INVOLVEMENT OF *Azotobacter vinelandii* HYPOTHETICAL PROTEIN Avin_16040 IN *Oryza sativa* ROOT ATTACHMENT

PAULINE LIEW WOAN YING

UNIVERSITI SAINS MALAYSIA

2013

INVOLVEMENT OF *Azotobacter vinelandii* HYPOTHETICAL PROTEIN Avin_16040 IN *Oryza sativa* ROOT ATTACHMENT

by

PAULINE LIEW WOAN YING

Thesis submitted in fulfillment of the requirements for the Degree of Doctor of Philosophy

July 2013

ACKNOWLEDGMENTS

First and foremost, I praise the Lord for bringing about the completion of this work.

To my supervisor Professor Dr Nazalan Najimudin, I am forever grateful for your guidance and continuous encouragement throughout my Ph.D. candidature.

To my co-supervisors Associate Professor Dr Amir Hamzah Ahmad Ghazali and Dr Khairuddin Abdul Rahim, thank you for your helps.

To Dr Jong Bor Chyan, thank you. Without you, this work would not have realized.

To my parents and siblings, thank you for always be there for me.

To Hok Chai, Kamariah, Suhaimi, Balqis, Puan Latiffah, thank you for all your helps.

Special thanks to Padiberas Nasional Berhad (BERNAS) for courteously providing the *Oryza sativa* MR 219 seeds.

Last but not least, I would like to acknowledge the Malaysian Ministry of Science, Technology and Innovation (MOSTI) for granting me the opportunity to pursue my Ph.D. under the HCD LDP program.

Also, I acknowledge the Malaysian Nuclear Agency (Nuclear Malaysia) for supporting my Ph.D. project by providing laboratory space, facilities and research materials.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS, ABBREVIATIONS AND ACRONYMS

Δ	deletion
β	beta
γ	gamma
%	percent
µ	micro
Å	Angstrom
®	copyright
™	trademark
X	times
≥	at least
+ end	positive end
+N	N-enriched
-N	N-free
1DE	single dimensional gel electrophoresis
2DE	two-dimensional gel electrophoresis
A_0	absorbance of the original bacterial culture
A_{10}	absorbance of the cell suspension after 10 min
A_{260}	spectrophotometric absorption at 260 nm wavelength
A_{260}/A_{280}	ratio of spectrophotometric absorption at 260 and 280 nm wavelength
AMS	Altered Murashige and Skoog
AMS+N	N-enriched AMS
AMS-N	N-free AMS
ARA	acetylene reduction assay
ATCC	American Type Culture Collection
BNF	biological nitrogen fixation
bp	base pairs
CFU/mL	colony forming unit per millilitre
BATH	adhesion of bacterial cells to hydrocarbon droplets
BLAST	Basic Local Alignment Search Tool
BSA	bovine serum albumin
cDNA	complementary DNA
cv.	cultivar variety
C_T	RT-PCR threshold value
DNA	deoxyribonucleic acid
DSM	Deutsche Sammlung von Mikroorganismen (German Collection of Microorganisms)
e^-	electron ion
ESI	electrospray ionization
et al.	and others
FTIC	fourier transform ion cyclotron
H	relative hydrophobicity of cells
IEF	isoelectric focusing
k	kilo
bp	base pairs
kDa	kilodalton
Km	kanamycin
L.	Linnaeus

LB	Luria Bertani
LSD	Least Significant Difference
m	milli
MΩ	megaohm
MALDI	matrix assisted laser desorption/ionization
MARDI	Malaysian Agricultural Research and Development Institute
mRNA	messenger ribonucleic acid
MS	mass spectrometry
MS medium	Murashige and Skoog medium
MS/MS	tandem mass spectrometry
n	nano
NCBI	National Center for Biotechnology Information
N_2	dinitrogen gas
OD_{600}	optical density at 600 nm wavelength
OD_{595}	optical density at 595 nm wavelength
ORF	open reading frame
p	probability value
PCR	polymerase chain reaction
pI	isoelectric point
psi	pounds per square inch pressure unit
RFU	relative fluorescence units
RNA	ribonucleic acid
rpm	revolutions per minute
rRNA	ribosomal RNA
RT-PCR	real-time polymerase chain reaction
SD	standard deviation
SDS-PAGE	sodium dodecyl sulphate-polyacrylamide gel electrophoresis
SEM	scanning electron microscopy
SNP	single nucleotide polymorphism
TEM	transmission electron microscopy
TOF	time-of-flight

