl et al., 1992). The supernatant was centrifuged at 120,000 x g for 1 hour to obtain the microsomes (Catalano et al., 2004).

### **3.5.8 Triton X-114 Partitioning**

The symbiosome pellet was resuspended in Tris-buffered saline (TBS) (50 mM Tris-Cl, pH 7.5, 150 mM NaCl) and mixed with 1/5 volume Triton X-114 and incubated on ice for 15 min with occasional mixing. The cell mixture was incubated at 37°C for 2 min and centrifuged for 10 minutes at 1000 x g at room temperature. The water phase and TX-114 phase were collected to be analyzed (Doering et al., 2001).

### **3.5.9 GUS Reporter assay**

The protocol was modified from the Arabidopsis book by Weigel (Weigel and Glazebrook, 2002). Transgenic nodules expressing the *GUS* gene were collected and incubated in 2-2.5 ml staining buffer (50 mM sodium phosphate buffer pH 7, 0.2% Triton X-100, 3 mM potassium ferricyanide, 3 mM potassium ferrocyanide, 20% methanol) on ice. The samples were infiltrated under vacuum for 20 min and incubated at 37°C. Samples were hand-sectioned by razor blade and viewed under a light microscope at 4X magnification.

## CHAPTER 4

### DNF2 SUBSTRATES: LYM1 AND LYM2

#### 4.1 Introduction

##### 4.1.1 LysM Domain Containing Proteins

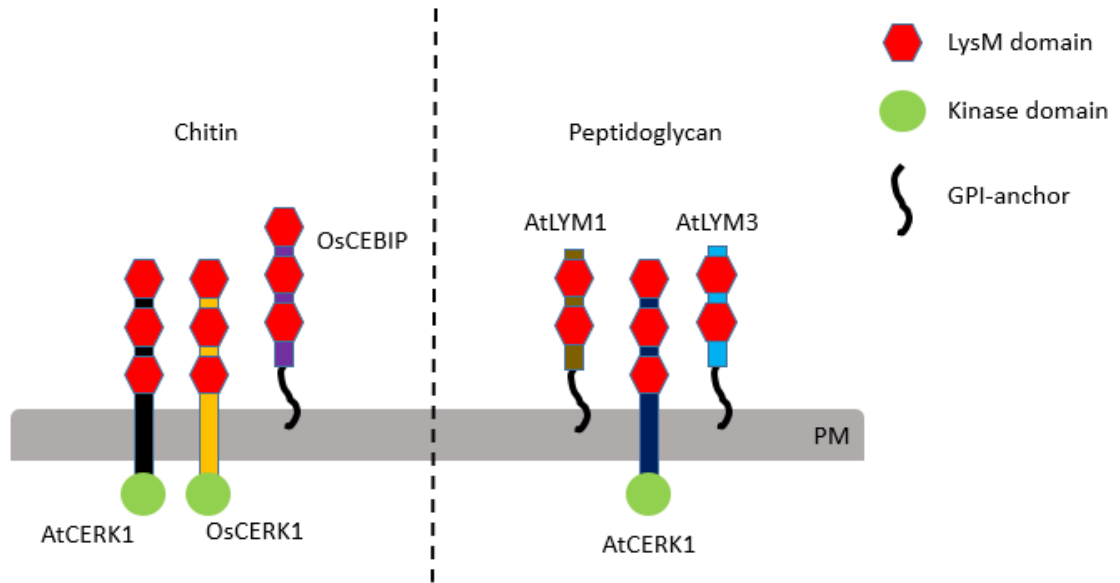
Plants associate with different populations of microbes including viruses, bacteria, and fungi (Mendes et al., 2011). Plants sense the microbe-derived signatures by plant pattern recognition receptors (PRRs) and discriminate between beneficial and detrimental microbes (Medzhitov and Janeway, 1997).

Two major types of PRRs exist in the plant innate immune system on the basis of whether they detect proteins or carbohydrates. While leucine-rich repeat receptor proteins (LRR-RPs) and leucine-rich repeat receptor kinases (LRR-RKs) interact with proteinaceous microbial signatures, such as bacterial flagellin and elongation factor Tu (Zipfel, 2008), lysin motif (LysM)-containing receptor proteins and receptor kinases bind N-acetylglucosamine (GlcNAc)-containing microbial glycans.

The LysM domain, containing one or more lysin motifs, is found in most living organisms except Archaea (Buist et al., 2008). A lysin motif typically contains from 44 to 65 amino acid residues and folds into  $\beta\alpha\alpha\beta$  secondary structure with two  $\alpha$  helices packing onto one side of a two-stranded antiparallel  $\beta$ -sheet (Bielnicki et al., 2006). LysM domains were originally identified in bacterial hydrolases that degrade the cell wall (Garvey et al., 1986). Since then, a variety of LysM domain proteins have been discovered in prokaryotes and eukaryotes.

Some plant LysM domain receptors induce immunity-associated defense responses upon recognition of bacterial peptidoglycans or fungal chitin (Fig. 4.1), whereas some detect Nod factors or Myc factors that trigger symbiotic interaction with rhizobia and fungi, respectively.

That these microbial glycans all possess a GlcNAc in their structures implies that different classes of LysM domain proteins evolved from a common ancestor binding to GlcNAc (Gust et al., 2012). Plant-specific LysM domain proteins are classified into three families: LysM receptor-like kinases (LysM-RLK), F-box LysM proteins and LYMs. LysM-RLKs are composed of one to three extracellular LysM domains, a transmembrane domain and an intracellular kinase domain. Among well-known LysM-RLKs is CERK1 of rice (*Oryza sativa*), which recognizes chitin and triggers the plant defense machinery against fungi (Kaku et al., 2006). F-box LysM proteins contain at least one F-box domain, and a protein-protein interaction motif, in addition to a LysM domain. F-box proteins are substrate-recognition components of SCF ubiquitin-ligase complexes marking proteins for ubiquitin-mediated degradation (Kipreos and Pagano, 2000). LYM proteins include a signal peptide, typically three LysM motifs and a GPI-anchor attachment site; however, they lack an intracellular signaling domain. The first LYM protein, named the chitin elicitor binding protein (CEBIP), was identified in rice (*Oryza sativa*) (Kaku et al., 2006). With the recognition of chitin, OsCEBIP activates plant defense response to fungal infection, which requires heterodimer formation with OsCERK1 (Shimizu et al., 2010). Other GPI-anchored LYM receptors, such as LYM1 and LYM3 in Arabidopsis, have similar structures to OsCEBIP. Nevertheless, rather than chitin, they bind peptidoglycan derived from the bacterial cell wall. The presence of both LYM1 and LYM3 are required to activate peptidoglycan-induced plant immunity against bacterial pathogens. LYM1 and LYM3 lack a signaling domain like OsCEBIP. Therefore, their functions in PGN-associated defense response depend on a LysM-RLK, called AtCERK1 (Willmann et al., 2011). Two CEBIP-like LYM receptors, named MtLYM1 and MtLYM2, were identified in *Medicago truncatula*. These proteins are glycosylated and GPI-anchored to the outer leaflet of the plasma membrane. (Fliegmann et al., 2011)



**Figure 4.1 LysM domain containing defense receptors**

Plant LysM domain containing receptors trigger a defense response recognizing bacterial peptidoglycans or fungal chitin. They can be tethered to the membrane via a transmembrane domain or GPI-anchor. Some LysM proteins have kinase activity. They can have different numbers of LysM domains.

#### 4.1.2 Research Question

Plants induce defense responses against invading organisms upon recognition. However, nitrogen-fixing symbiosis in which legumes harbor the rhizobia within their cells is exceptional. Rhizobial invasion occurs in many steps that are controlled by the host plant. Legumes first distinguish compatible rhizobia among root microbiota by their symbiotic receptors detecting Nod factors, and allow them to infect their roots. To move the rhizobia through root cortical cells, legumes form infection thread networks in which the bacteria grow and divide. Once the infection threads reach the root cortex, host cells engulf individual bacteria in a symbiosome. Host proteins that are delivered to the symbiosome facilitate the maintenance and differentiation of intracellular rhizobia. Presumably, legumes attenuate defense responses against the rhizobia in each step, otherwise they would recognize the invading bacteria as an enemy and destroy them. How legumes shut down their defense mechanisms without risking being susceptible to pathogens needs to be answered to understand better symbiotic nitrogen fixation.

Our studies demonstrated that DNF2 is required to inhibit host immunity against intracellular bacteria. DNF2 is a symbiosome-localized enzyme cleaving GAPs from the peribacteroid membrane. A recent study reported the presence of two LysM domain-containing receptors, called LYM1 and LYM2 in *Medicago truncatula*. LYMs are GPI-anchored proteins that recognize bacterial peptidoglycan, and activate the host defense system.

Here I found that LYM1 and LYM2 localize to the infection thread membrane. Once bacteria are endocytosed, DNF2 removes LYM1 and LYM2 from the peribacteroid membrane, hence inactivating the defense response.

## 4.2 Experimental Strategy

In this study, I discovered that two LysM-domain containing GPI-anchored proteins in *M. truncatula*, LYM1 and LYM2, are processed by DNF2 to shutdown host defense response inside symbiosomes, which is essential to maintain the intracellular bacterial symbiont. I first analyzed the expressions of *LYM1* and *LYM2* using the *M. truncatula* gene expression atlas. By Western Blot analysis, I assessed the production of LYM1 and LYM2 in nodules and whether these proteins are processed by DNF2. I further investigated the localization of LYM1 and LYM2 by confocal microscopy. Finally, I examined the roles of LYM1 or LYM2 in nodulation and the significance of their regulation by DNF2 for symbiotic nitrogen fixation.

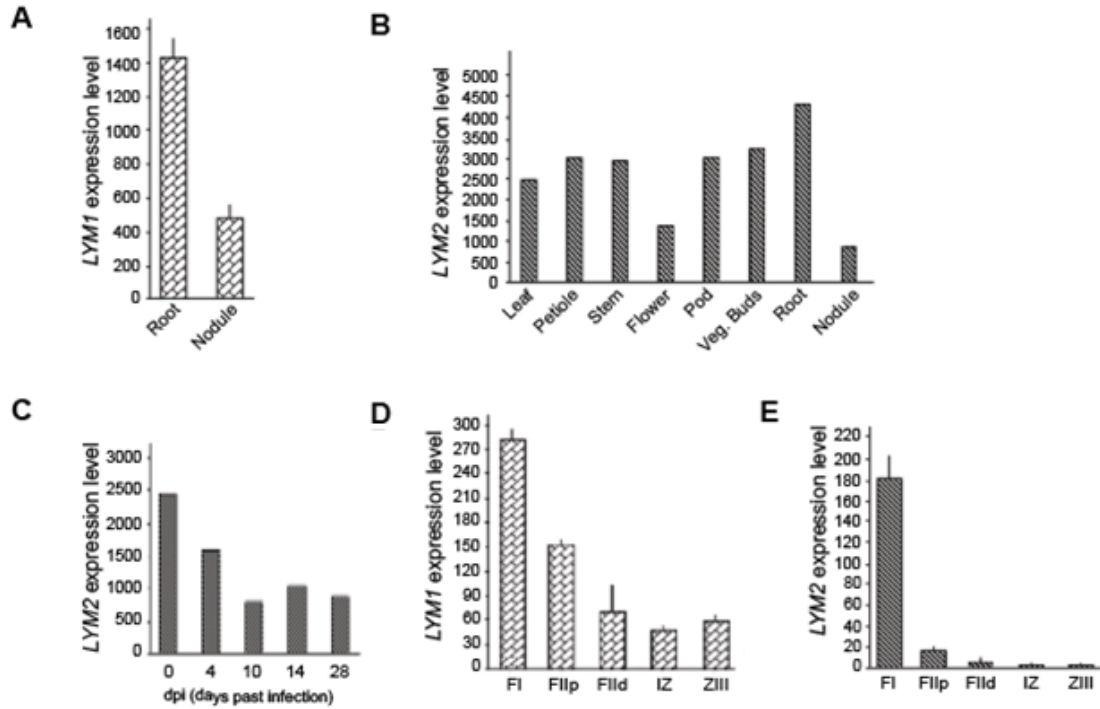
## 4.3 Results

### 4.3.1 LYM1 and LYM2 are found in nodules

I studied the expressions of *LYM1* and *LYM2* using the *Medicago truncatula* gene expression atlas (Benedito et al., 2008) and the Symbimics database (<https://iant.toulouse.inra.fr/symbimics/>). *LYM1* is expressed at higher levels in roots than nodules (Fig. 4.2A). Similarly, *LYM2* is transcribed in each tissue, while the expression level in nodules is the lowest (Fig. 4.2B). The decrease in *LYM2* expression with rhizobial inoculation (Fig. 4.2C), and the downregulation of *LYM1* and *LYM2* in the nodule zones housing the rhizobia (Fig. 4.2D&E) suggested that the presence of LYM1 and LYM2 may have negative impact on nodulation.

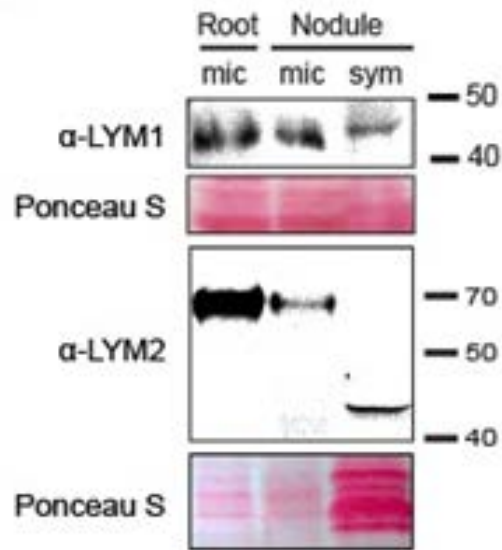
By using LYM1 and LYM2 antibodies, I detected both proteins in the symbiosome and on the plasma membrane, but with higher sizes than the expected molecular weights, consistent with a report claiming that LYMs are glycosylated. The presence of a symbiosome-specific form of LYM2, which is apparently distinct from its plasma membrane counterpart, suggests that this protein may have unique modifications and biological functions inside the symbiosomes (Fig

4.3). To verify this result, I obtained *Tnt1* insertional mutant lines for LYM1 and LYM2 in the *M. truncatula* R108 background, and confirmed the *Tnt1* insertion in the *LYM1* and *LYM2* genes by performing PCR with a *Tnt1* primer and primers designed for *LYM1* and *LYM2* genomic sequences (Fig 4.4). The bands detected by LYM1 and LYM2 antibodies were lost in *lym1* and *lym2* mutants, respectively, indicating that the antibodies specifically detect LYM1 and LYM2 (Fig. 4.5). The bands were also absent in root samples that underwent the identical symbiosome isolation procedure (root control), and in cultured rhizobia, confirming the presence of LYM1 and LYM2 in symbiosomes (Fig. 4.6).



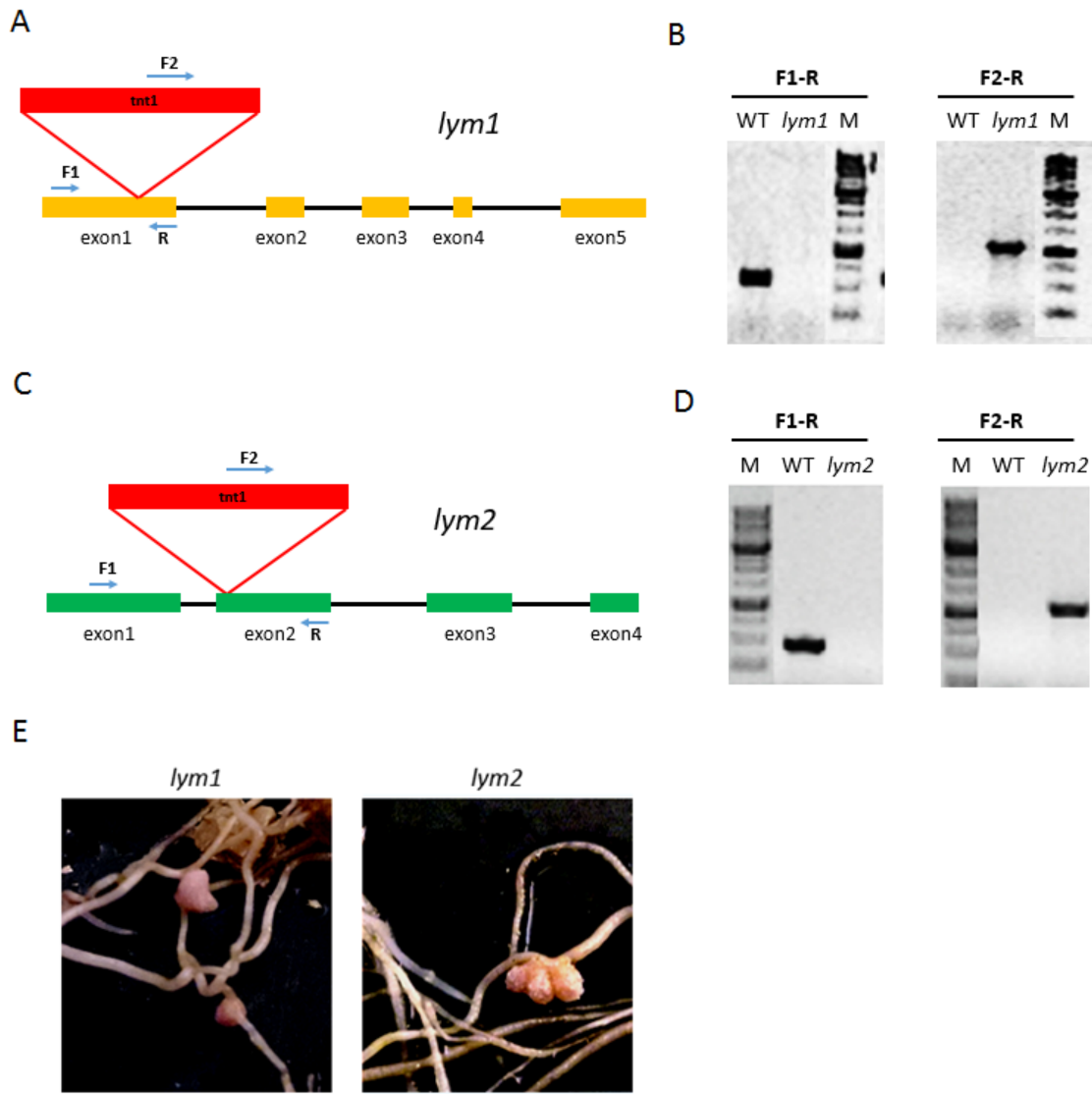
**Figure 4.2** *LYM1* and *LYM2* are downregulated in nodules

**(A)** *LYM1* expression level in root and nodule. **(B)** *LYM2* expression level in different tissues. **(C)** Temporal profile of *LYM2* expression in nodule. Spatial expression pattern of *LYM1* **(D)** and *LYM2* **(E)**. I, zone I; IId, zone II distal; IIp, zone II proximal; IZ, interzone II-III; III, zone III. Values in (B) and (C) are levels of Affymetrix probe signal based on microarray data from the MtGEA database (<http://mtgea.noble.org>). Values in (A), (D) and (E) are from RNA-seq data from Symbimics (<https://iant.toulouse.inra.fr/symbimics>).



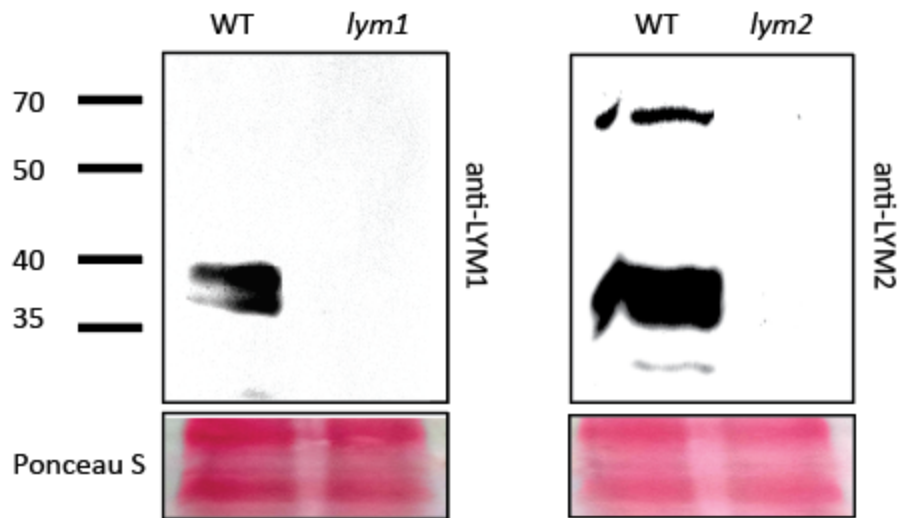
**Figure 4.3 LYM1 and LYM2 are detected in microsomes and symbiosomes**

LYM1 and LYM2 protein levels in microsome (mic) and symbiosome (sym) extracts of roots and nodules. Bacteria: free living *Sinorhizobium medicae* strain ABS7. Ponceau S-stained membrane is shown as a loading control.



