# Understanding the Mechanism of Symbiosome-Secreted Proteins at Nodulation in *Medicago truncatula*

by

Ceren ÖZDEMİR

A Dissertation Submitted to the
Graduate School of Sciences and Engineering
in Partial Fulfillment of the Requirements for
the Degree of

Master of Science

in

Molecular Biology and Genetics


September 3th, 2025

# Understanding the Mechanism of Symbiosome-Secreted Proteins at Nodulation in *M. truncatula*

Koç University

Graduate School of Sciences and Engineering

This is to certify that I have examined this copy of a master's thesis by

**Ceren ÖZDEMİR**

and have found that it is complete and satisfactory in all respects,
and that any and all revisions required by the final
examining committee have been made.

Committee Members:

---

Assist. Prof. Onur ÖZTAŞ

---

Assist . Prof. Hasan DEMİRCİ

---

Assist. Prof. Seçkin EROĞLU

Date: September 3th, 2025

*For my Mama and Pops…*

# ABSTRACT


**Understanding the Mechanism of Symbiosome-Secreted Proteins at Nodulation in *Medicago truncatula***

**Ceren ÖZDEMİR**

**Master of Science in Molecular Biology and Genetics**

**September 3th, 2025**

*Medicago truncatula*, commonly known as barrel medic, is an annual legume indigenous to the Mediterranean region and widely used as a model organism for legume research. It establishes symbiotic interactions with nitrogen-fixing Sinorhizobium meliloti in a specialized root organ called the nodule. Within nodules, rhizobia are endocytosed by host cells and enclosed in symbiosomes. Inside this compartment, the bacteria differentiate into their nitrogen-fixing form and provide ammonia to the host. This developmental transition requires host proteins secreted into the symbiosome, which regulate bacterial survival and differentiation.

In this project, we aimed to identify symbiosome-secreted host proteins and assess their effects on nitrogen fixation in *M. truncatula*. At first part late nodulins called Nodule Cysteine Rich (NCR) peptides and their effects has been discovered when it binds to bacterial genome. For this symbiosomes were purified by Percoll density gradient centrifugation, and DNA-binding proteins were subsequently isolated. Proteomic characterization was carried out using an Orbitrap platform for Mass Spectrometry. From this analysis, we detected three nodule cysteine-rich (NCR) peptides—NCRAL7, MtNCR418, and MtNCRM81—that bound directly to the rhizobial genome and effects rhizobial differentiation. They are directly secreted to symbiosome or located at ER. Their interaction with bacterial DNA likely alters transcription and contributes to the differentiation of free-living rhizobia into nitrogen-fixing bacteroids. However, the precise DNA regions targeted by these peptides remain to be determined, and their direct effects on transcriptional repression or activation are not yet fully resolved.

In a parallel set of experiments, it has been focused on early nodulins. Among these, ENOD16 was of particular interest due to its predicted glycosylphosphatidylinositol (GPI) anchor, which enables its association with the symbiosome membrane. For analysis Tnt1 insertion mutants related to the protein used for phenotyping and rescue experiments has been run on these mutants. Expression analysis showed that MtENOD16 is localized specifically to nodules, with distinct spatial distribution across developmental zones. Loss-of-function experiments revealed that intracellular rhizobia fail to persist inside nodules lacking ENOD16. These results indicate that MtENOD16 is required for effective nodule function and nitrogen fixation. Still, the precise molecular mechanism by which MtENOD16 stabilizes rhizobia within symbiosomes remains unclear.

Taken together, our findings highlight two critical aspects of host–symbiont crosstalk: (i) DNA-binding NCR peptides such as NCRAL7, NCR418, and NCRM81 may directly reprogram the rhizobial transcriptome, and (ii) ENOD16 is necessary for intracellular rhizobial survival during nodule development. While these results provide new insight into the molecular dialogue within symbiosomes, additional mechanistic studies will be required to define the precise regulatory pathways. Our work underscores the importance of identifying and characterizing host factors to better understand—and potentially optimize—symbiotic nitrogen fixation.

# ÖZETÇE


***Medicago truncatula'da* Nodülasyonda Simbiyozoma Salgılanan Proteinlerin Mekanizmasının Anlaşılması**

**Ceren ÖZDEMİR**

**Moleküler Biyoloji ve Genetik , Yüksek Lisans**

**3 Eylül 2025**

*Medicago truncatula*, yaygın olarak varil yoncası olarak bilinen, Akdeniz bölgesine özgü yıllık bir baklagil türüdür ve baklagil araştırmaları için yaygın olarak model organizma olarak kullanılır. Bu tür, nodül adı verilen özel bir kök organında azot bağlayan *Sinorhizobium meliloti* ile simbiyotik etkileşimler kurar. Nodüllerin içinde, rizobiyalar konak hücreler tarafından endositozlanır ve simbiyozomlar içinde hapsolur. Bu bölmede bakteriler azot bağlayıcı formlarına farklılaşır ve konağa amonyak sağlar. Bu gelişimsel geçiş, bakterilerin hayatta kalmasını ve farklılaşmasını düzenleyen konak proteinlerinin simbiyozom içine salgılanmasını gerektirir.

Bu projede, simbiyozom tarafından salgılanan konak proteinlerini tanımlamayı ve bunların *M. truncatula*'da azot bağlanmasına etkilerini değerlendirmeyi amaçladık. İlk aşamada, Nodule Cysteine Rich (NCR) peptidleri olarak adlandırılan geç nodulinler ve bunların bakteriyel genomla bağlandıklarında etkileri keşfedilmiştir. Bunun için simbiyozomlar Percoll yoğunluk gradient santrifüjleme ile saflaştırılmış ve ardından DNA bağlayıcı proteinler izole edilmiştir. Proteomik karakterizasyon, Kütle Spektrometresi için Orbitrap platformu kullanılarak gerçekleştirilmiştir. Bu analizden, rizobiyal genomuna doğrudan bağlanan ve rizobiyal farklılaşmayı etkileyen üç nodül sistein bakımından zengin (NCR) peptid tespit edildi: NCRAL7, MtNCR418 ve MtNCRM81. Bunlar doğrudan simbiyozoma salgılanmakta veya Endoplazmik Retikulum'da bulunmaktadır. Bakteriyel DNA ile etkileşimleri muhtemelen transkripsiyonu değiştirir ve serbest farklılaşmaya katkıda bulunur. Ancak, bu peptitlerin hedef aldığı kesin DNA bölgeleri henüz belirlenmemiştir ve transkripsiyon baskılanması veya aktivasyonu üzerindeki doğrudan etkileri henüz tam olarak çözülmemiştir.

Paralel bir dizi deneyde, erken nodulinlere odaklanılmıştır. Bunlar arasında, simbiyozom membranıyla birleşmesini sağlayan öngörülen glikosilfosfatidilinositol (GPI) ankrajı nedeniyle ENOD16 özellikle ilgi çekiciydi. Analiz için, fenotipleme ve kurtarma deneylerinde kullanılan proteinle ilgili Tnt1 insersiyon mutantları bu mutantlar üzerinde çalıştırılmıştır. Ekspresyon analizi, MtENOD16'nın özellikle nodüllere lokalize olduğunu ve gelişim bölgeleri arasında belirgin bir uzamsal dağılım gösterdiğini ortaya koymuştur. Fonksiyon kaybı deneyleri, ENOD16'nın bulunmadığı nodüllerin içinde hücre içi rizobiyaların kalıcı olamadığını ortaya koymuştur. Bu sonuçlar, MtENOD16'nın etkili nodül fonksiyonu ve azot fiksasyonu için gerekli olduğunu göstermektedir. Yine de, MtENOD16'nın simbiyozomlar içinde rizobiyaları stabilize ettiği kesin moleküler mekanizma hala belirsizdir.
Sonuçlarımız, konak-simbiyont etkileşiminin iki kritik yönünü vurgulamaktadır: (i) NCRAL7, NCR418 ve NCRM81 gibi DNA bağlayıcı NCR peptitleri, rizobiyal transkriptomu doğrudan yeniden programlayabilir ve (ii) ENOD16, nodül gelişimi sırasında hücre içi rizobiyal hayatta kalması için gereklidir. Bu sonuçlar simbiyozomlar içindeki moleküler diyalog hakkında yeni bilgiler sağlasa da, kesin düzenleyici yolakları tanımlamak için ek mekanizma çalışılması gerekecektir. Çalışmamız, simbiyotik azot

fiksasyonunu daha iyi anlamak ve potansiyel olarak optimize etmek için konakçı faktörleri tanımlamanın ve karakterize etmenin önemini vurgulamaktadır.

# ACKNOWLEDGEMENTS

I would like to express my deepest gratitude to my supervisor, Asst. Prof. Onur ÖZTAŞ for his invaluable guidance, encouragement, and patience throughout this study. His expertise and insightful feedback have been fundamental in shaping this work.

I am also grateful to Öztaş Lab members for their support, collaboration, and for creating a positive and inspiring research environment. Special thanks to Janani KALAIVANI SIVAKUMAR for her assistance in my weird questions and for the many discussions that enriched my understanding of the academia, dealing together.

I am particularly grateful to my friends and colleagues outside of the laboratory Deniz, Kaan and Gökçe who have offered not only intellectual inspiration but also friendship, encouragement, and understanding during challenging times. Their presence has made the journey more enjoyable and meaningful.

Finally, I owe an immeasurable debt of gratitude to my family. Their unconditional love, endless patience, and unwavering belief in me have been the foundation on which I could pursue my academic ambitions. This work is as much theirs as it is mine.

# TABLE OF CONTENTS

# LIST OF TABLES

# LIST OF FIGURES

# ABBREVIATIONS

| | |
|---|---|
| AGP | Arabinogalactan Proteins |
| AMP | Antimicrobial Peptides |
| BNF | Biological Nitrogen Fixation |
| CDS | Coding Sequence |
| DB-NCR | DNA-Binding Nodule Cysteine Rich Peptide |
| DNA | Deoxyribonucleic Acid |
| dpi | Days post inoculation |
| DTT | Dithiothreitol |
| ENOD16 | Early Nodulin-16 |
| EPS | Exopolysaccharide |
| ER | Endoplasmic Reticulum |
| GFP | Green Fluorescent Protein |
| GPI | Glycosylphosphatidylinositol |
| GRP | Glycine-Rich Peptides |
| GUS | β-glucuronidase |
| HEPES | 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid |
| HRGP | Hydroxyproline-rich Glycoproteins |
| Hyp | Hydroxyproline |
| IRLC | Inverted Repeat Lacking Clade |
| KOH | Potassium Hydroxide |
| LCO | Lipochitooligosaccharides |
| LPS | Lipopolysaccharide |
| MES | 2-(N-morpholino)ethanesulfonic acid |
| MgCl2 | Magnesium Chloride |
| MS | Murashige and Skoog |
| $N_2$ | Dinitrogen |
| NaCl | Sodium Chloride |
| Na-P | Sodium Phosphate |

| | |
|---|---|
| NCR | Nodule Cysteine Rich |
| NER | Nucleotide Excision Repair |
| NFR | Nod Factor Receptors |
| $NH_3$ | Ammonium |
| NifH | Nitrogenase Iron Protein |
| PCR | Polymerase Chain Reaction |
| PGPR | Plant growth-promoting rhizobacteria |
| PI | Propidium Iodide |
| PIC | Protease Inhibitor Cocktail |
| RNA | Ribonucleic Acid |
| ROS | Reactive Oxygen Species |
| SDS | Sodium Dodecyl Sulphate |
| PAGE | Polyacrylamide Gel Electrophoresis |
| SNF | Symbiotic Nitrogen Fixation |
| SNM | Standard Nodulation Medium |
| SPC | Signal Peptidase Complex |
| TF | Transcription Factor |
| TX | Triton-X |
| UNIPROT | Universal Protein Knowledgebase |
| UV | Ultraviolet |
| WT | Wildtype |
| YEB | Yeast Extract Beef |

# CHAPTER 1:
# INTRODUCTION TO NITROGEN FIXATION

In the past two decades, robust progress in plant genomics has transformed our understanding of gene function and regulation in plants. These tools have not only helped with the identification of genes related to many important biological processes but have also allowed researchers to map out complex networks. Yet, one of the crucial challenges that remains is deciphering the molecular regulators that orchestrate these genetic programs, particularly in dynamic biological contexts like symbiotic interactions.

## *1.1 The Nitrogen Fixation Process*

Nitrogen is a vital element for life, essential for the production of proteins, RNA and DNA, but as known it is approximately 70% of the atmosphere. The amount of inorganic nitrogen available to plants in the soil is limited and is mainly present in the form of nitrate and ammonium. Agricultural production is often limited by nitrogen source levels. (Pankievicz et al., 2019) . In the early 1900's, Fritz Haber and Carl Bosch created a process for industrial nitrogen fixation, which converts dinitrogen into ammonia; especially important for fertilizer production in agricultural consumption (Erfani et al., 2024). This process requires high amount of energy, consuming about 1–2% of the world's energy supply and it takes approximately one ton of natural gas to produce one ton of synthetic nitrogen fertilizers, and all of these excuses makes the cost extremely expensive (Pankievicz et al., 2019).

Atmospheric dinitrogen ($N_2$) participates in biochemical cycling processes that lead to losses through volatilization, requires continuous replacement of bioavailable nitrogen to maintain ecological nitrogen resources. (Figure 1) (Takai, 2019).

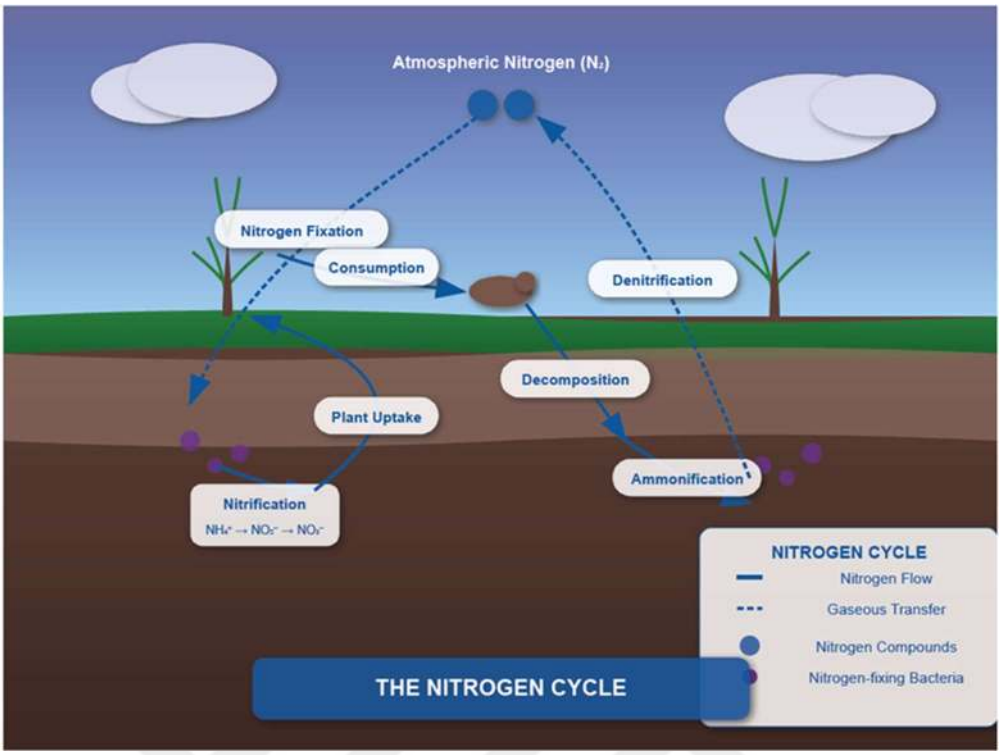



Figure 1.1 The nitrogen cycle.

The conversion of $N_2$ to ammonia ($NH_3$) as known as biological nitrogen fixation (BNF) and is carried out by specialized diazotrophic prokaryotes that possess the nitrogenase enzyme complex (Pankievicz et al., 2019). Optimal nitrogen fixation efficiency occurs within highly evolved symbiotic relationships between these nitrogen-fixing prokaryotes and higher plant partners—predominantly leguminous species and in these symbiotic associations. The host plant provides vital elements as energy resources; while creating low-oxygen microenvironments that protect the sensitive nitrogenase enzyme, facilitating optimal nitrogen fixation (Udvardi & Poole, 2013). Symbiotic nitrogen fixation (SNF) represents a significantly more efficient mechanism than free-living or dependent nitrogen fixation systems. It contributes significantly to agricultural productivity and to the nitrogen cycle. These plant-microbe interactions have evolved specialized anatomical structures as root nodules, that create ideal conditions for metabolically demanding pathways(Mus et al., 2016).

## *1.2 Symbiotic Nitrogen Fixation*

### *1.2.1 Symbiotic Plant-Microbe Interactions*

Plant growth-promoting rhizobacterias (PGPR) respond to root extudates through chemotaxis and colonize the rhizosphere without invading plant cells (Compant et al., 2010; Santi et al., 2013). Nitrogen-fixing PGPRs are found in bacilli and proteobacteria, and their recognition by roots improves plant growth by increasing the uptake of nitrogen, phosphorus, and essential minerals (Ahemad & Kibret, 2014).

Nodulation in plants have originated around 100 million years ago, leading to the development of nodulation in approximately 70% of legume species (van Velzen et al., 2017). Additionally, this has emerged in several plant lines, particularly in actinorhizal plants, which are found across orders of *Angiosperms* (Werner et al., 2014). Throughout time, nitrogen fixing ability in legumes has diversed among couple species within the *alpha-Proteobacteria* and *beta-Proteobacteria* groups, known as rhizobia (Remigi et al., 2016). Phylogenomic methods are uncovering the genetic innovations that enabled plants to form nitrogen-fixing symbiotic relationships with rhizobia (Masson-Boivin & Sachs, 2018).

### *1.2.2 Legume – Rhizobia Interactions*

Legumes have the property to form a symbiotic relationship with nitrogen-fixing soil bacteria known as rhizobia, which leads to the formation of specialized root organs called nodules. In these organs, the bacteria convert atmospheric dinitrogen ($N_2$) into ammonia ($NH_3$) that plants can actively use (Galloway et al., n.d.). This process is called symbiotic nitrogen fixation (SNF) and it offers an environmentally sustainable alternative to synthetic nitrogen fertilizers, whose overuse contributes to water pollution, greenhouse gas emissions, and soil degradation (Oldroyd & Dixon, 2014).

The molecular dialogue that explains SNF is highly complex and involves the perception of microbial signals, transcriptional reprogramming, cellular remodeling, and metabolic coordination (Brewin, 2004; Oldroyd, 2013; Radutoiu et al., 2003; Singh et al., 2014; Udvardi & Poole, 2013). Model legumes such as *Medicago truncatula* and *Lotus japonicus* have been vital in discovering the plant genes involved in nodule formation and

bacterial accommodation (Rose, 2008). But the main difference that sets apart indeterminate nodule structures such as *M. truncatula* is the terminal differentiation of Rhizobia into enlarged, polyploid, nitrogen-fixing bacteroids. This developmental transition is ran by the host and holds a central step in successful symbiosis (Kondorosi et al., 2013).

### *1.2.3 Nodulation Signalling*

When rhizobium enters to the host plant, the plant secretes phenolic compounds like flavonoids, which will took the attention of the bacteria and start chemotaxis under nitrogen deficency. Flavaonoids are effective in triggering nodule development in legume roots, even when rhizobium not exist. The interaction between plants and rhizobia leads to noticeable changes in the plant, such as the curling of root hairs and the formation of early nodule structures(Gautrat et al., 2019; Roy et al., 2020). When rhizobium captured by the host plant, plant secretes phenolic compounds like flavonoids, which will attract the bacteria and also start chemotaxis during nitrogen deficiency. These compounds have the ability to induce nodule formation in the roots of legumes even in the distinct scenario of absence of rhizobium (Luu et al., 2022; Rutten et al., 2020). Nod factors are specific to host-rhizobium pairs and initiate the interaction by being recognized by Nod factor receptors (NFRs) on the plant's cell membrane, further facilitating symbiosis (Kelly et al., 2017; Roy et al., 2020) (Figure 1.2).