PENGLIBATAN PROTEIN HIPOTESIS Avin_16040 *Azotobacter vinelandii* DALAM PELEKAPAN KEPADA AKAR *Oryza sativa*

ABSTRAK

Dalam kajian ini, perubahan dinamik ke atas proteom *Azotobacter vinelandii* ATCC 12837 sebagai tindakbalas kepada *Oryza sativa* L. cv. MR 219, varieti beras tempatan, telah diteliti. Analisis 2DE MS/MS mendedahkan beberapa bintik protein yang menunjukkan kehadiran berbeza apabila *A. vinelandii* ATCC 12837 ditumbuh dalam media yang diperkaya nitrogen (+N) atau bebas nitrogen (-N). Persediaan eksperimen yang sama juga digunakan untuk mengkaji profil proteom apabila strain bakteria ini didedah kepada akar. Protein intrasel yang menunjukkan kehadiran berbeza telah dikenalpasti melalui analisis MALDI-TOF/TOF and ini termasuk protein yang terlibat dalam pengangkutan dan penyimpanan ion logam, respirasi, tindak balas terhadap stres, struktur, pengawalaturan, sintesis asid amino, dan pengangkutan elektron. Protein ekstrasel yang menunjukkan perbezaan pula termasuk protein pengangkutan membran, lipoprotein dan protein yang bertindak balas terhadap tegasan oksigen. Selain itu, beberapa protein hipotetikal yang tidak diketahui fungsinya juga menunjukkan perbezaan kehadiran. Beberapa protein bakteria menunjukkan ekspresi teraruh di bawah keadaan kekurangan N. Terdapat juga beberapa protein yang lain yang menunjukkan kehadiran berbeza apabila *A. vinelandii* ATCC 12837 terdedah kepada akar padi berbanding dengan yang tidak terdedah. Protein hipotetikal Avin_16040 yang menunjukkan kehadiran eksklusif apabila *A. vinelandii* ATCC 12837 melakukan pengkolonian permukaan akar telah dipilih untuk analisis yang selanjutnya. Untuk mengkaji fungsi biologi bagi Avin_16040, satu mutan delesi untuk gen *Avin_16040* telah dijana melalui

rekombinasi berhomolog. Mutan Δ*Avin_16040* mempamerkan morfologi koloni terubah, pengurangan pembentukan biofilem, penurunan kehidrofobikan permukaan sel serta penurunan pelekapan pada permukaan akar berbanding dengan jenis liar. Pengklonan dan ekspresi gen *Avin_16040* dalam *Esherichia coli* DH5α menghasilkan sel berfilamen dengan struktur berupa tiub yang lutsinar. Analisis carian penjajaran jujukan dilakukan terhadap protein Avin_16040 menunjukkan bahawa ia berkongsi identiti 39% dengan protein S-lapisan tetragon parahablur *Aeromonas hydrophila*.

INVOLVEMENT OF *Azotobacter vinelandii* HYPOTHETICAL PROTEIN Avin_16040 IN *Oryza sativa* ROOT ATTACHMENT

ABSTRACT

In this study, the dynamic changes of *Azotobacter vinelandii* ATCC 12837's proteome in response to *Oryza sativa* L. cv. MR 219, a local rice variety, was observed. Analysis by 2DE MS/MS revealed various protein spots which showed differential presence when *A. vinelandii* ATCC 12837 was grown in the N-enriched (+N) or N-free (-N) media. Similar experimental setup was also applied to study the profiles when the bacterial strain was exposed to root. The differentially-present intracellular proteins were identified by MALDI-TOF/TOF analysis and these included those involved in metal ion transport and storage, respiration, stress response, structure, regulatory, amino acid synthesis, and electron transport. The identified differentiated extracellular proteins were membrane protein transporters, lipoprotein and oxygen stress response proteins. Besides that, several hypothetical proteins with unknown function were also differentiated. Some of these bacterial proteins demonstrated induced expression under N deficient conditions. There are also a number of other proteins that showed differential presence when *A. vinelandii* ATCC 12837 was exposed to rice roots compared to those unexposed. The hypothetical protein Avin_16040, which showed exclusive presence when *A. vinelandii* ATCC 12837 colonized root surface, was further analyzed. To investigate the biological function of Avin_16040, a deletion mutant of the *Avin_16040* gene was generated by homologous recombination. The Δ*Avin_16040* mutant exhibited altered colony morphology, reduced biofilm formation, decreased cell surface hydrophobicity as well as declined root attachment when compared to the wild type.