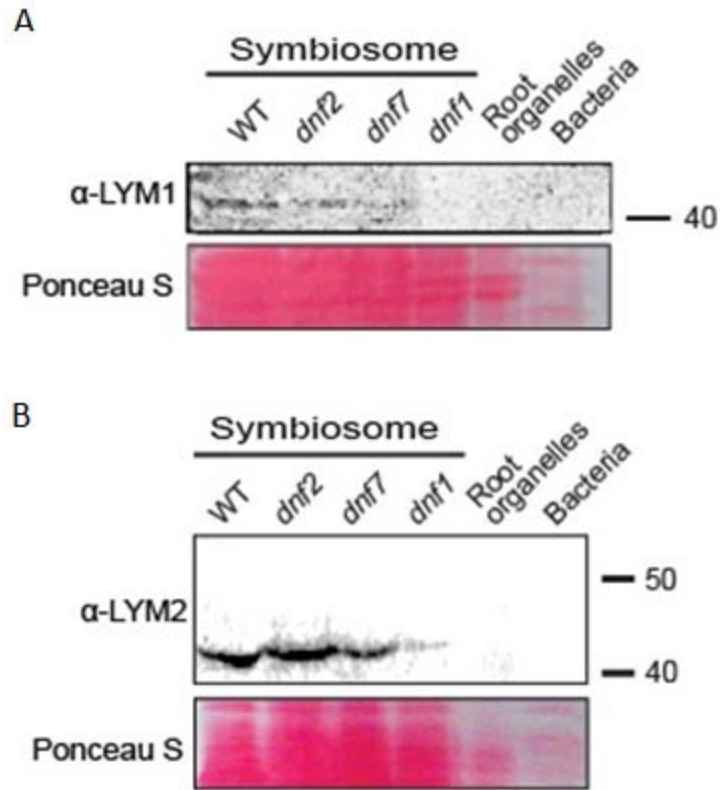
**Figure 4.4 Genotyping and nodule phenotype of *lym1* and *lym2* mutants**

**(A)** Schematic diagram of *Lym1* sequence containing the Tnt1 insertion site in the 1st exon (NF21014). Position and orientation of the primers used for genotyping are shown with arrows. **(B)** PCR products amplified with F1-R and F2-R primer pairs by using WT and *lym1* tnt1 insertion line as templates. **(C)** Schematic diagram of *Lym2* sequence containing the Tnt1 insertion site in the 2nd exon (NF9836). Position and orientation of the primers used for genotyping are shown with arrows. **(D)** PCR products amplified with F1-R and F2-R primer pairs by using WT and *lym1* tnt1 insertion line as templates. M: 1 kb DNA ladder (Caisson Labs) **(E)** Nodule morphology of *lym1* and *lym2* mutant. Plants were inoculated with *Sinorhizobium medicae* strain ABS7.



**Figure 4.5 Testing LYM1 and LYM2 antibodies**

**(A)** Detection of LYM1 in WT and *lym1* extracts containing roots and nodules. **(B)** Detection of LYM2 in WT and *lym2* extracts containing roots and nodules. Ponceau S-stained membrane is shown as a loading control.



**Figure 4.6 Detection of LYM1 and LYM2 in symbiosomes**

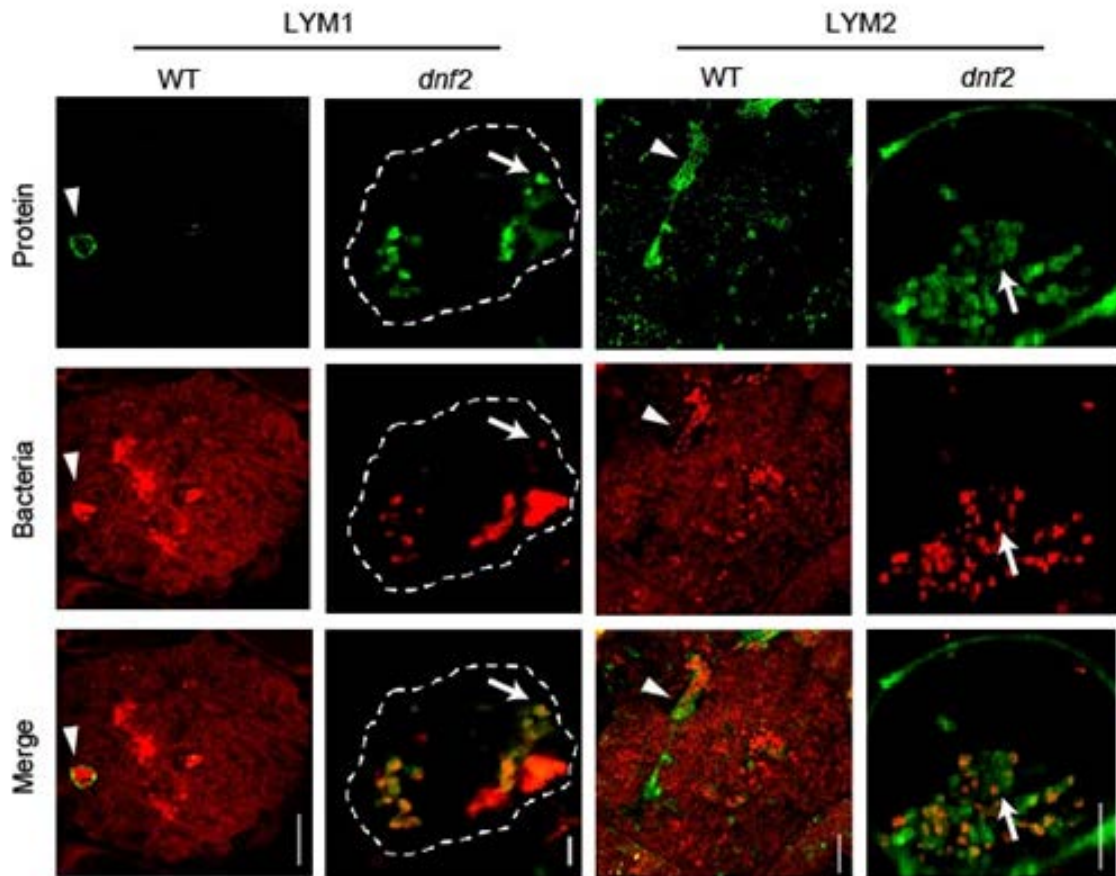
Detection of LYM1 (**A**) and LYM2 (**B**) in the symbiosome extracts of WT, *dnf2*, *dnf7*, and *dnf1*. Bacteria: free living *Sinorhizobium medicae* strain ABS7. Ponceau S-stained membrane is shown as a loading control. Root organelles: the root samples that underwent the identical symbiosome isolation procedure.

#### **4.3.2 LYM1 and LYM2 localize to the infection thread**

To determine the subcellular localization, I expressed GFP-fused LYM1 and LYM2 in WT nodules under the control of their own promoters. I observed LYM1 and LYM2 in the infection thread, as well as in vesicle-like cytosolic bodies, but not in the symbiosomes. However, when I produced GFP-LYM1/2 in *dnf2* nodules, LYM1 and LYM2 localized to the symbiosomes (Fig. 4.7), suggesting that LYM1 and LYM2 are cleaved by DNF2. Presumably, after cleavage, GFP-LYM1/2 diffuse inside WT symbiosomes, thereby the GFP signal is below the detection limit under confocal microscopy, explaining the reason I did not observe LYM1 and LYM2 inside symbiosomes unlike in concentrated WT symbiosome extracts on western blots (Fig. 4.3).

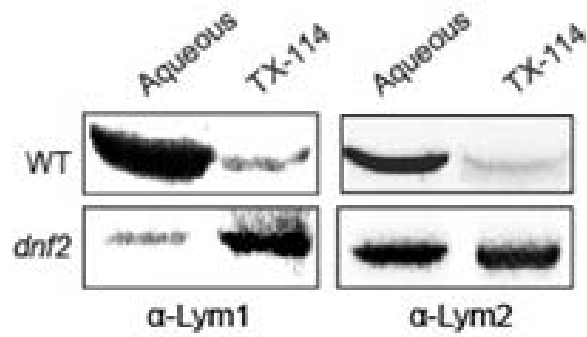
#### **4.3.3 LYM1 and LYM2 are cleaved by DNF2**

To determine whether DNF2 cleaves LYM1 and LYM2, I used TritonX-114 partitioning to enrich intact and cleaved forms of GPI-anchored proteins in the TritonX-114 and aqueous phase, respectively. In WT symbiosomes, most LYM1 and LYM2 were found in the aqueous phase, whereas LYM1 and LYM2 were concentrated in the Triton X-114 phase in *dnf2*, indicating that symbiosome-localized LYM1 and LYM2 are cleaved by DNF2 (Fig. 4.8).



**Figure 4.7 LYM1 and LYM2 localize to the infection thread**

Localization of GFP-LYM1 and GFP-LYM2 in nodule cells of WT and *dnf2*, inoculated with *Sinorhizobium meliloti* 1021 carrying *hemA::mCherry*. The arrowhead and arrow indicate infection thread and symbiosome, respectively. Scale bar, 10  $\mu$ m.



**Figure 4.8 LYM1 and LYM2 are cleaved by DNF2**

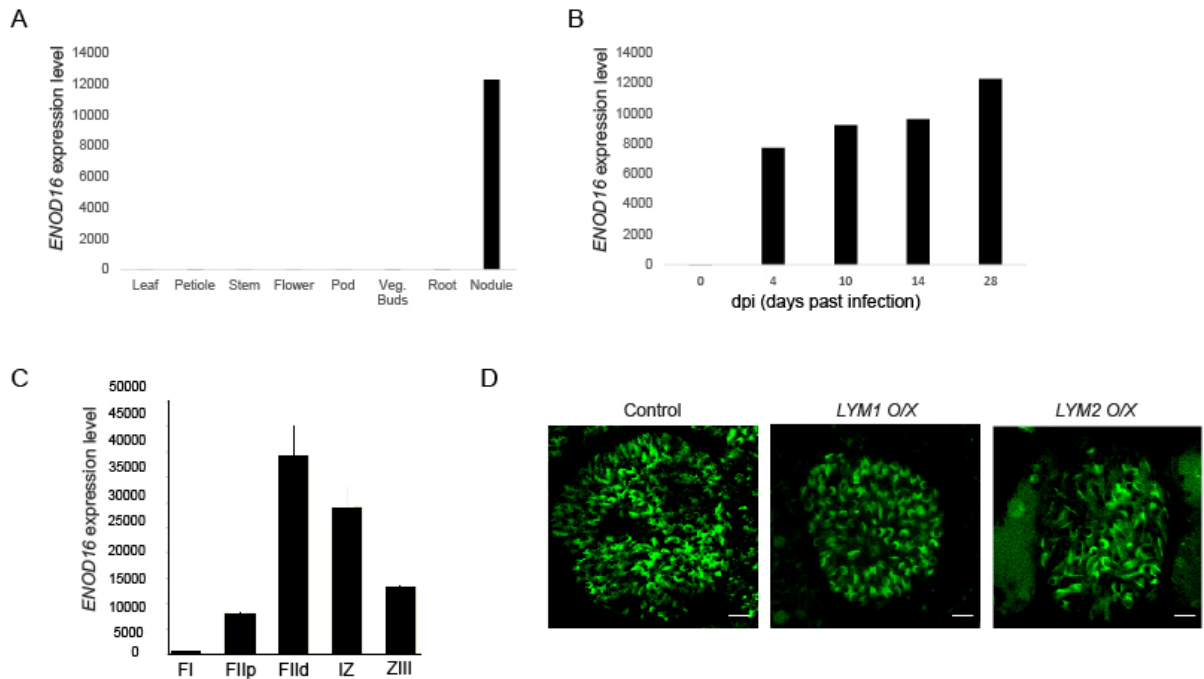
The levels of cleaved and uncleaved LYM proteins in WT and *dnf2* by TX-114 partitioning and Western Blot by LYM1 and LYM2 antibodies. Aqueous phase contains cleaved forms, while TX-114 phase includes uncleaved forms. The fractions were loaded at equivalent proportions of the original extract by volume.

#### 4.3.4 Cleavage of LYM1 and LYM2 are required for nodule development

Transcriptional downregulation of *LYMs* in nodules, as well as their DNF2-mediated removal from the peribacteroid membrane suggest that LYM1 and LYM2 may have negative impact on nitrogen-fixing symbiosis. Consistent with this hypothesis, *lym1* and *lym2* mutant plants form normal nodules (Fig. 4.4E).

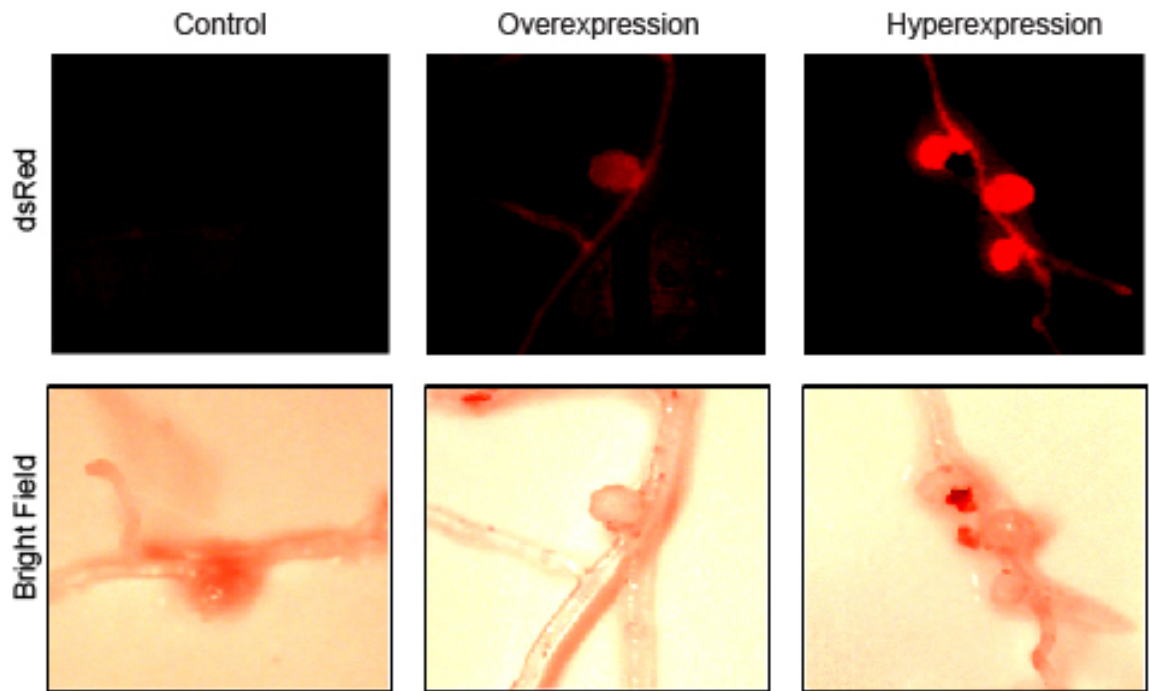
LYMs are defense-related receptors that sense bacterial peptidoglycan and induce defense response. Presumably, the bacterial death phenotype in *dnf2* results from the persistence of LYMs on the peribacteroid membrane. To test this hypothesis, I overexpressed *GFP-LYM1/2* in WT nodules using the promoter of *ENOD16*, which is abundantly expressed in nodules (Fig. 4.9). When *LYM1* and *LYM2* were moderately overexpressed, they were cleaved and localize to the symbiosome space in normal looking nodules. However, the nodules that hyper-accumulated LYMs (Fig. 4.10) were small and lacked pink-colored cells (Fig4. 11A), did not express leghemoglobin or NifH (Fig4. 11B) and lacked fully infected cells in the nitrogen fixation zone (Fig4. 11C). With a similar appearance to the *dnf2* nodules, the highly-elevated levels of LYM proteins inside symbiosomes caused the death of intracellular bacteria, showing that hyper-accumulation of LYM proteins can phenocopy the *dnf2* mutant.

To test whether the uncleaved LYM receptors on the symbiosome membrane caused the *dnf2* phenotype, I silenced *LYM1*, *LYM2*, and *both* simultaneously in the *dnf2* nodules. In all three cases the *dnf2* phenotype was partially rescued, and infected cells with live, differentiated bacteroids were resulted (Fig. 4.12). These data suggest that LYM1 and LYM2 are first produced as GPI-anchored defense receptors in the infection thread and they are later inactivated via cleavage by DNF2 in the symbiosome.



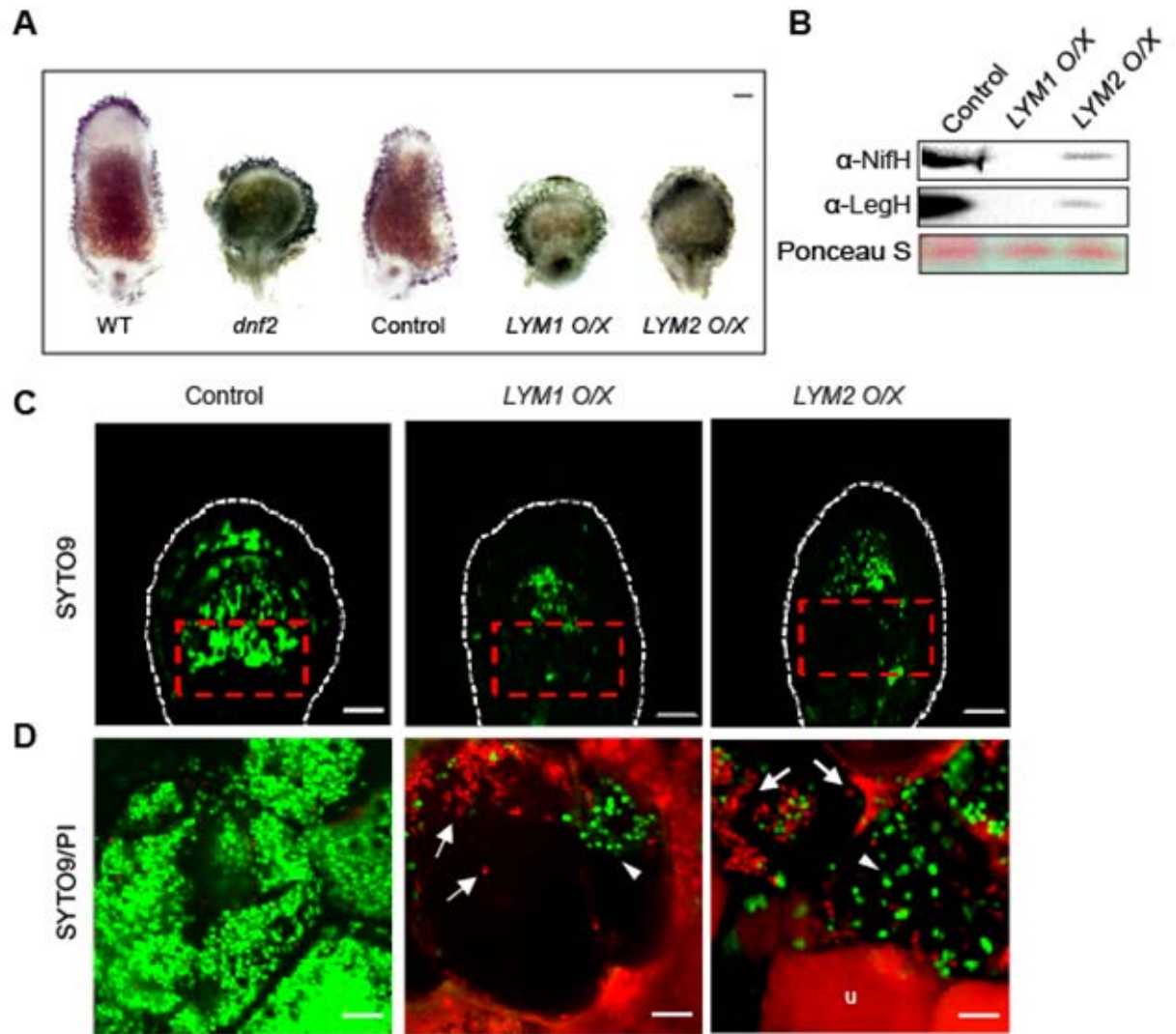
**Figure 4.9** The overexpression of LYM1 and LYM2 using *ENOD16* promoter

**(A)** *ENOD16* expression level in different tissues. **(B)** Temporal profile of *ENOD16* expression in nodule. **(C)** Spatial expression pattern of *ENOD16*. I, zone I; IId, zone II distal; IIp, zone II proximal; IZ, interzone II-III; III, zone III. Values in (A) and (B) are levels of Affymetrix probe signal based on microarray data from the *Medicago truncatula* Gene Expression Atlas. Values from (C) are from RNA-seq data from Symbimics. **(D)** Localization of GFP-tagged LYM1/2, under the control of *ENOD16* promoter, with *enod16* signal sequence in WT nodule cells.