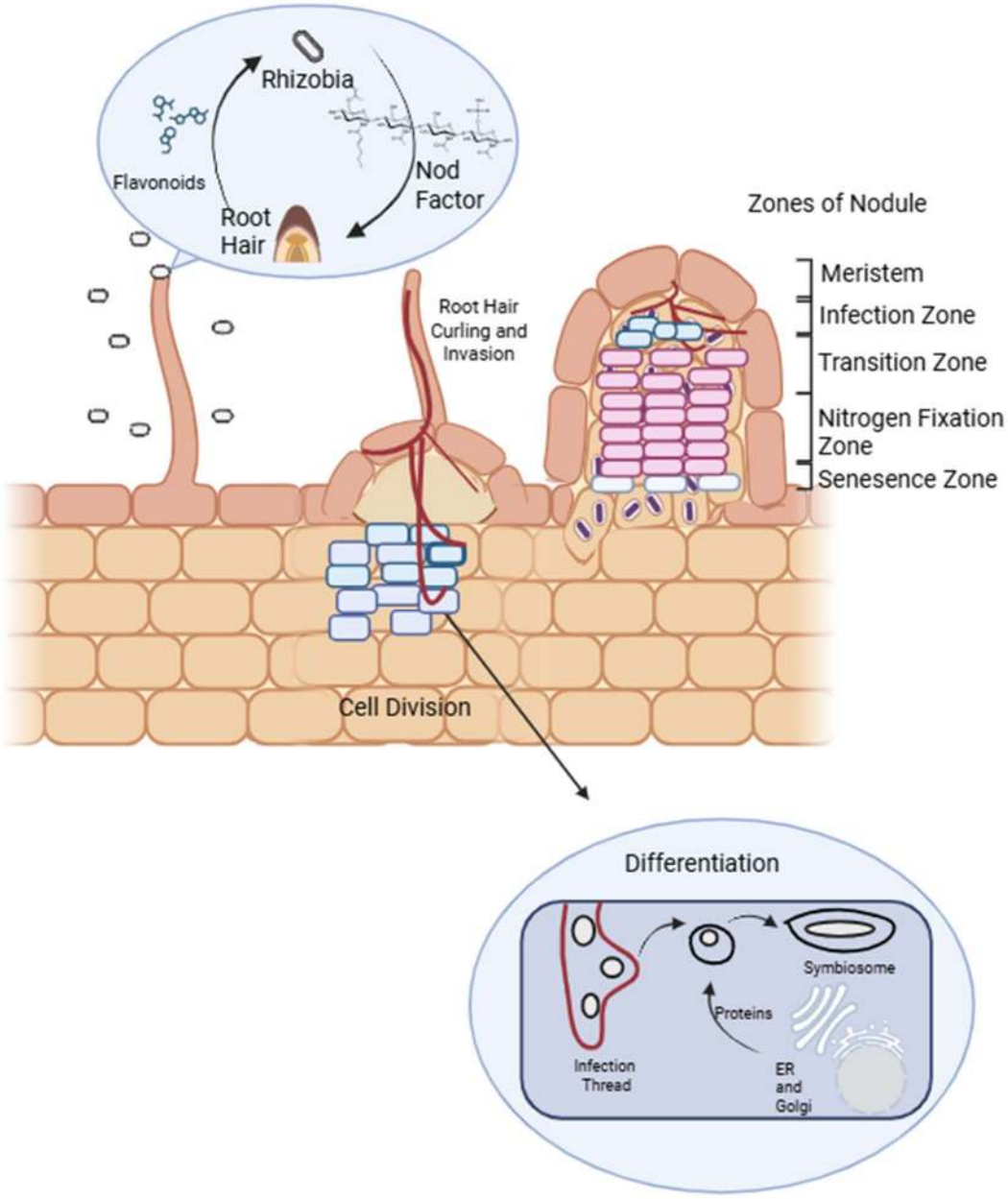



Figure 1.2: Root signalling throughout nodule formation

Nod factors also effect calcium flux in the host and it will lead to curling of root hair and the formation of infection pockets. Rhizobia can invade the plant through two main pathways: infection thread formation or crack entry between epidermal cells (Mendoza-Suárez et al., 2021; Mus et al., 2016). As they go through the infection thread, rhizobia reach the nodule primordium, where they undergo cell division and become trapped, preparing for nitrogen fixation. In this process, reactive oxygen species (ROS) release is

inhibited, allowing the bacteria to successfully colonize and develop in the nodule (Chang et al., 2009).

Inside plant cells, rhizobia are released via infection threads and exposed to NCR peptides, which are secreted by the plant and processed by the ER and Golgi. These peptides induce terminal differentiation of rhizobia into bacteroids, which are housed in symbiosomes, specialized compartments that facilitate nitrogen fixation using the nitrogenase enzyme (Figure 1.2, Figure 1.3) (Oztas, 2017). Nitrogenase is sensitive to oxygen and it collaborates with the oxygen level regulator protein in nitrogen fixation zone, leghemoglobin (Mus et al., 2016). Essential nutrients like phosphate, iron, copper, cobalt and sulphur play critical roles in effective nitrogen fixation, when nitrogen turned into ammonia, it becomes available for plant use and it will enhance the plant's growth (Oztas, 2017; Roy et al., 2020).

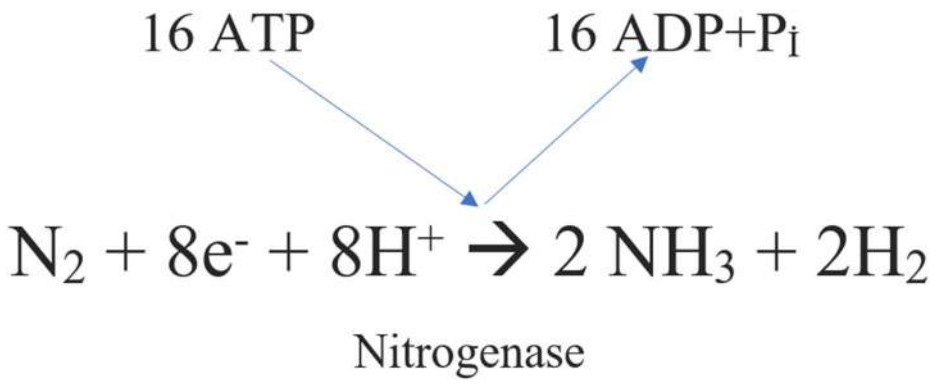



Figure 1.3: Enzymatic reaction of nitrogenase of microbes

The fate of the bacteria in the senescence zone depends on various biological and environmental factors like drought conditions, and soil nitrate levels. Reactive oxygen species (ROS) are significant indicators of senescence, as they can result from the release of iron from leghemoglobin, leading to a loss of oxygen capture capability (Roy et al., 2020). Ultimately, the successful completion of nitrogen fixation and the subsequent transition of bacteria to the senescence zone are influenced by these complex interactions and environmental conditions.

# CHAPTER 2:
# EFFECTS OF NODULE CYSTEINE RICH (NCR) PEPTIDES ON TRANSCRIPTION

## *2.1 Introduction*

### *2.1.1 An Overview to NCR Peptides*

The terminal differentiation of Rhizobia is strictly controlled by a family of plant-specific peptides known as nodule cysteine-rich (NCR) peptides. The peptides are located into the symbiosome, where they lead bacterial physiology (Kereszt et al., 2018; Kondorosi et al., 2013). NCRs are unique to a certain set of legumes of the Inverted Repeat-Lacking Clade (IRLC) (Mergaert et al., 2006) and show structural features similar to antimicrobial peptides (AMPs), including a conserved cysteine motif and a short peptide length (30–50 amino acids) (Downie & Kondorosi, 2021) (Figure 3). More than 600 NCR genes are found in the genome of *M. truncatula*, representing an important fraction (~5%) of the nodule transcriptome, with most being nodule-specific and induced during Rhizobium infection (Mergaert et al., 2006). Studies have shown that these peptides can induce membrane permeabilization, block bacterial division, and control the Rhizobium-to-bacteroid transition (Van de Velde et al., 2010).

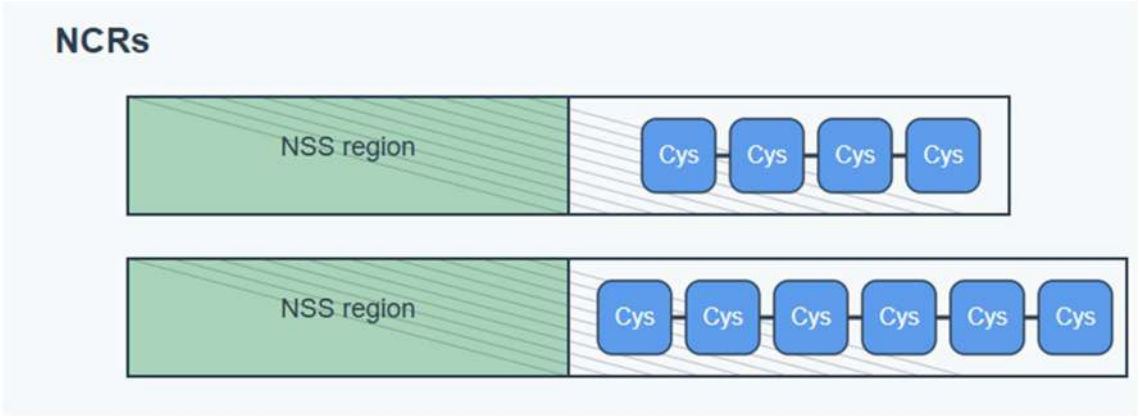



Figure 2.1: NCR peptide structure

The functional relevance of NCRs is highly important however the reason behind the massive expansion and sequence diversity of this gene family remains mysterious. While a few NCRs, such as NCR247 (Farkas et al., 2014), NCR211 (Kim et al., 2015), and NCR335 (Tiricz et al., 2013), have been functionally characterized—revealing roles in antimicrobial activity, bacteroid stability, and translation modulation—most NCRs remain poorly understood. Models for NCR working mechanism, generally describes them as effectors that induce bacterial differentiation, but literature evidence points at a more complex regulatory role.

A comparative transcriptomic analysis between *M. truncatula* and *L. japonicus* (which does not induce terminal bacteroid differentiation) further reinforces the centrality of NCRs (Kondorosi et al., 2013). Researchers have shown that plant genes encoding NCR and glycine-rich (GRP) peptides were strongly expressed in *M. truncatula* nodules, but were absent in *L. japonicus*, suggesting that NCRs are key drivers of terminal differentiation (Alunni et al., 2007). The absence of NCR homologs in legumes that do not undergo bacteroid terminal differentiation (e.g., *L. japonicus*) highlights the evolutionary development of NCRs in indeterminate nodule-forming legumes such as *Medicago*, *Pisum*, and *Vicia* (Kondorosi et al., 2013).

### *2.1.2 Secretion Pathway of NCR peptides*

NCR peptides are essential for the terminal differentiation of rhizobia into nitrogen-fixing bacteroids within indeterminate nodules (Kouchi et al., 2010). These peptides, encoded by host genes, are structurally similar to defensin-type antimicrobial proteins and contain a signal peptide that directs them into the symbiosome secretory pathway. The secretion of NCR peptides is tightly regulated by the host plant through a specialized nodule-specific protein secretory system, ensuring their proper delivery to the symbiosomes, the organelle-like structures housing bacteroids (Van de Velde et al., 2010).

One of the critical component of this pathway is the signal peptidase complex (SPC). SPC processes NCR peptides before their way into bacteroids. DNF1 protein is a key SPC component at it is essential for signalling (Kouchi et al., 2010; Wang et al., 2010). dnf1 tnt1 insertion mutant fail to target NCR peptides and resulting in an inability to trigger

bacteroid differentiation or develop nitrogen-fixing symbiosomes (Van de Velde et al., 2010). This defect highlights the essential role of the SPC in ensuring the successful delivery and activation of NCR peptides.

Once delivered to the symbiosomes, NCR peptides interact with bacteroids and trigger terminal differentiation. In indeterminate nodules, like the ones in *M. truncatula*, this differentiation is seperated by morphological changes in bacteroids (Figure 2.2). These changes making sure that bacteroids become functional by nitrogen-fixing entities but lose their viability as a necessary adaptation for the symbiotic relationship (Mergaert et al., 2003; Wang et al., 2010). The role of NCR peptides in directing bacteroid differentiation underscores the complex regulatory mechanisms evolved by legumes to optimize nitrogen fixation.

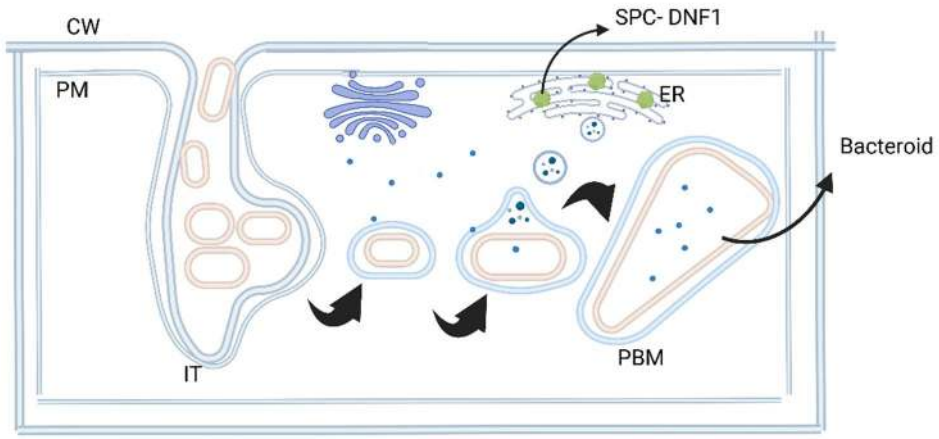



Figure 2.2: A model for a nodule-specific secretory pathway of NCR peptides. A signal peptidase complex (SPC), DNF1 is essential for NCR peptides to the symbiosome and those are incorporated into bacteroids, leading to terminal differentiation of bacteroids. CW, cell wall; PM, plasma membrane; ER, endoplasmic reticulum; IT, infection thread, PBM: Peribacteroid Membrane respectively.

Determinate nodules formed by *L. japonicus*, do not possess NCR genes (Kondorosi et al., 2013). This suggests that alternative mechanisms are responsible for bacteroid differentiation in these legumes. The absence of NCR peptides in determinate nodules highlights evolutionary divergence in the strategies legumes use to regulate their symbiotic interactions with rhizobia (Kondorosi et al., 2013; Shimoda et al., 2019).

In summary, the NCR peptide secretion and activation pathway represents a unique mechanism by which legumes control rhizobial differentiation and nitrogen fixation. This pathway regulates the crosstalk between host plants and their symbiotic partners and it is showcasing the evolutionary adaptations that for this symbiotic relationship.

### *2.1.3 Cationic NCR peptides and properties of DNA-Binding NCR (DB-NCR) Peptides*

In early symbiotic nodule cells, where endosymbionts actively multiply, only a limited number of non-cationic NCR genes are expressed. As the bacterial population reaches a critical density, differentiation begins, marked by cell division arrest and cell enlargement (Oztas, 2017) These changes, including alterations in the bacterial cell envelope and increased membrane permeability will be caused by cationic NCRs and facilitate metabolite exchange between the plant and bacterium (Kondorosi et al., 2013).

Cationic NCRs can bind to bacterial membranes, leading to their entry into the cytosol either through pore formation or by causing membrane damage and subsequent cell lysis. When inside the rhizobia, NCRs induce widespread transcriptional changes because of their DNA-binding properties and interact with various bacterial proteins, collectively disrupting essential cellular functions (Lima et al., 2020). To counteract the lethal effects of NCRs, symbiotic bacteria utilize protective mechanisms involving specific proteins such as BacA, HrrP, SMc03872, and polysaccharides like EPS and LPS, which protects the bacterial membranes and carries cellular integrity against NCR-mediated damage (Lima et al., 2020).

A key function of NCR peptides is to finally inhibit bacterial cell division. Studies have shown that cationic NCR peptides, such as NCR035 and NCR247, increase membrane permeability, induce cell elongation, amplify DNA, and kill the threats (Van de Velde et al., 2010). For NCR 247, its activation in nodules coincides with the onset of bacteroid differentiation, leading to cell division arrest and elongation. With higher concentrations NCR247 broke down the integrity of bacterial membranes and led to lethality (Farkas et al., 2014; Mikuláss et al., 2016). At sublethal concentrations, NCR247 can enter the cytosol without forming pores, triggering global transcriptional changes that affect nearly quarter of the bacterial genome (Penterman et al., 2014; Tiricz et al., 2013). Moreover, NCR247 was found to associate with bacterial RNA polymerase, implying a

possible role in blocking transcription through bulky DNA adduct formation and effecting the decision between differentiation and cell death (Farkas et al., 2014). Other experimental example for DB-NCRS take place at transcriptional level will be NCR343, which is needed for the viability of the rhizobia (Gao et al., 2023). Such adducts could prevent transcriptional elongation, leading to gene silencing—a mechanism reminiscent of small-molecule DNA inhibitors.

This transcription-blocking hypothesis may also explain the complicated roles of individual NCRs. For example, the diverse effects of NCR247—ranging from translation inhibition to altered morphology—could stem from the repression of different gene subsets in Rhizobia (Farkas et al., 2014; Penterman et al., 2014). The divergence in amino acid sequences through NCRs can bring an evolutionary strategy by the host plant to control over multiple ways of bacterial metabolism and development. However, the symbiosis may not be entirely on a single direction. It is clear that Rhizobia may counteract host-driven processes through their own regulatory arsenal, including the secretion of effector proteins (Fauvart & Michiels, 2008) (Figure 2.3). These bacterial effectors may modulate plant signaling pathways or gene expression in an attempt to evade host control or optimize their intracellular environment.

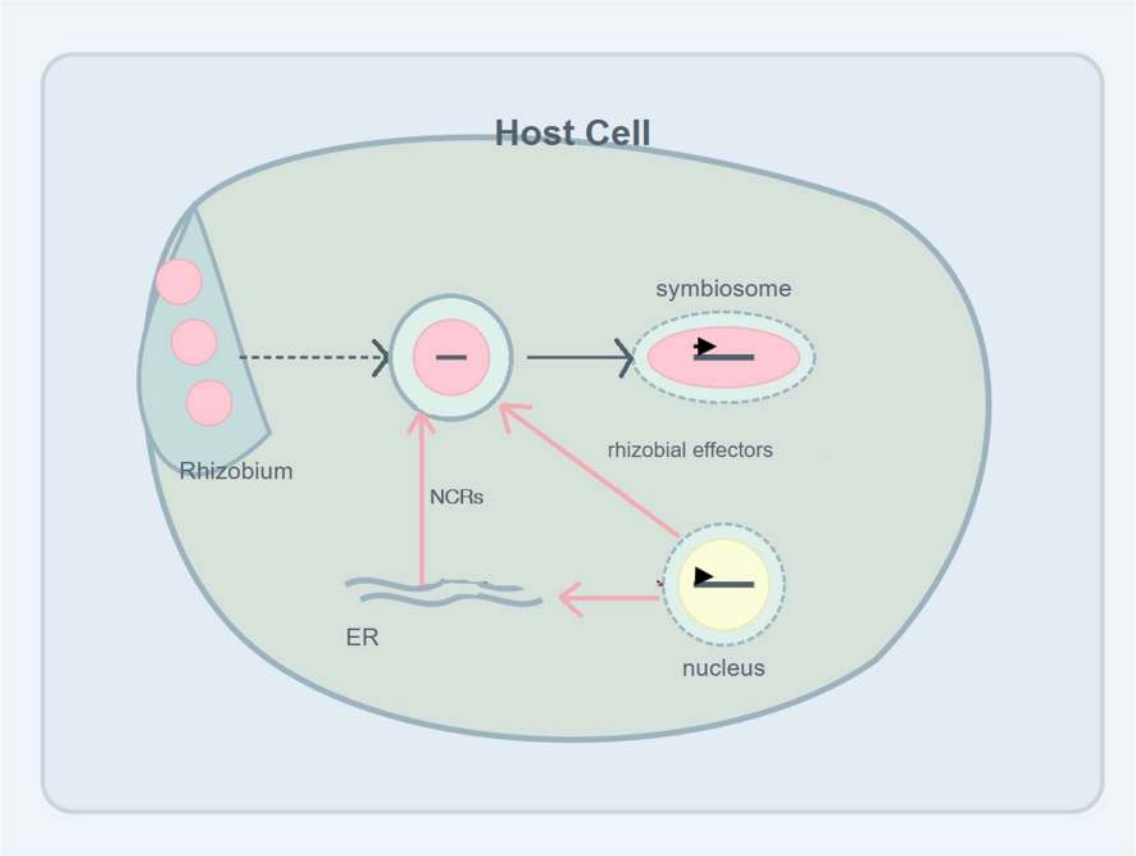



Figure 2.3: Role of NCRs in transcriptional regulation. Each rhizobium in infection thread is internalized by plant host cell inside an organelle, known as symbiosome. Rhizobium inside symbiosome differentiate into nitrogen-fixing form and provides ammonia to host plant in exchange of carbon source. Host cell and intracellular rhizobia manipulate each other's transcriptome by NCR peptides and rhizobial effectors, respectively.