Cloning and expression of the *Avin_16040* gene into *Escherichia coli* DH5α produced filamentous cells with transparent tube-like structures. A sequence alignment search analysis performed on the Avin_16040 protein showed that it shared 39% identity with the paracrystalline tetragonal S-layer protein of *Aeromonas hydrophila*.

CHAPTER 1

INTRODUCTION

Azotobacter vinelandii is an obligate aerobe, Gram-negative, free-living bacteria which inhabit the soil environment. The bacterium belongs to the *Gammaproteobacteria* bacterial class. It is one of the bacterial strains able to fix atmospheric dinitrogen molecule in its free-living condition, un-associated with plant or plant tissue. *A. vinelandii* has been a bacterium of interest particularly for its several features including its ability to morphologically differentiate to desiccation-resistant cyst, nitrogen fixation activity and special feature of housing several oxygen-sensitive mechanisms while being an obligate aerobic bacterium. Accredited by these characteristics, *A. vinelandii* has received vast interest from researchers worldwide with a record of exceeding 100 years of research. To date, comprehensive works have been carried out on the bacterium including completion of the whole genome sequencing of *A. vinelandii* DJ (GenBank accession no. NC_012560). Despite these, there is still limited information with regards to the bacterium's interaction with the natural environment and its beneficial relationship with plants, particularly the rhizosphere. Studying the interactive biology of root-bacteria association could generate new knowledge and useful information to enhance the beneficial effects of plant-growth-promoting bacteria to plant and vice versa. This study aims to examine the interaction between *A. vinelandii* and plant rhizosphere in a laboratory setting. This study will encompass the differential protein response of an *A. vinelandii* strain, specifically ATCC 12837, during interaction with *O. sativa* L. cv. MR 219, a national rice variety in Malaysia. The approaches employed in this study

include a proteomic investigation by two-dimensional gel electrophoresis (2DE) and a real-time reverse transcription-polymerase chain reaction (RT-PCR) analysis. The analysis identified a hypothetical protein of *A. vinelandii* ATCC 12837, which showed an upregulated expression when colonizing *O. sativa* MR 219 root. This protein was designated as Avin_16040 in the *A. vinelandii* DJ genome. In order to decipher the role of Avin_16040 in *A. vinelandii* ATCC 12837, the structural gene of Avin_16040 was knocked out from the bacterial genome by homologous recombination. The deletion mutant showed differential phenotypic performances when examined by hydrophobicity test, root attachment assay, biofilm assay and plant growth assay.

In summary, the objectives of this study are:

I. To evaluate the differential response of *A. vinelandii* proteome during plant-microbe interaction with *O. sativa* (test model plant).

II. To perform gene expression analysis on the gene coding for a protein of interest, namely the hypothetical protein Avin_16040, using real-time reverse transcription-polymerase chain reaction (RT-PCR).

III. To elucidate the role of Avin_16040 by constructing a deletion mutant of its gene and characterizing it.

CHAPTER 2

LITERATURE REVIEW

2.1 *Azotobacter vinelandii*

Azotobacter vinelandii is an obligatory aerobic, rod-shape, Gram-negative bacterium that belongs to the γ-Proteobacteria group. The bacterium is well known for its relatively large size which is comparable to the yeast cell and its signature of yellowish green pigmentation (Wilson and Knight, 1952). *A. vinelandii* is one of the six bacterial species in the genus *Azotobacter*. All *Azotobacter* species are known to fix and convert atmospheric N_2 into ammonia, a nitrogen compound readily utilizable by plant as nutrient. *A. vinelandii* exists as free-living bacteria in the soil and are able to perform dinitrogen fixation in its free-living condition. In contrast to the symbiotic *Rhizobium*, *A. vinelandii* and other *Azotobacter* species cannot exist in symbiotic association with plants. However, the bacterium fixes atmospheric nitrogen when grown in the plant root rhizosphere. During its interaction with plant roots, the bacterium provides fixed nitrogen to the plant while acquiring sugars and other nutrients that leak from the roots (Sommers *et al.*, 2004; Gray and Smith, 2005; Bais *et al.*, 2006).

A. vinelandii has adapted several mechanisms to survive environmental stress. One unique feature is its ability to differentiate from vegetative cell to desiccation-resistant cyst under unfavourable environment of carbon source deficiency (Sadoff, 1975). Carbon sources such as n-butanol or β-hydroxybutyrate (Lin and Sadoff, 1969; Page and Sadoff, 1975) also induced *A. vinelandii* to form cyst. During encystment, *A. vinelandii* is not able to perform N_2 fixation. The bacterium usually

reverts to vegetative cell form when it was provided with favourable growth condition containing utilizable carbon source.