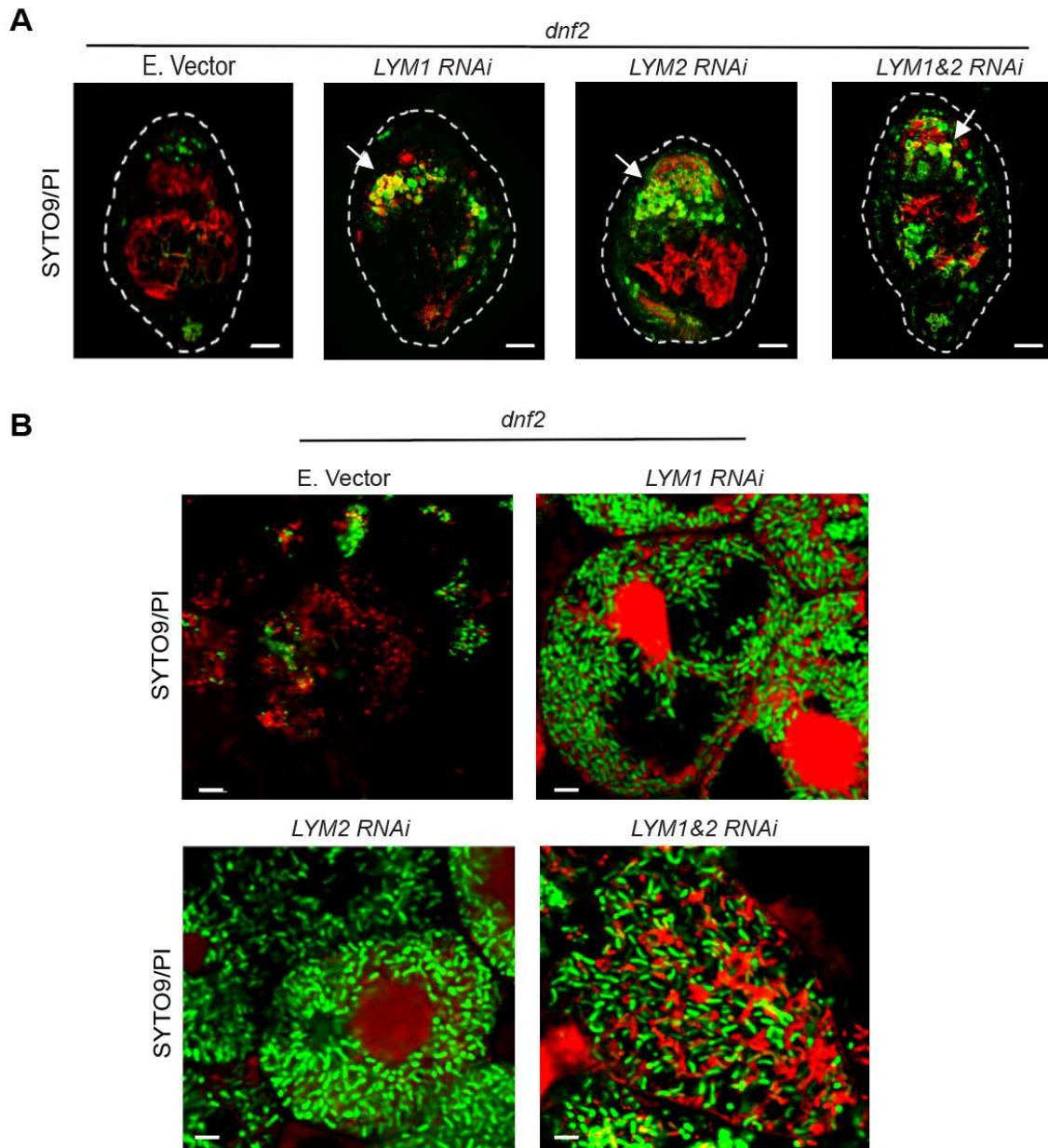
**Figure 4.10 Detection of LYM overexpression level**

The overexpression levels of LYMs were determined using DsRED1 fluorescence marker. The overexpression level is proportional to intensity of the DsRed1 fluorescence.



**Figure 4.11 Hyperexpression of LYM1 and LYM2 caused *dnf2* phenotype**

(A) Phenotypes of LYM1 and LYM2 overexpressed WT nodules. Control: overexpressed GFP protein. Scale bar: 0.1 mm. (B) Detection of the Leghemoglobin and NifH proteins in the nodule extracts of LYM1- and LYM2- overexpressed WT nodules. (C) SYTO9 staining of LYM1 and LYM2 overexpressed WT nodules. Scale bar: 0.1 mm. Green color from SYTO 9 indicates live bacteria. The rectangle shows the nitrogen fixation zone. (D) Live/Dead staining of LYM1 and LYM2 overexpressed WT nodules. Green color from SYTO 9 indicates live bacteria (arrowhead) and red color from propidium iodide (PI) indicates dead bacteria (arrow). U stands for the uninfected cells expressing mCherry fluorescence marker. Scale bar, 10  $\mu$ m. Plants were inoculated with *Sinorhizobium medicae* strain ABS7.



**Figure 4.12 Silencing of *LYM1* and *LYM2* partially rescued *dnf2* phenotype**

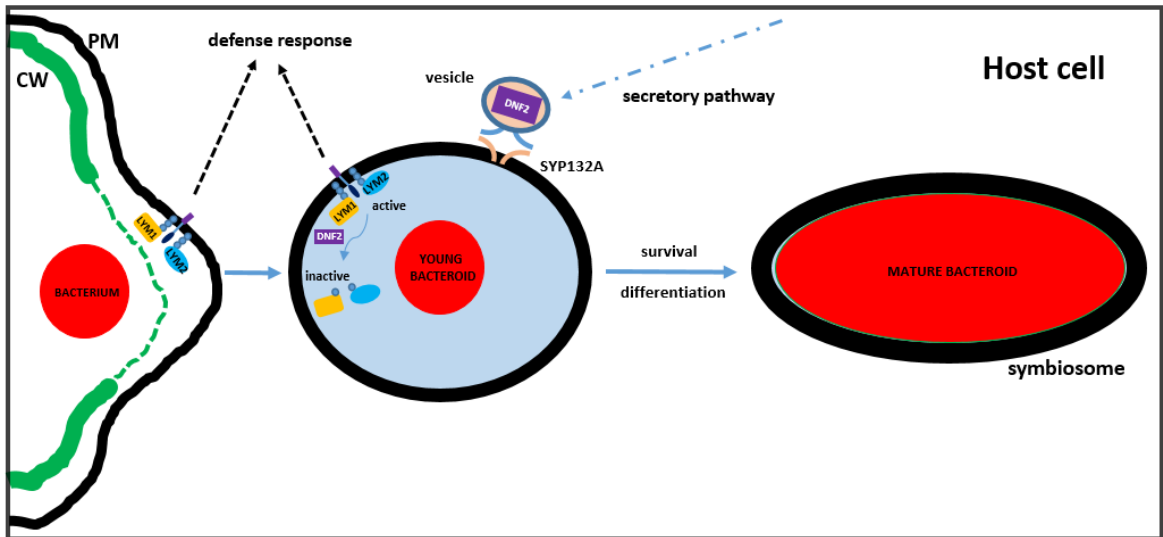
**(A)** Live-dead staining of *LYM1*-, *LYM2*-, and both-, *LYM1* and *LYM2*, silenced *dnf2* nodules. Scale bar, 0.1 mm **(B)** High magnification images of nodules in **(A)**. Scale bar, 10  $\mu$ m. Plants were inoculated with *Sinorhizobium medicae* strain ABS7.

#### 4.4 Discussion

In this study, I explained how host cell shut down their defense mechanism to maintain the bacterial partner inside symbiosomes. I identified two LysM domain containing GPI-anchored receptors called LYM1 and LYM2 on the infection thread membrane. These receptors are endocytosed onto the peribacteroid membrane when bacteria are internalized into a host cell. LYM proteins could detect peptidoglycan of invading rhizobia as a danger signal and trigger a host immune response, therefore their inactivation is crucial to mediate the survival of intracellular bacteria. I found that host cells secrete DNF2 enzymes into the symbiosomes to remove rhizobia-sensing LYM1 and LYM2 from the peribacteroid membrane to inactivate the host defense response. Thus, bacteria can survive inside symbiosomes and differentiate into nitrogen-fixing bacteroids (Fig. 4.13).

Silencing of *LYM1* or *LYM2* rescued the bacterial death phenotype of *dnf2*, denoting that both proteins are required to activate a defense mechanism, similar to LYM1 and LYM3 receptors in Arabidopsis. The requirement of both receptors for plant immunity demonstrates that LYM1 and LYM2 likely form a receptor complex, which needs to be confirmed by protein binding assays. In addition, the fact that both proteins lack an intracellular signaling domain implies the presence of a LysM-RLK, like AtCERK1, in this receptor complex.

Our studies showed that DNF2 functions inside the symbiosome after bacterial internalization, implying the presence of active LYM1/2 on the infection thread membrane. During bacterial invasion, the cell wall is degraded in the zone where bacterial entry happens, increasing the possibility of interaction between LYMs and microbial signatures. Somehow the defense response is not induced in this step. The mechanisms that protect invading bacteria before DNF2 activity wait to be identified by future studies.



**Figure 4.13 DNF2-mediated inactivation of LYM1 and LYM2 blocks host defense mechanism**

LYM1 and LYM2 are GPI-anchored receptors that detect bacterial peptidoglycan and trigger defense response. Once the bacteria are internalized, DNF2 removes the LYMs on the peribacteroid membrane and shut down the recognition of the bacteroid. Hence, intracellular bacteria can survive inside the symbiosomes.

## 4.5 Materials and Methods

### 4.5.1 Plant Growth Conditions and Inoculation

The legume model plant *Medicago truncatula* ecotypes *Jemalong A17* and *R108* were used for this study. Plants were inoculated with *Sinorhizobium medicae* strain ABS7 and *Sinorhizobium meliloti* strain *Rm1021-mCherry* in ½X BNM medium to obtain nodules (21 dpi) in different experiments. Plants were grown under a 16-hour, 22°C/8-hour, 18°C light/dark cycle.

### 4.5.2 Molecular Cloning

For *pLYM1::GFP-LYM1* construct, 35S promoter of the *pMDC43* vector was replaced with the region including 1146 bp of *LYM1* promoter and the first 66 bp of *LYM1* coding sequence. Then, *LYM1* coding sequence (67-1239 bp) was recombined with the vector by using Gateway technology (Invitrogen). For *pLYM2::GFP-LYM2* construct, 35S promoter of the *pMDC43* vector was replaced with the region including 1966 bp of *LYM1* promoter and the first 93 bp of *LYM2* coding sequence. Then, *LYM2* coding sequence (94-1119 bp) was recombined with the vector using Gateway technology (Invitrogen).

To overexpress *LYM1* and *LYM2*, 35S promoter of the *pMDC43* vector was replaced with the region including 1849 bp of *ENOD16* promoter and the first 60 bp of *ENOD16* coding sequence. *LYM1* coding sequence (67-1239 bp) and *LYM2* coding sequence (94-1119 bp) were recombined with the vector using Gateway technology (Invitrogen) to prepare *pENOD16::GFP-LYM1* and *pENOD16::GFP-LYM2* construct, respectively.

For RNA interference, the region between 358-660 bp of the *LYM1* coding sequence and 377-749 bp of the *LYM2* coding sequence (for *LYM1* and *LYM2*, respectively) were first cloned into the *pCR8/GW/TOPO* vector and recombined into the *pHellsGate8* vector modified with a 35S promoter-driven mCherry fluorescent marker for plant transformation (Helliwell and

Waterhouse, 2005) . To co-silence *LYM1* and *LYM2*, the region between 358-660 bp of the *LYM1* coding sequence and 377-749 bp of the *LYM2* coding sequence were joined by fusion PCR.

The vectors used in this study were modified with a 35S promoter-driven *mCherry* fluorescent marker for plant transformation.

#### **4.5.3 Hairy Root Transformation**

The introduction of plasmids into plants was carried out by hairy root transformation (Haney and Long, 2010). The seeds of *A17* plants were sterilely germinated to obtain seedlings. The root tips of 1 day old seedlings were cut with a razor blade and incubated with *A. rhizogenes* strain ARqua1 containing the desired plant expression vector on Fahraeus medium (FÅHRAEUS, 1957) for one week. After all emerging roots were removed; plants were transferred to selective media with either 25 mg/L kanamycin (remained on selective media for 15 days) or 10 mg/L hygromycin (for 10 days). Then, the plants were transferred to a mixture of turface and sand (1:1) with 1/2× Gamborg's B5 Basal Salt medium (Sigma-Aldrich) for one week. Transgenic roots were selected based on antibiotic resistance or mCherry fluorescence.

#### **4.5.4 Microscopy**

For confocal microscopy, nodules were hand-sectioned using double-edged razor blades and mounted on microscope slides. Samples were observed under an Olympus FLUOVIEW FV1000 confocal laser scanning microscope. GFP signal was detected using excitation with a 485 nm laser and emission with a 490-540 nm band pass filter, whereas mCherry signal was detected using excitation with a 587 nm laser and emission with a 575-675 nm band pass filter.

#### **4.5.5 Live/Dead Staining**

For live/dead staining, nodule slices were stained in 5  $\mu$ M SYTO9 and propidium iodide (Life Technologies) with 50 mM Tris-HCl (pH 7.0) and 25 mg/ml sucrose for 15 minutes.

#### **4.5.6 Protein Analysis**

The antibodies used in this study were diluted in TBS-Tween (1%) supplemented with 5% powdered milk or 3% Bovine Serum Albumin (BSA) as follows: anti-GFP (JL-8, Clontech) 1/10000; anti-mouse (Sigma-Aldrich) 1/10000; anti-rabbit (GenScript) 1/10000; anti-chicken (Fisher) 1/10000; anti-LYM1(Fliegmann et al., 2011) 1/10000, anti-LYM2(Fliegmann et al., 2011) 1/10000. anti-NifH (1:2500); and anti-Leghemoglobin (1:10000). The signals were detected by using ECL-detection kit (Thermo Scientific).

#### **4.5.7 Microsome and Symbiosome Isolation**

The nodules were collected in extraction buffer (0.5 M sucrose, 10 mM DTT, 50 mM Tris-HCl pH 7.5) including protease inhibitor cocktail (Sigma P9599) on ice. The tissue was grinded by mortar and pestle and filtered through 4 layers of Miracloth (Calbiochem) by centrifugation for 10 min at 200 x g to remove debris. The flow through was centrifuged for 15 min at 1600 x g to collect the rough symbiosomes(Rosendahl et al., 1992). The supernatant was centrifuged at 120,000 x g for 1 hour to obtain the microsomes (Catalano et al., 2004).

#### **4.5.8 Triton X-114 Partitioning**

The symbiosome pellet was resuspended in Tris-buffered saline (TBS) (50 mM Tris-Cl, pH 7.5, 150 mM NaCl) and mixed with 1/5 volume Triton X-114 and incubated on ice for 15 minutes with occasional mixing. The cell mixture was incubated at 37°C for 2 minutes and centrifuged for

10 minutes at 1000 x g at room temperature. The water phase and TX-114 phase were collected to be analyzed (Doering et al., 2001).

#### **4.5.9 PCR reverse-screening for *Tnt1* insertion mutants**

Plant genomic DNA was extracted from leaves using CTAB and ethanol precipitation. *Medicago truncatula* NF21014 and NF9836 mutant lines were screened for the presence of insertion in *LYM1* and *LYM2* by PCR using a forward primer specific to the *Tnt1* sequence and a reverse primer specific to the genomic sequences of *LYM1* or *LYM2* genomic sequences

## CHAPTER 5

### DNF2 SUBSTRATES: ENOD16 AND ENOD20

#### 5.1 Introduction

##### 5.1.1 Arabinogalactan Proteins

Arabinogalactan proteins (AGPs) and extensins are two large families of hydroxyproline (Hyp)-rich glycoproteins implicated in many aspects of plant growth and development (Schultz et al., 1998). On the basis of backbone sequence, AGPs are classified as classical and non-classical (Neumann et al., 2008). Classical AGPs possess an N-terminal signal peptide, a core protein and, typically, a putative C-terminal signal sequence for GPI-anchor attachment. The core protein varying in length and domain complexity consists of Ala-Pro, Pro-Ala, Ser-Pro, Thr-Pro, Val-Pro, and Gly-Pro dipeptide repeats. These motifs are extremely O-glycosylated through the Hyp residues by galactan- and arabinose-containing carbohydrate complexes (Ellis et al., 2010), that consist of  $\beta$ -1,3-galactan backbones attached to varying  $\beta$ -1,6-galactose side chains decorated by  $\alpha$ -arabinose,  $\beta$ -(methyl)glucuronic acid,  $\alpha$ -rhamnose, and  $\alpha$ -fucose (Showalter and Basu, 2016). Different than AGPs, extensins possess contiguous Hyp residues, such as Ser-Hyp<sub>2</sub> or Ser-Hyp<sub>4</sub>, which are arabinosylated. AGPs can form crosslinks with pectin and arabinoxylan<sup>87</sup>, whereas extensins undergo intra- and intermolecular cross-linking leading to formation of extensin networks in the cell wall (Brady et al., 1998).

AGPs are difficult to purify due to their high content of complex carbohydrates, also preventing the recognition by protein-specific antibodies. The large size of the AGP families, in addition, brings about functional redundancy limiting study of individual AGPs. Among the tools to study AGPs is  $\beta$ -Yariv reagent, a dye that specifically interacts with AGPs and disrupt their functions by sequestering them into large complexes. Despite the fact that their epitope specificities are often questionable, the monoclonal antibodies reacting with carbohydrate content on AGPs,

such as MAC204 and PN16.4B4, also enabled to gain information about the distribution, localization, and function of AGPs (Seifert and Roberts, 2007).

Even if the biochemical function of AGPs are yet to be known, studies involving transcriptional analysis and plant mutant lines, as well as  $\beta$ -Yariv reagent and antibodies, have described various roles of AGPs in many biological processes (Ellis et al., 2010; Seifert and Roberts, 2007).

The treatment of plant cell suspension cultures, such as rose and tobacco, with the  $\beta$ -Yariv reagent either activated programmed cell death or only inhibited cell division, implying that AGPs participate in regulation of cell viability and promotion of cell division (Ellis et al., 2010).

That their composition alters during development in several plant tissues suggests that AGPs play a role in plant developmental processes. In immature carrot seeds, AGPs have roles in somatic embryogenesis, whereby a single or group of somatic cells forms an embryo (Van Hengel et al., 2002). A gibberellin induced-AGP in cucumber hypocotyls is involved in stem elongation (Park et al., 2003). Furthermore, the fact that the growth medium containing soluble AGPs, derived from embryogenic suspension cultures of carrot, triggers the embryogenesis of non-embryonic cells suggests that AGPs can act as signaling molecules (Schultz et al., 1998). An AGP of *Zinnia elegans*, xylogen, brings about an inductive cell-cell interaction required for the development of tracheary elements, the water- and mineral salts- conducting xylem cells (Motosse et al., 2004). An Arabidopsis root-specific AGP, AGP30, functions in the abscisic acid (ABA) response and is required for root generation and regulation of seed germination timing (van Hengel and Roberts, 2003). In addition, AGPs have roles during salt stress. The induction of AGPs in response to salt stress implies that AGPs plays a role in conferring salt tolerance (Lampert et al., 2006). SOS5 is a putative cell surface adhesion protein containing AGP-like domains and putative cell adhesion domains, named fascillin-like domains. Abnormal cell expansion, thinner cell walls and reduced pectin layers during salt stress in the absence of SOS5

implicate a role for SOS5 in cell expansion in Arabidopsis (Shi et al., 2003). Besides, AGPs participate in initiation of female gametogenesis, pollen tube growth and guidance (Cheung et al., 1995; Jiao et al., 2016; Wu et al., 1995), pollen grain development, and self-incompatibility in pollen (Ellis et al., 2010).