This reciprocal transcriptional reprogramming suggests a molecular tug-of-war inside the nodule, where host and microbe continually adjust each other's transcriptomes to maintain functional symbiosis (Nelson & Sadowsky, 2015).

## *2.2 Research Question*

Based on this evidence, can host-derived NCR peptides function as DNA-binding transcriptional repressors on the rhizobial genome, inhibiting transcription of key bacterial genes and what is DNA-Binding NCRs' role and NF mechanism? A novel model has been proposed for which the host plant reprograms the intracellular symbiont by deploying DNA-binding NCR (DBNCR) peptides that act as transcriptional inhibitors. These DBNCRs may form peptide-DNA adducts at specific loci on the rhizobial genome, preventing gene expression and forcing the bacterium into a non-dividing, nitrogen-fixing state. In parallel, the differentiated bacteroids may release effectors to counteract or fine-tune host cellular responses, potentially manipulating host transcriptional networks to ensure their survival. I hypothesized that host-derived NCR peptides function as DNA-binding transcriptional repressors on the rhizobial genome, inhibiting transcription of key bacterial genes, and that intracellular rhizobia, in turn, secrete effectors that modulate host gene expression to balance this regulatory influence.

Together, these investigations revealed the molecular basis of transcriptional cross-talk between plant and microbe, providing novel insight into how hosts and symbionts co-regulate each other to sustain a productive symbiotic relationship.

## *2.3 Material and Methods*

### *2.3.1 Plant material and inoculation*

*M. truncatula* wild-type ecotype R108 seeds were scarified using sandpaper for removal of seed coat. The seeds were then imbibed in sterile water and cold-treated overnight at 4 °C for breaking down the dormancy. Following stratification, seeds were placed in the dark on wet Whatman filter paper in Petri dishes and allowed to germinate overnight at room temperature. Germinated seedlings were transferred to small pots containing a 1:1 (v/v) mixture of vermiculite and perlite. Plants were grown under controlled environmental conditions (24 °C, 16/8 h light/dark photoperiod).

Seven days after planting, seedlings were inoculated with *Sinorhizobium meliloti* strain FSM-MA. The bacterial culture was grown overnight at 28 °C in 10 mL of Yeast Extract Beef (YEB; beef extract 5 g, yeast extract 1 g, peptone 5 g, 14.6 mM sucrose , 1.22 mM $MgSO_4 \cdot 7H_2O$ , agar 20 g, and distilled water to 1000 mL, 121 C- 15 min) medium with shaking. On the day of inoculation, the culture was optimized to an optical density ($OD_{600}$) of 0.01 using Standard Nodulation Medium (SNM) as described at the related reference (Crook et al., 2012). Each plant received an appropriate volume of the diluted rhizobial suspension at the root zone. Plants were harvested and analyzed 21 days post-inoculation (dpi).

### *2.3.2 Determining the DBNCR pool*

The pool of DBNCRs was initially identified using a mass spectrometry-based approach. Symbiosomes were isolated from 21-day-old nodules through Percoll density gradient centrifugation prepared with MES-Tris buffer (25 mM, pH 7.2) supplemented with 600 mOsmol/L with mannitol. Layers included %30, %60 an %80 Percoll bacteroids isolated between %30 and %60. Cleanup of the bacteroids from done in 1:5 ratio with MES-Tris buffer supplemented with 600 mOsmol/L mannitol.

To isolate DNA-binding proteins for the half of the samples crosslinking was performed by incubating the symbiosome samples in 1% formaldehyde inside of crosslinking buffer (0.4 M Sucrose, 50 mM HEPES-KOH, pH 7.5,5 mM DTT, PIC) at room temperature quenched with 125 mM glycine for 10 minutes, centrifuged at 4000 G

for 20 minutes. Resulted pellet exposed to extraction buffer-2 (0.25 M Sucrose, 10 mM Tris-Cl, pH 8, 10 mM MgCl2, 1% Triton X-100, 5 mM DTT, PIC) and extraction buffer-3 (1.7 M Sucrose, 10 mM Tris-Cl, pH 8, 2 mM $MgCl_2$, 0.15% Triton X-100, 5 mM DTT, PIC) and centrifuged at maximum for 20 and 40 minutes respectively. Resulted pellet dissolved at Na-P (50mM, pH 7.5, 150 mM NaCl, and 0.1% Triton X-100).

The resulting protein fraction—enriched for rhizobial genome-associated proteins—was subjected to %10 SDS-PAGE for separation. Protein bands were then analyzed using mass spectrometry-based proteomics to identify and characterize the DBNCRs and other DNA-binding proteins potentially involved in transcriptional regulation within bacteroids (Figure 2.4).

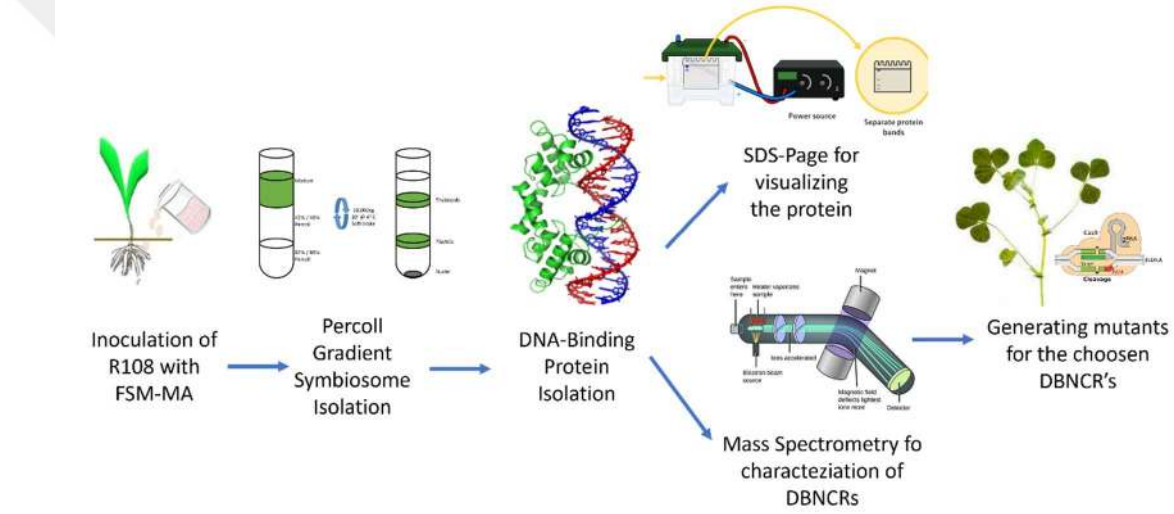



Figure 2.4: Workflow for Isolation and Detection of DB-NCRs

### *2.3.3 Mass Spectrometry*

Protein samples (optimized for 15 µg) were prepared for mass spectrometry analysis following an in-solution trypsin digestion protocol as followed up in reference (Medzihradszky, 2005). To minimize contamination, all reagents were prepared fresh or stored as aliquots at -20°C, and ultra-pure MilliQ water was used for all solutions.

LC-MS/MS analysis was performed using a high-resolution Q-Exactive Orbitrap Mass Spectroscopy Analytical System (Thermo Scientific) coupled to an EASY-nLC 1200 system for peptide separation. Mass spectra were acquired using a data-dependent acquisition (DDA) method. Raw data that are observed from run analyzed by Max Quant

(Cox & Mann, 2008) with *M. truncatula* reference proteome (UP000002051) for characterizing the proteins, and downstream applications ran by Perseus (Tyanova et al., 2016). As final result 1469 protein was a hit and from all ; 363 as plant proteins and 1106 as bacteria proteins. 6 NCR peptides has been annotated, 3 most abundant has been choosen for this research.

### *2.3.4 Choosing the NCR Candidates*

As first analysis plant proteins and bacterial proteins have been grouped and after that NCRs has been found and separately grouped. In the results candidates has been chosen by each of their abundance level especially in non-crosslinked experimental groups based on hits th at we obtained from MS. From our data, 3 candidates were available as the highest hits in all of the repeats and especially non- crosslinked experiment groups which are NCR418, NCRAL7, NCRM81. These samples ran from RoseTTAFold2NA (Baek et al., 2024) to prove the DNA-binding property in silico which implicates our data in on-going analyses (See Figure 2.9).

### *2.3.5 Generation of the Mutants*

To check whether candidate DBNCRs are involved in root-nodule symbiosis, *Medicago truncatula* Tnt1 insertion mutants has been used, which were kindly provided by Oklahoma State University. These mutants are generated in the background of R108 ecotype with insertion of Tnt1 transposon randomly. Insertion sites has been selected to be located to exon 1 to eliminate the possibilities for splicing. All of them have been named as *al7*, *m81* and *418*. To confirm homozygous mutants of DBNCRs, PCR-based genotyping performed using primers mentioned (Table B.1) Homozygous mutants are selected for intended use and F2 generation of the seeds used for next experiments.

### *2.3.6 Phenotyping of the mutants*

Mutants inoculated by *Sinorhizobium mellilloti*, FSM-MA strain. Initially, bacterial culture grown overnight inside YEB Medium. 7 days old plants will be inoculated with bacteria OD 0.01 adjusted with Standard Nodulation Medium (SNM). At 14- and 21-dpi plants checked for the color and size of the nodules under stereo microscope; above ground parts phenotype.

### *2.3.7 Determining the function of the candidate DBNCRs*

For determining the localization of the proteins p*NCR418::NCR418_mCherry* and *pNCRAL7::NCRAL7_mCherry* vectors has been constructed with the pMDC43 backbone and related primers (Figure A.1, Figure A.2, Figure A.3, Table B.1). Hairy root transformation is done to R108 wildtype plants (Boisson-Dernier et al., 2001). For all of the transformed plants at 7 days *Sinorhizobium* inoculation will be done as explained above (See 2.3.1). For mCherry constructs inoculation will be done with RM1021-GFP strain of *S. melilotti*, confocal imaging was done in KUTTAM Imaging Core with Leica DMI8.

## *2.4 Results*

### *2.4.1 Percoll Gradient with crosslinking is an effective technique for isolation of the DNA-Binding Protein*

At 21th dpi of the plant, symbiosomes isolated via Percoll Gradient technique. Layers preapared as %30,%60 and %90 and symbiosomes were expected between %30 and %60 (Figure 2.5). Throughout layering of the symbiosome there has been two layers (Which might be related to developmental phases of rhizobia) observed between %30 and %60 with pellet on down.

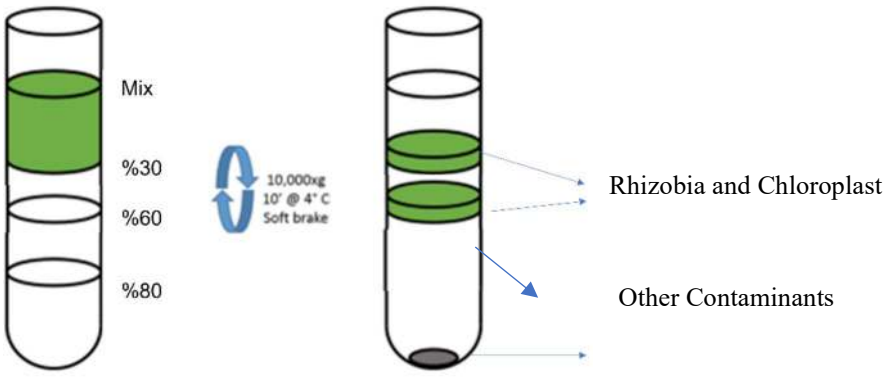



Figure 2.5: Percoll Layering of the Rhizobia

Firstly for the proof of presence of the bacteria 1 ul of the isolated symbiosome layer checked under flourescence microscope with SYTO9/PI staining. According to the

images, Percoll gradient techniques is effective for full and intact isolation of the symbiosome (Figure 2.6).

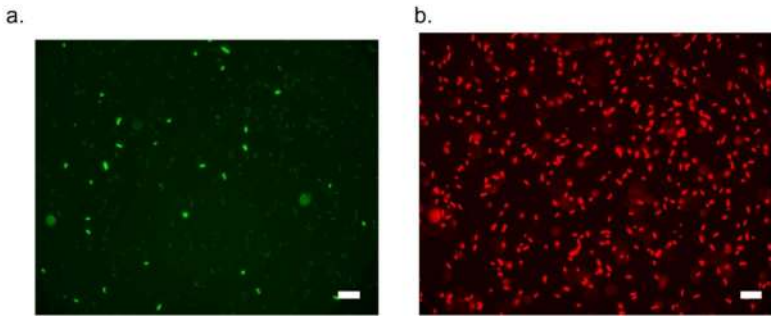



Figure 2.6: Verification of symbiosomes by fluorescence imaging. Dyes used a. Sytox Green for live bacteria, b. Propidium Iodide for dead bacteria. Scale bar was set at 30 um.

After the symbiosome isolation for verification of the isolation of the DNA Binding Protein Pool, %10 SDS Page has been run. For all of the layers has been isolated and for + and – crosslinked products protein was present and there has been differences between protein profiles according to band patterns (Figure 2.7).

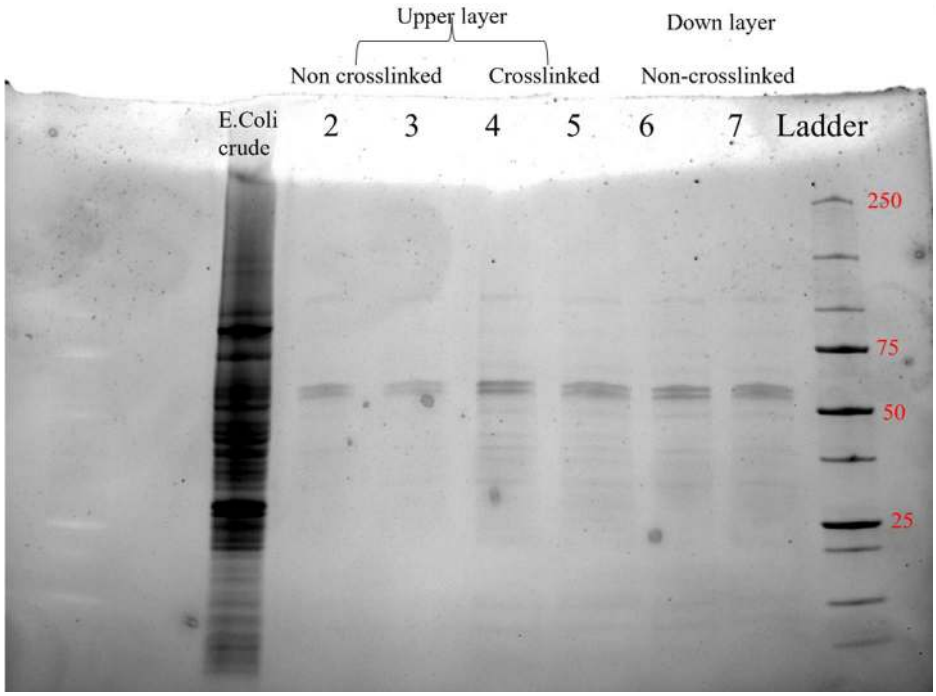



Figure 2.7: SDS-Page image of total DNA-binding proteins of symbiosomes. *M. truncatula* R108 line inoculated by *S. mellilotti* FSM-MA. Samples have been collected 21 days post inoculation (dpi).

In order to verify what are the layers and is there any contamination, PCR screening approach has been used. For contaminants for symbiosome and determination of the layers; presence of mitochondria, nuclei and chloroplast genome screened by PCR with selected housekeeping genes spesifically expressed for selected organelles and NifH spesific for nitrogen-fixing bacteria. According to the PCR results, isolated symbiosome only contains the contamination of chloroplast genome when it compared to Percoll layers and pellet (Figure 2.8).

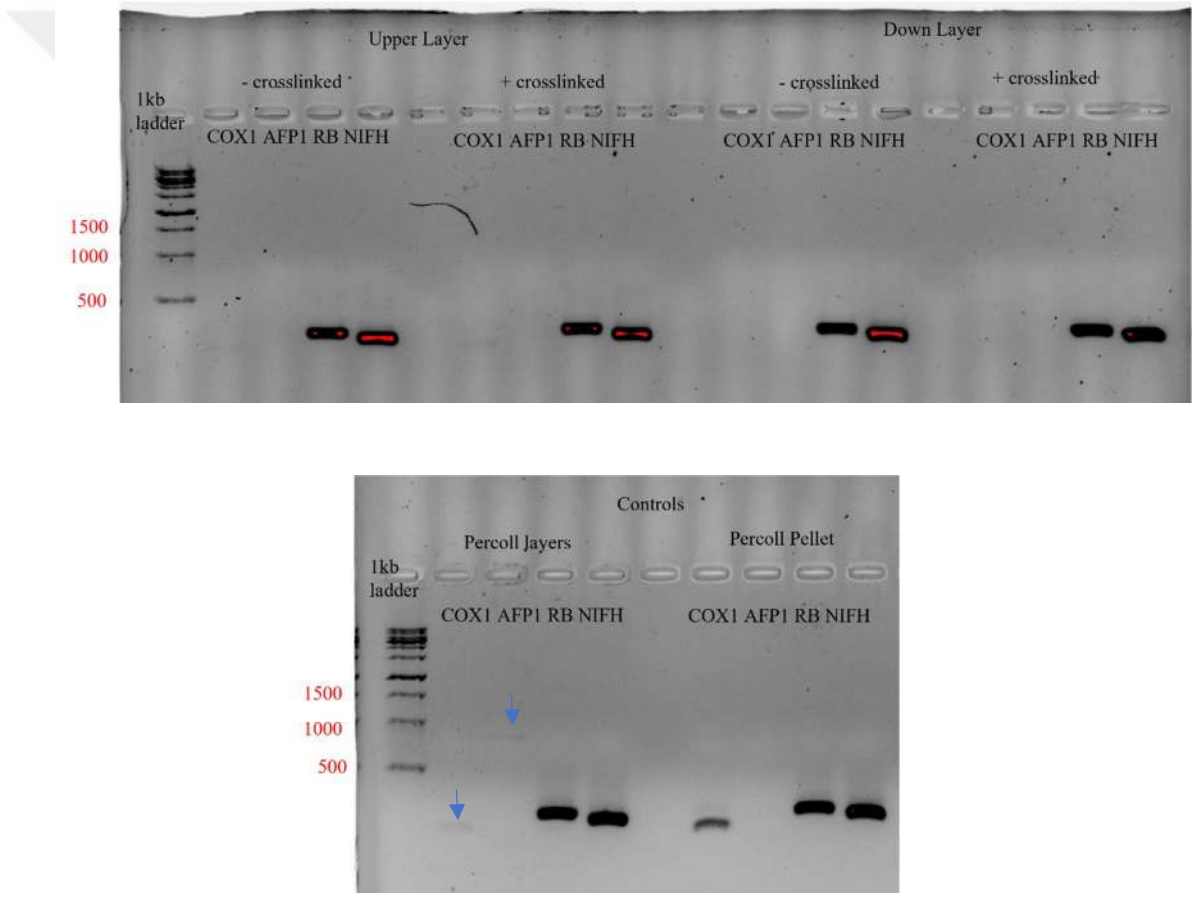



Figure 2.8 : Agarose gel image of total DNA-binding proteins of symbiosomes. *M. truncatula* R108 line inoculated by *S. mellilotti* FSM-MA. Samples have been collected 21 days post inoculation (dpi).

*2.3.2 Mass Spectrometry verifies that NCRs can bind to DNA*

DB-Protein Pool that has been collected from 21 dpi analyzed by ORBITRAP Mass Spectrometry, digested through in-solution digestion (Medzihradszky, 2005) technique

with trypsin. According to our results run through Max Quant Software (Cox & Mann, 2008) for the annotation of the pool, according to abundance in non-crosslinked experiment groups 3 candidates for DB-NCR's and DB-Proteins has been selected. In our NCR candidates has been shown direct binding to DNA but for the major protein candidates has been chosen there is no binding to DNA. (Table 2.1, Table 2.2).