In addition to N_2 fixation, *A. vinelandii* biosynthesizes the extracellular polysaccharide alginate, the intracellular polyester poly-β-hydroxybutyrate (PHB) and siderophores compounds that were reported to have multiple biotechnology and biomedical applications (Diaz-Barrera and Soto, 2010). These applications include siderophores as drug delivery (Möllmann *et al.*, 2009), antimicrobial (Upadhyay and Srivastava, 2008) and soil bioremediation agents (Braud *et al.*, 2009), alginate for control release of medical drugs (Yao *et al.*, 2009). Other applications include alginate as food additives (thickener, stabilizer, gelling agent and emulsifier), and polyhydroxybutyrate (PHB) for development of biodegradable and biocompatible thermoplastics (Diaz-Barrera and Soto, 2010).

2.1.1 *Azotobacter vinelandii* Lipman ATCC 12837

Azotobacter vinelandii Lipman ATCC 12837 (also designated as DSM87) is the earliest *A. vinelandii* strain studied (Robson *et al.*, 1984). The studies include physiological interaction with plant in the field through inoculation to plant roots for N_2 fixation, amino acid and vitamins (Rodelas *et al.*, 1999), growth of ATCC 12837 in soil distillate and different media compositions (Gonzalez-Lopez *et al.*, 1983), and production of PHB (Vargas-Garcia *et al.*, 2002). Patents of *A. vinelandii* ATCC 12837 mutants with enhanced PHB production was obtained by Page *et al.* in 1991 (Patent Number 5,059,536). The productions of vitamins and amino acids under different media conditions and with various media supplements have also been studied by Gonzalez-Lopez *et al.* (1986) and Yoneyama *et al.* (2011). Molecular investigation of other characteristic of interest, for instance siderophores (Menhart *et*

al., 1991), was also conducted.

2.1.2 *Azotobacter vinelandii* interaction with plants

Azotobacter vinelandii is a common soil bacterium. The interactions between *A. vinelandii* and plants were reviewed since decades ago. The occurrence of *A. vinelandii* at the rice rhizospheres was well versed for various characteristics, most commonly dinitrogen fixation (Ueda *et al.*, 1995), and the production of indole-3-acetic acid (IAA) (Torres-Rubio *et al.*, 2000). The genus *Azotobacter* was also reported to release growth hormones and antibiotics which improved soil fertility and agriculture crop productivity (Forlain *et al.*, 1995; Kumar *et al.*, 2001; Shafeek *et al.*, 2004).

2.2 *Oryza sativa* L. cv. MR 219

Oryza sativa L. cv. MR 219 is an indica rice variety generated from a cross between the MR 137 and MR 151 by the Malaysian Agricultural Research and Development Institute (MARDI). The rice variety was officially released in 2001. Since then, it has become one of the most popular rice varieties in Malaysia. In general, the rice variety was developed by means of a direct seeding planting system (http://agromedia.mardi.gov.my/magritech/tech_detail_fdcrop.php?id=346, accession date 6th February 2013). It has a maturation period of 105 to 111 days, produced larger grain size and higher yield by 12% when compared to rice varieties MR 84 and MR 220. This rice variety constituted approximately 70% and 40-50% of the total rice granary areas in the year 2002 and 2007, respectively.

2.3 Nitrogen in agriculture

Nitrogen gas constitutes 78.08% by volume of the Earth's atmosphere and it is colourless, odourless, tasteless and most often exists as inert diatomic N_2. Nitrogen is the most important cell-building element in all the living organisms. It is a crucial constituent element of amino acids (proteins) and nucleic acids DNA and RNA.

Nitrogen is often the limiting factor for growth and biomass production when a plant is cultivated under suitable climate and sufficient water supply. It is one of the three primary nutrients required for plant growth, in addition to phosphorus (P) and potassium (K). These three primary nutrients are represented in the fertilizer formulas by the numbers on fertilizer container labels for N:P:K. Other essential elements required by plants include carbon (C), hydrogen (H) and oxygen (O). Secondary nutrients include calcium (Ca), magnesium (Mg) and sulphur (S). Nutrients required in trace amounts include boron (B), chlorine (Cl), copper (Cu), iron (Fe), manganese (Mn), molybdenum (Mo) and zinc (Zn) (http://www.albrightseed.com/nitrogen.htm, accession date 6th February 2013).