### 5.1.2 ENOD16 and ENOD20 proteins

The proteins that are expressed in the early steps of nodule formation are named Early Nodulins (ENODs). Most ENODs are hydroxyproline (Hyp)-rich glycoproteins (HRGPs) that are secreted to the cell wall and extracellular matrix (Cassab, 1998), and presumably remodel plant cell wall during nodule development (Brewin, 2004). A class of HRGPs, including *Medicago truncatula* ENOD16 and ENOD20, contains a signal peptide, a plastocyanin-like domain and a GPI anchor domain (Fig. 5.1). These proteins possess motifs for arabinogalactosylation as well (Vernoud et al., 1999). Similar proteins called Early nodulin-like (ENODL) exist in non-nodulating plant species such as Arabidopsis, rice, maize or poplar. These proteins are considered phytoalexins, a plant-specific subfamily of copper-binding proteins involved in electron transport; however, some of them lack amino acids for copper binding, suggesting that they have a different function (Mashiguchi et al., 2009). Information about the roles of ENODLs in biological processes is largely lacking. A recent study reported that ENODLs in Arabidopsis are involved in pollen tube recognition (Hou et al., 2016).

ENOD16 and ENOD20 possess very similar amino acid sequences. Nevertheless, they differ in the length of their hydroxyproline-rich domains including motifs for glycosylation (Fig. 5.1). ENOD16 contain only one AGP motif, whereas ENOD20 consists of several AGP motifs, and an additional arabinosylation motif (Fig 5.2). Both proteins show homology to pea/vetch ENOD5 and soybean N315/ENOD55 that are expressed in infected nodule cells. The expression of

*Ps/VsENOD5* was also shown in root cortical cells carrying infection threads (Greene et al., 1998). These GPI-anchored extracellular glycoproteins probably act as cell wall structural proteins during bacterial invasion and nodule development. However, experimental studies demonstrating this claim are lacking.

### **5.1.3 Research Question**

I only identified LYM1 and LYM2 proteins as substrates of DNF2; however, DNF2 presumably has multiple substrates inside symbiosomes. ENOD16 and ENOD20 are predicted to be hydroxyproline-rich arabinogalactan proteins that are expressed in the early steps of nodule development, which makes them potential substrates of DNF2. I showed that a GFP protein with GPI anchor domain of ENOD16 is cleaved by DNF2 (Fig. 3.10). ENOD16 and ENOD20 possess the same GPI-attachment sequence (Fig. 5.1), denoting that ENOD20 is also a GPI-anchored protein that is cleaved by DNF2. That DNF2 most likely regulates ENOD16 and ENOD20 during nodulation made me curious about these proteins. Nevertheless, any information about the functions of ENOD16 and ENOD20 are lacking.

Here, I found that ENOD16 and ENOD20 are exclusively expressed in nodules and localize to the symbiosome. Both proteins take role in the establishment of symbiotic nitrogen fixation; however, they participate in the different steps of nodule development.

A

```

ENOD16  MASSSPILLMIIFSMWLLISHSESTDYLGDSHNSWKVPLPSRRAFARWASAHEFTVGD
ENOD20  MSSSSPILLMFIFSIWMLISYSESTDYLVGDSSENSKFFPLPTRHALTRWASNYQFIVGDT
          *:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*
          *:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*

ENOD16  ILFEYDNETESVHEVNEHDYIMCHTNGEHVEHHDGNTKVVLDKIGVYHFISGTRHCKMG
ENOD20  ITFQYNNKTESVHEVEEEDYDRCGIRGEHVDHYDGNTMVLKKTGIHFIHSGKKRHCRLG
          *:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*
          *:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*

ENOD16  LKLAVVQNKHDLVLPPL-----ITMPMP-----
ENOD20  LKLAVVMVAPVLSPPPPSPPTPRSTPIPHPPRRSLPSPSPSPSPSPSPSPSPSPSPSP
          ***** * **                               ::* *

ENOD16  -----PSPSPS
ENOD20  STPIPHPRKRSPASPSPSLSKSPSPSESPSLAPSPSDSVASLAPSSSPSDESPPAPS
          ****:*

ENOD16  PNSSGNGGGAAGLGFIMWLGVSVMFLI*-
ENOD20  PSSSGSKGGGAGHGFL--EVSIAMMFLIF*
          *:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*:*
          ↑

```

B



**Figure 5.1 ENOD16 and ENOD20 proteins**

**(A)** Sequence alignment of ENOD16 and ENOD20 using Clustal Omega (<http://www.ebi.ac.uk/Tools/msa/clustalo/>). The red line represents the GPI-anchor attachment site predicted by big-PI Plant Predictor (Eisenhaber et al., 2001). **(B)** The domain organization of ENOD16 and ENOD20 proteins. SP: signal peptide.

ENOD16

MASSSPILLMIIFSMWLLISHSESTDYLGDSHNSWKVPLPSRRAFARWASAHEFTVGDITLF  
EYDNETESVHEVNEHDYIMCHTNGEHVEHHDGNTKVLDKIGVYHFISGTKRHCKMGLKL  
AVVVQNKHDLVLPPLITMPMPSPSPSPNSSGNKGGGAAGLGFIMWLGVSVMFLI\*

ENOD20

MSSSSPILLMFIFSIWMLISYSESTDYLVGDSENSWKFLPTRHALTRWASNYQFIVGDTITF  
QYNNKTESVHEVEEEDYDRCGIRGEHVDHYDGNTMVVLKKTGIHHFISGKKRHCRLGLKL  
AVVVMVAPVLSPPPPPSPTPRSSTPIPHPPRRSLPSPSPSPSPSPSPSPRSTPIP  
HPRKRSPASPSPSLSKSPSPSESPSLAPSPSDSVASLAPSSPSDESPPAPSPSSSG  
SKGGGAGHGFLEVSIAMMMFLIF\*

**Figure 5.2 ENOD16 and ENOD20 contain glycosylation motifs**

ENOD16 and ENOD20 proteins possess motifs for glycosylation. The motifs for arabinogalactan and arabinose addition were shown as yellow and red, respectively.

## 5.2 Experimental Strategy

In this study, I found that ENOD16 and ENOD20 are symbiosome-localized proteins involved in maintaining nodule development. I first analyzed the expressions of *ENOD16* and *ENOD20* using the *M. truncatula* gene expression atlas. I investigated the localization of ENOD16 and ENOD20 under a confocal microscope. Finally, I examined the possible roles of ENOD16 ENOD20 in nodulation.

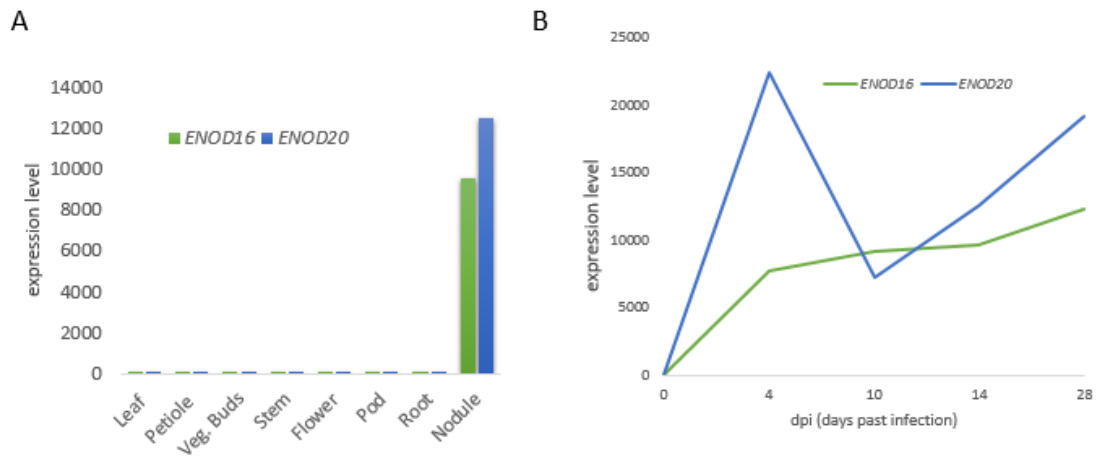
## 5.3 Results

### 5.3.1 ENOD16 and ENOD20 are exclusively expressed in nodules.

I investigated the expressions of *ENOD16* and *ENOD20* using the *Medicago truncatula* gene expression atlas (Benedito et al., 2008) and the Symbimics database (Roux et al., 2014). *ENOD16* and *ENOD20* are only expressed in nodules (Fig. 5.3A), implying a nodule-specific function. Both genes are induced with bacterial inoculation (Fig. 5.3B) and expressed in all nodule zones. While *ENOD20* reached its highest expression level in the infection zone, *ENOD16* expression peaked somewhat later in the transition zone (Fig. 5.4).

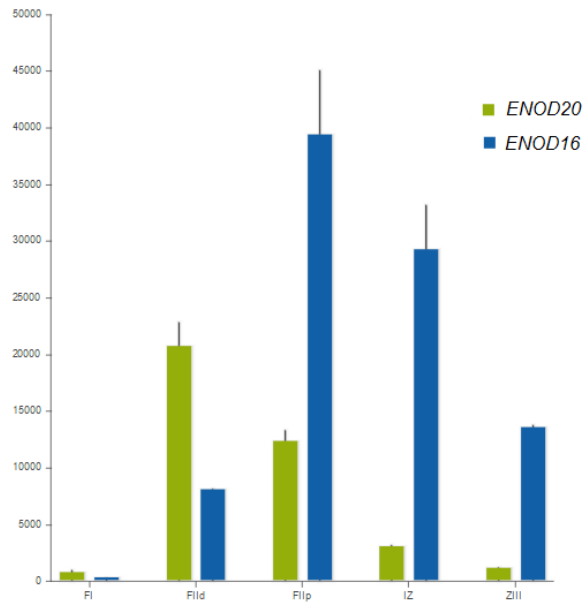
### 5.3.2 ENOD16 and ENOD20 localize to the symbiosomes

To determine their subcellular localization, I constructed GFP-tagged ENOD16 and ENOD20 proteins. Because these proteins have a GPI-anchor domain in their C-termini, I placed the GFP near the N-terminus after the signal peptide. I produced GFP-fused ENOD16 and ENOD20 in WT nodules under the control of their own promoters. ENOD16 and ENOD20 both localized in the symbiosome surrounding young bacteroids, as well as differentiated (mature) bacteroids. In *dnf2* nodules, both proteins are found in the symbiosomes as well. (Fig. 5.5&5.6)



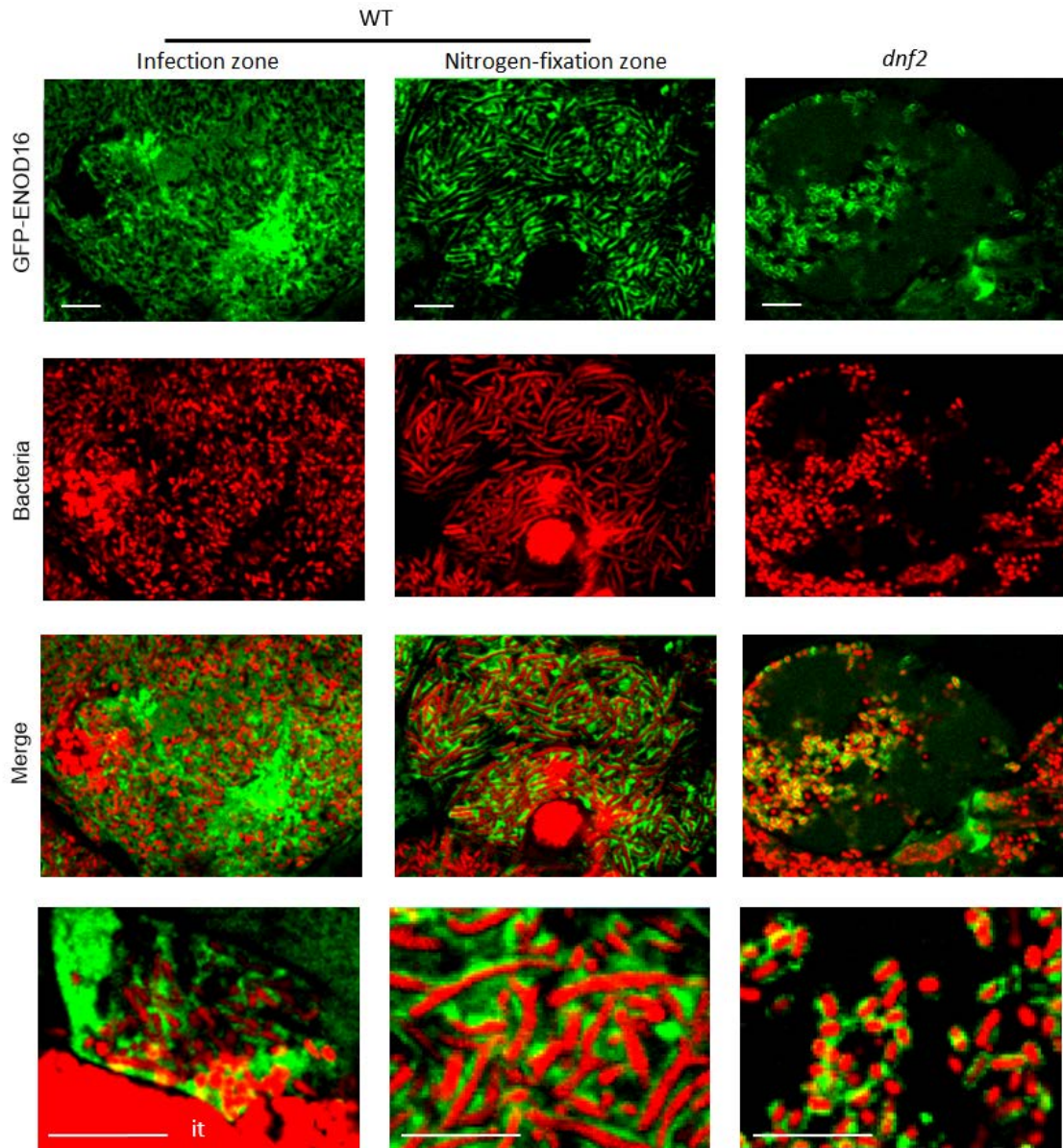
**Figure 5.3** *ENOD16* and *ENOD20* have nodulation-specific expression.

**(A)** *ENOD16* and *ENOD20* expression levels in different tissues. **(B)** Temporal profile of *ENOD16* and *ENOD20* expression in nodule.



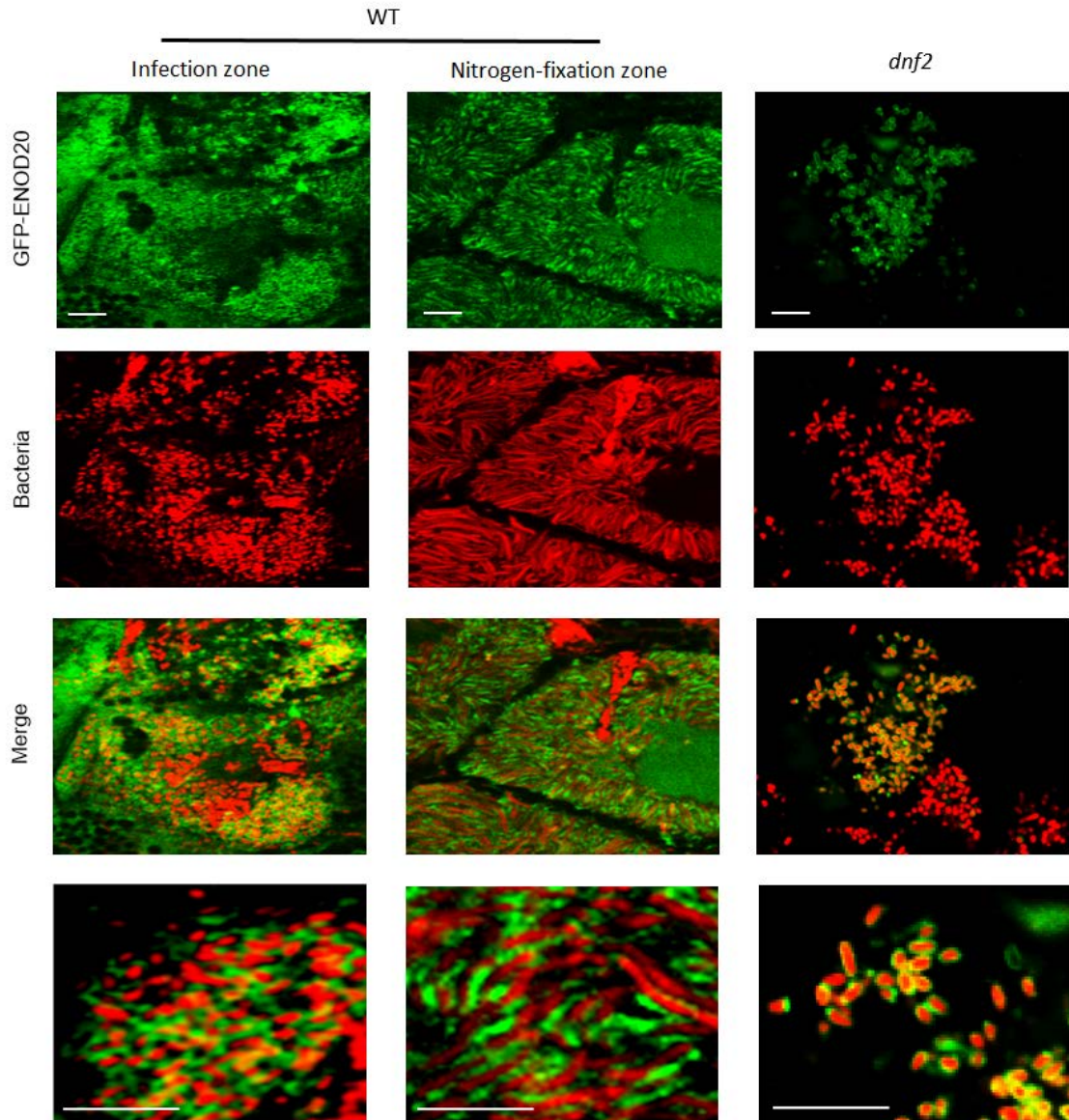
**Figure 5.4 ENOD16 and ENOD20 are expressed in all nodule zones**

Spatial expression pattern of *ENOD16*. I, zone I; Ild, zone II distal; Ilp, zone II proximal; IZ, interzone II-III; III, zone III. Values are from RNA-seq data from Symbimics(Roux et al., 2014)



**Figure 5.5 ENOD16 localizes to the symbiosome**

Localization of GFP-ENOD16 in nodule cells of WT and *dnf2*, inoculated with *Sinorhizobium meliloti* 1021 carrying *hemaA::mCherry*. Scale bar: 10  $\mu$ m. It: infection thread.



**Figure 5.6 ENOD20 localizes to the symbiosome**

Localization of GFP-ENOD20 in nodule cells of WT and *dnf2*, inoculated with *Sinorhizobium meliloti* 1021 carrying *hema::mCherry*. Scale bar: 10  $\mu$ m.