Table 2.1 NCR candidates chosen from DNA- Binding Protein pool generated by Mass Spectrometry. Upstream applications are done By MaxQuant*(Cox & Mann, 2008)* software and 3 NCRs that has best coverage has been choose to be a candidate.

| Accession | Description | Sum PEP Score | Coverage [%] | # AAs | MW [kDa] | Found in Sample | Dna-Binding (RoseTTAFold2DNA) |
|---|---|---|---|---|---|---|---|
| G7JBC9 | Nodule Cysteine-Rich (NCR) secreted peptide OS=*Medicago truncatula* OX=3880 GN=MTR_3g084820 PE=2 SV=1 | 5,731 | 19 | 75 | 8,4 | High | + |
| G7JM81 | Nodule Cysteine-Rich (NCR) secreted peptide OS=*Medicago truncatula* OX=3880 GN=MTR_4g057120 PE=2 SV=1 | 5,641 | 33 | 60 | 7,1 | High | + |
| G7JAL7 | Nodule Cysteine-Rich (NCR) secreted peptide OS=*Medicago truncatula* OX=3880 GN=MTR_3g069870 PE=4 SV=1 | 2,808 | 16 | 77 | 8,9 | High | + |

Table 2.2: Major protein candidates chosen from DNA- Binding Protein pool generated by Mass Spectrometry. Upstream applications are done By MaxQuant *(Cox & Mann, 2008)* software and 3 NCRs that has best coverage has been choose to be a candidate.

| Accession | Description | Layer | Signal Sequence | Dna-Binding |
|---|---|---|---|---|
| A0A072TQB9 | Brefeldin A-inhibited guanine | All Layers | No | No |

| | | | | |
|---|---|---|---|---|
| | nucleotide-exchange protein OS=*Medicago* truncatula OX=3880 GN=25501014 PE=4 SV=1 | | | |
| A0A072UZU4 | Phosphate-responsive 1 family protein OS=*Medicago* truncatula OX=3880 GN=25489481 PE=3 SV=1 | ? | No | No |
| A0A072W1H5 | ATP synthase subunit beta OS=*Medicago* truncatula OX=3880 GN=25485440 PE=3 SV=1 | Bottom | No | No |

In our NCR candidates has been shown direct binding to DNA but for the major protein candidates has been chosen there is no direct binding to DNA.

### *2.4.3 Targeted DB-NCRs has DNA-binding and metal ion binding motifs*

Targeted DB-NCR's protein sequences checked and verified by genomic sequences form R108 annotation and UNIPROT (Apweiler, 2004) sequences (Figure 2.9).


Figure 2.9: Protein sequence of the targeted NCRs. All of the sequences are highlighted for DNA-binding and metal ion binding motifs with signal sequences. Blue refers metal binding. Yellow refers signal sequence, green refers transmembrane domain and pink refers DNA binding motifs.

Targeted NCR's has metal-ion binding domains verified by UniProt (Apweiler, 2004) and according to the computational analysis repeating CXXC motifs and zinc finger motif shows DNA binding properties. For AL7 other than the NCRs in the list that has signal sequence, a transmembrane domain has been observed for possible membrane interactions (Figure 2.10). With that, when looked at 3D structure predicted from AlphaFold (Jumper et al., 2021), all 3 has their catalytic sites closer to C terminal and they all share similar alpha fold – beta sheet structures relatively at N and C terminal (Figure 2.10).

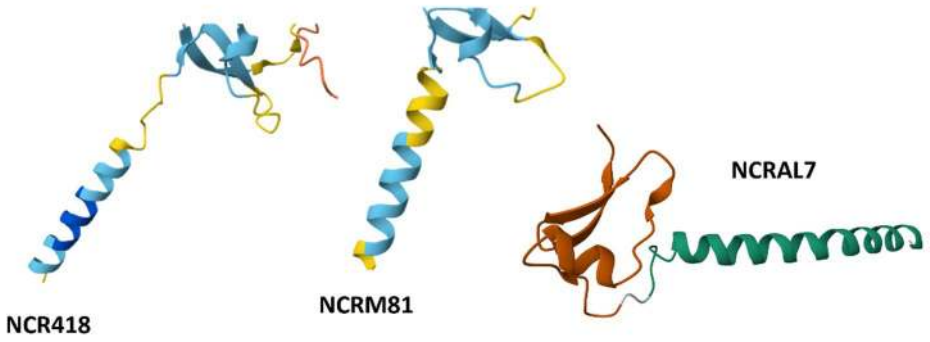



Figure 2.10: 3D structures that has been generated from AlphaFold2. Generally catalytic sites are (closely) located to the C terminal (Jumper et al., 2021).

### *2.4.4 DBNCRs have nodule-specific expressions*

When relative expression levels checked with gene ID individually for each NCR from Symbimics (Roux et al., 2014) database, all of them expressed specifically at nodules and when checked for nodule zone expression they mostly expressed around transition and nitrogen fixation zones (Figure 2.11).

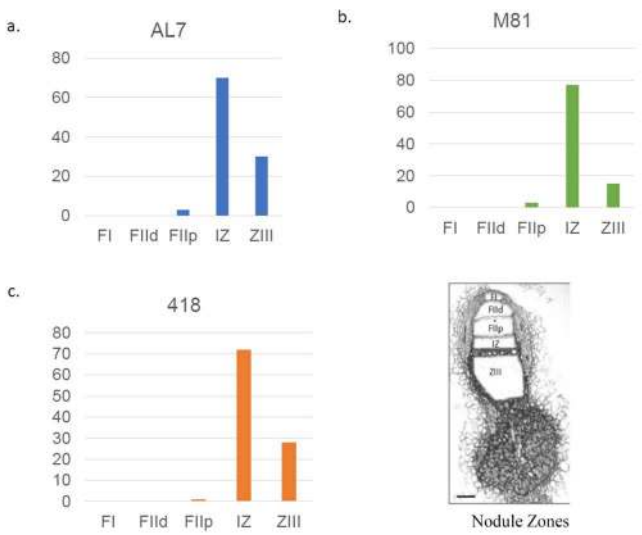



Figure 2.11: Relative Expression levels of a. NCRAL7, b. NCRM81,c. NCR418 across different nodule developmental zones (meristem, infection, differentiation, $N_2$-fixation, and maturation)using data from the Symbimics database (Roux et al., 2014). The schematic of the nodule zones is shown alongside.

### *2.4.5 Genotyping of the **Tnt1** insertion DBNCR mutants*

Genotyping of the Tnt1 insertion mutants has been done by PCR based approach. Since NCR CDS sequences are short sequences, for seeing the band difference better we did PCR from promoter sequences to Tnt1 sequence with related primers (Figure 2.12, Table B.1)

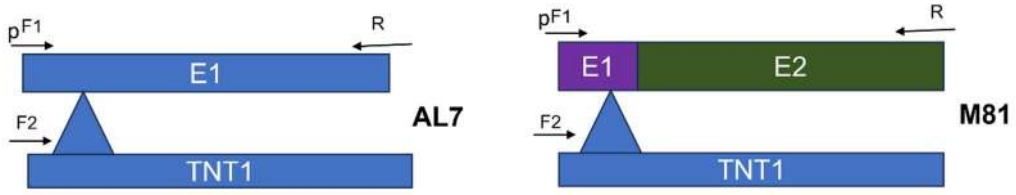



Figure 2.12: Genotyping of NCR TNT1 insertion mutants by PCR amplification. a. Scheme of PCR amplification of each plant. Amplification has been done by each from NCRs promoter (pForward) or start of TNT sequence (F2) to CDS(Reverse).

After running the results of the PCR for genotyping results, there was 2 homozygous plants for *al7* mutant, 1 homozygous plant for *m81* mutant which verified by F2R band without pF1R band (Figure 2.13).

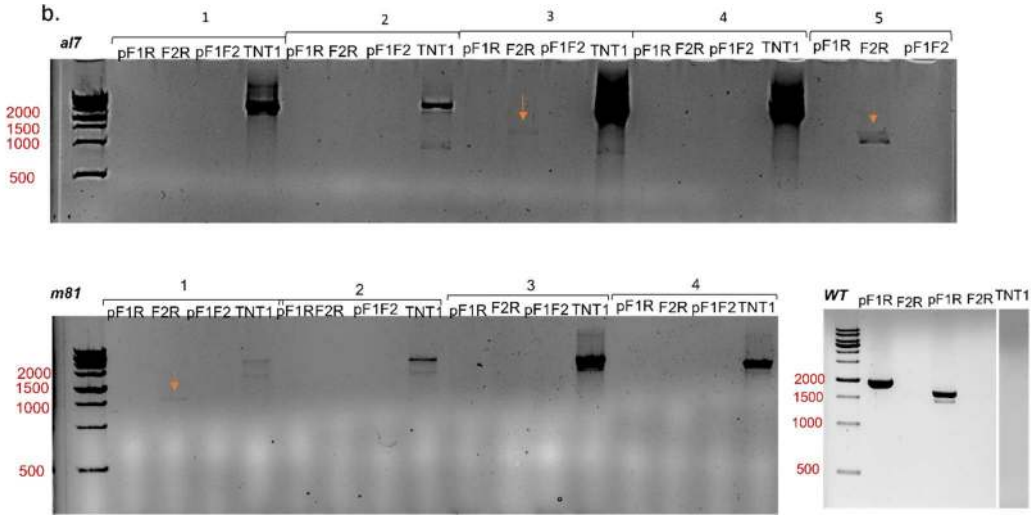



Figure 2.13: PCR products amplified with pF1-R and F2-R primer pairs by using WT and al7, m81 tnt1 insertion line as templates.

### *2.4.6 NCR Tnt1 homozygous mutant plants show fix- phenotype*

For checking the deficency of the NCRs that has been chosen we detected the phenotype nodule and whole plant for nitrogen deficency at 14 dpi. Roots did not give any branching and shorter in length comparatively to WT. At the same time plants were not able to reach the height that WT upper ground parts has been reached and leaves were relatively yellower comparatively to WT. When looked at the nodules; at both of the plants, nodules they were white and smaller; when looked at the WT it is bigger in size and has pinkish hue which shows the signs for nitrogen fixation all along. According the information given, related homozygous NCR mutants showed fix- phenotype (Figure 2.14).

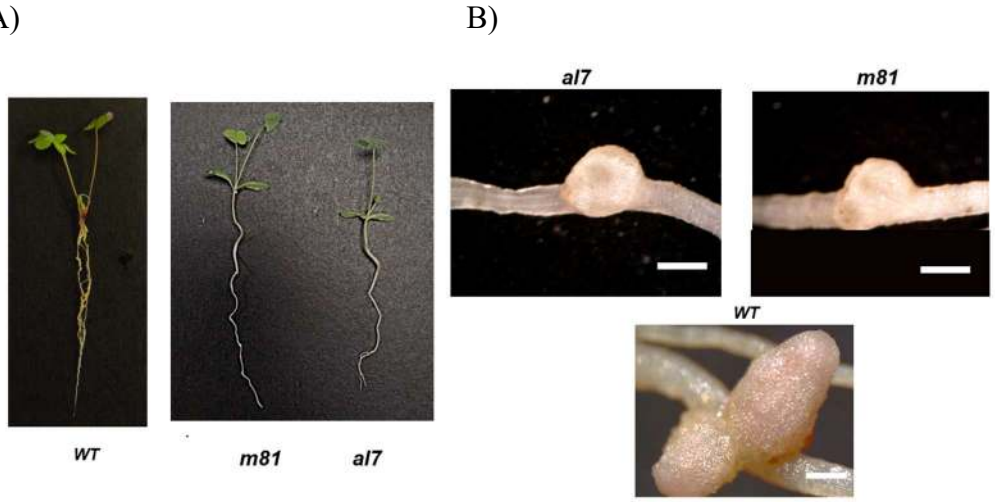



Figure 2.14: Phenotyping results of DBNCR mutants. A) Whole plant image of al7, m81 and WT. B) Whole nodule images of *al7, m81* and WT, 14 dpi; scale bar was set at 500 um. Samples observed under S9i Stereo Microscope (Leica, Germany).

### *2.4.7 NCR418 localizes to the symbiosome at infection thread*

For DBNCR 418 localization purposes p418::418-mCherry construct transformed to R108 ecotype plants. Plants inoculated with *S. melliloti* Rm1021- GFP for observing the localization of the protein over nodules via confocal microscopy. In full nodule, it is observed that the protein expression expanded through nodule. When the symbiosomes has been observed, high expression and colocalization shown at infection thread. In individual rhizobia analysis has been showed the full colocalization of protein which indicates to protein localization to the symbiosome (Figure 2.15).

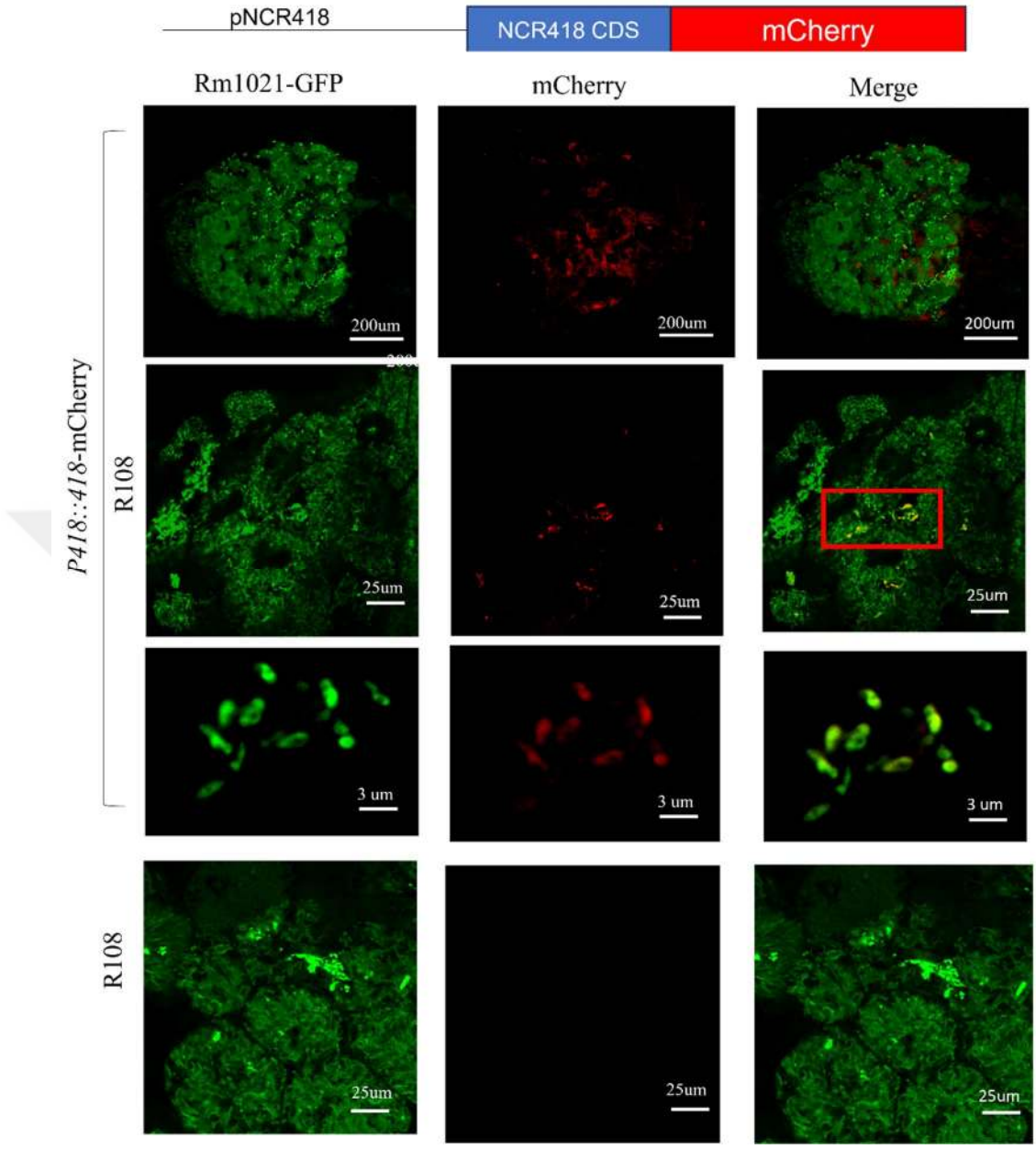



Figure 2.15: Localization of NCR418. WT R108 plants transformed with p418::418-mCh construct and inoculated with *Sinorhizobium* mellilotti Rm1021-GFP, at 21 dpi. mCherry signal was detected using excitation with a 561 nm laser and emission with a 570-630 nm band pass filter, whereas GFP signal was detected using excitation with a 488 nm laser and emission with a 490-550 nm band pass filter.

### *2.4.8 NCRAL7 mostly localizes to the ER*

For DBNCR 418 localization purposes pAL7::AL7-mCherry construct transformed to R108 ecotype plants. Plants inoculated with *S. melliloti* Rm1021- GFP for observing the localization of the protein over nodules via confocal microscopy. In full nodule, it is observed that the protein expression expanded through nodule. When the symbiosomes has been observed , we couldn't observe colocalization see throughout symbiosome. In individual rhizobia we were able to observe the of protein around the symbiosome instead of colocalization. When it compared with the localization of ER by mCherry tagged ER marker *wak4* transgenic plant, where protein located hits the same location as the transgenic plant. As a result, it has been overcome that localization of AL7 is mostly to ER (Figure 2.16).

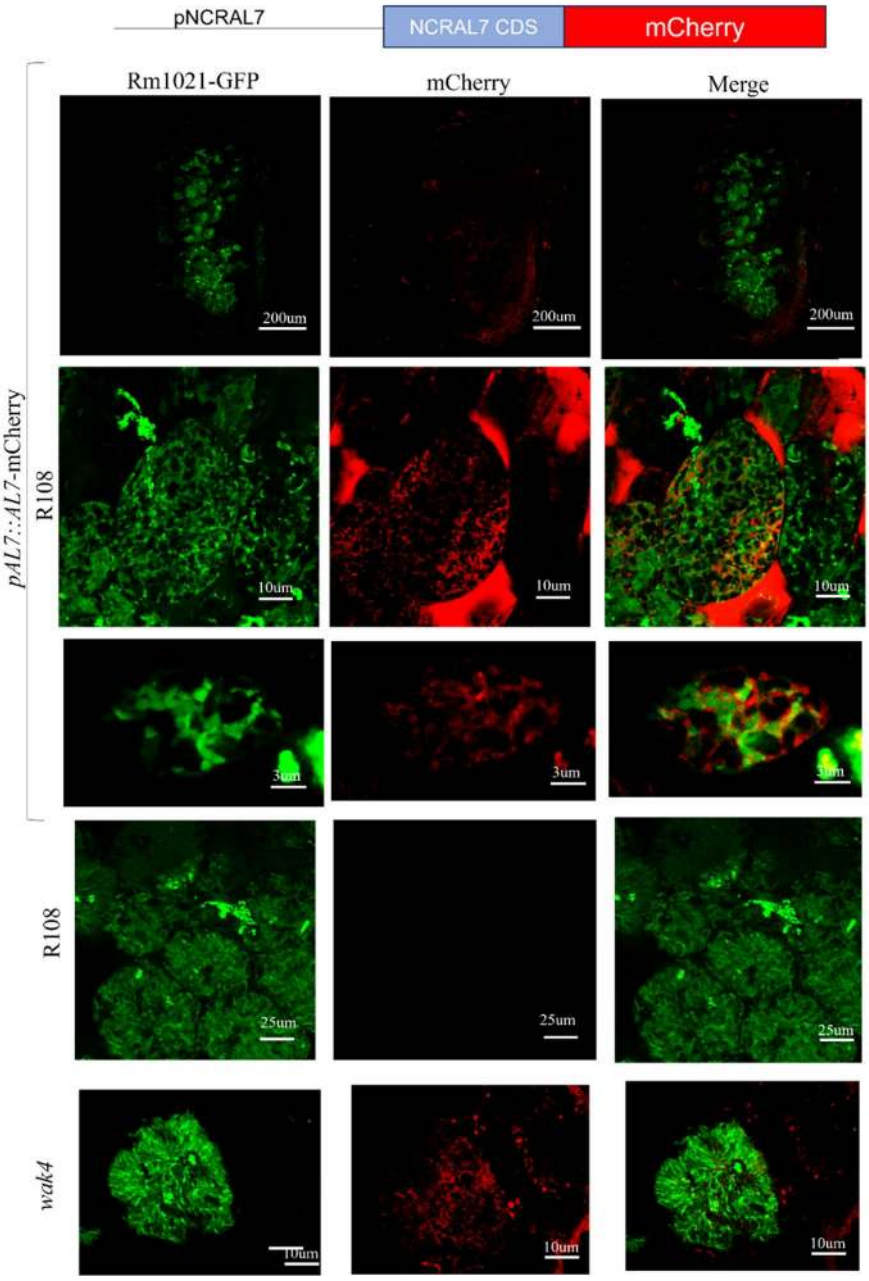



Figure 2.16: Localization of NCRAL7. WT R108 plants transformed with pAL7::AL7-mCh and inoculated with *S. mellilotti* Rm1021-GFP, at 21 dpi. mCherry signal was detected using excitation with a 561 nm laser and emission with a 570-630 nm band pass filter, whereas GFP signal was detected using excitation with a 488 nm laser and emission with a 490-550 nm band pass filter.