Industrially, nitrogen is converted to its plant-accessible fertilizer formats of ammonia and nitrate through the expensive Haber-Bosch process, which requires a lot of energy in terms of pressure and temperature. On the other hand, natural processes exist which convert atmospheric nitrogen to nitrates through lightning storms and nitrogen-fixing bacteria. The nitrates are acquired by plants to construct amino acids, DNA and proteins. The natural process of acquisition and conversion of nitrogen from the atmosphere to nitrates, fertilization of plants, consumption of plants by animals and the returns of nitrogen from plants and animals to the soil and atmosphere is demonstrated in the nitrogen cycle below (Figure 2.1).

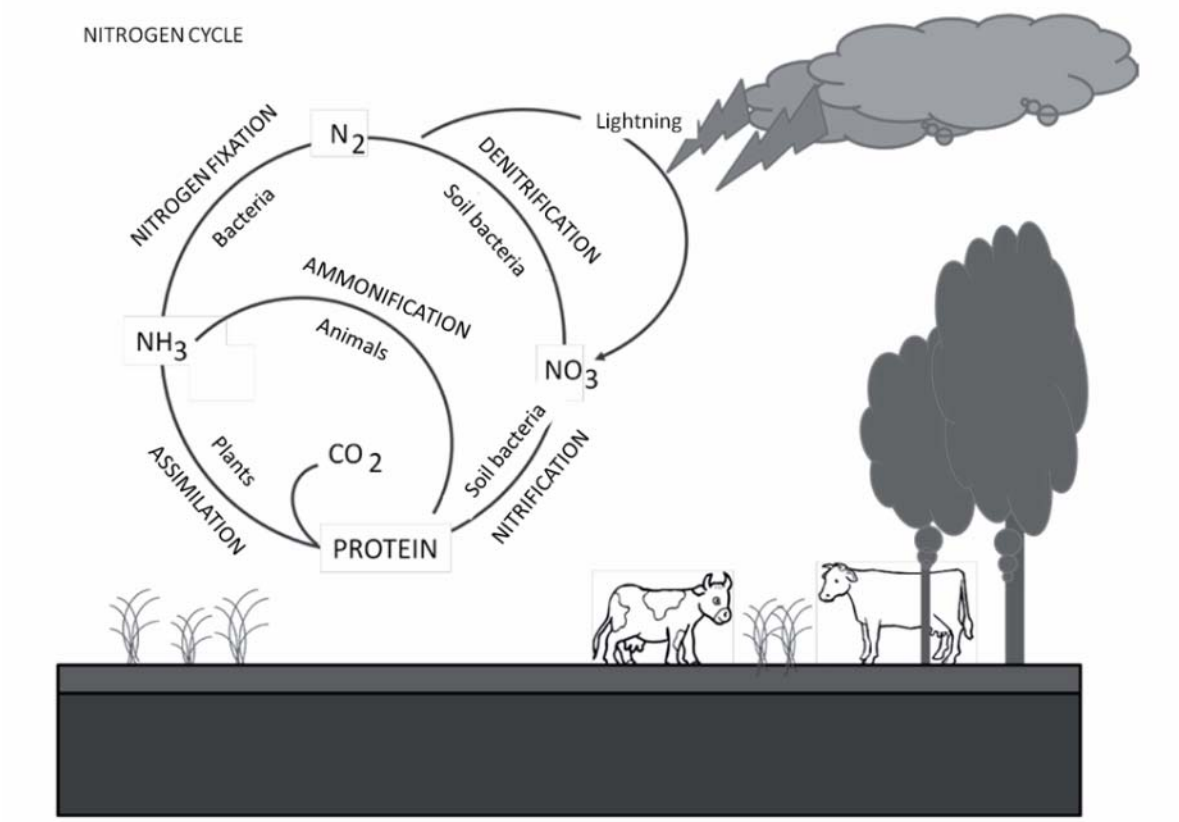

Figure 2.1 The nitrogen cycle (modified from the Environmental Protection Agency, Nitrogen Element Facts, http://www.chemicool.com/elements/nitrogen.html, accession date 6th February 2013).

2.4 Nitrogen fixation

Despite its high composition of more than 78% in the earth atmosphere, nitrogen is not accessible by plant. In its atmospheric form, gas N_2, the two nitrogen atoms are connected by a triple covalent bond which is unbreakable by the higher plants for their uses. Therefore, the labile N_2 needs to be reduced to nitrate or ammonia before it could be assimilated by the plants. The process of reducing the atmospheric N_2 molecules to nitrate and ammonia is called nitrogen fixation or dinitrogen fixation. Generally, there are three commonly known means of nitrogen fixation, as in the following:

i) Spontaneous fixation of atmospheric N_2 by lightning and photochemical reactions which contributes to 10% of naturally fixed nitrogen;

ii) Industrial fixation of N_2 through the Haber-Bosch process mediated by iron catalyst. The process is expensive and involves high energy in terms of high pressure and high temperature; and

iii) Biological fixation of atmospheric N_2 by nitrogen-fixing bacteria, termed diazotrophs. Biological nitrogen fixation (BNF) contributes to about 60% of the total nitrogen fixed by all three processes. It contributes to 90% of the naturally fixed nitrogen.