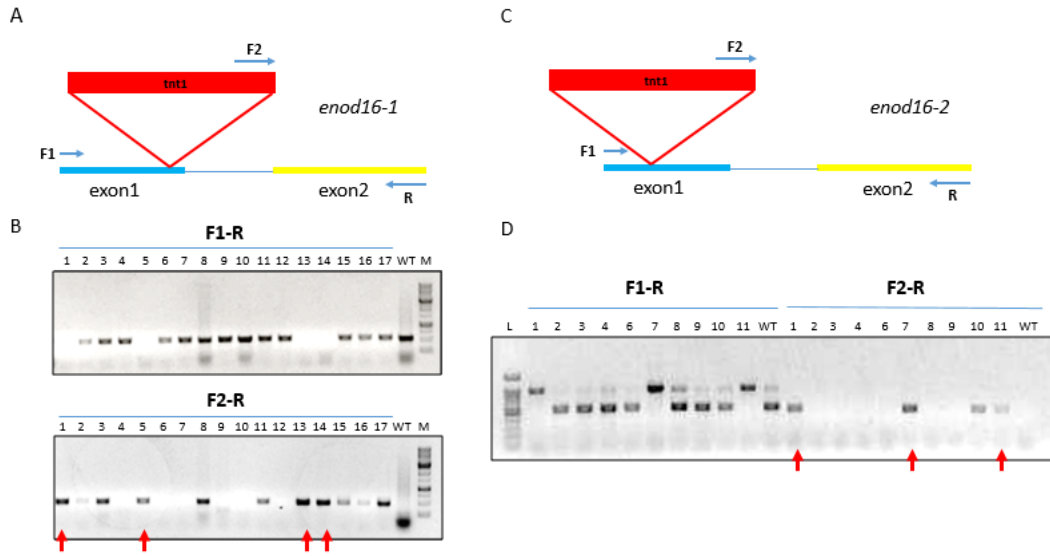
### 5.3.3 ENOD16 and ENOD20 are crucial for nodule development

I obtained *Tnt1* insertional mutant lines for ENOD16 and ENOD20 in the *M. truncatula* R108 background, and confirmed the *Tnt1* insertion sites in the *ENOD16* and *ENOD20* genes by performing PCR with a *Tnt1* primer and primers designed for *ENOD16* and *ENOD20* genomic sequences. I have two mutant lines with different insertion sites for ENOD16 named *enod16-1* and *enod16-2*, as well as for ENOD20 called *enod20-1* and *enod20-2* (Fig. 5.7&5.8).

To investigate the roles of ENOD16 and ENOD20 in nodulation, I inoculated the *enod16* and *enod20* plants, as well as WT control, with rhizobia. The *enod20* plants were smaller than WT and had pale, yellow leaves, a sign of nitrogen-starvation. The *enod16-2* plants showed a similar phenotype to *enod20* plants; however, the *enod16-1* appeared similar to WT (Fig. 5.9). I further stained the nodules with SYTO9 and examined them under a confocal microscope. I detected red autofluorescence in the *enod16* and *enod20* mutants (Fig. 5.10), implying the accumulation of defense response-related phenolics, as in the *dnf2* mutant (Fig. 3.3E). I verified the presence of phenolics in the mutant nodules by potassium permanganate staining (Vasse et al., 1993) (Fig. 5.11). I observed that the infection and early differentiation occurred in *enod16* nodules. Nevertheless, these nodules did not have a well-developed, multi-cell layer fixation zone, which likely resulted from the fact that differentiated bacteria did not persist in *enod16-1* nodules. Consistent with the aerial phenotype, the *enod16-1* nodules had milder defects than the *enod16-2* nodules. In *enod20* mutants, nodule development stopped at an early stage. To sum up, ENOD20 and ENOD16 play roles in early and late stages of nodule development, respectively. The results are consistent with their transcriptional profiles suggesting that ENOD20 acts earlier than ENOD16.

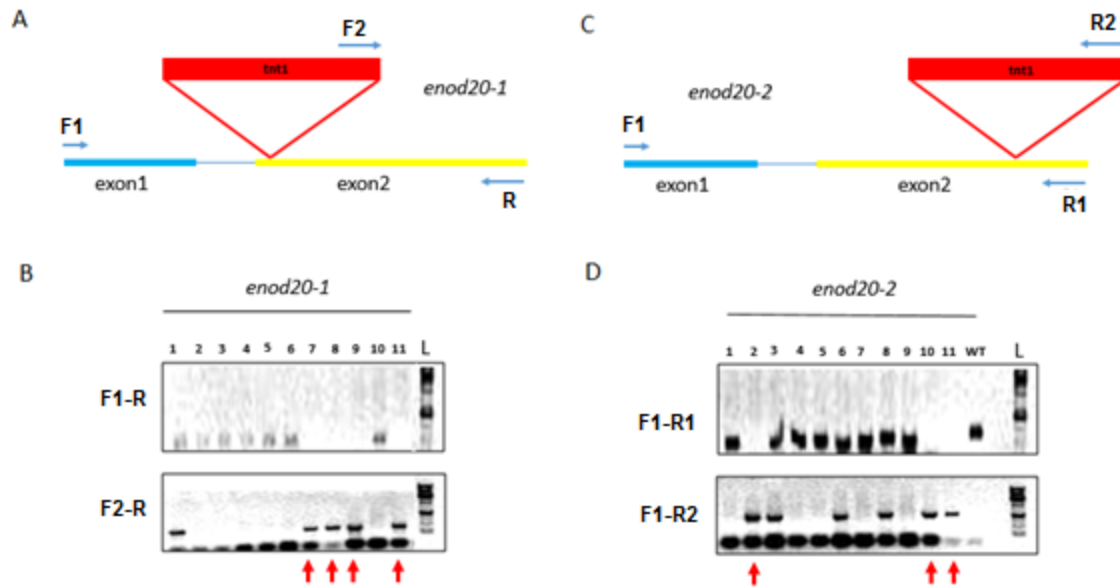
When expressing GFP-ENOD16 and GFP-ENOD20 in the *dnf2* nodules to determine their localization, I was surprised to observe an effect in certain nodules: some symbiosomes

contained differentiated bacteria (Fig. 5.12), implying that some *dnf2* bacteroids can survive long enough to differentiate under high levels of peribacteroid membrane-bound ENOD16/20 proteins. These symbiosomes expanded in volume (Fig. 3.3D). Nevertheless, the expression of ENOD16 and ENOD20 in *dnf2* did not produce functional nodules.



**Figure 5.7 Genotyping of *enod16* mutant lines**

Schematic diagrams of *Tnt1* insertion mutant line NF15343 named *enod16-1* (A), and NF3825 named *enod16-2* (C). (B and D) PCR products amplified with F1-R and F2-R primer pairs using genomic DNA from WT and *enod16 Tnt1* insertional line as templates. M: 1 kb DNA ladder (Caisson Labs). The numbers label the genotyped individuals. Red arrows indicate the individuals that are homozygous for the *Tnt1* insertion.



**Figure 5.8 Genotyping of *enod20* mutant lines**

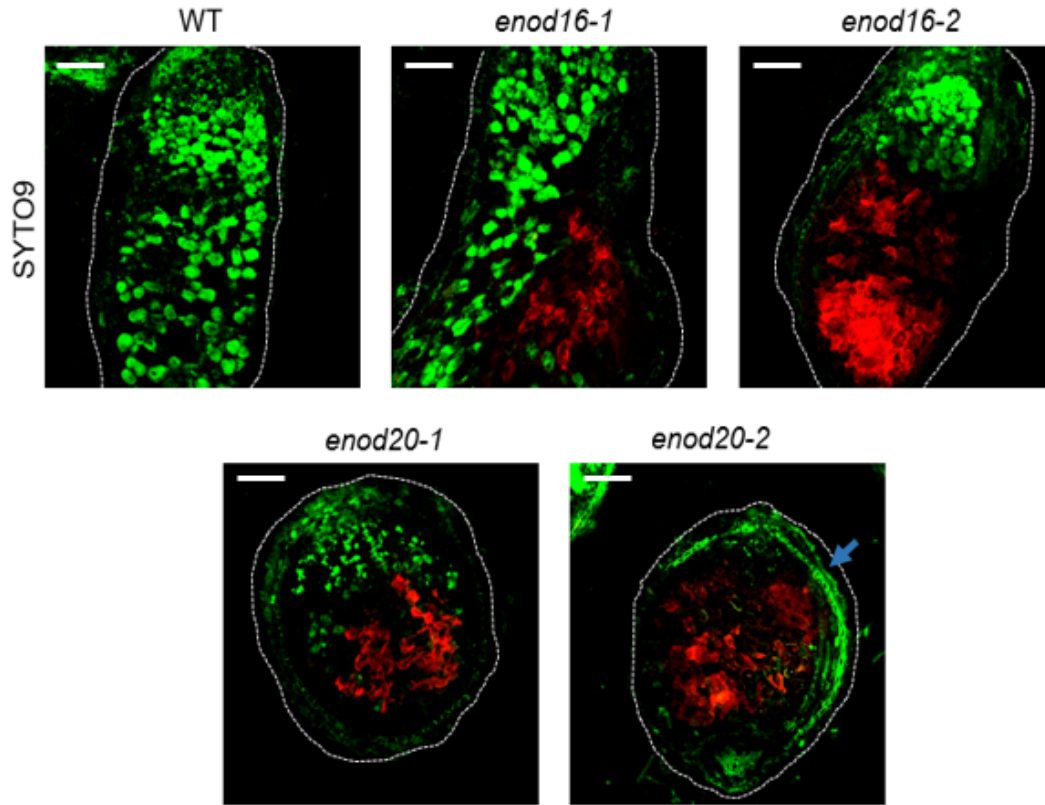
Schematic diagrams of *Tnt1* insertional mutant line NF9226 named *enod20-1* (A), and NF12701 named *enod20-2* (C). (B and D) PCR products amplified with F1-R and F2-R primer pairs using genomic DNA from WT and *enod20* *Tnt1* insertional line as templates. M: 1 kb DNA ladder (Caisson Labs). The numbers label the genotyped individuals. Red arrows indicates the individuals that are homozygous for the *Tnt1* insertion.



WT      *enod16-1*      *enod16-2*      *enod20-1*      *enod20-2*

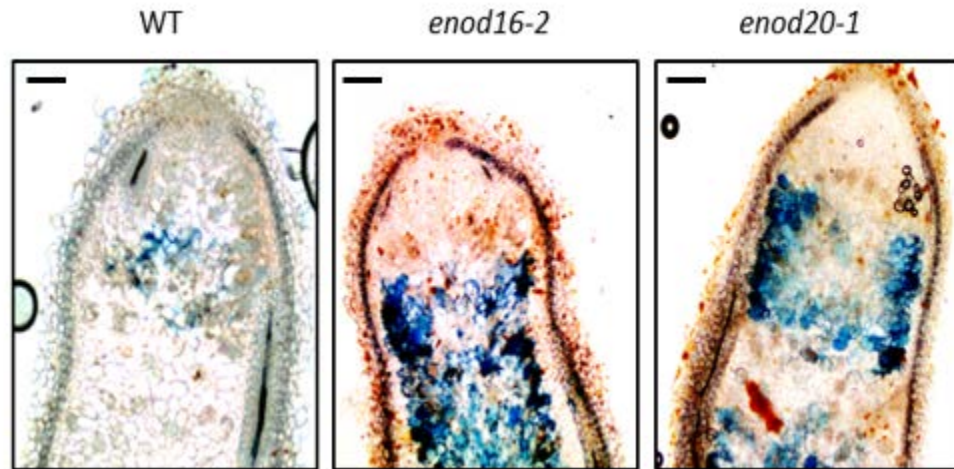
**Figure 5.9 Growth phenotypes of rhizobia-inoculated *enod16* and *enod20* plants**

Growth phenotypes of WT, *enod16* and *enod20* that were inoculated with *Sinorhizobium medicae* strain ABS7. The plants were examined 4 weeks after bacterial inoculation.



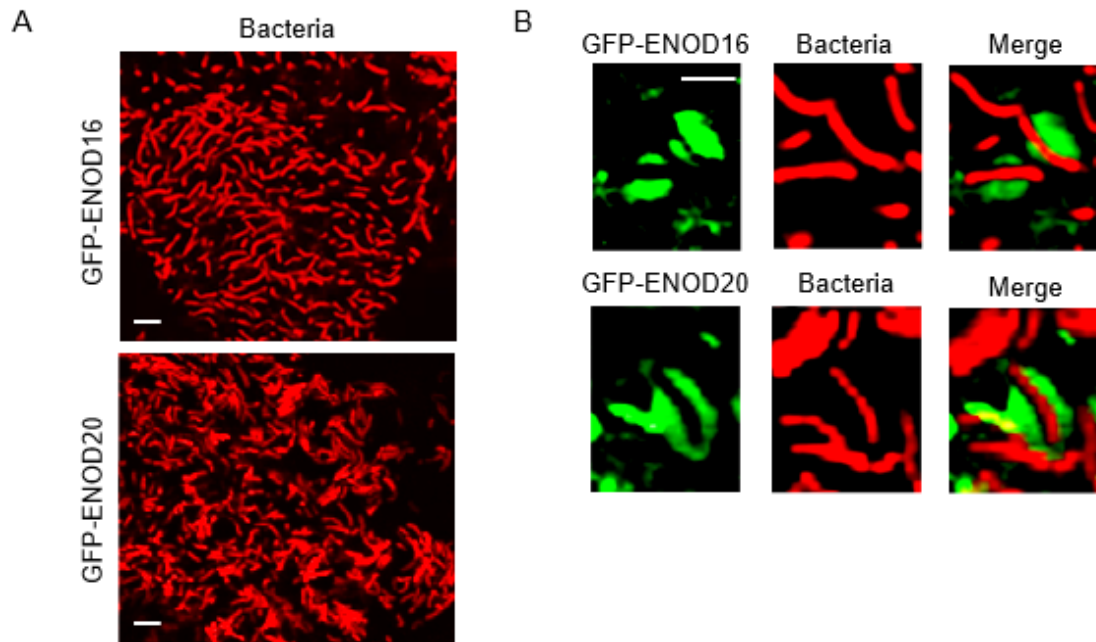
**Figure 5.10 *enod16* and *enod20* plants produce defective nodules**

SYTO9 staining of 4 weeks old WT, *enod16*, and *enod20* nodules. While green color indicates live bacteria, red color indicates autofluorescence from phenolics accumulation. Blue arrow shows green autofluorescence from nodule vasculature. Scale bar: 0.1 mm.



**Figure 5.11 Phenolic accumulation is observed in *enod16* and *enod20* nodules**

Staining of phenolics by potassium permanganate in six week old WT, *enod16*, and *enod20* nodules. Blue color indicates the presence of phenolics. Scale bar: 0.1 mm.



**Figure 5.12 The *dnf2* bacteroid phenotype is rescued by the expression of GFP-ENOD16 and GFP-ENOD20 in *dnf2* nodules**

GFP-ENOD16 and GFP-ENOD20 were expressed in *dnf2* nodules, inoculated with *Sinorhizobium meliloti* 1021 carrying *hemA::mCherry*. (A) Bacteroids in nodules. (B) High resolution image of bacteroids. Scale bar: 5  $\mu$ m.

## 5.4 Discussion

Here I described two putative arabinogalactan proteins, ENOD16 and ENOD20, that are potential substrates of DNF2. They are exclusively expressed in nodules and localize to the symbiosome after bacterial internalization (Fig. 5.13). These proteins play roles in different steps of nodule development.

ENOD16 and ENOD20 were predicted to be GPI-anchored proteins. I showed that a GFP protein with the ENOD16 GPI anchor domain are cleaved by DNF2, implying that ENOD16 is a GPI-anchored protein (Fig. 3.10). Since ENOD16 and ENOD20 have similar GPI-attachment sequence, ENOD20 is likely GPI-anchored as well. However, this is still an indirect proof. That ENOD16/20 are GPI-anchored should be shown *in vivo* by TritonX-114 extraction and partitioning in future studies.

*ENOD16* and *ENOD20* are expressed in the early steps of nodulation. Both genes are expressed in root cortical cells carrying infection threads and in infected nodule cells. In addition, *ENOD20* is expressed in the nodule primordium (Vernoud et al., 1999). Our transcription analysis showed that *ENOD16/20* genes have nodule-specific expressions and are expressed in all the zones except meristem. However, the expression profiles of *ENOD16* and *ENOD20* differ inside nodules. In the infection zone, *ENOD20* expression is higher than *ENOD16*. Then, *ENOD20* is downregulated in the transition and nitrogen-fixation zones, while *ENOD16* is upregulated. The difference in the expression patterns of *ENOD20* and *ENOD16* demonstrates that ENOD20 and ENOD16 function in different steps of nodule development.

ENOD16 and ENOD20 are both located in the symbiosome space, surrounding young and mature bacteroids. I observed both proteins in the *dnf2* symbiosomes as well, suggesting that ENOD16 and ENOD20 are most likely cleaved into the symbiosome space by DNF2.

ENOD16 and ENOD20 proteins have been known for a long time; however, our understanding of their roles in nodulation is limited. I found that the absence of ENOD16 or ENOD20 causes phenolic accumulation in nodules, implying the activation of host defense responses as in *dnf2*. Nevertheless, *enod16* and *enod20* nodules had different phenotypes. The differentiation of bacteria did not happen in *enod20* nodules, implying that nodule development stopped at an early stage. In *enod16* nodules, the infection and early differentiation took place; however, these nodules did not have a well-developed, multi-cell layer fixation zone due to a failure of differentiated bacteria to persist. I observed a milder phenotype in *enod16-1* compared to *enod16-2*. This might result from the difference in the Tnt1 insertion sites. A partially functional protein might be produced in *enod16-1*. The mutant analysis demonstrates that ENOD20 acts earlier than ENOD16 in nodule development which is consistent with their expression profiles in the nodule zones.

Presumably DNF2 mediates the release of ENOD16/20 on the peribacteroid membrane into symbiosome space. Whether the soluble or intact protein is the active form for ENOD16 and ENOD20 have yet to be known. Unexpectedly, the expression of GFP- ENOD16 and GFP-ENOD20 in the *dnf2* mutant resulted in the formation of symbiosomes with differentiated bacteroids, which partially rescued *dnf2* bacteroid phenotype. This implies that some *dnf2* bacteroids can survive long enough to differentiate under high levels of peribacteroid membrane-bound ENOD16/20 proteins.

ENOD16 and ENOD20 are predicted to be hydroxyproline-rich arabinogalactan proteins that act as cell wall structural proteins. Nevertheless, the biochemical functions of arabinogalactan proteins in general have not been identified yet. ENOD20 consist of more motifs for arabinogalactan addition than ENOD16, and also contains an additional motif for

arabinylation. Whether glycosylation profile is crucial for the functions of these proteins has yet to be known either.

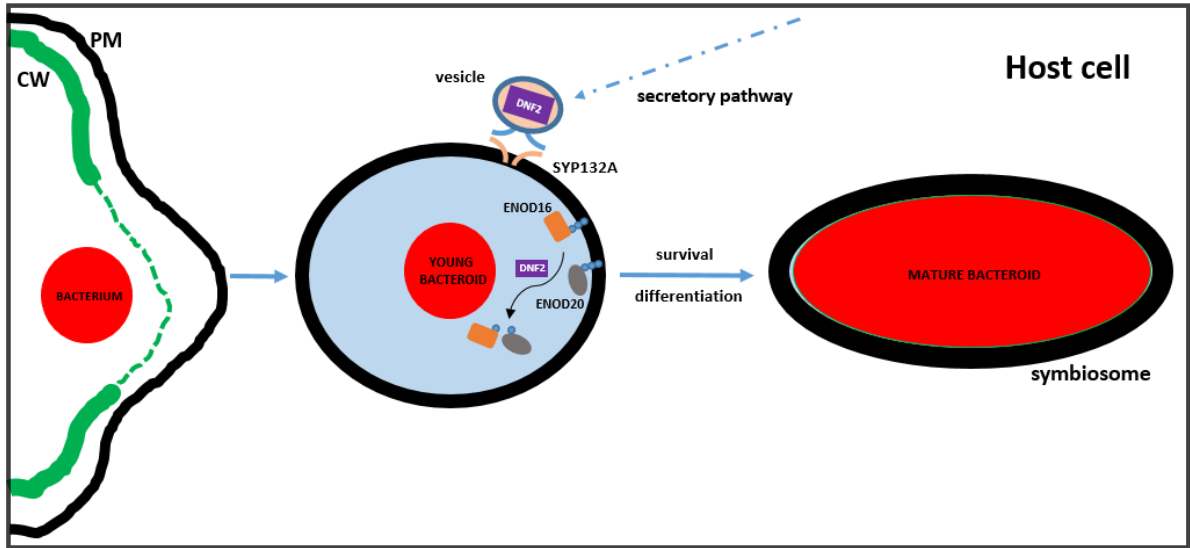


Figure 5.13 DNF2 cleaves ENOD16 and ENOD20 inside the symbiosomes.

ENOD16 and ENOD20 are secreted to the symbiosomes once bacteria are endocytosed into host cell. These hydroxyproline-rich GPI-anchored proteins, that are cleaved by DNF2 inside the symbiosomes, participate in the different steps of nodule development.