## *2.5 Discussion*

While the role of plant transcription factors (TFs) in symbiosis has been increasingly studied, particularly those regulating early nodulation genes, only a small fraction of TFs have been functionally characterized—about 10% in *A. thaliana* and less than 1% in legumes (Moreau et al., 2011). Transcriptional reprogramming during nodulation is known to involve hundreds of genes, as revealed by transcriptome analyses comparing wild-type and mutant nodules affected at different developmental stages (Libault, 2018). However, the reciprocal influence of host on rhizobia gene expression remains largely unexplored.

In this study we showed that how DB-NCR pool can be created and how DB-NCR's will act on nodulation. We identified 3 NCRs that might have role on bloking transcription by creating DNA adducts. So in a future model, it will be proposed that RNA-polymerase II is causing NCR peptide-caused bulky DNA adducts at rhizobial genome and transcription of these genes are inhibited that (Figure 2.17).

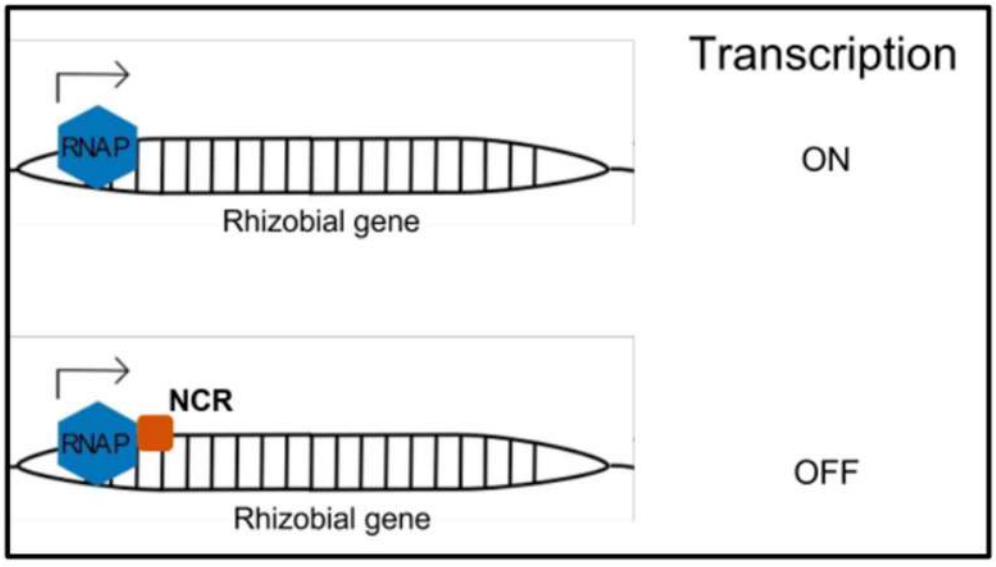



Figure 2.17: Final Model for transcriptional reprogramming with NCR peptides.

More support for this model comes from the observation of DNA-NCR peptide adducts can be excised by the bacterial UvrABC nucleotide excision repair (NER) complex. It is suggesting that these interactions are recognized as DNA damage. Bacteria protect their genome from DNA damage using repair mechanisms like nucleotide

excision repair (NER), which removes bulky adducts caused by UV light and chemicals such as benzo[a]pyrene (Li et al., 2017; Oztas et al., 2018). These adducts block transcription by stalling RNA polymerase II (Perlow et al., 2002). In bacteria, NER involves UvrA, UvrB, UvrC, and UvrD proteins, which detect and excise damaged DNA, followed by gap filling by polymerase and ligase (Kisker et al., 2013). When adducts occur on transcribed strands, the RNA polymerase stalls and MfD recruits NER proteins (Adebali et al., 2017). In *S. meliloti*, NER genes (UvrB, UvrC, UvrD, MfD) are downregulated in nodules, suggesting the host suppresses NER in bacteroids to prevent removal of NCR peptide-induced DNA damage. Aromatic residues and high cysteine content—a hallmark of NCRs—are features known to facilitate stable and bulky DNA interactions (Hwa Yun et al., 2020). Given the diversity of NCR sequences and their varied isoelectric points (Pan & Wang, 2017), it is plausible that different NCRs selectively bind distinct genomic regions, thereby reprogramming the rhizobial transcriptome in a gene-specific manner.

Hence, we identified their DNA binding properties and their role on symbiotic nitrogen fixation by phenotyping the mutants as starter research. Some major proteins have been identified together but none DNA binding properties has been found since it might be a result of crosslinking with formaldehyde, and they might have an indirect binding of DNA which will be investigated on future aspect, so in results detection of protein-protein interactions will be needed for a proper investigation of DNA interactions.

In this research experiment has been moved with the novel aspect for identifying the DNA-Binding proteins with an adapted Chip-SEQ approach. But throughout this process still loss of the protein profile is unknown while cleanup of the DNA. With this approach, it should be identified by mass spectrometry profile comparison. And model is not supportive enough since all of this experiment has been done in R108 ecotype and proteome data is not available for this ecotype. A17 ecotype work will support all of this experiment with its available database in proteome and genome.

## *2.6 Future Direction*

Same set of experiments will be run for A17 ecotype and analyzed in the same way for proteomics library availability and candidates will be verified by comparison in

other proteome. For the layering of the Percoll sample, protein profile of the differnt deveoplental stages of the rhizobia will be compared and analyzed.

Cross-linking of the protein was not suitable for this analysis because cross-linking is done with DNA, and this situation will not alter unless it will be removed with heat treatment. Heat treatment should be optimized prior to the protection of the protein from degardation. After that remaining genomic material should be cleaned up for proper read in Mass Spectrometry. For that purpose after heat treatment with Bensonase nuclease treatment will be done to samples.

For verifying the DNA-Binding properties of the NCRs' Chip-SEQ will be done by FLAG tag construct atransformed into WT plants and rescue experiments with mCh constructs should be designed with mutants has been detected to see the NCRs' effect on nitrogen fixation. For that missing construct, pM81::M81-mCh will be prepared, and 418 mutant plants will be genotyped to complete the genotyping of all 3 NCRs. After the CHIP-SEQ optimization and A17-R108 RNA-Seq results will be used for the connect the results of the sites detected in the CHIP-Seq for transcription related events.

# CHAPTER 3:
# INVESTIGATING THE ROLE OF EARLY NODULIN-16 (ENOD16) AT SYMBIOTIC NITROGEN FIXATION

## *3.1 Introduction*

### *3.1.1 Arabinogalactan Proteins*

Arabinogalactan proteins (AGPs) are a group of hydroxyproline-rich glycoproteins (HGP) that play important roles in plant development, and stress responses(Schultz et al., 1998). With extensins, AGPs are components of the plant cell wall, contributing to its structure and function (Brady et al., 1998). These glycoproteins are categorized into classical and non-classical types based on their backbone sequences (Figure 3.1). The core protein is composed of dipeptide repeats such as Ala-Pro and Pro-Al etc. , which are extensively O-glycosylated at hydroxyproline (Hyp) residues (Ellis et al., 2010). The complexes attached to these residues include β-1,3-galactan backbones with β-1,6-galactose side chains, further decorated with α-arabinose, β-(methyl)glucuronic acid and more (Showalter & Basu, 2016).

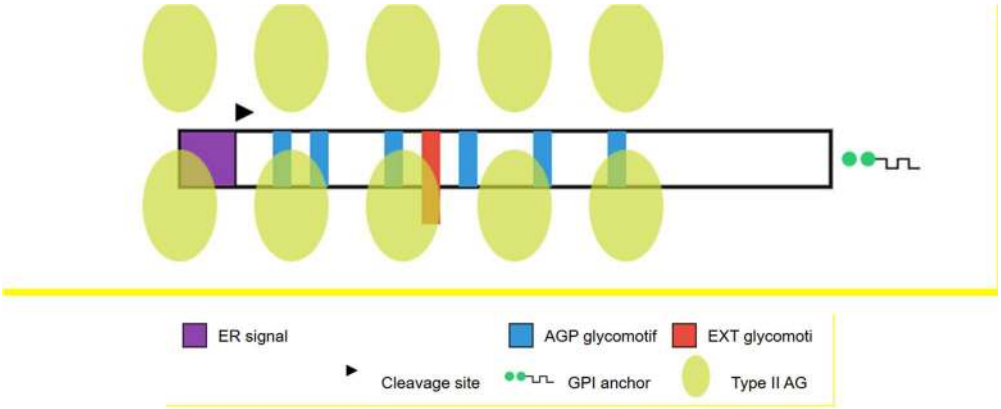



Figure 3.1: Classical AGP Motif

There are main differences between extensins and AGPs featuring; adjacent Hyp residues at extensins and while AGPs can do croslinking with arabinoxylan, but when it comes to extensins it will do crosslinking intra- and intermolecular to create extensin paths within

the cell wall. These structural differences highlight the distinct roles of AGPs and extensins in maintaining cell wall integrity and function (Brady et al., 1998). Although the biochemical functions of AGPs is unknown, research has revealed their involvement in a range of biological processes. For instance, AGPs regulate cell viability and stimulate cell division, as shown by experiments on plant cell suspension cultures (Ellis et al., 2010). AGPs also facilitate specialized cell differentiation. Xylogen triggers the intercellular communication necessary for the development of tracheal elements, especially making xylem cells responsible for water and mineral transport. (Motose et al., 2004). In *Arabidopsis*, related AGPs from the root is crucially influences root growth and seed germination by effecting abscisic acid recognition (Van Hengel et al., 2002). In addition to their functions in development, AGPs have a critical role in stress responses. Induction under salt stress accentuates that they also play a significant role in enhancing salt tolerance. In SOS5 absence throughout salt stress, *Arabidopsis* cells show irregular growth, weaker cell walls and weaker pectin layers; shows its importance in cell wall integrity under stress (Shi et al., 2003).

Studying AGPs brings challenges due to their high carbohydrate content, which complicates purification and antibody recognition. Monoclonal antibodies are targeting AGP carbohydrate content, such as MAC204. 4B4, have also provided looks into their function, localization and concerns about epitope specificity (Seifert & Roberts, 2007).

In conclusion, AGPs are versatile glycoproteins with diverse roles in plant biology, ranging from cell wall interactions and signaling to stress tolerance and reproductive development. Despite the difficulties encountered in studying their biochemical functions, ongoing research is on its way to enlighten their importance in plant growth and adaptation.

### *3.1.2 ENOD16*

The ENOD16 protein is a member of the Early Nodulin (ENOD) family, which plays the role at early nodulation in legumes. They are structurally hydroxyproline-rich glycoproteins (HRGP) that are secreted into the extracellular matrix. It takes role at survival and differentiation of bacteria and for that it creates structural modifications at nitrogen-fixing bacteria (Brewin, 2004; Cassab, 1998). From its family, ENOD16 stands out due to its unique structural features and functions in plant development and stress responses.

ENOD16 consists of three domains which are signal peptide, plastocyanin-like domain, and GPI anchor domain (Figure 3.2). The signal peptide ensures the secretion of ENOD16 into the extracellular matrix, allowing it to localize correctly within the plant cell wall (Greene et al., 1998).The plastocyanin-like domain resembles copper-binding proteins involved in electron transport. The GPI anchor domain attaches ENOD16 to the plasma membrane or extracellular matrix, enabling interactions with other cell wall components and signaling molecules.(Greene et al., 1998) Interestingly, proteins called ENOD-like has been identified in nodulating lacking plant species, such as *Arabidopsis* and rice, suggesting different roles beyond nodulation(Mashiguchi et al., 2009). It provides completely different structure including lacking of copper ion binding properties. In *Arabidopsis* they have different roles including pollen tube recognition and guidance during fertilization (Hou et al., 2016).


ENOD16

MASSSPILLMIIFSMWLLISHSESTDYLIGDSHNSWKVPLPSRRAFARWASAHEFTVGDTILF
EYDNETESVHEVNEHDYIMCHTNGEHVEHHDGNTKVVLDKIGVYHFISGTKRHCKMGLKL
AVVVQNKHDLVLPPLITMPMPPSPSPSPNSSGNKGGAAGLGFIMWLGVSLVMMMFLI*

Figure 3.2 Structure of ENOD16. Yellow color indicates AGP motif.

Additionally, ENOD16 contains a single arabinogalactan protein (AGP) motif, a feature that is characteristic of hydroxyproline-rich glycoproteins.(Greene et al., 1998) The motif can get glycosylated and it can contribute to the structural complexity of the protein and its functional diversity(Strasser et al., 2021). Glycosylation of ENOD16 is hypothesized to play a role in multiple events. Their interactions with other cell wall components like arabinoxylan and signaling in plant development can serve as examples for this (Figure 3.2). (Hou et al., 2016)

ENOD16 is also homologous with pea/vetch ENOD5 and soybean ENOD55/N315 proteins, both localized to infected nodule cells. As these proteins have been suggested to play roles in cell wall remodeling during bacterial infection, the implication is that ENOD16 performs similar roles (Greene et al., 1998). Experimental verification of these roles is limited, however, and additional work is needed.

ENOD16 is very close in sequence to other nodulins, particularly ENOD20, but differs in structure as in AGP motifs. Unlike its counterpart ENOD20, which contains multiple AGP motifs and additional arabinosylation sites, ENOD16 has only a single AGP motif (Greene et al., 1998). This structural simplicity may reflect functional specialization, with ENOD16 focusing on specific roles in nodulation and stress responses, while ENOD20 may have broader applications in cell wall dynamics. Although the exact biochemical functions of ENOD16 remain unclear, studies suggest it plays diverse roles in plant development, stress responses, and symbiotic interactions. During nodule formation, ENOD16 appears to act as a structural protein, facilitating bacterial invasion and infection thread development. It is believed to enhance cell wall flexibility, allowing bacterial colonies to establish themselves during symbiosis.

## *3.2 Research Question*

According to this information, will ENOD16 take role at bacterial survival and differentiation (by cell wall differentiation) and what is the role of ENOD16 at nitrogen fixation? It has been hypothesized that ENOD16 is an AGP that takes role as early-stage player at nodule development and nitrogen fixation. Expression pattern and localization of ENOD16 under the lacking of SNARE proteins and their interactions are unknown. Here it has been investigated this arabinogalactan protein's role on nitrogen fixation and bacteria survival.

## *3.3 Material and Method*

### *3.3.1 The legume model plant*

*M. truncatula* ecotypes R108 were used for this study. Plants were inoculated with *S. medicae* strain ABS7 in Standard Nodulation Medium (SNM) as described in the related reference (Crook et al., 2012) to obtain 4 weeks old nodules. Plants grown under 16/8 hour day/night period at 22 C in controlled environment.

### *3.3.2 Choosing the Mutants*

To check whether candidate ENOD16 is involved in symbiosis, *Medicago truncatula* Tnt1 insertion mutants has been used, which were kindly provided by Oklahoma State University. These mutants are generated in the background of R108 ecotype with insertion of Tnt1 transposon randomly. Insertion site has been selected to be located differently to check whether Tnt1 insertion site can effect the phenotype. Two mutnats has been seleceted and have been named as *enod16-1* and *enod16-2*. Homozygous mutants are selected for intended use and inoculated with *Sinorhizobium mel.* ABS7 strain. Plants has been phenotyped at 28 dpi for nodule and full plant.

### *3.3.3 Molecular Cloning*

For pENOD16::ENOD16-GFP and pENOD16::ENOD16-HA constructs, ENOD16 native promoter and the cloned by restriction digestion and T4 ligation overnight at 16$^{o}$C into pMDC43 vector (Figure A.1, Figure A.4, Figure A.5). ENOD16 coding sequence were recombined with the vector by using Gateway technology (#CatNo11791020, Invitrogen) by using manufacturer's instructions.

### *3.3.4 Hairy Root Transformation*

To transform the plants with related plasmids an altered hairy root transformation protocol has been used (Boisson-Dernier et al., 2001). The root tips of one day old seedlings were cut with a razor blade and and inoculated the plant by touching the tip to the culture. Plant has been incubated with *A. rhizogenes* strain ARqua1 containing the desired plant expression vector embed into sand with distilled water for 3 weeks.

### *3.3.5 Live/Dead Staining*

Nodule slices were stained by Cell-Check™ Viability/Cytotoxicity Kit for Bacteria Cells (#CatNoA018, ABP Biosciences). For preparation of 100x dye 1,5 µl SYTO9 and 1,5 µl Propidium Iodide mixed with 7 µl of 0.85% NaCl solution and diluted to 1X. Each slice covered by 20 µl 1X solution and incubated in dark at room temperature for 20 minutes and washed for 3 times with same amount of distilled water. Samples are left in water and just mounted by the cover slide.

### *3.3.6 GUS Staining for Promoter Analysis*

For GUS promoter analysis plats has been inoculated with Sinorhizobium melliloti Rm1021 strain expressing pNifH::GUS. Since NifH is a part of the nitrogenase complex, it is a known marker of nitrogen fixation presence (Oztas, 2017). At the 28 dpi promoter analysis has been done.

For promoter analysis, sections of nodules from 28 dpi wild type roots were stained with 1 mM X-Gluc in PBS supplemented with 5 mM potassium ferricyanide and 5 mM potassium ferrocyanide under vacuum for 2 h and 37 ◦C for 16 h. Subsequently, they were rinsed in 1.5% glutaraldehyde in 1X PBS (pH 7.4) three times for 5 min.

### *3.3.7 Imaging*

For confocal microscopy, nodules were sectioned around 10um using razor blades and mounted on microscope slides. Samples were observed live under Leica SP8 (Leica, Germany) confocal laser scanning microscope provided by KUTTAM, Turkiye. For GFP and SYTO9 signal , excitation with a 485 nm laser and emission with a 490-540 nm band pass filter has been used; for mCherry and Propidium Iodide signal excitation with a 587 nm laser and emission with a 575-675 nm band pass filter has been used. For imaging of GUS Staining and nodule phenotyping Leica S9i Stereo Microscope (Leica, Germany) has been used and provided by KUYTAM, Turkiye.

## *3.4 Results*

### *3.4.1 ENOD16 are spesifically expressed in nodules*

ENOD16 is a gene that shows nodule-specific expression in *M. truncatula*, as demonstrated by analyses using the *M. truncatula* gene expression atlas (Benedito et al., 2008) and the Symbimics database (Roux et al., 2014). The gene is induced following bacterial inoculation, suggesting its role in nodule development or function. While ENOD16 is expressed throughout all nodule zones, its expression reaches its peak level in the transition zone. This expression pattern implies that ENOD16 may have specific functions related to nodule development or metabolism that are particularly important during the transition phase of nodule development (Figure 3.3).

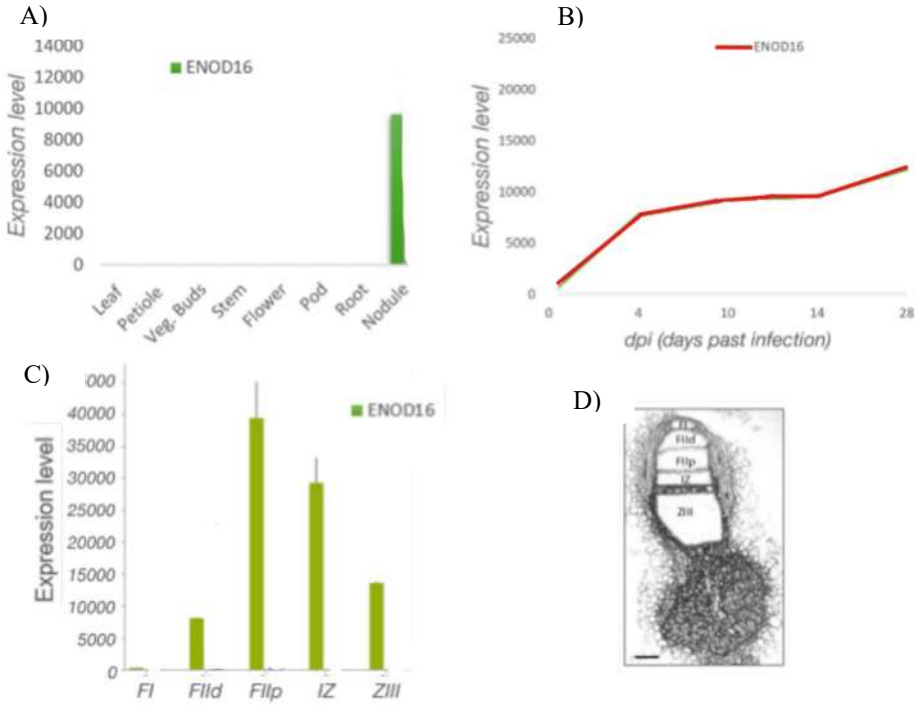



Figure 3.3: Expression profiles of ENOD16. (A) ENOD16 and expression levels in different tissues. (B) Temporal profile of ENOD16 expression in nodules. (C) ENOD16 expression in different nodule zones. (D) Nodule zones. FI-meristem, FII2d-infection, FIIp- differentiation, IZ- nitrogen fixation, ZIIImaturation. (Roux et al., 2014)

### *3.4.2 ENOD16 is essential for nitrogen fixation*

Tnt1 insertion *enod16* mutant plants screened with the PCR approach at a before work for homozygous plants (Oztas, 2017). After inoculation at 28 days, plants checked for their whole plant and nodule phenotypes (Figure 3.4).