Among the three nitrogen fixation means, biological N_2 fixation poses the highest potential for prolonged application simply because it is a natural and sustainable process.

2.4.1 Nitrogen-fixing bacteria

Generally, there are two categories of nitrogen-fixing bacteria (diazotrophs), specifically the free-living and symbiotic nitrogen-fixers. Examples of free-living nitrogen-fixers include aerobic bacteria from the genus *Azotobacter*, *Beijerinckia*, *Klebsiella*, and anaerobic bacteria from genus *Desulfovibrio* and *Clostridium*. The symbiotic nitrogen-fixers are the diazotrophs which maintain symbiotic associations with plants, i.e. endophytic in plant cells and colonization of plant (legume) root nodules. Examples of symbiotic nitrogen-fixers include legume symbiont *Rhizobium* and non-legume symbiont *Frankia* and *Azospirillum*.

2.4.2 Mechanism of nitrogen fixation

The process of nitrogen fixation involves conversion of one mole of N_2 molecule to 2 moles of ammonia. The process utilizes 16 moles of ATP and a supply of electrons and protons (hydrogen ions). The interaction is represented by the following equation:

$$N_2 + 8H^+ + 8e^- + 16ATP = 2NH_3 + H_2 + 16ADP + 16Pi$$

A typical nitrogen fixation process is mediated by the nitrogenase enzyme complex which consists of Fe and FeMo proteins (Figure 2.2). Typically, an N_2 molecule is bound by the nitrogenase enzyme complex. The Fe protein is reduced by electrons which were donated by ferredoxin. The reduced Fe protein then binds ATP and reduces the FeMo protein. The process donated electrons for the reduction of N_2 to HN=NH, then $H_2N–NH_2$, and finally $2NH_3$. The ammonia is thus assimilated by plants by first converting it into amino acids glutamine and glutamate (Meeks *et al.*,

1978; Helber *et al.*, 1988), and then into nucleic acids.

The nitrogenase enzyme complex is highly sensitive to oxygen. The enzyme is inactivated upon exposure to oxygen due to the interaction between oxygen and iron in the protein. This characteristic poses a problem to the aerobic nitrogen-fixing bacteria such as the soil bacteria *Azotobacter*, *Beijerinckia*, and the photosynthetic cyanobacteria. However, these bacteria have developed various methods to overcome the problem. For instance, *Azotobacter* acquired high respiratory metabolism rate to eliminate oxygen concentration in the cells. Through the production of extracellular polysaccharide, *Azotobacter* and *Rhizobium* minimized the diffusion of oxygen into the cells. For the symbiotic nodule bacteria such as *Rhizobium*, the root nodules of legume contain oxygen-scavenging molecules, for instance leghaemoglobin, which regulate the oxygen concentration in root nodule tissue.

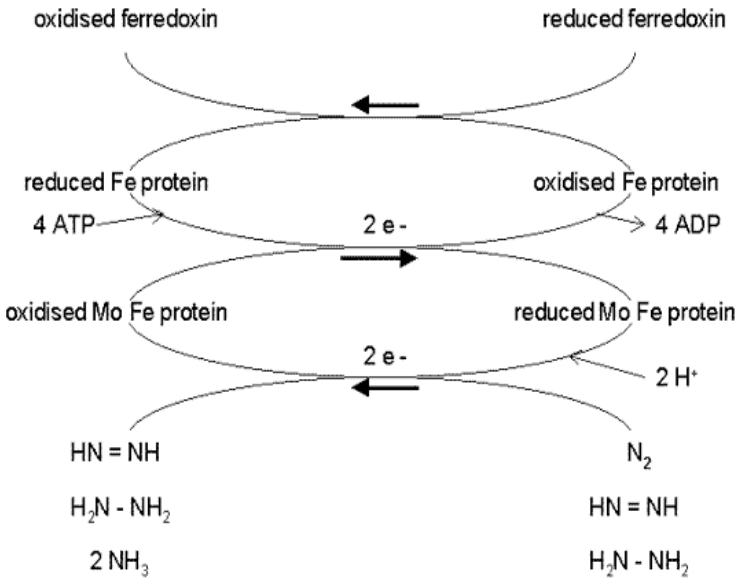

Figure 2.2 A typical nitrogen fixation mechanism (adapted from "The Microbial World: The Nitrogen Cycle and Nitrogen Fixation, by John Deacon", accessed from www.biology.ed.ac.uk/archive/jdeacon/microbes/nitrogen.htm, February 11, 2013).