## 5.5 Materials and Methods

### 5.5.1 Plant Growth Conditions and Inoculation

The legume model plant *Medicago truncatula* ecotypes *Jemalong A17* and *R108* were used for this study. Plants were inoculated with *Sinorhizobium medicae* strain ABS7 in ½X BNM medium to obtain 4 weeks old nodules in different experiments. Plants were grown under a 16-hour, 22°C/8-hour, 18°C light/dark cycle.

### 5.5.2 Molecular Cloning

For *pENOD16::GFP-ENOD16* construct, 35S promoter of the *pMDC43* vector was replaced with the region including 1849 bp of *ENOD16* promoter and the first 60 bp of *ENOD16* coding sequence. For *pENOD20::GFP-ENOD20* construct, 35S promoter of the *pMDC43* vector was replaced with the region including 2694 bp of *ENOD20* promoter and the first 75 bp of *ENOD20* coding sequence. *ENOD16* and *ENOD20* coding sequences were recombined with the vector by using Gateway technology (Invitrogen).

### 5.5.3 Hairy Root Transformation

The introduction of plasmids into plants was carried out by hairy root transformation (Haney and Long, 2010). The seeds of *A17* plants were sterilely germinated to obtain seedlings. The root tips of 1 day old seedlings were cut with a razor blade and incubated with *A. rhizogenes* strain ARqua1 containing the desired plant expression vector on Fahraeus medium (FÅHRAEUS, 1957) for one week. After all emerging roots were removed; plants were transferred to selective media with 10 mg/L hygromycin (for 10 days). Then, the plants were transferred to a mixture of turface and sand (1:1) with 1/2× Gamborg's B5 Basal Salt medium (Sigma-Aldrich) for one week.

#### **5.5.4 Microscopy**

For confocal microscopy, nodules were hand-sectioned using double-edged razor blades and mounted on microscope slides. Samples were observed under an Olympus FLUOVIEW FV1000 confocal laser scanning microscope. GFP signal was detected using excitation with a 485 nm laser and emission with a 490-540 nm band pass filter, whereas mCherry signal was detected using excitation with a 587 nm laser and emission with a 575-675 nm band pass filter.

#### **5.5.5 SYTO9 Staining**

Nodule slices were stained in 5  $\mu$ M SYTO9 (Life Technologies) with 50 mM Tris-HCl (pH 7.0) and 25 mg/ml sucrose for 15 minutes.

#### **5.5.6 PCR reverse-screening for *Tnt1* insertion mutants**

Plant genomic DNA was extracted from leaves using CTAB and ethanol precipitation. NF15343 and NF3825 mutant lines with a *Tnt1* insertion in *ENOD16*, and NF9226 and NF12701 mutant lines with a *Tnt1* insertion in *ENOD20* were screened by PCR using a forward primer specific to the *Tnt1* sequence and a reverse primer specific to the genomic sequences of *ENOD16* and *ENOD20* genomic sequences.

#### **5.5.7 Phenolics Staining**

The sectioned nodules were fixed in 0.04% KMnO<sub>4</sub>, 10 mM PIPES, pH 7.2, for 1 h at room temperature. They were rinsed with 10 mM PIPES, pH 7.2, and stained in 0.01% methylene blue for 10 min. After they were rinsed with water, nodule sections were photographed using bright-field microscope (Vasse et al., 1993)

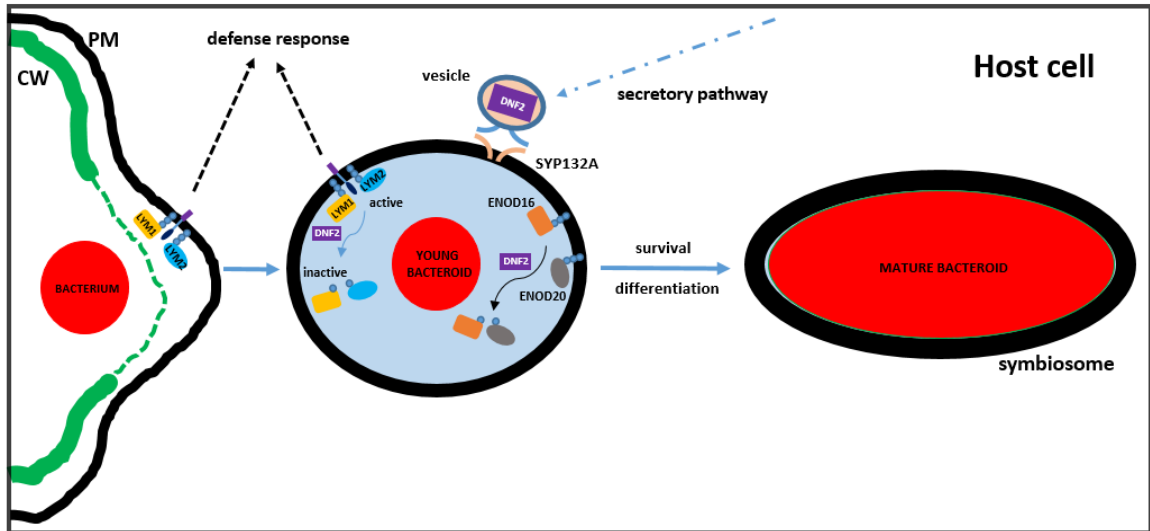
## CHAPTER 6

### SUMMARY AND FUTURE DIRECTIONS

#### 6.1 Summary

Legumes overcome nitrogen limitation by forming a symbiotic interaction with nitrogen-fixing bacteria. This interaction begins with a signaling dialogue between host plant and the rhizobia on the root surface. The bacteria penetrate the nodule through root hairs in a structure called the infection thread. Once the infection thread reaches the cortical cells, the bacteria are endocytosed into symbiosomes and differentiate into the nitrogen-fixing form.

In my thesis, I described the host mechanisms that maintain the intracellular rhizobia in the *Medicago truncatula*-*Sinorhizobium meliloti* symbiotic interaction. After rhizobial entry, the host cell secretes a PI-PLC enzyme, called DNF2 to the symbiosomes through a nodule-specific protein secretory pathway. Trafficking to the symbiosome is mediated by a symbiotic-t-SNARE protein on the peribacteroid membrane produced via alternative cleavage and polyadenylation of the *Medicago truncatula* *SYNTAXIN132* gene. Inside symbiosomes, DNF2 cleaves two defense-related GPI-anchored receptors, LYM1 and LYM2, to inactivate host defense response against intracellular rhizobia, and two hydroxyproline-rich proteins, ENOD16 and ENOD20, participating in different steps of nodule development. In this way, intracellular rhizobial bacteria can survive and differentiate into the nitrogen-fixing form inside symbiosomes (Fig. 6.1).



**Figure 6.1 The host cell maintains intracellular bacteria by secreting DNF2 enzyme into the symbiosomes**

After bacterial entry, the host cell secretes a PI-PLC enzyme called DNF2 into the symbiosomes to remove two defense-related GPI-anchored receptors, LYM1 and LYM2, to inactivate its own defense mechanism. DNF2 also cleaves two hydroxyproline-rich proteins, ENOD16 and ENOD20, participating in different steps of nodule development. In this way, intracellular bacteria can survive and differentiate into the nitrogen-fixing form inside symbiosomes.

## 6.2 Future Directions

My study described the molecular mechanism of how a host cell maintains intracellular symbiotic bacteria in symbiotic nitrogen fixation. In Chapter 2, I showed that the host cell regulates protein trafficking to symbiosomes by alternative cleavage and polyadenylation of a *t-SNARE* gene, *SYNTAXIN132*. The alternative protein isoform, *SYP132A*, is required to target host proteins to the symbiosome. An earlier report revealed that two v-SNARE proteins, called *VAMP721d* and *VAMP721e*, have nodulation specific functions. Future examinations should clarify whether *SYP132A* and *VAMP721d/e* form a SNARE complex to enable the fusion of symbiosome-targeted vesicles to the peribacteroid membrane. In addition, other components, which are involved in the symbiosome-specific protein delivery, need to be explored. I found that two isoforms, *Syp132A* and *Syp132C*, show different localizations. While *Syp132A* is found on the peribacteroid membrane, *Syp132C* localize to the plasma membrane. These isoforms differ in their transmembrane domains. Whether distinct localizations of these isoforms stem from different transmembrane domains remains to be addressed. Also, additional studies are required for a complete understanding of APA regulation of *SYP132* gene.

In Chapter 3, I found that a secreted host protein, *DNF2*, is required for the survival of symbiont bacteria inside symbiosomes. I showed that *DNF2* is a PI-PLC which cleaves GPI-anchored proteins (GAPs) *in vivo*. It would be desirable to show the cleavage of GAPs by the purified *DNF2* protein *in vitro*. I showed that *LYM1* and *LYM2* proteins are substrates of *DNF2*. In addition, *ENOD16* and *ENOD20* proteins are presumably cleaved by *DNF2* inside symbiosomes. Whether additional *DNF2* substrates exist and their functions need to be explored by future studies. Also, much remains to be learned about the biochemical properties and the regulation of *DNF2* protein.

In Chapter 4, I explained how host cells shut down the defense system against symbiont bacteria. LYM proteins which sense the peptidoglycan of bacteria are removed from the peribacteroid membrane by DNF2, leading to the survival of symbiont bacteria inside symbiosomes. However, whether *Medicago* LYM proteins induce the defense response and their binding to peptidoglycan are yet to be shown experimentally. LYM proteins do not possess a signaling domain, suggesting that they likely form a complex with a kinase, like AtCERK1. Future studies should identify components of the LYM-dependent signaling pathway. I explained how bacteria are prevented from the host defense inside symbiosomes, however how symbiont bacteria escape LYM recognition inside the infection thread remain to be addressed. I observed that the absence of DNF2 causes the accumulation of phenolics, which is a well-known marker for the activation of host defense response. Future examinations should determine other defense responses against symbiont bacteria inside nodules.

In Chapter 5, I found that ENOD16 and ENOD20 proteins are presumably the substrates of DNF2, which needs be proved by Triton X-114 partitioning. I showed that ENOD16 and ENOD20 are required for the maintenance of symbiont bacteria; however biochemical functions of these proteins remain to be explored. Their cleavage by DNF2 brings about two possible forms of ENOD16 and ENOD20: soluble and membrane-bound. Which form is the active one should be shown by future experiments.

## APPENDIX A

### A SYMBIOTIC SNARE PROTEIN GENERATED BY ALTERNATIVE TERMINATION OF TRANSCRIPTION

nature  
plants

ARTICLES

PUBLISHED: 11 JANUARY 2016 | ARTICLE NUMBER: 15197 | DOI: 10.1038/NPLANTS.2015.197

## A symbiotic SNARE protein generated by alternative termination of transcription

Huairong Pan<sup>1†</sup>, Onur Oztas<sup>1,2†</sup>, Xiaowei Zhang<sup>3</sup>, Xiaoyi Wu<sup>4</sup>, Christina Stonoha<sup>1,5</sup>, Ertao Wang<sup>3</sup>, Bin Wang<sup>4</sup> and Dong Wang<sup>1\*</sup>

Many microbes interact with their hosts across a membrane interface, which is often distinct from existing membranes. Understanding how this interface acquires its identity has significant implications. In the symbiosis between legumes and rhizobia, the symbiosome encases the intracellular bacteria and receives host secretory proteins important for bacterial development. We show that the *Medicago truncatula* *SYNTAXIN 132* (*SYP132*) gene undergoes alternative cleavage and polyadenylation during transcription, giving rise to two target-membrane soluble NSF attachment protein receptor (t-SNARE) isoforms. One of these isoforms, SYP132A, is induced during the symbiosis, is able to localize to the peribacteroid membrane, and is required for the maturation of symbiosomes into functional forms. The second isoform, SYP132C, has important functions unrelated to symbiosis. The SYP132A sequence is broadly found in flowering plants that form arbuscular mycorrhizal symbiosis, an ancestral mutualism between soil fungi and most land plants. *SYP132A* silencing severely inhibited arbuscule colonization, indicating that SYP132A is an ancient factor specifying plant-microbe interfaces.

## APPENDIX B

### PRIMER TABLE

PRIMER	SEQUENCE	DESCRIPTION
GPlanc_F	GATATCACAAAGTTGTACAAAAAAGC	to amplify GPI-anchor sequence of <i>ENOD16</i>
GPlanc_R	CCCTTAAGGTACCTCATATTAAG	to amplify GPI-anchor sequence of <i>ENOD16</i>
Syp_RNAi_F	ACAAAAGCCTGGCTGTGAGA	to target both isoforms
Syp_RNAi_R	AAGAATTGCTGCTGGAAGA	to target both isoforms
SypC_RNAi_F	CAATGCAGTCGATCATGTCC	to target SypC isoform
SypC_RNAi_R	AATGGCAACAAAACCAGCTT	to target SypC isoform
SypA_RNAi_F	GAAGTCGGGTAAACGATGCTC	to target SypA isoform
SypA_RNAi_R	CCATGAACCTTGTAAACTGC	to target SypA isoform
LYM2RNAi_F	ACGCAATAGCCAGGGTTAGA	to target <i>LYM2</i> gene
LYM2_R1	CTGGAAGGGGAACATCAAGA	to target <i>LYM2</i> gene
LYM1RNAi_F	TCCATTACGGTGGTTTGGT	to target <i>LYM1</i> gene
LYM1RNAi_R	TGATGCACAGGCAGGAATAG	to target <i>LYM1</i> gene
LYM2RNAi_F3	GGACACTGTGTGCAGTGTAGACGCAATAGCCAGGGTTAGA	to fuse <i>LYM1</i> RNAi construct with <i>LYM2</i> RNAi construct
PLYM1-HindIII	AAGCTTTTTGGTGTGGCACTCTGTTAATG	to amplify <i>LYM1</i> promoter + signal sequence
PLYM1-KpnI	GGTACCTTTTTGAGTTGTTGTTATTAAG	to amplify <i>LYM1</i> promoter + signal sequence
PLYM2-HindIII2	AAGCTTTTTGTGGTGATCCAACGTCAC	to amplify <i>LYM2</i> promoter + signal sequence
PLYM2-KpnI	GGTACCTTTTTGAGCTTCTGTTACTGTTATTC	to amplify <i>LYM2</i> promoter + signal sequence
pE16-HindIII	AAGCTTACTCCGCTAACTGAATGAAAC	to amplify <i>ENOD16</i> promoter + signal sequence
pE16-KpnI	GGTACCTTGGAATTAGCAGCCACATTG	to amplify <i>ENOD16</i> promoter + signal sequence
pE20-HindIII	AAGCTTTCCTTTGATTTAATTGGTCC	to amplify <i>ENOD20</i> promoter + signal sequence
pE20-KpnI	GGTACCTTTGTGGATTACAGAGTAGGAAATT	to amplify <i>ENOD20</i> promoter + signal sequence
Tnt1_Fw	CATCCAGAGGAGGTAGCACTGTC	for genotyping
Tnt1_Rev	CACAGGTTCTGCTCGTTCACTG	for genotyping
LYM1_SP_F3	ACAATAGAACCATGCACAAC	to amplify <i>LYM1</i> CDS without signal sequence
LYM1_Last	TTACAATGCTACAGGGATCATC	to amplify <i>LYM1</i> CDS without signal sequence
LYM2_SP_F	CCAGAAGCAAACCTCAAGTG	to amplify <i>LYM2</i> CDS without signal sequence
LYM2_Last	CTACAAAAGATAGACAAATAAG	to amplify <i>LYM2</i> CDS without signal sequence
E16_SP_F	CACTCTGAATCCACAGATTATC	to amplify <i>ENOD16</i> CDS without signal sequence
E16_Last	TCATATTAAGAACATCATCATCACC	to amplify <i>ENOD16</i> CDS without signal sequence
E20_SP_F	GATTATCTGGTCGGAGACAG	to amplify <i>ENOD20</i> CDS without signal sequence
E20_Last	CTAGAAAATTAAGAACATCATCATAGCG	to amplify <i>ENOD20</i> CDS without signal sequence
Syp_Start	ATGAACGACCTTCTCACTGAT	to amplify SypA and SypC CDS without signal sequence
SypA_Last	CTTCTTCATGGTTTCAACAAC	to amplify SypA CDS without signal sequence
SypC_Last	AGAACTCTTCCAAGGTTTGAG	to amplify SypC CDS without signal sequence

## APPENDIX C

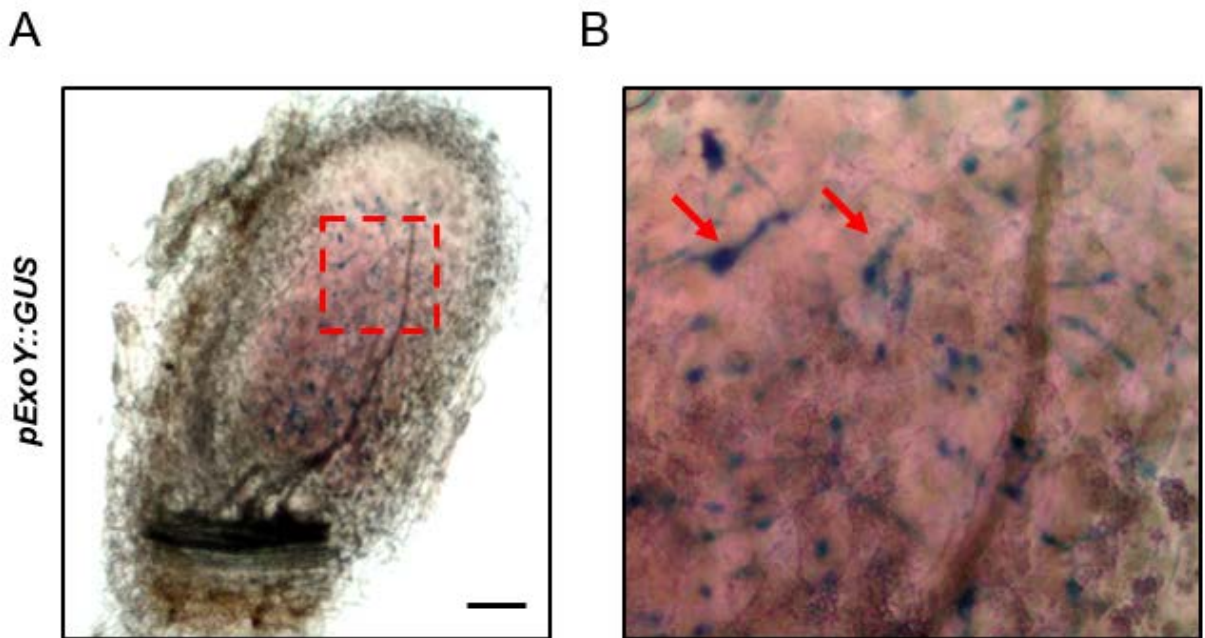
### DOES SUCCINOGLYCAN SHIELD SYMBIONT BACTERIA FROM HOST IMMUNITY?

Succinoglycan is an acidic exopolysaccharide (EPS) produced by *Sinorhizobium meliloti* strain *Rm1021*. This polymer contains repeating octasaccharide subunits with modified glucose and galactan molecules, which are synthesized by proteins encoded by *Exo* genes. Among them is the *ExoY* gene which produces the galactosyl-1-P transferase required for the first step of succinoglycan synthesis. *ExoY* mutants do not produce succinoglycan (Reuber and Walker, 1993).

Succinoglycan plays a significant role in the invasion of alfalfa root nodules by *Sinorhizobium meliloti* strain *Rm1021* (Glucksmann et al., 1993). Succinoglycan deficiency leads to the formation of smaller nodules lacking bacteria. The absence of succinoglycan also leads to the accumulation of antimicrobial phenolics and phytoalexins, and the upregulation of host defense genes, implying the role of succinoglycan in dampening host immunity (Gibson et al., 2008).