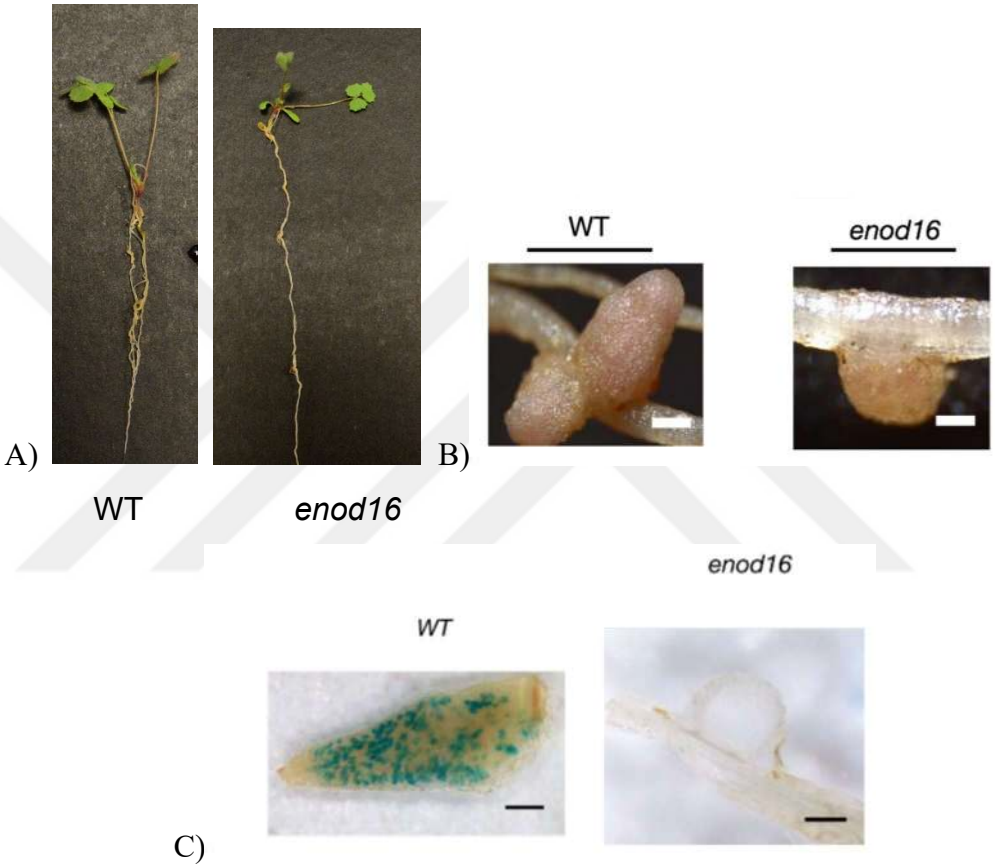



Figure 3.4: a. Growth phenotypes of R108 (WT), *enod16* mutant plants that were inoculated with *S. medicae* strain ABS7. The plants were examined 28 days post inoculation (dpi), b. Nodule phenotypes of WT, enod16 that were inoculated with *S. meliloti* strain Rm1021 expressing pNifH::GUS at 28 dpi (Scale bar: 500 um).

Roots were shorter in length compared to WT, at the same time, plants were not able to reach the height that WT upper ground parts have reached, and leaves were showing the symptoms of nitrogen starvation more compared to WT. When looked at the nodules were white and smaller. When looking at the WT, it is bigger in size and has a pink color which shows the signs for nitrogen fixation all along. All of this information confirmed

with GUS staining at nodule sectons inoculated with Rm1021 containing pNifH:::GUS, with promoter assay mutant showed no signal of nitrogen fixation. According to the information given, related homozygous *enod16* mutants showed fix- phenotype.

In the rescue experiment for confirming the role on nitrogen fixation and nodule development that has been identified, pE16::E16 constructs transformed to *enod16* mutant plants. As result enod16 phenotype fully saved after transformation at 28 dpi and showed WT phenotype as pink nodules compared to non-treated mutant plants (Figure 3.5). When checked with confocal images undifferenciated symbiosomes are fully rescued in transformed plants and their appearance became elongated compared to dot-like phenotpe at mutant plant rhizobia (Figure 3.6). Expression of the ENOD16 confirmed by a western blot experiment done with GFP antibody at 21 dpi and band was at the expected size at 37 kDa (Figure 3.7).


Figure 3.5: Rescue experiment of ENOD16 with *enod16-2* mutant plants at 28 dpi. A) pE16::E16-HA construct., B) Nodule images of WT, *enod16* with or without ENOD16 expression. Plants have been inoculated with *S. melliloti* strain ABS7.

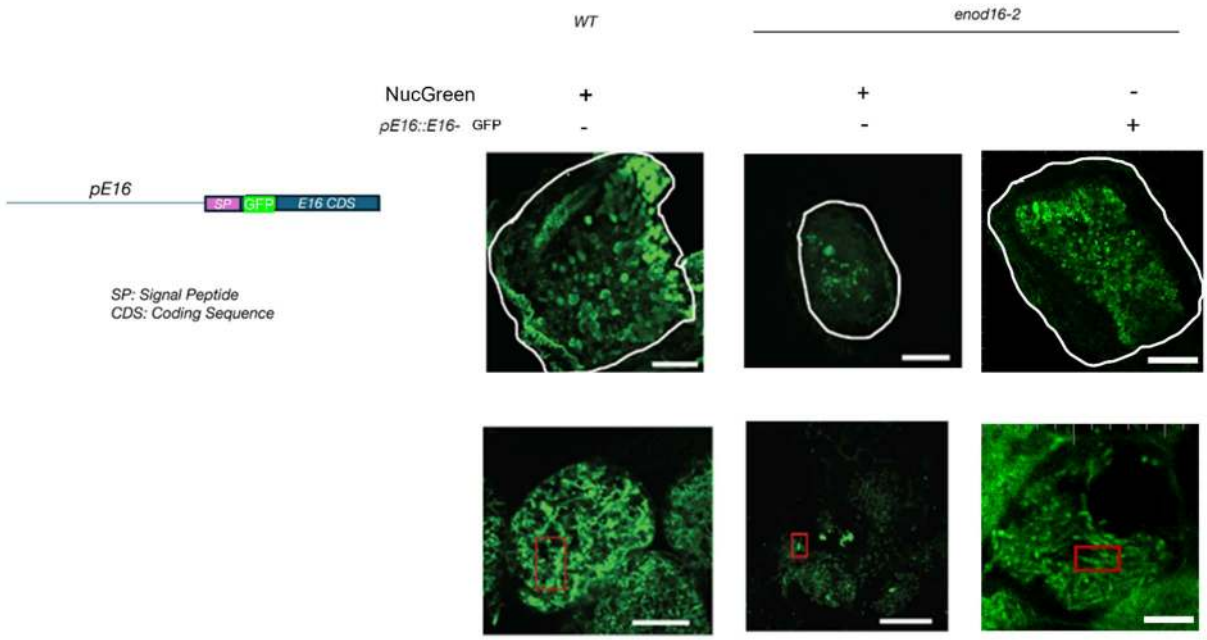



Figure 3.6: Confocal imaging of rescue experiment for ENOD16. GFP-ENOD16 construct transformed to *enod16-2* Tnt1 insertion mutants and get inoculated with *Sinorhizobium medicae* strain ABS7. Nodule images of WT, *enod16-2* with or without expression of GFP-ENOD16.(scale bar: 200um) at first row, symbiosome images of WT, *enod16-2* with or without expression of GFP-ENOD16 at second row. Scale bar was set at 20 um.

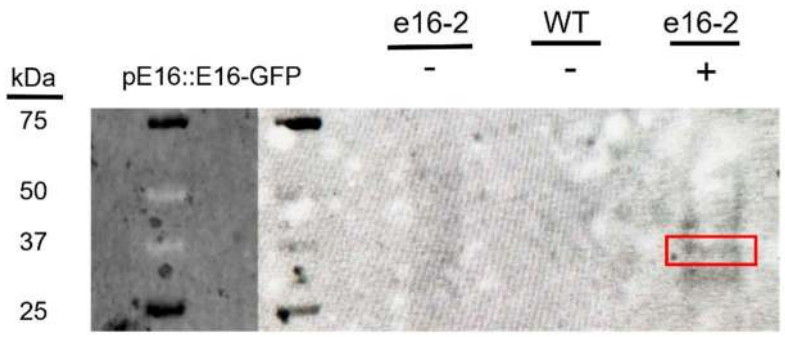



Figure 3.7: Western Blot results of rescue experiment for ENOD16. GFP-ENOD16 construct transformed to *enod16-2* Tnt1 insertion mutants and get inoculated with Sinorhizobium medicae strain ABS7.

### *3.4.3 ENOD16 is essential for the survival of rhizobia*

For observation of rhizobia differentiation and survival, bacterial LIVE/DEAD staining has been done on 10 um nodule sections. NucGreen marks the alive bacteria while Propidium Iodide marks dead bacteria. At full nodule as expected *enod16-2* showed fully dead symbiosomes but at *enod16-1* did not form proper symbiosoms (Figure 3.8). When we look at the close up on symbiosomes we cannot see any differentiated rhizobia around symbiosome dead or alive. But at enod16-1 we can see differentiated rhizobia with some vitality which shows a phenotype as WT (Figure 3.8). Which all comes to conclusion that without ENOD16 bacteria cannot differentiate or survive, thus it will end up in improper nitrogen fixation.

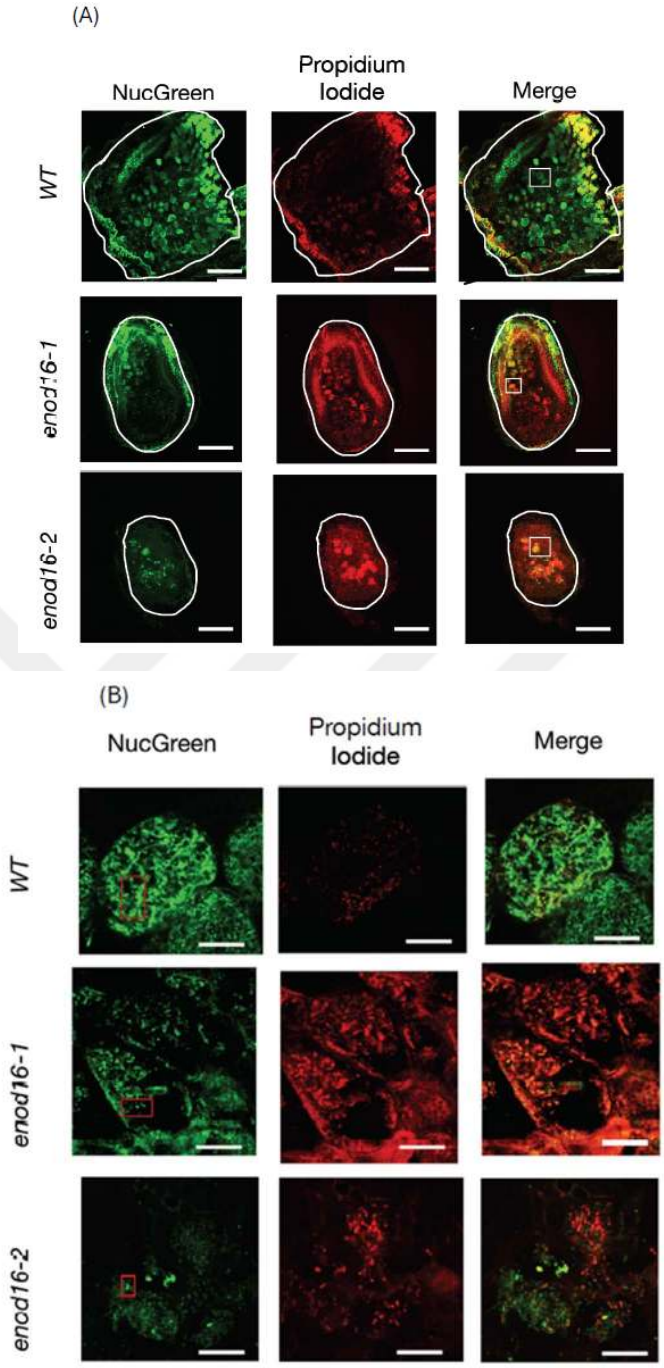



Figure 3.8: Live/dead staining images of *enod16-1, enod16-2* and WT. A) Whole nodule sections. Scale bar was set at 200 um. B) Symbiosomes stained with PI/NucGreen, 28 dpi . PI signal was detected using excitation with a 561 nm laser and emission with a 570-630 nm band pass filter, whereas NucGreen signal was detected using excitation with a 488 nm laser and emission with a 490-550 nm band pass filter. Scale bar was set at 20 um.

### *3.4.4 ENOD16 is localized to symbiosome membrane*

DNF1 is part of signal peptidase complex that acts in delivery of NCR peptides. DNF1 help with the secretion of DNF2 from ER and DNF2 will solubilize ENOD16 by cleaving GPI-anchor prior to experiments. It actively interacts with DNF2 for rhizobia survival. *dnf1* Tnt1 mutants shows fix- phenotype and share similarity with *enod16* mutant such as undifferentiated symbiosomes- rhizobia and improper nitrogen fixation. And as known The behavior of ENOD16 has been checked when DNF1's absence by transforming pE16::E16-GFP to *dnf1* mutants. While *dnf1* mutants shows undifferentiated symbiosomes we can see the ENOD16 signal anchored to the symbiosome membrane whether at infection thread at higher expression levels, or through the nodule at undifferentiated symbiosome. However, unexpectedly in some of the expression patterns we observed unexpected differentiation of the bacteria, which we can say it partially rescued the phenotype (Figure3.9). This helps with the hypothesis that in the absence of DNF1, ENOD16 cannot solubilize and will be localized to the symbiosome membrane.

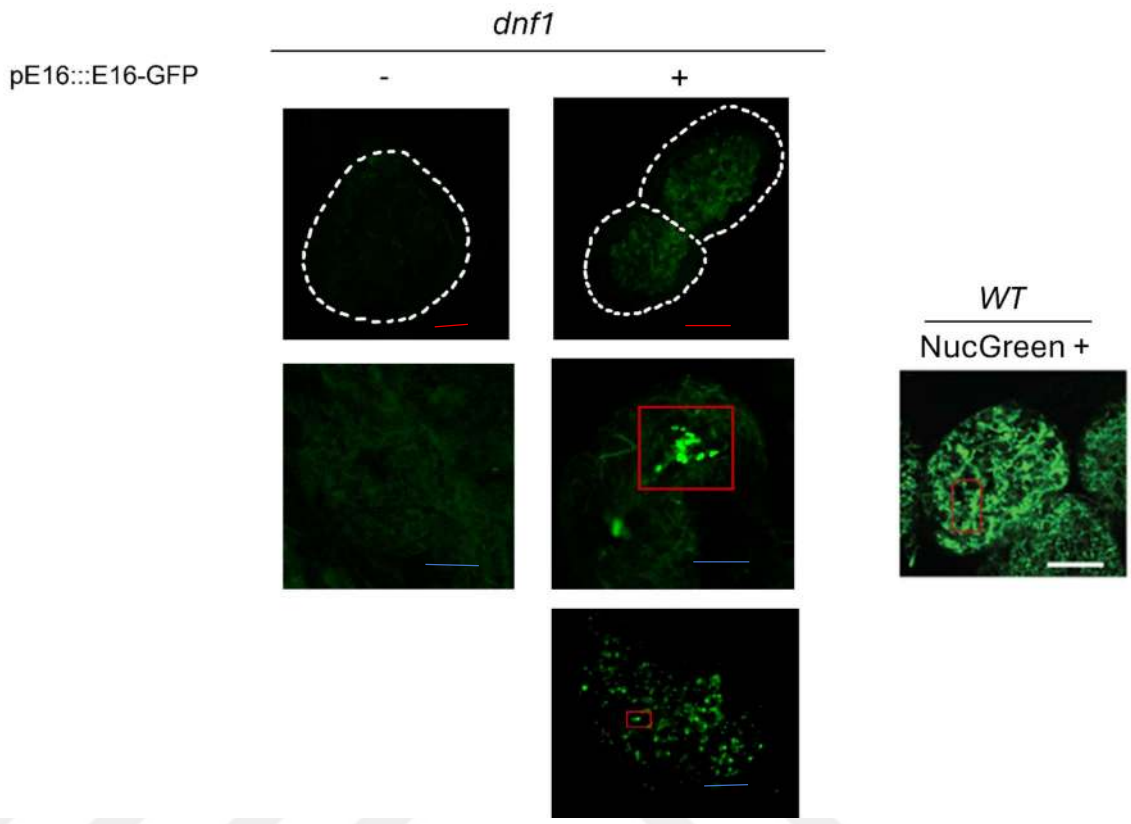


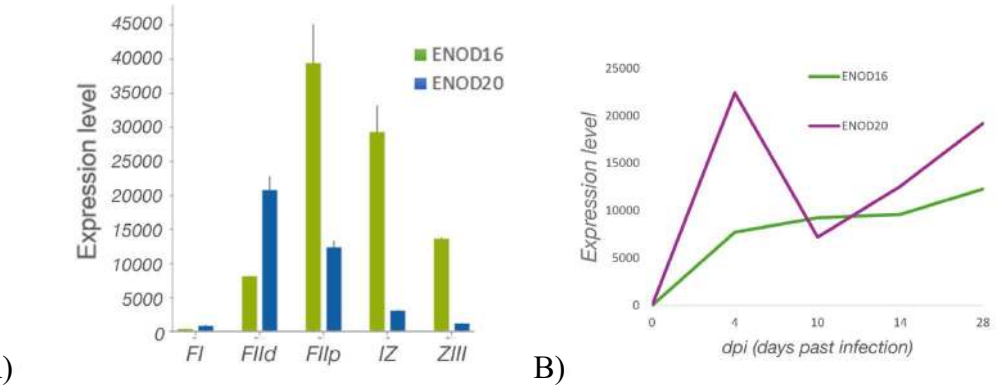

Figure 3.9: Imaging of transformation of *dnf1* Tnt1 insetion mutants with GFP-ENOD16 at 28 dpi. Controls will include untreated *dnf1* mutant and WT plant stained with NucGreen. (red scale bar: 200um ,blue and white scale bar:20 um)

## *3.5 Discussion*

In this research an arabinogalactan protein called ENOD16 has been described and explained. It is only expressed in nodules and have different properties in different steps of nodule development and nitrogen fixation.

In previous experiments and with glycosylation motifs shown in amino acid sequence it has been proven that ENOD16 GPI anchor get cleaaved by DNF2, which showes that ENOD16 is a GPI-anchored protein but it is an indirect proof since solubilizatiıon of the protein did not verified (Oztas, 2017) (Figure 3.2). Solubilization of the protein by GPI-anchor excition should be shown in vivo by TritonX-114 partitioning with both *dnf1* and *dnf2* mutants in next experiments.

ENOD16 is expressed at the early times of nodulation phase , in root cortical cells with infection threads and nodule cells. ENOD16 is only expressed in nodules and has the highest expression levels transition and nitrogen-fixation zones(Roux et al., 2014) (Figure 3.3). But there is a distinct expression pattern between ENOD16 and close companion ENOD20, which indicates to their different roles at different nodulation processes which should be examined in future (Figure 3.10).




A) B)

Figure 3.10: Expression profiles of ENOD16 and ENOD20. A) Temporal profile of ENOD16 and ENOD20 expression in nodules. B) ENOD16 and ENOD20 expression in different nodule zones. (Roux et al., 2014)

ENOD16 located in the symbiosome space, surrounding young and mature bacteroids (Oztas, 2017). I observed ENOD16 in the *dnf1* symbiosomes, suggesting that ENOD16 are most likely cleaved into the symbiosome space with a relation of DNF1 but does have a partial rescue, which needs to be uncovered for reasons and might mean *dnf1* bacteroids can survive until entering differentiation.

ENOD16 protein has been explained long time before (Greene et al., 1998); however, their phenotypes throughout nodulation did not shown. We found that the absence of ENOD16 causes, improper maturation of the rhizobia and stops the survival throughout the nodulation process which is causing nitrogen fixation deficiency. We observed a phenotype that is closer to WT in *enod16-1* compared to *enod16-2*, this is a result of the different Tnt1 insertion sites.