2.4.3 Nitrogen fixation genes and their affiliates

The diazotrophs produce a special enzyme complex called nitrogenase, which can break the triple covalent bond of N_2 molecule. These nitrogenase enzymes are typically encoded by *nif* genes that are mediated by both the iron (Fe) and iron-molybdenum (FeMo) cofactors. Alternative genes coding for nitrogenase include *vnf* and *anf* genes which are mediated by vanadium and iron cofactors, respectively. There are approximately 20 *nif* genes involved in encoding the nitrogenase complex. The genes are listed in Table 2.1. In the beneficial relationships between the plant and nitrogen-fixing bacteria, the bacteria supply the plant with reduced (fixed) nitrogen while the plant provides the nitrogen-fixing bacteria with protection and source of carbon for energy and hydrogen for reduction of atmospheric nitrogen.

The expression and regulation of *nif* also involved other genes. For instance, the transcription of *nif* genes is activated when there is nitrogen stress, when there is not enough nitrogen supply for the plant's use. Nitrogen stress first triggers the expression of nitrogen sensitive activator protein called NifA, and this process is mediated by NtrC which interacts with the sigma factor RpoN to allow RNA polymerase to express the *nifA* gene. NifA then activates the transcription of the other *nif* genes to produce nitrogenase. Under nitrogen excess, NifL protein is activated. The protein inhibits NifA activity thus terminating the production of nitrogenase (Triplett, 2000).

In the symbiotic N_2 fixation model, the N_2 fixers participate in intimate interaction with plant by residing inside special plant structure namely the root nodules. This intimate relationship between symbiotic N_2 fixers and plant requires a series of genes which were affiliated to the N_2 fixation mechanism. Through proteome and transcriptome analyses, Resendis-Antonio *et al.* (2011) identified 415

Table 2.1 The *nif* genes and symbiosis gene clusters involved in nitrogen fixation.

GENES/ GENE CLUSTERS	IDENTITY/ROLE	SOURCE/ REFERENCE
***nif* genes**		
nifH	Dinitrogenase reductase. Obligate electron donor to dinitrogenase during nitrogenase turnover. Also is required for FeMo-co biosynthesis and apodinitrogenase maturation.	
nifD	α subunit of dinitrogenase. Forms an $\alpha_2\beta_2$ tetramer with β subunit. FeMo-co, the site of substrate reduction. Is present buried within the α subunit of dinitrogenase.	
nifK	β subunit of dinitrogenase. P-clusters are present at the β subunit interface.	
nifT	Unknown.	
nifY	In *K. pneumonia,* aids in the insertion of FeMo-co into apodinitrogenase.	
nifE	Forms $\alpha_2\beta_2$ tetramer with NifN. Required for FeMo-co synthesis. Proposed to function as a scaffold on which FeMo-co is synthesized.	
nifN	Required for FeMo-co synthesis.	Adapted from Triplett (2000); Schübbe *et al.* (2009); Dodsworth and Leigh (2006); Enkh-Amgalan *et al.* (2006)
nifX	Involved in FeMo-co synthesis. Specific role is not known.	
nifU	Involved in mobilization of Fe for Fe-S cluster synthesis and repair.	
nifS	Involved in mobilization of S for Fe-S cluster synthesis and repair.	
nifV	Homocitrate synthase. Involved in FeMo-co synthesis.	
nifW	Involved in stability of dinitrogenase. Proposed to protect dinitrogenase from O_2 inactivation.	
nifZ	Unknown.	
nifM	Required for the maturation of NifH.	
nifF	Flavodoxin. Physiologic electron donor to NifH.	
nifL	Negative regulatory element.	
nifA	Positive regulatory element.	
nifB	Required for FeMo-co synthesis. Metabolic product, NifB-co is the specific Fe and S donor to FeMo-co.	
fdxN	Flavodoxin in *R. capsulatus,* it serves as electron donor to nitrogenase.	
nifQ	Involved in FeMo-co synthesis. Proposed to function in early MO_4^{2-} processing.	
nifJ	Pyruvate:flavodoxin (ferredoxin) oxidoreductase. Involved in electron transport to nitrogenase.	
nifR	A repressor binding site between the promoter of the *nifRLA* operon and *nifL* gene.	
nifI	Posttranslational regulation of nitrogenase, or switch-off.	
Symbiosis genes/ gene clusters involved in N_2 fixation		
nod/ nol/ noe	Synthesis and regulation of the Nod factor (root nodulation) during symbiotic N_2 fixation. The studied operons include *nodABCSUILZnoeCHOP* with *nodD1* located downstream (*Azorhizobium caulinodans*), *nodD1-YABCSUIJ* (*Bradyrhizobium japonicum*), and *nodDABCIJ* (*Rhizobium leguminosarum*, *Sinorhizobium meliloti*).	Reviewed by Black *et al.* (2012)
fix	Involved in the regulation and metabolism of oxygen in symbiotic N_2 fixation (*B. japonicum*). *fixABCX* involved	