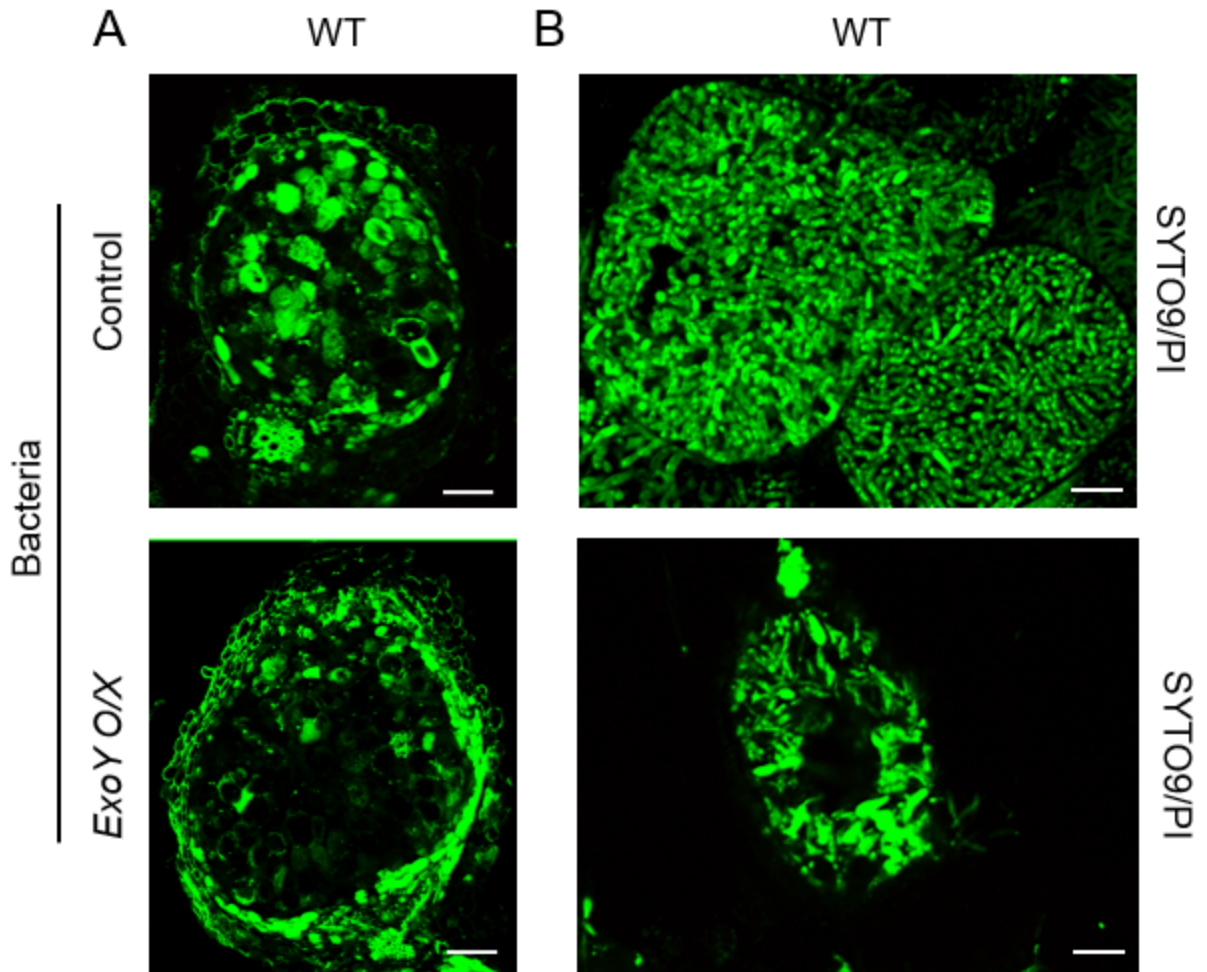
To analyze the temporal and spatial pattern of *exoY* gene expression during nodulation, I used the *GUS* reporter gene under the control of an *exoY* promoter. I inoculated WT plants with the *pExoY::GUS-expressing Rm1021* strain. *ExoY* was only expressed by bacteria in infection threads, but not in symbiosomes (Fig. C.1), implying that succinoglycan is not produced by intracellular bacteria. To understand the effect of succinoglycan production inside symbiosomes, I used a *Rm1021* strain expressing the *exoY* gene under the control of a constitutive promoter (Jones, 2012), referred to as *ExoY O/X*. As a control, I used a *Rm1021* strain including empty vector. Being compared to control, WT plants inoculated with *ExoY O/X* bacteria contained fewer infected host cells, and these cells possessed aberrant-shaped bacteroids (Fig. C.2), demonstrating the inhibition of succinoglycan production in intracellular bacteria is critical for symbiotic nitrogen fixation.

LYM proteins localize to the infection thread and the symbiosome. DNF2 blocks the host defense response against intracellular rhizobia inside symbiosomes by removing LYM proteins from the peribacteroid membrane; however how these bacteria overcome LYM recognition in infection threads and during internalization is unknown. Succinoglycan deficiency in invading bacteria and the presence of LYM proteins on the peribacteroid membrane in *dnf2* plants lead to similar phenotypes such as phenolics accumulation and the induction of host defense response, suggesting that succinoglycan might shield symbiont bacteria from LYM recognition inside infection threads. To test whether LYM proteins underlie the *exoY* phenotype, I inoculated WT, *lym1* and *lym2* plants with *exoY* bacteria and checked nodule formation, using WT bacteria as a control. WT bacteria-inoculated plants had nodules; however, the inoculation with *exoY* bacteria did not result in nodule formation in WT, *lym1* and *lym2* plants (Fig. C.3), implying that the *exoY* phenotype during nodulation is not LYM1 or LYM2 dependent. Furthermore, I tested whether succinoglycan production inside symbiosomes prevent rhizobia from LYM recognition, and thereby eliminate the *dnf2* phenotype. I inoculated *dnf2* plants with *ExoY O/X* bacteria and examined nodules by Live/Dead staining. *ExoY O/X*-inoculated *dnf2* plants possessed infected cells with elongated bacteria, and showed less, or no phenolics accumulation (Fig. C.4), demonstrating that succinoglycan inhibits LYM-dependent host immunity, and partially rescues the *dnf2* phenotype.



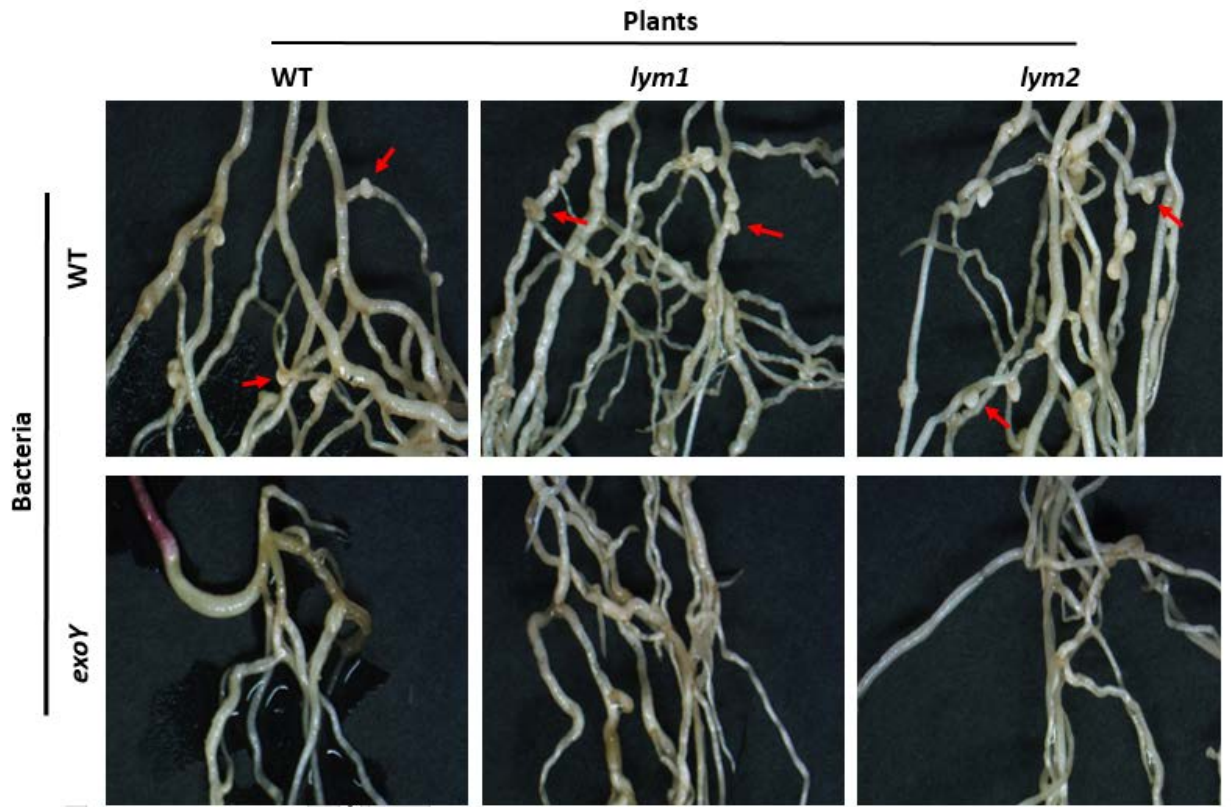
**Figure C.1** *ExoY* is only expressed by infection thread-localized bacteria inside nodules

**(A)** *ExoY* promoter activity in 21 dpi nodules inoculated with *Rm1021* carrying *pExoY::GUS*. GUS activity is shown in blue. Scale bar, 100  $\mu\text{m}$ . **(B)** Magnification of the region outlined (dashed line) in A. Arrows indicate *ExoY*-expressing bacteria in infection threads. Pink color results from leghemoglobin proteins.



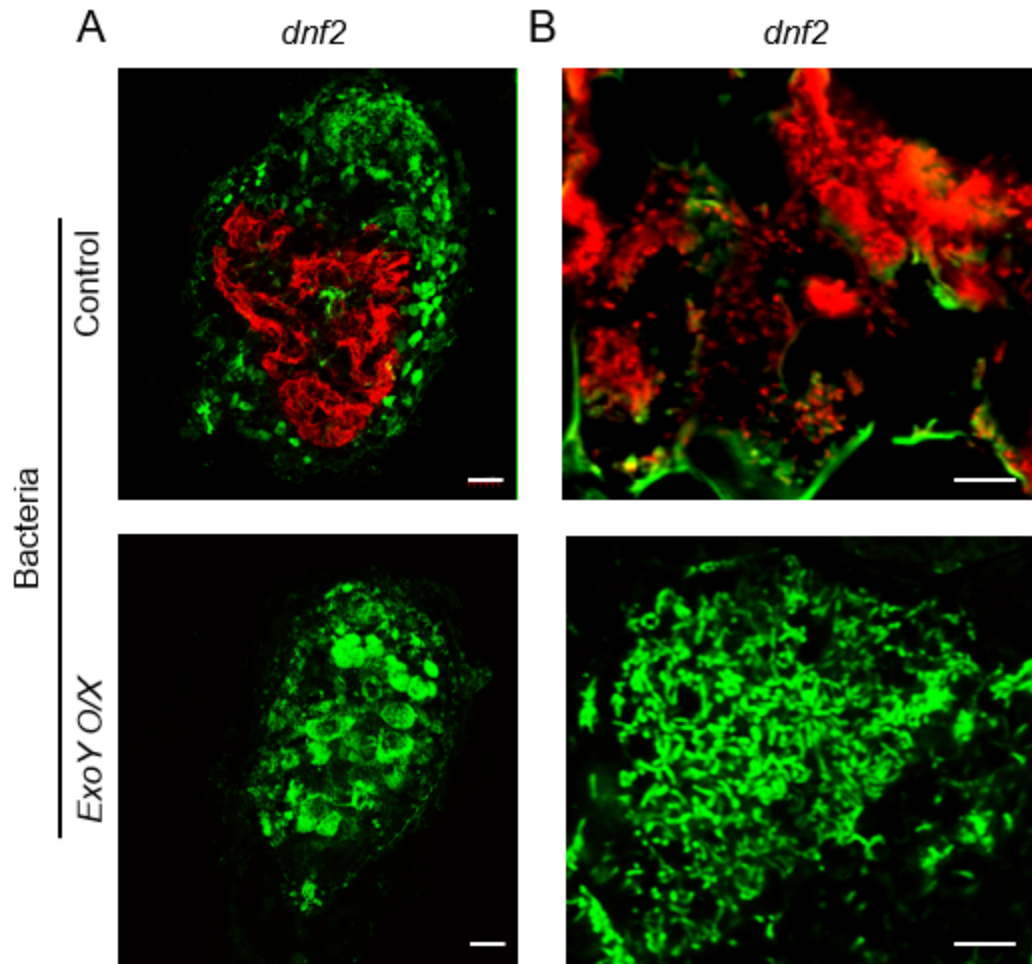
**Figure C.2 Bacteria constitutively expressing *ExoY* cause defective nodules.**

Live/dead staining of WT nodules inoculated with *ExoY*-overexpressing and control bacteria (*Sinorhizobium meliloti* strain *Rm1021*). The images of (A) nodules and (B) nitrogen-fixation zones. Scale bar is 100  $\mu\text{m}$  in A and 5  $\mu\text{m}$  in B. Control: bacteria with empty vector. Green color from SYTO 9 indicates live bacteria and red color from propidium iodide (PI) indicates dead bacteria.



**Figure C.3 *ExoY* mutant bacteria do not produce nodules in *lym1* and *lym2* plants**

Nodule formation by WT, *lym1* and *lym2* plants (*Medicago truncatula* R108 ecotype) inoculated with *exoY* and WT bacteria (*Sinorhizobium meliloti* strain Rm1021). Arrows indicate nodules.



**Figure C.4** *ExoY*-overexpression in bacteria partially rescues the *dnf2* phenotype

Live/dead staining of *dnf2* nodules inoculated with *ExoY*-overexpressing and control bacteria (*Sinorhizobium meliloti* strain *Rm1021*). The images of **(A)** nodules and **(B)** nitrogen-fixation zones. Scale bar is 100  $\mu\text{m}$  in A and 5  $\mu\text{m}$  in B. Control: bacteria with empty vector. Green color from SYTO 9 indicates live bacteria and red color from propidium iodide (PI) indicates dead bacteria.

## BIBLIOGRAPHY

- Albus, U., Baier, R., Holst, O., Pühler, A., and Niehaus, K. (2001). Suppression of an elicitor-induced oxidative burst reaction in *Medicago sativa* cell cultures by *Sinorhizobium meliloti* lipopolysaccharides. *New Phytol.* *151*, 597–606.
- Ané, J.-M., Kiss, G.B., Riely, B.K., Penmetsa, R.V., Oldroyd, G.E.D., Ajax, C., Lévy, J., Debelle, F., Baek, J.-M., Kalo, P., et al. (2004). *Medicago truncatula* DMI1 required for bacterial and fungal symbioses in legumes. *Science* *303*, 1364–1367.
- Auclair, S.M., Bhanu, M.K., and Kendall, D.A. (2012). Signal peptidase I: Cleaving the way to mature proteins. *Protein Sci. Publ. Protein Soc.* *21*, 13–25.
- Benedito, V.A., Torres-Jerez, I., Murray, J.D., Andriankaja, A., Allen, S., Kakar, K., Wandrey, M., Verdier, J., Zuber, H., Ott, T., et al. (2008). A gene expression atlas of the model legume *Medicago truncatula*. *Plant J.* *55*, 504–513.
- Bielnicki, J., Devedjiev, Y., Derewenda, U., Dauter, Z., Joachimiak, A., and Derewenda, Z.S. (2006). *B. subtilis* ykuD protein at 2.0 Å resolution: insights into the structure and function of a novel, ubiquitous family of bacterial enzymes. *Proteins* *62*, 144–151.
- Bonifacino, J.S., and Glick, B.S. (2004). The mechanisms of vesicle budding and fusion. *Cell* *116*, 153–166.
- Boulos, L., Prévost, M., Barbeau, B., Coallier, J., and Desjardins, R. (1999). LIVE/DEAD BacLight : application of a new rapid staining method for direct enumeration of viable and total bacteria in drinking water. *J. Microbiol. Methods* *37*, 77–86.
- Brady, J.D., Sadler, I.H., and Fry, S.C. (1998). Pulcherosine, an oxidatively coupled trimer of tyrosine in plant cell walls: its role in cross-link formation. *Phytochemistry* *47*, 349–353.
- Brewin, N.J. (2004). Plant Cell Wall Remodelling in the Rhizobium–Legume Symbiosis. *Crit. Rev. Plant Sci.* *23*, 293–316.
- Buist, G., Steen, A., Kok, J., and Kuipers, O.P. (2008). LysM, a widely distributed protein motif for binding to (peptido)glycans. *Mol. Microbiol.* *68*, 838–847.
- Cassab, G.I. (1998). Plant Cell Wall Proteins. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* *49*, 281–309.
- Catalano, C.M., Lane, W.S., and Sherrier, D.J. (2004). Biochemical characterization of symbiosome membrane proteins from *Medicago truncatula* root nodules. *Electrophoresis* *25*, 519–531.
- Chen, Y.A., and Scheller, R.H. (2001). SNARE-mediated membrane fusion. *Nat. Rev. Mol. Cell Biol.* *2*, 98–106.

- Chen, G., Snyder, C.L., Greer, M.S., and Weselake, R.J. (2011). Biology and Biochemistry of Plant Phospholipases. *Crit. Rev. Plant Sci.* *30*, 239–258.
- Cheung, A.Y., Wang, H., and Wu, H.M. (1995). A floral transmitting tissue-specific glycoprotein attracts pollen tubes and stimulates their growth. *Cell* *82*, 383–393.
- Curtis, M.D., and Grossniklaus, U. (2003). A gateway cloning vector set for high-throughput functional analysis of genes in planta. *Plant Physiol.* *133*, 462–469.
- D’Antuono, A.L., Ott, T., Krusell, L., Voroshilova, V., Ugalde, R.A., Udvardi, M., and Lepek, V.C. (2008). Defects in rhizobial cyclic glucan and lipopolysaccharide synthesis alter legume gene expression during nodule development. *Mol. Plant-Microbe Interact. MPMI* *21*, 50–60.
- Dénarié, J., Debellé, F., and Promé, J.C. (1996). Rhizobium lipo-chitooligosaccharide nodulation factors: signaling molecules mediating recognition and morphogenesis. *Annu. Rev. Biochem.* *65*, 503–535.
- D’Haeze, W., and Holsters, M. (2002). Nod factor structures, responses, and perception during initiation of nodule development. *Glycobiology* *12*, 79R–105R.
- Doering, T.L., Englund, P.T., and Hart, G.W. (2001). Detection of Glycophospholipid Anchors on Proteins. In *Current Protocols in Protein Science*, (John Wiley & Sons, Inc.), p.
- Dun, A.R., Rickman, C., and Duncan, R.R. (2010). The t-SNARE complex: a close up. *Cell. Mol. Neurobiol.* *30*, 1321–1326.
- Durgo, H., Klement, E., Hunyadi-Gulyas, E., Szucs, A., Kereszt, A., Medzihradzky, K.F., and Kondorosi, E. (2015). Identification of nodule-specific cysteine-rich plant peptides in endosymbiotic bacteria. *PROTEOMICS* *15*, 2291–2295.
- Eisenhaber, B., Bork, P., and Eisenhaber, F. (2001). Post-translational GPI lipid anchor modification of proteins in kingdoms of life: analysis of protein sequence data from complete genomes. *Protein Eng.* *14*, 17–25.
- Elkon, R., Ugalde, A.P., and Agami, R. (2013). Alternative cleavage and polyadenylation: extent, regulation and function. *Nat. Rev. Genet.* *14*, 496–506.
- Ellis, M., Egelund, J., Schultz, C.J., and Bacic, A. (2010). Arabinogalactan-Proteins: Key Regulators at the Cell Surface? *Plant Physiol.* *153*, 403–419.
- Esseling, J.J., Lhuissier, F.G.P., and Emons, A.M.C. (2003). Nod factor-induced root hair curling: continuous polar growth towards the point of nod factor application. *Plant Physiol.* *132*, 1982–1988.
- FÅHRAEUS, G. (1957). The Infection of Clover Root Hairs by Nodule Bacteria Studied by a Simple Glass Slide Technique. *Microbiology* *16*, 374–381.