ENOD16 might function as structural protein in cell walls which can take role in cell wall differentiation as AGPs. However, the specific biochemical functions of AGP remain unexplained in literature. It is still unclear whether the glycosylation pattern or other post translational modifications plays a important role in selection the functions of this protein.

All of these leads to us our final model of ENOD16 is solubilized by DNF2 (Oztas, 2017)which will be triggered by DNF1. After solubilization bacteria will enter the survival mode and start to differenciate actively fixing nitrogen (Figure 3.11).

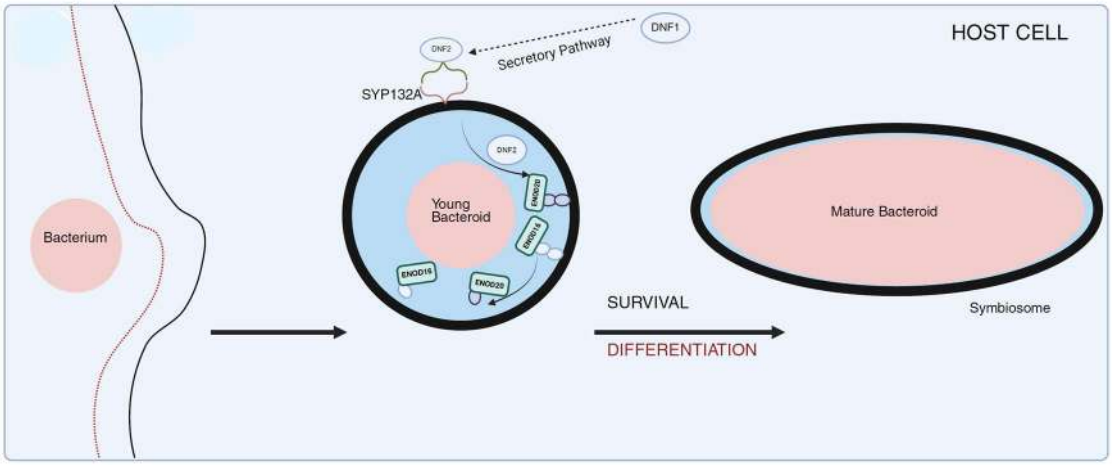



Figure 3.11: Final model for bacteroid survival and differentiation

## *3.6 Future Directions*

For future experiments, GPI-anchor cleavage by DNF2 must be verified in vivo by TX-114 phasing of the proteins, since evidence is very limited. At the same time, for final verification of the structure, crystallization will be run on purified protein from the tobacco expression system. It will be determined by X-Ray Diffraction imaging.

For their distinct expression levels in different nodule layers and nodulation phases, ENOD16 and ENOD20 show different phenotypes, which can make their role different in nodulation. For further investigation of these double knockouts, the proteins should be prepared by CRISPR/CAS technology, and phenotype rescue should be applied for each protein separately on the mutant.

The specific biochemical functions of arabinogalactan proteins remain largely unidentified. The specific role of the glycosylation motif of ENOD16 is unknown, as are post-translational modifications. The role of glycosylation and other post-translational modifications will be identified by proteomics approaches and will be verified by reverse genetics.

# CHAPTER 4:
# CONCLUSION

In order to combat nitrogen shortage, legumes form symbiotic relationships with rhizobia. It begins with a conversation between the rhizobia and the host plant over the root. Through a mechanism known as the infection thread, the bacteria enters the root through the root hairs. Following the infection thread's extension into the cortical cells, the bacteria are endocytosed and placed in membrane-bound spaces known as symbiosomes, where they undergo nitrogen-fixing form differentiation.

It has been described that the two of the mechanisms that maintain the intracellular rhizobia in the *M. truncatula- S. meliloti* symbiotic interaction. First part described the portion focusing on the role of DNA-binding Nodule Cysteine- Rich (NCR) peptides from the host and bacterial effectors in modifying each others transcriptomes. It has been shown that host NCR peptides can bind to DNA function as transcription inhibitors. Three NCR peptides that can bind to DNA has been identified by proteomics approaches and they have an active role on symbiotic nitrogen fixation.

At second part, experiments has been done for Early Nodulins (ENODs), proteins expressed during the initial stages of nodule formation. They are secreted to participate in the remodeling of the plant cell wall during nodule development, which helps in the bacterial survival. ENOD16 are expressed exclusively in nodules, and here it has been proved that ENOD16 is a nodule specific protein that is essential for nitrogen fixation.

As early and late nodulins examined in this research, symbiosome secreted proteins together will complete the process of survival and differentiation of rhizobia and will led to efficient nitrogen fixation.

# VITA

Ceren ÖZDEMİR graduated with distinction from Yıldız Technical University, where she earned her Bachelor of Science degree. Her thesis was on tissue culture techniques for the increasing the yield of production of secondary metabolites from medicinal plants and under the title of "Establishment of In Vitro Callus Cultures in Salvia sclarea" and supported by TUBITAK 2209-A program. Her current area of interest is understanding symbiotic nitrogen fixation pathways and their molecular intricacies. Now on pursuing a MSc at Koç University, she focuses on unraveling the mysteries of these symbiotic relationships at a molecular level and check out for the two of the main events in symbiotic nitrogen fixation: differentiation and survival. Her interdisciplinary approach in her project at the other side, incorporating proteomics analyses with the transcriptomic data and enlightens the unknown role of the NCR peptides in the host - rhizobia interactions to resolve the main flow of the nitrogen-fixation at plants.

# BIBLIOGRAPHY

Adebali, O., Chiou, Y. Y., Hu, J., Sancar, A., & Selby, C. P. (2017). Genome-wide transcription-coupled repair in Escherichia coli is mediated by the Mfd translocase. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(11), E2116–E2125. https://doi.org/10.1073/pnas.1700230114

Ahemad, M., & Kibret, M. (2014). Mechanisms and applications of plant growth promoting rhizobacteria: Current perspective. In *Journal of King Saud University - Science* (Vol. 26, Issue 1, pp. 1–20). Elsevier B.V. https://doi.org/10.1016/j.jksus.2013.05.001

Alunni, B., Kevei, Z., Redondo-Nieto, M., Kondorosi, A., Mergaert, P., & Kondorosi, E. (2007). Genomic Organization and Evolutionary Insights on GRP and NCR Genes, Two Large Nodule-Specific Gene Families in Medicago truncatula. *Molecular Plant-Microbe Interactions MPMI*, *20*(9), 1138–1148. https://doi.org/10.1094/MPMI

Apweiler, R. (2004). UniProt: the Universal Protein knowledgebase. *Nucleic Acids Research*, *32*(90001), 115D – 119. https://doi.org/10.1093/nar/gkh131

Baek, M., McHugh, R., Anishchenko, I., Jiang, H., Baker, D., & DiMaio, F. (2024). Accurate prediction of protein–nucleic acid complexes using RoseTTAFoldNA. *Nature Methods*, *21*(1), 117–121. https://doi.org/10.1038/s41592-023-02086-5

Benedito, V. A., Torres-Jerez, I., Murray, J. D., Andriankaja, A., Allen, S., Kakar, K., Wandrey, M., Verdier, J., Zuber, H., Ott, T., Moreau, S., Niebel, A., Frickey, T., Weiller, G., He, J., Dai, X., Zhao, P. X., Tang, Y., & Udvardi, M. K. (2008). A gene expression atlas of the model legume Medicago truncatula. *Plant Journal*, *55*(3), 504–513. https://doi.org/10.1111/j.1365-313X.2008.03519.x

Boisson-Dernier, A., Chabaud, M., Garcia, F., Bécard, G., Rosenberg, C., & Barker, D. G. (2001). Agrobacterium rhizogenes-Transformed Roots of Medicago truncatula for the Study of Nitrogen-Fixing and Endomycorrhizal Symbiotic Associations. In */ 695 MPMI* (Vol. 14, Issue 6).

Brady, J. D., Sadler?, I. H., & Frye, S. C. (1998). *PULCHEROSINE, AN OXIDATIVELY COUPLED TRIMER OF TYROSINE IN PLANT CELL WALLS: ITS ROLE IN CROSS-LINK FORMATION* (Vol. 47, Issue 3).

Brewin, N. J. (2004). Plant cell wall remodelling in the rhizobium-legume symbiosis. In *Critical Reviews in Plant Sciences* (Vol. 23, Issue 4, pp. 293–316). https://doi.org/10.1080/07352680490480734

Cassab, G. I. (1998). PLANT CELL WALL PROTEINS. In *Annu. Rev. Plant Physiol. Plant Mol. Biol* (Vol. 49). www.annualreviews.org

Chang, C., Damiani, I., Puppo, A., & Frendo, P. (2009). Redox changes during the legume-rhizobium symbiosis. In *Molecular Plant* (Vol. 2, Issue 3, pp. 370–377). Oxford University Press. https://doi.org/10.1093/mp/ssn090

Compant, S., Clément, C., & Sessitsch, A. (2010). Plant growth-promoting bacteria in the rhizo- and endosphere of plants: Their role, colonization, mechanisms involved and prospects for utilization. In *Soil Biology and Biochemistry* (Vol. 42, Issue 5, pp. 669–678). https://doi.org/10.1016/j.soilbio.2009.11.024

Cox, J., & Mann, M. (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nature Biotechnology*, *26*(12), 1367–1372. https://doi.org/10.1038/nbt.1511

Crook, M. B., Lindsay, D. P., Biggs, M. B., Bentley, J. S., Price, J. C., Clement, S. C., Clement, M. J., Long, S. R., & Griffitts, J. S. (2012). Rhizobial Plasmids That Cause Impaired Symbiotic Nitrogen Fixation and Enhanced Host Invasion. *Molecular Plant-Microbe Interactions MPMI*, *25*(8). https://doi.org/10.1094/MPMI

Downie, J. A., & Kondorosi, E. (2021). Why Should Nodule Cysteine-Rich (NCR) Peptides Be Absent From Nodules of Some Groups of Legumes but Essential for Symbiotic N-Fixation in Others? In *Frontiers in Agronomy* (Vol. 3). Frontiers Media S.A. https://doi.org/10.3389/fagro.2021.654576

Ellis, M., Egelund, J., Schultz, C. J., & Bacic, A. (2010). Arabinogalactan-proteins: Key regulators at the cell surface? *Plant Physiology*, *153*(2), 403–419. https://doi.org/10.1104/pp.110.156000

Farkas, A., Maróti, G., Dürgo, H., Györgypál, Z., Lima, R. M., Medzihradszky, K. F., Kereszt, A., Mergaert, P., & Kondorosi, É. (2014). Medicago truncatula symbiotic peptide NCR247 contributes to bacteroid differentiation through multiple

mechanisms. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(14), 5183–5188. https://doi.org/10.1073/pnas.1404169111

Fauvart, M., & Michiels, J. (2008). Rhizobial secreted proteins as determinants of host specificity in the rhizobium-legume symbiosis. In *FEMS Microbiology Letters* (Vol. 285, Issue 1, pp. 1–9). https://doi.org/10.1111/j.1574-6968.2008.01254.x

Galloway, J. N., Dentener, F. J., Capone, D. G., Boyer, E. W., Howarth, R. W., Seitzinger, S. P., Asner, G. P., Cleveland, C. C., Green, P. A., Holland, E. A., Karl, D. M., Michaels, A. F., Porter, J. H., Townsend, A. R., Vo¨ro¨smarty, C. J., Vo¨ro, V., & Vo¨ro¨smarty, V. (n.d.). *Nitrogen cycles: past, present, and future*.

Gao, F., Yang, J., Zhai, N., Zhang, C., Ren, X., Zeng, Y., Chen, Y., Chen, R., & Pan, H. (2023). NCR343 is required to maintain the viability of differentiated bacteroids in nodule cells in Medicago truncatula. *New Phytologist*, *240*(2), 815–829. https://doi.org/10.1111/nph.19180

Gautrat, P., Mortier, V., Laffont, C., De Keyser, A., Fromentin, J., Frugier, F., & Goormachtig, S. (2019). Unraveling new molecular players involved in the autoregulation of nodulation in Medicago truncatula. *Journal of Experimental Botany*, *70*(4), 1407–1417. https://doi.org/10.1093/jxb/ery465

Greene, E. A., Erard, M., Dedieu, A., & Barker, D. G. (1998). MtENOD16 and 20 are members of a family of phytocyanin-related early nodulins. In *Plant Molecular Biology* (Vol. 36).

Hwa Yun, B., Guo, J., Bellamri, M., & Turesky, R. J. (2020). DNA adducts: Formation, biological effects, and new biospecimens for mass spectrometric measurements in humans. In *Mass Spectrometry Reviews* (Vol. 39, Issues 1–2, pp. 55–82). John Wiley and Sons Inc. https://doi.org/10.1002/mas.21570

Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S. A. A., Ballard, A. J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., … Hassabis, D. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature*, *596*(7873), 583–589. https://doi.org/10.1038/s41586-021-03819-2

Kelly, S., Radutoiu, S., & Stougaard, J. (2017). Legume LysM receptors mediate symbiotic and pathogenic signalling. In *Current Opinion in Plant Biology* (Vol. 39, pp. 152–158). Elsevier Ltd. https://doi.org/10.1016/j.pbi.2017.06.013

Kereszt, A., Mergaert, P., Montiel, J., Endre, G., & Kondorosi, É. (2018). Impact of plant peptides on symbiotic nodule development and functioning. In *Frontiers in Plant Science* (Vol. 9). Frontiers Media S.A. https://doi.org/10.3389/fpls.2018.01026

Kim, M., Chen, Y., Xi, J., Waters, C., Chen, R., & Wang, D. (2015). An antimicrobial peptide essential for bacterial survival in the nitrogen-fixing symbiosis. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(49), 15238–15243. https://doi.org/10.1073/pnas.1500123112

Kisker, C., Kuper, J., & Van Houten, B. (2013). Prokaryotic nucleotide excision repair. *Cold Spring Harbor Perspectives in Biology*, *5*(3). https://doi.org/10.1101/cshperspect.a012591

Kondorosi, E., Mergaert, P., & Kereszt, A. (2013). A Paradigm for endosymbiotic life: Cell differentiation of rhizobium bacteria provoked by host plant factors. *Annual Review of Microbiology*, *67*, 611–628. https://doi.org/10.1146/annurev-micro-092412-155630

Kouchi, H., Imaizumi-Anraku, H., Hayashi, M., Hakoyama, T., Nakagawa, T., Umehara, Y., Suganuma, N., & Kawaguchi, M. (2010). How many peas in a pod? Legume genes responsible for mutualistic symbioses underground. *Plant and Cell Physiology*, *51*(9), 1381–1397. https://doi.org/10.1093/pcp/pcq107

Li, W., Hu, J., Adebali, O., Adar, S., Yang, Y., Chiou, Y. Y., & Sancar, A. (2017). Human genome-wide repair map of DNA damage caused by the cigarette smoke carcinogen benzo[a]pyrene. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(26), 6752–6757. https://doi.org/10.1073/pnas.1706021114

Libault, M. (2018). Transcriptional reprogramming of legume genomes: Perspective and challenges associated with single-cell and single cell-type approaches during nodule development. *Frontiers in Plant Science*, *871*. https://doi.org/10.3389/fpls.2018.01600

Lima, R. M., Kylarová, S., Mergaert, P., & Kondorosi, É. (2020). Unexplored Arsenals of Legume Peptides With Potential for Their Applications in Medicine and Agriculture. In *Frontiers in Microbiology* (Vol. 11). Frontiers Media S.A. https://doi.org/10.3389/fmicb.2020.01307

Luu, T. B., Ourth, A., Pouzet, C., Pauly, N., & Cullimore, J. (2022). A newly evolved chimeric lysin motif receptor-like kinase in Medicago truncatula spp. tricycla R108 extends its Rhizobia symbiotic partnership. *New Phytologist*, *235*(5), 1995–2007. https://doi.org/10.1111/nph.18270

Mashiguchi, K., Asami, T., & Suzuki, Y. (2009). Genome-wide identification, structure and expression studies, and mutant collection of 22 early nodulin-like protein genes in arabidopsis. *Bioscience, Biotechnology and Biochemistry*, *73*(11), 2452–2459. https://doi.org/10.1271/bbb.90407

Masson-Boivin, C., & Sachs, J. L. (2018). Symbiotic nitrogen fixation by rhizobia — the roots of a success story. In *Current Opinion in Plant Biology* (Vol. 44, pp. 7–15). Elsevier Ltd. https://doi.org/10.1016/j.pbi.2017.12.001

Medzihradszky, K. F. (2005). In-solution digestion of proteins for mass spectrometry. In *Methods in Enzymology* (Vol. 405, pp. 50–65). Academic Press Inc. https://doi.org/10.1016/S0076-6879(05)05003-2

Mendoza-Suárez, M., Andersen, S. U., Poole, P. S., & Sánchez-Cañizares, C. (2021). Competition, Nodule Occupancy, and Persistence of Inoculant Strains: Key Factors in the Rhizobium-Legume Symbioses. In *Frontiers in Plant Science* (Vol. 12). Frontiers Media S.A. https://doi.org/10.3389/fpls.2021.690567

Mergaert, P., Nikovics, K., Kelemen, Z., Maunoury, N., Vaubert, D., Kondorosi, A., & Kondorosi, E. (2003). A novel family in Medicago truncatula consisting of more than 300 nodule-specific genes coding for small, secreted polypeptides with conserved cysteine motifs. *Plant Physiology*, *132*(1), 161–173. https://doi.org/10.1104/pp.102.018192

Mergaert, P., Uchiumi, T., Alunni, B., Naë Lle Evanno, G., Lique Cheron, A., Catrice, O., Mausset, A.-E., Dé Rique Barloy-Hubler, F., Galibert, F., Kondorosi, A., & Kondorosi, E. (2006). *Eukaryotic control on bacterial cell cycle and differentiation in the Rhizobium-legume symbiosis*. https://www.pnas.org

Mikuláss, K. R., Nagy, K., Bogos, B., Szegletes, Z., Kovács, E., Farkas, A., Váró, G., Kondorosi, É., & Kereszt, A. (2016). Antimicrobial nodule-specific cysteine-rich peptides disturb the integrity of bacterial outer and inner membranes and cause loss of membrane potential. *Annals of Clinical Microbiology and Antimicrobials*, *15*(1). https://doi.org/10.1186/s12941-016-0159-8

Moreau, S., Verdenaud, M., Ott, T., Letort, S., de Billy, F., Niebel, A., Gouzy, J., de Carvalho-Niebel, F., & Gamas, P. (2011). Transcription reprogramming during root nodule development in Medicago truncatula. *PLoS ONE*, *6*(1). https://doi.org/10.1371/journal.pone.0016463

Motose, H., Sugiyama, M., & Fukuda, H. (2004). A proteoglycan mediates inductive interaction during plant vascular development. *Nature*, *429*(6994), 873–878. https://doi.org/10.1038/nature02613

Mus, F., Crook, M. B., Garcia, K., Costas, A. G., Geddes, B. A., Kouri, E. D., Paramasivan, P., Ryu, M. H., Oldroyd, G. E. D., Poole, P. S., Udvardi, M. K., Voigt, C. A., Ané, J. M., & Peters, J. W. (2016). Symbiotic nitrogen fixation and the challenges to its extension to nonlegumes. In *Applied and Environmental Microbiology* (Vol. 82, Issue 13, pp. 3698–3710). American Society for Microbiology. https://doi.org/10.1128/AEM.01055-16

Nelson, M. S., & Sadowsky, M. J. (2015). Secretion systems and signal exchange between nitrogen-fixing rhizobia and legumes. In *Frontiers in Plant Science* (Vol. 6, Issue JULY). Frontiers Media S.A. https://doi.org/10.3389/fpls.2015.00491

Oldroyd, G. E. D. (2013). Speak, friend, and enter: Signalling systems that promote beneficial symbiotic associations in plants. In *Nature Reviews Microbiology* (Vol. 11, Issue 4, pp. 252–263). https://doi.org/10.1038/nrmicro2990

Oldroyd, G. E. D., & Dixon, R. (2014). Biotechnological solutions to the nitrogen problem. In *Current Opinion in Biotechnology* (Vol. 26, pp. 19–24). https://doi.org/10.1016/j.copbio.2013.08.006

Oztas, O. (2017). *THE KEY QUESTION IN SYMBIOTIC NITROGEN FIXATION: HOW DOES HOST MAINTAIN A BACTERIAL SYMBIONT?* [UNIVERSITY OF MASSACHUSETTS AMHERST]. https://doi.org/10.7275/9588108.0

Oztas, O., Selby, C. P., Sancar, A., & Adebali, O. (2018). Genome-wide excision repair in Arabidopsis is coupled to transcription and reflects circadian gene expression

patterns. *Nature Communications*, *9*(1). https://doi.org/10.1038/s41467-018-03922-5