- Ferguson, M.A. (1999). The structure, biosynthesis and functions of glycosylphosphatidylinositol anchors, and the contributions of trypanosome research. *J. Cell Sci.* *112 ( Pt 17)*, 2799–2809.
- Ferguson, M.A., Kinoshita, T., and Hart, G.W. (2009). Glycosylphosphatidylinositol Anchors. In *Essentials of Glycobiology*, A. Varki, R.D. Cummings, J.D. Esko, H.H. Freeze, P. Stanley, C.R. Bertozzi, G.W. Hart, and M.E. Etzler, eds. (Cold Spring Harbor (NY): Cold Spring Harbor Laboratory Press), p.
- Fliegmann, J., Uhlenbroich, S., Shinya, T., Martinez, Y., Lefebvre, B., Shibuya, N., and Bono, J.-J. (2011). Biochemical and phylogenetic analysis of CEBiP-like LysM domain-containing extracellular proteins in higher plants. *Plant Physiol. Biochem.* *49*, 709–720.
- Foussard, M., Garnerone, A.M., Ni, F., Soupène, E., Boistard, P., and Batut, J. (1997). Negative autoregulation of the *Rhizobium meliloti* fixK gene is indirect and requires a newly identified regulator, FixT. *Mol. Microbiol.* *25*, 27–37.
- Frugier, F., Kosuta, S., Murray, J.D., Crespi, M., and Szczyglowski, K. (2008). Cytokinin: secret agent of symbiosis. *Trends Plant Sci.* *13*, 115–120.
- Garvey, K.J., Saedi, M.S., and Ito, J. (1986). Nucleotide sequence of *Bacillus* phage phi 29 genes 14 and 15: homology of gene 15 with other phage lysozymes. *Nucleic Acids Res.* *14*, 10001–10008.
- Geurts, R., Fedorova, E., and Bisseling, T. (2005). Nod factor signaling genes and their function in the early stages of *Rhizobium* infection. *Curr. Opin. Plant Biol.* *8*, 346–352.
- Gibson, K.E., Kobayashi, H., and Walker, G.C. (2008). Molecular Determinants of a Symbiotic Chronic Infection. *Annu. Rev. Genet.* *42*, 413–441.
- Gilles-Gonzalez, M.A., Ditta, G.S., and Helinski, D.R. (1991). A haemoprotein with kinase activity encoded by the oxygen sensor of *Rhizobium meliloti*. *Nature* *350*, 170–172.
- Glucksmann, M.A., Reuber, T.L., and Walker, G.C. (1993). Family of glycosyl transferases needed for the synthesis of succinoglycan by *Rhizobium meliloti*. *J. Bacteriol.* *175*, 7033–7044.
- Goñi, F.M., Montes, L.-R., and Alonso, A. (2012). Phospholipases C and sphingomyelinases: Lipids as substrates and modulators of enzyme activity. *Prog. Lipid Res.* *51*, 238–266.
- Greene, E.A., Erard, M., Dedieu, A., and Barker, D.G. (1998). MtENOD16 and 20 are members of a family of phycocyanin-related early nodulins. *Plant Mol. Biol.* *36*, 775–783.
- Griffith, O.H., and Ryan, M. (1999). Bacterial phosphatidylinositol-specific phospholipase C: structure, function, and interaction with lipids. *Biochim. Biophys. Acta* *1441*, 237–254.
- Gust, A.A., Willmann, R., Desaki, Y., Grabherr, H.M., and Nürnberger, T. (2012). Plant LysM proteins: modules mediating symbiosis and immunity. *Trends Plant Sci.* *17*, 495–502.

- Haney, C.H., and Long, S.R. (2010). Plant flotillins are required for infection by nitrogen-fixing bacteria. *Proc. Natl. Acad. Sci. U. S. A.* *107*, 478–483.
- Heinz, D.W., Essen, L.O., and Williams, R.L. (1998). Structural and mechanistic comparison of prokaryotic and eukaryotic phosphoinositide-specific phospholipases C. *J. Mol. Biol.* *275*, 635–650.
- Helliwell, C.A., and Waterhouse, P.M. (2005). Constructs and Methods for Hairpin RNA-Mediated Gene Silencing in Plants. B.-M. in *Enzymology*, ed. (Academic Press), pp. 24–35.
- van Hengel, A.J., and Roberts, K. (2003). AtAGP30, an arabinogalactan-protein in the cell walls of the primary root, plays a role in root regeneration and seed germination. *Plant J. Cell Mol. Biol.* *36*, 256–270.
- Hooper, N.M. (2001). Determination of glycosyl-phosphatidylinositol membrane protein anchorage. *Proteomics* *1*, 748–755.
- Hou, Y., Guo, X., Cyprys, P., Zhang, Y., Bleckmann, A., Cai, L., Huang, Q., Luo, Y., Gu, H., Dresselhaus, T., et al. (2016). Maternal ENODLs Are Required for Pollen Tube Reception in Arabidopsis. *Curr. Biol. CB* *26*, 2343–2350.
- Hunt, A.G. (2011). RNA regulatory elements and polyadenylation in plants. *Front. Plant Sci.* *2*, 109.
- Ivanov, S., Fedorova, E.E., Limpens, E., De Mita, S., Genre, A., Bonfante, P., and Bisseling, T. (2012). Rhizobium–legume symbiosis shares an exocytotic pathway required for arbuscule formation. *Proc. Natl. Acad. Sci. U. S. A.* *109*, 8316–8321.
- Jiao, J., Mizukami, A.G., Sankaranarayanan, S., Yamguchi, J., Itami, K., and Higashiyama, T. (2016). Structure-Activity Relation of AMOR Sugar Molecule that Activates Pollen-Tubes for Ovular Guidance. *Plant Physiol.* pp.01655.2016.
- Jones, K.M. (2012). Increased Production of the Exopolysaccharide Succinoglycan Enhances *Sinorhizobium meliloti* 1021 Symbiosis with the Host Plant *Medicago truncatula*. *J. Bacteriol.* *194*, 4322–4331.
- Kaku, H., Nishizawa, Y., Ishii-Minami, N., Akimoto-Tomiya, C., Dohmae, N., Takio, K., Minami, E., and Shibuya, N. (2006). Plant cells recognize chitin fragments for defense signaling through a plasma membrane receptor. *Proc. Natl. Acad. Sci. U. S. A.* *103*, 11086–11091.
- Kawaharada, Y., Kelly, S., Nielsen, M.W., Hjuler, C.T., Gysel, K., Muszyński, A., Carlson, R.W., Thygesen, M.B., Sandal, N., Asmussen, M.H., et al. (2015). Receptor-mediated exopolysaccharide perception controls bacterial infection. *Nature* *523*, 308–312.
- Kelley, L.A., and Sternberg, M.J.E. (2009). Protein structure prediction on the Web: a case study using the Phyre server. *Nat. Protoc.* *4*, 363–371.

- Kim, S.-J., and Brandizzi, F. (2012). News and Views into the SNARE Complexity in Arabidopsis. *Front. Plant Sci.* *3*.
- Kim, M., Chen, Y., Xi, J., Waters, C., Chen, R., and Wang, D. (2015). An antimicrobial peptide essential for bacterial survival in the nitrogen-fixing symbiosis. *Proc. Natl. Acad. Sci. U. S. A.* *112*, 15238–15243.
- Kipreos, E.T., and Pagano, M. (2000). The F-box protein family. *Genome Biol.* *1*, reviews3002.1-reviews3002.7.
- Kjellbom, P., Snogerup, L., Stöhr, C., Reuzeau, C., McCabe, P.F., and Pennell, R.I. (1997). Oxidative cross-linking of plasma membrane arabinogalactan proteins. *Plant J. Cell Mol. Biol.* *12*, 1189–1196.
- Lampart, D.T.A., Kieliszewski, M.J., and Showalter, A.M. (2006). Salt stress upregulates periplasmic arabinogalactan proteins: using salt stress to analyse AGP function. *New Phytol.* *169*, 479–492.
- Leong, S.A., Williams, P.H., and Ditta, G.S. (1985). Analysis of the 5' regulatory region of the gene for delta-aminolevulinic acid synthetase of *Rhizobium meliloti*. *Nucleic Acids Res.* *13*, 5965–5976.
- Lévy, J., Bres, C., Geurts, R., Chalhoub, B., Kulikova, O., Duc, G., Journet, E.-P., Ané, J.-M., Lauber, E., Bisseling, T., et al. (2004). A Putative Ca<sup>2+</sup> and Calmodulin-Dependent Protein Kinase Required for Bacterial and Fungal Symbioses. *Science* *303*, 1361–1364.
- Lohar, D.P., Sharopova, N., Endre, G., Peñuela, S., Samac, D., Town, C., Silverstein, K.A.T., and VandenBosch, K.A. (2006). Transcript Analysis of Early Nodulation Events in *Medicago truncatula*. *Plant Physiol.* *140*, 221–234.
- Lohar, D.P., Haridas, S., Gantt, J.S., and VandenBosch, K.A. (2007). A transient decrease in reactive oxygen species in roots leads to root hair deformation in the legume-rhizobia symbiosis. *New Phytol.* *173*, 39–49.
- López-Lara, I.M., Sohlenkamp, C., and Geiger, O. (2003). Membrane lipids in plant-associated bacteria: their biosyntheses and possible functions. *Mol. Plant-Microbe Interact. MPMI* *16*, 567–579.
- Mandel, C.R., Bai, Y., and Tong, L. (2008). Protein factors in pre-mRNA 3'-end processing. *Cell. Mol. Life Sci. CMLS* *65*, 1099–1122.
- Mashiguchi, K., Asami, T., and Suzuki, Y. (2009). Genome-wide identification, structure and expression studies, and mutant collection of 22 early nodulin-like protein genes in *Arabidopsis*. *Biosci. Biotechnol. Biochem.* *73*, 2452–2459.

- Mathesius, U., Schlaman, H.R., Spaink, H.P., Of Sautter, C., Rolfe, B.G., and Djordjevic, M.A. (1998). Auxin transport inhibition precedes root nodule formation in white clover roots and is regulated by flavonoids and derivatives of chitin oligosaccharides. *Plant J. Cell Mol. Biol.* *14*, 23–34.
- Mayor, S., and Riezman, H. (2004). Sorting GPI-anchored proteins. *Nat. Rev. Mol. Cell Biol.* *5*, 110–120.
- McConville, M.J., and Ferguson, M.A. (1993). The structure, biosynthesis and function of glycosylated phosphatidylinositols in the parasitic protozoa and higher eukaryotes. *Biochem. J.* *294*, 305–324.
- Medzhitov, R., and Janeway, C.A. (1997). Innate immunity: the virtues of a nonclonal system of recognition. *Cell* *91*, 295–298.
- Mendes, R., Kruijt, M., Bruijn, I. de, Dekkers, E., Voort, M. van der, Schneider, J.H.M., Piceno, Y.M., DeSantis, T.Z., Andersen, G.L., Bakker, P.A.H.M., et al. (2011). Deciphering the Rhizosphere Microbiome for Disease-Suppressive Bacteria. *Science* *332*, 1097–1100.
- Mergaert, P., Uchiumi, T., Alunni, B., Evanno, G., Cheron, A., Catrice, O., Mausset, A.-E., Barloy-Hubler, F., Galibert, F., Kondorosi, A., et al. (2006). Eukaryotic control on bacterial cell cycle and differentiation in the *Rhizobium*-legume symbiosis. *Proc. Natl. Acad. Sci. U. S. A.* *103*, 5230–5235.
- Miller, J.B., and Oldroyd, G.E.D. (2012). The Role of Diffusible Signals in the Establishment of Rhizobial and Mycorrhizal Symbioses. In *Signaling and Communication in Plant Symbiosis*, S. Perotto, and F. Baluška, eds. (Springer Berlin Heidelberg), pp. 1–30.
- Motose, H., Sugiyama, M., and Fukuda, H. (2004). A proteoglycan mediates inductive interaction during plant vascular development. *Nature* *429*, 873–878.
- Neumann, K.-H., Kumar, A., and Sopory, S.K. (2008). *Recent Advances in Plant Biotechnology and Its Applications: Prof. Dr. Karl-Hermann Neumann Commemorative Volume (I. K. International Pvt Ltd)*.
- Nozaki, M., Ohishi, K., Yamada, N., Kinoshita, T., Nagy, A., and Takeda, J. (1999). Developmental abnormalities of glycosylphosphatidylinositol-anchor-deficient embryos revealed by Cre/loxP system. *Lab. Investig. J. Tech. Methods Pathol.* *79*, 293–299.
- Oldroyd, G.E.D. (2013). Speak, friend, and enter: signalling systems that promote beneficial symbiotic associations in plants. *Nat. Rev. Microbiol.* *11*, 252–263.
- Oldroyd, G.E.D., and Downie, J.A. (2008). Coordinating nodule morphogenesis with rhizobial infection in legumes. *Annu. Rev. Plant Biol.* *59*, 519–546.
- Oldroyd, G.E.D., Murray, J.D., Poole, P.S., and Downie, J.A. (2011). The rules of engagement in the legume-rhizobial symbiosis. *Annu. Rev. Genet.* *45*, 119–144.

- Pan, H., Oztas, O., Zhang, X., Wu, X., Stonoha, C., Wang, E., Wang, B., and Wang, D. (2016). A symbiotic SNARE protein generated by alternative termination of transcription. *Nat. Plants* 2, 15197.
- Park, M.H., Suzuki, Y., Chono, M., Knox, J.P., and Yamaguchi, I. (2003). CsAGP1, a gibberellin-responsive gene from cucumber hypocotyls, encodes a classical arabinogalactan protein and is involved in stem elongation. *Plant Physiol.* 131, 1450–1459.
- Paulick, M.G., and Bertozzi, C.R. (2008). The Glycosylphosphatidylinositol Anchor: A Complex Membrane-Anchoring Structure for Proteins†. *Biochemistry (Mosc.)* 47, 6991–7000.
- Prell, J., and Poole, P. (2006). Metabolic changes of rhizobia in legume nodules. *Trends Microbiol.* 14, 161–168.
- Ramu, S.K., Peng, H.-M., and Cook, D.R. (2002). Nod factor induction of reactive oxygen species production is correlated with expression of the early nodulin gene *rip1* in *Medicago truncatula*. *Mol. Plant-Microbe Interact. MPMI* 15, 522–528.
- Reuber, T.L., and Walker, G.C. (1993). Biosynthesis of succinoglycan, a symbiotically important exopolysaccharide of *Rhizobium meliloti*. *Cell* 74, 269–280.
- Richard, P., and Manley, J.L. (2009). Transcription termination by nuclear RNA polymerases. *Genes Dev.* 23, 1247–1269.
- Roche, P., Debelle, F., Maillet, F., Lerouge, P., Faucher, C., Truchet, G., Dénarié, J., and Promé, J.C. (1991). Molecular basis of symbiotic host specificity in *Rhizobium meliloti*: *nodH* and *nodPQ* genes encode the sulfation of lipo-oligosaccharide signals. *Cell* 67, 1131–1143.
- Rosenberg, E., and Gophna, U. (2011). *Beneficial Microorganisms in Multicellular Life Forms* (Springer Science & Business Media).
- Rosendahl, L., Dilworth, M.J., Glenn, A.R., Rosendahl, L., Dilworth, M.J., and Glenn, A.R. (1992). Exchange of metabolites across the peribacteroid membrane in pea root nodules. *J Plant Physiol* 139, 635–638.
- Roux, B., Rodde, N., Jardinaud, M.-F., Timmers, T., Sauviac, L., Cottret, L., Carrère, S., Sallet, E., Courcelle, E., Moreau, S., et al. (2014). An integrated analysis of plant and bacterial gene expression in symbiotic root nodules using laser-capture microdissection coupled to RNA sequencing. *Plant J. Cell Mol. Biol.* 77, 817–837.
- Sachs, A. (1990). The role of poly(A) in the translation and stability of mRNA. *Curr. Opin. Cell Biol.* 2, 1092–1098.
- Schultz, C., Gilson, P., Oxley, D., Youl, J., and Bacic, A. (1998). GPI-anchors on arabinogalactan-proteins: implications for signalling in plants. *Trends Plant Sci.* 3, 426–431.
- Seifert, G.J., and Roberts, K. (2007). The biology of arabinogalactan proteins. *Annu. Rev. Plant Biol.* 58, 137–161.

- Shi, H., Kim, Y., Guo, Y., Stevenson, B., and Zhu, J.-K. (2003). The Arabidopsis SOS5 locus encodes a putative cell surface adhesion protein and is required for normal cell expansion. *Plant Cell* 15, 19–32.
- Shimizu, T., Nakano, T., Takamizawa, D., Desaki, Y., Ishii-Minami, N., Nishizawa, Y., Minami, E., Okada, K., Yamane, H., Kaku, H., et al. (2010). Two LysM receptor molecules, CEBiP and OsCERK1, cooperatively regulate chitin elicitor signaling in rice. *Plant J. Cell Mol. Biol.* 64, 204–214.
- Showalter, A.M., and Basu, D. (2016). Extensin and Arabinogalactan-Protein Biosynthesis: Glycosyltransferases, Research Challenges, and Biosensors. *Front. Plant Sci.* 7.
- Sigurbjörnsdóttir, S., Mathew, R., and Leptin, M. (2014). Molecular mechanisms of de novo lumen formation. *Nat. Rev. Mol. Cell Biol.* 15, 665–676.
- Takagaki, Y., Seipelt, R.L., Peterson, M.L., and Manley, J.L. (1996). The polyadenylation factor CstF-64 regulates alternative processing of IgM heavy chain pre-mRNA during B cell differentiation. *Cell* 87, 941–952.
- Taylor, D.R., and Hooper, N.M. (2011). GPI-Anchored Proteins in Health and Disease. In *Post-Translational Modifications in Health and Disease*, C.J. Vidal, ed. (Springer New York), pp. 39–55.
- Tian, B., and Manley, J.L. (2016). Alternative polyadenylation of mRNA precursors. *Nat. Rev. Mol. Cell Biol.* *advance online publication*.
- Udvardi, M., and Poole, P.S. (2013). Transport and metabolism in legume-rhizobia symbioses. *Annu. Rev. Plant Biol.* 64, 781–805.
- Van de Velde, W., Zehirov, G., Szatmari, A., Debreczeny, M., Ishihara, H., Kevei, Z., Farkas, A., Mikulass, K., Nagy, A., Tiricz, H., et al. (2010). Plant peptides govern terminal differentiation of bacteria in symbiosis. *Science* 327, 1122–1126.
- Van Hengel, A.J., Van Kammen, A., and De Vries, S.C. (2002). A relationship between seed development, Arabinogalactan-proteins (AGPs) and the AGP mediated promotion of somatic embryogenesis. *Physiol. Plant.* 114, 637–644.
- Vasse, J., de Billy, F., and Truchet, G. (1993). Abortion of infection during the *Rhizobium meliloti*—alfalfa symbiotic interaction is accompanied by a hypersensitive reaction. *Plant J.* 4, 555–566.
- Verma, D.P.S., Gu, X., and Hong, Z. (1995). Biogenesis of the Peribacteroid Membrane in Root Nodules: Roles of Dynamev and Phosphatidyl-Inositol 3-Kinase. In *Nitrogen Fixation: Fundamentals and Applications*, I.A. Tikhonovich, N.A. Provorov, V.I. Romanov, and W.E. Newton, eds. (Springer Netherlands), pp. 467–470.

- Vernoud, V., Journet, E.-P., and Barker, D.G. (1999). MtENOD20, a Nod Factor-Inducible Molecular Marker for Root Cortical Cell Activation. *Mol. Plant. Microbe Interact.* *12*, 604–614.
- Wang, D., Griffitts, J., Starker, C., Fedorova, E., Limpens, E., Ivanov, S., Bisseling, T., and Long, S. (2010). A nodule-specific protein secretory pathway required for nitrogen-fixing symbiosis. *Science* *327*, 1126–1129.
- Weaver, C.D., Shomer, N.H., Louis, C.F., and Roberts, D.M. (1994). Nodulin 26, a nodule-specific symbiosome membrane protein from soybean, is an ion channel. *J. Biol. Chem.* *269*, 17858–17862.
- Weigel, D., and Glazebrook, J. (2002). *Arabidopsis: A Laboratory Manual* (CSHL Press).
- Willmann, R., Lajunen, H.M., Erbs, G., Newman, M.-A., Kolb, D., Tsuda, K., Katagiri, F., Fliegmann, J., Bono, J.-J., Cullimore, J.V., et al. (2011). Arabidopsis lysin-motif proteins LYM1 LYM3 CERK1 mediate bacterial peptidoglycan sensing and immunity to bacterial infection. *Proc. Natl. Acad. Sci. U. S. A.* *108*, 19824–19829.
- Wu, H.M., Wang, H., and Cheung, A.Y. (1995). A pollen tube growth stimulatory glycoprotein is deglycosylated by pollen tubes and displays a glycosylation gradient in the flower. *Cell* *82*, 395–403.
- Yamaji-Hasegawa, A., and Tsujimoto, M. (2006). Asymmetric distribution of phospholipids in biomembranes. *Biol. Pharm. Bull.* *29*, 1547–1553.
- Zipfel, C. (2008). Pattern-recognition receptors in plant innate immunity. *Curr. Opin. Immunol.* *20*, 10–16.
- UNEP YEAR BOOK 2014.