Pan, H., & Wang, D. (2017). Nodule cysteine-rich peptides maintain a working balance during nitrogen-fixing symbiosis. In *Nature Plants* (Vol. 3). Palgrave Macmillan Ltd. https://doi.org/10.1038/nplants.2017.48

Pankievicz, V. C. S., Irving, T. B., Maia, L. G. S., & Ané, J. M. (2019). Are we there yet? The long walk towards the development of efficient symbiotic associations between nitrogen-fixing bacteria and non-leguminous crops. In *BMC Biology* (Vol. 17, Issue 1). BioMed Central Ltd. https://doi.org/10.1186/s12915-019-0710-0

Penterman, J., Abo, R. P., De Nisco, N. J., Arnold, M. F. F., Longhi, R., Zanda, M., & Walker, G. C. (2014). Host plant peptides elicit a transcriptional response to control the Sinorhizobium meliloti cell cycle during symbiosis. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(9), 3561–3566. https://doi.org/10.1073/pnas.1400450111

Perlow, R. A., Kolbanovskii, A., Hingerty, B. E., Geacintov, N. E., Broyde, S., & Scicchitano, D. A. (2002). DNA adducts from a tumorigenic metabolite of benzo[a]pyrene block human RNA polymerase II elongation in a sequence- and stereochemistry-dependent manner. *Journal of Molecular Biology*, *321*(1), 29–47. https://doi.org/10.1016/S0022-2836(02)00593-4

Radutoiu, S., Madsen, L. H., Madsen, E. B., Felle, H. H., Umehara, Y., Grønlund, M., Sato, S., Nakamura, Y., Tabata, S., Sandal, N., & Stougaard, J. (2003). *Plant recognition of symbiotic bacteria requires two LysM receptor-like kinases*. www.nature.com/nature

Remigi, P., Zhu, J., Young, J. P. W., & Masson-Boivin, C. (2016). Symbiosis within Symbiosis: Evolving Nitrogen-Fixing Legume Symbionts. In *Trends in Microbiology* (Vol. 24, Issue 1, pp. 63–75). Elsevier Ltd. https://doi.org/10.1016/j.tim.2015.10.007

Rose, R. J. (2008). Medicago truncatula as a model for understanding plant interactions with other organisms, plant development and stress biology: Past, present and future. In *Functional Plant Biology* (Vol. 35, Issue 4, pp. 253–264). https://doi.org/10.1071/FP07297

Roux, B., Rodde, N., Jardinaud, M. F., Timmers, T., Sauviac, L., Cottret, L., Carrère, S., Sallet, E., Courcelle, E., Moreau, S., Debellé, F., Capela, D., De Carvalho-Niebel, F., Gouzy, J., Bruand, C., & Gamas, P. (2014). An integrated analysis of plant and bacterial gene expression in symbiotic root nodules using laser-capture microdissection coupled to RNA sequencing. *Plant Journal*, *77*(6), 817–837. https://doi.org/10.1111/tpj.12442

Roy, S., Liu, W., Nandety, R. S., Crook, A., Mysore, K. S., Pislariu, C. I., Frugoli, J., Dickstein, R., & Udvardi, M. K. (2020). Celebrating 20 Years of Genetic Discoveries in Legume Nodulation and Symbiotic Nitrogen Fixation[OPEN]. In *Plant Cell* (Vol. 32, Issue 1, pp. 15–41). American Society of Plant Biologists. https://doi.org/10.1105/tpc.19.00279

Rutten, L., Miyata, K., Roswanjaya, Y. P., Huisman, R., Bu, F., Hartog, M., Linders, S., van Velzen, R., van Zeijl, A., Bisseling, T., Kohlen, W., & Geurts, R. (2020). Duplication of symbiotic lysin motif receptors predates the evolution of nitrogen-fixing nodule symbiosis. *Plant Physiology*, *184*(2), 1004–1023. https://doi.org/10.1104/pp.19.01420

Santi, C., Bogusz, D., & Franche, C. (2013). Biological nitrogen fixation in non-legume plants. In *Annals of Botany* (Vol. 111, Issue 5, pp. 743–767). https://doi.org/10.1093/aob/mct048

Schultz, C., Gilson, P., Oxley, D., Youl, J., & Bacic, A. (1998). GPI-anchors on arabinogalactan-proteins: implications for signalling in plants. *Trends in Plant Science*, *3*(11), 426–431. https://doi.org/10.1016/S1360-1385(98)01328-4

Seifert, G. J., & Roberts, K. (2007). The biology of arabinogalactan proteins. In *Annual Review of Plant Biology* (Vol. 58, pp. 137–161). https://doi.org/10.1146/annurev.arplant.58.032806.103801

Shi, H., Kim, Y. S., Guo, Y., Stevenson, B., & 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*(1), 19–32. https://doi.org/10.1105/tpc.007872

Shimoda, Y., Nishigaya, Y., Yamaya-Ito, H., Inagaki, N., Umehara, Y., Hirakawa, H., Sato, S., Yamazaki, T., & Hayashi, M. (n.d.). *The rhizobial autotransporter*

*determines the symbiotic nitrogen fixation activity of Lotus japonicus in a host-specific manner*. https://doi.org/10.1073/pnas.1913349117/-/DCSupplemental

Showalter, A. M., & Basu, D. (2016). Extensin and arabinogalactan-protein biosynthesis: Glycosyltransferases, research challenges, and biosensors. In *Frontiers in Plant Science* (Vol. 7, Issue JUNE2016). Frontiers Research Foundation. https://doi.org/10.3389/fpls.2016.00814

Singh, S., Katzer, K., Lambert, J., Cerri, M., & Parniske, M. (2014). CYCLOPS, A DNA-binding transcriptional activator, orchestrates symbiotic root nodule development. *Cell Host and Microbe*, *15*(2), 139–152. https://doi.org/10.1016/j.chom.2014.01.011

Strasser, R., Seifert, G., Doblin, M. S., Johnson, K. L., Ruprecht, C., Pfrengle, F., Bacic, A., & Estevez, J. M. (2021). Cracking the "Sugar Code": A Snapshot of N- and O-Glycosylation Pathways and Functions in Plants Cells. *Frontiers in Plant Science*, *12*. https://doi.org/10.3389/fpls.2021.640919

Takai, K. (2019). The nitrogen cycle: A large, fast, and mystifying cycle. *Microbes and Environments*, *34*(3), 223–225. https://doi.org/10.1264/jsme2.ME3403rh

Tiricz, H., Szücs, A., Farkas, A., Pap, B., Lima, R. M., Maróti, G., Kondorosi, É., & Kereszt, A. (2013). Antimicrobial nodule-specific cysteine-rich peptides induce membrane depolarization-associated changes in the transcriptome of Sinorhizobium meliloti. *Applied and Environmental Microbiology*, *79*(21), 6737–6746. https://doi.org/10.1128/AEM.01791-13

Tyanova, S., Temu, T., Sinitcyn, P., Carlson, A., Hein, M. Y., Geiger, T., Mann, M., & Cox, J. (2016). The Perseus computational platform for comprehensive analysis of (prote)omics data. In *Nature Methods* (Vol. 13, Issue 9, pp. 731–740). Nature Publishing Group. https://doi.org/10.1038/nmeth.3901

Udvardi, M., & Poole, P. S. (2013). Transport and metabolism in legume-rhizobia symbioses. In *Annual Review of Plant Biology* (Vol. 64, pp. 781–805). https://doi.org/10.1146/annurev-arplant-050312-120235

Van de Velde, W., Zehirov, G., Szatmari, A., Debreczeny, M., Ishihara, H., Kevei, Z., Farkas, A., Mikulass, K., Nagy, A., Tiricz, H., Satiat-Jeunemaître, B., Alunni, B., Bourge, M., Kucho, K., Abe, M., Kereszt, A., Maroti, G., Uchiumi, T., Kondorosi, E., & Mergaert, P. (2010). Plant Peptides Govern Terminal Differentiation of

Bacteria in Symbiosis. *Science*, *327*(5969), 1122–1126. https://doi.org/10.1126/science.1184057

Van Hengel, A. J., Van Kammen, A., & De Vries, S. C. (2002). *A relationship between seed development, Arabinogalactan-proteins (AGPs) and the AGP mediated promotion of somatic embryogenesis*.

van Velzen, R., Holmer, R., Bu, F., Rutten, L., van Zeijl, A., Liu, W., Santuari, L., Cao, Q., Sharma, T., Shen, D., Roswanjaya, Y. P., Wardhani, T. A. K., Kalhor, M. S., Jansen, J., van den Hoogen, D. J., Güngör, B., Hartog, M., Hontelez, J., Verver, J., … Geurts, R. (2017). *Parallel loss of symbiosis genes in relatives of nitrogen-fixing non-legume* Parasponia. https://doi.org/10.1101/169706

Wang, D., Griffitts, J., Starker, C., Fedorova, E., Limpens, E., Ivanov, S., Bisseling, T., & Long, S. (2010). A nodule-specific protein secretory pathway required for nitrogen-fixing symbiosis. *Science*, *327*(5969), 1126–1129. https://doi.org/10.1126/science.1184096

Werner, G. D. A., Cornwell, W. K., Sprent, J. I., Kattge, J., & Kiers, E. T. (2014). A single evolutionary innovation drives the deep evolution of symbiotic N2-fixation in angiosperms. *Nature Communications*, *5*. https://doi.org/10.1038/ncomms5087

# Appendix A:

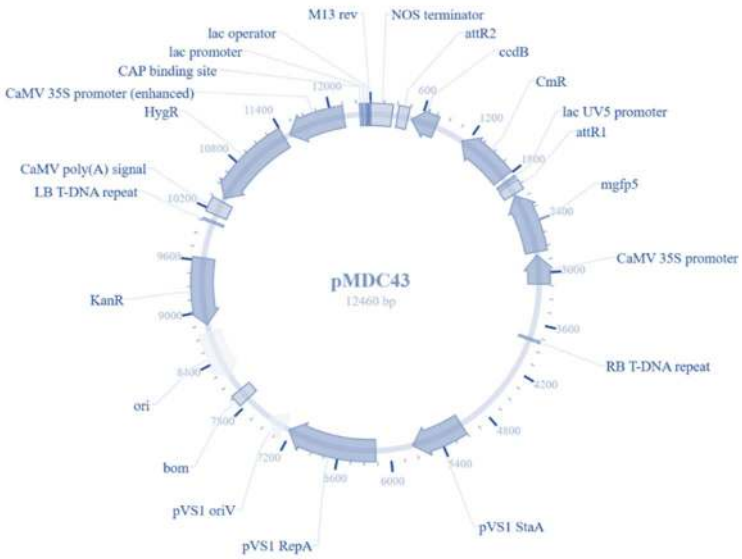



Figure A.1: Plasmid map of pMDC43

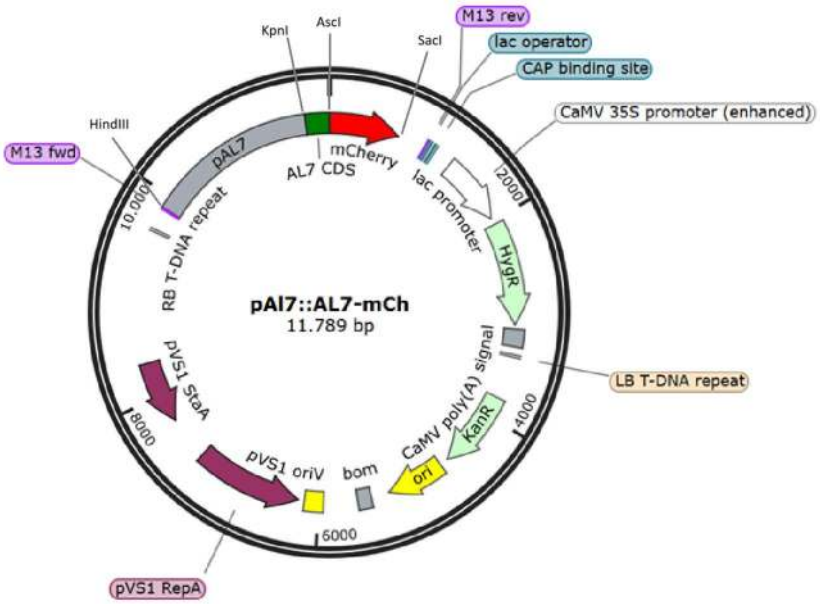



Figure A.2: Plasmid map of pAL7::AL7-mCh.

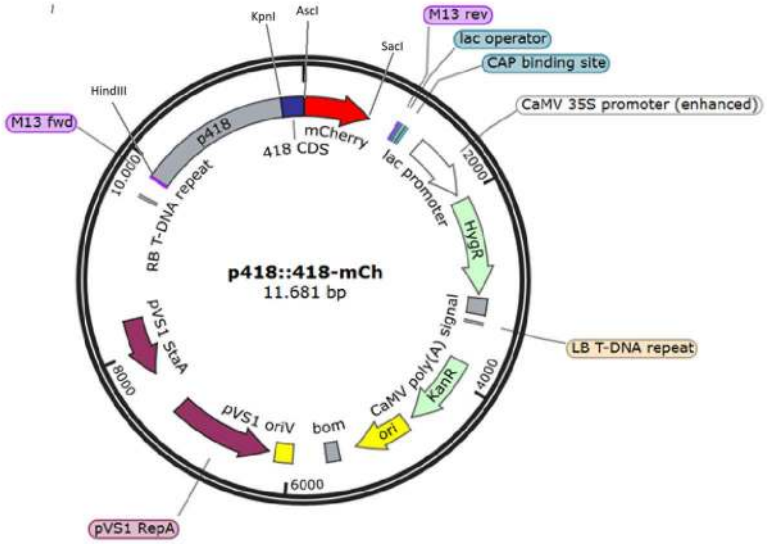



Figure A.3: Plasmid map of p418::418-mCh.

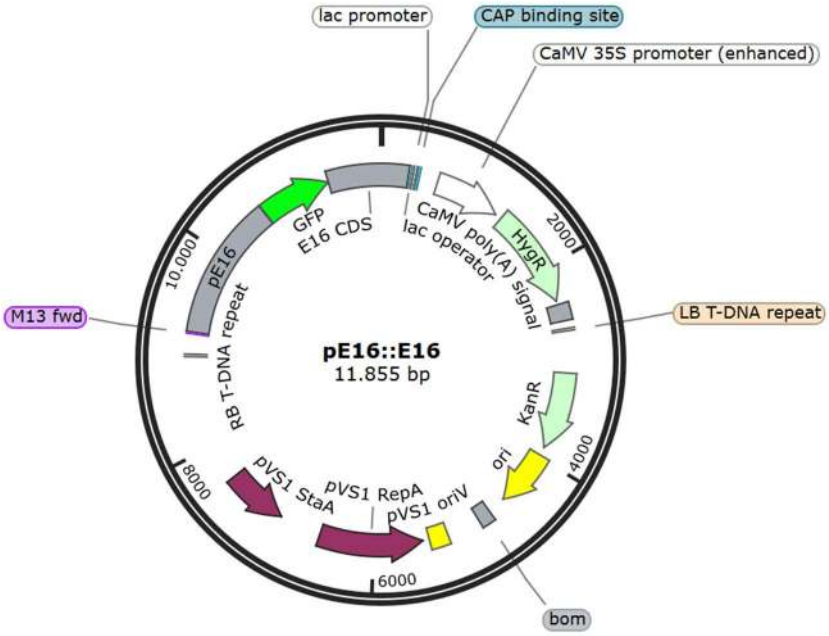



Figure A.4: Plasmid map of pE16::E16

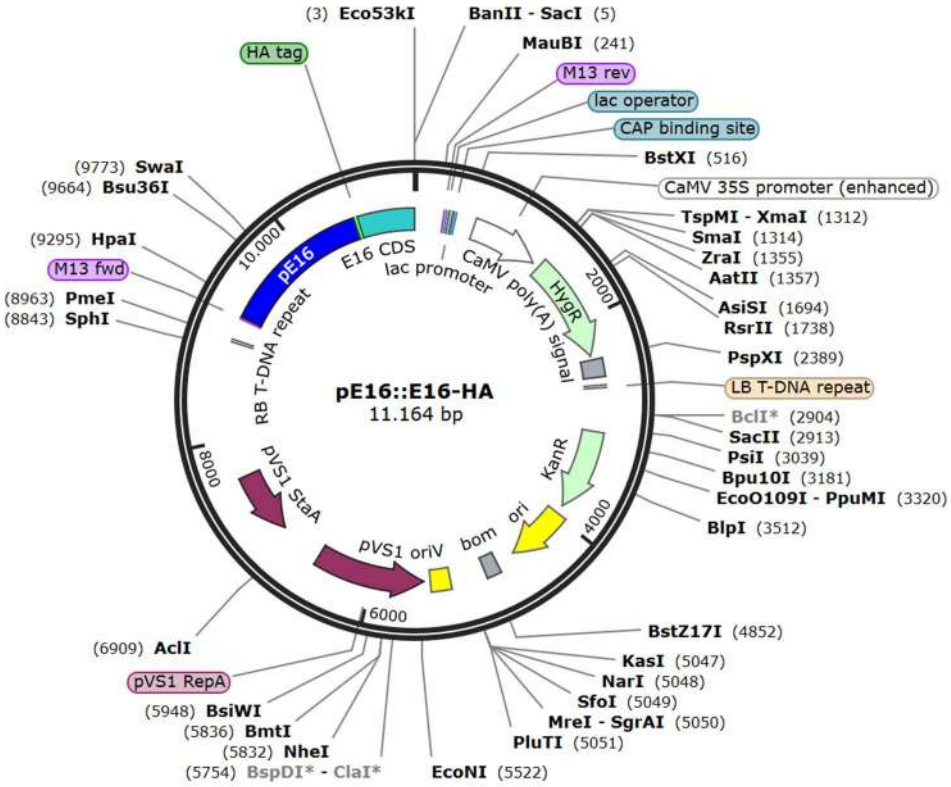



Figure A.5: Plasmid map of pE16::E16-HA

# APPENDIX B:

Table B.1: Primer List

| Primer Name | Sequence | Aim |
|---|---|---|
| Tnt1_F | CATCCAGAGGAGGTAGCACTGTC | Genotyping |
| Tnt1_R | CACAGGTTCTGCTCGTTCACTG | Genotyping |
| RBS_F1 | CGAATCCCCGTTGCTTATGT | Rubisco Large Subunit (Chloroplast marker) |
| RBS_R1 | ACGCATAAATGGTTGGGAGT | Rubisco Large Subunit (Chloroplast marker) |
| COX_F1 | AGGGTCTGTGCATTCCTTCT | COX1 (mitochondrial marker) |
| COX_R1 | TATTGAGGGAAAGCAACGCG | COX1 (mitochondrial marker) |
| AFP1.8_F1 | CGGAATGTGAGCAGGTGTTT | AFP1-8ch (nuclear marker) |
| AFP1.8_R1 | CTCCGCCGTTAAAACCAACA | AFP1-8ch (nuclear marker) |
| NCRG7JAL7p_F2 | tatAAGCTTATTAAGATGTCTTGTTGATGATG | Promoter, HindIII, R108, Genotyping |
| NCRG7JAL7p_R | agatatGGTACCATAGAAAAAAATTATTTTATCGCTTTA | Promoter, KpnI |
| NCRG7JAL7_F2 | atGGTACCATGACTAAAATTGTCAAGTTTATT | CDS, KpnI, R108 |
| NCRG7JAL7_R3 | tGGCGCGCCAAAGAGAATGCATTTGCATGTTT | CDS, AscI, WITHOUT STOP CODON, R108 |
| NCR418p_F2 | tatAAGCTTTTGTTTACTGAATTAAATAAAGAAG | Promoter, HindIII, R108 |
| NCR418p_R2 | atGGTACCATTTTTTCTATATTTGTAAAAAATAGAATATG | Promoter, KpnI, R108 |
| NCR418_F2 | atGGTACCATGGCTGGAAATCTCAACTTT | CDS, KpnI, R108 |
| NCR418_R3 | tGGCGCGCCTTTGTGTACACTATTGGGTG | CDS, AscI, WITHOUT STOP CODON, R108 |
| NCRG7JM81p_R | gaatatGGTACCATTTTTCCTCTTTTGTATGTATTAAG | Promoter, KpnI |
| NCRG7JM81p_F2 | tatAAGCTTTTTTTTTTAAGGATGAATATTGATTATGC | Promoter, HindIII, R108, Genotyping |
| NCRG7JM81_F2 | atGGTACCTGAATCGTATTCTCATGTTTGTTT | CDS, KpnI, R108 |

| | | |
|---|---|---|
| NCRG7JM81_R | gcatatGGCGCGCCCTCAATTTCTTCGACAAGGT | CDS, AscI, WITHOUT STOP CODON |