Table 2.1. Continued.

	in regulation of gene transcription under low oxygen concentration. *fixGHIS* is required for the initial construction of cytochrome *cbb3* oxidase complex. *fixNOPQ* encodes the cytochrome *cbb3* complex which mediates electron exchange and synthesis of ATP.	
exo	Synthesis of exopolysacchaaride EPS I (*S. meliloti*). Involved in bacterial invasion of root and nodulation.	
pss	Synthesis of exopolysacchaaride EPS I (*R. leguminosarum*). Involved in bacterial invasion of root and nodulation.	
pps	Synthesis of exopolysacchaaride EPS II (*B. japonicum*, *Mesorhizobium* sp.). Involved in bacterial invasion of root and nodulation.	
tat	Transportation of pre-folded proteins to the periplasmic space. Genus specific protein secretion system.	
arp/ hly/ prt	Transportation of targeted proteins across bacterial membrane to the extracellular space. Common Type 1 secretion system.	
gsp	Excreting proteins into extracellular space. Type II secretion system (*Bradyrhizobium*, *M. meliloti*, *Sinorhizobium*).	
sec	Excreting proteins into extracellular space. Common Type II secretion system.	
ysc/ fli/ hrc	Involved in the production of nodulation outer proteins Nops (*B. japonicum*, *Mesorhizobium loti*, *R. etli*, *Sinorhizobium* sp.). Type III secretion system.	
vir/ trb	Involved in virulence and conjugal transfer. Ubiquitous Type IV secretion system. F-type protein family.	
cpa/ tab/ pli	Adaptation from flagella proteins. Type IV secretion system. P-type protein family. Ubiquitous, except in *A. caulinodans*.	
aut	Protein translocating outer membrane porins (*R. leguminosarum*, *Mesorhizobium* sp.). Type V secretion system.	
iscN	Fe-S cofactor nitrogenase synthesis protein (*Rhizobium etli*). Co-transcribed with *nifU* and *nifS*.	Reviewed by Resendis-Antonio *et al.* (2011)
icd	Isocitrate dehydrogenase gene. Influence N_2 fixation (*S. meliloti*).	
pckA	PEP carboxykinase gene. Essential for symbiotic N_2 fixation (*R. etli*).	
fbaB	Bisphosphate aldolase gene. Essential for symbiotic N_2 fixation (*R. etli*).	
idhA/ iolB	Production of myo-inositol protein in root nodules. Influence symbiotic N_2 fixation (*R. etli*).	
purB	Catalyze the biosynthesis of 5-aminoimidazole-4-carboxamide ribonucleotide. Involved in nodule invasion (*Lotus japonicus*). Essential for symbiosis.	
purH	Involved in nodule development (*L. japonicus*). Essential for symbiotic N_2 fixation.	
ndvF	Essential in nodule development and N_2 fixation (*R. meliloti*). Gene locus contains four genes *phoCDET* which encode ABC-type phosphate transport system.	Bardin *et al.* (1996)

proteins and 689 upregulated genes in the symbiosis of *Rhizobium etli* and *Phaseolus vulgaris* (bean plant). The gene clusters involved in the construction of the plant-microbe symbiosis machineries such as root nodules (*nod/ nol/ noe* operons), as well as regulation of oxygen metabolism (*fix operons*), bacterial invasion of plant root (*exo*, *pss*), and bacterial protein secretion systems (Type 1 to VI) (Black *et al.*, 2012). These gene clusters were located on symbiotic plasmids, or existed as laterally transferrable genomic (symbiotic) islands. Some of the symbiosis gene clusters are also listed in Table 2.1.

2.5 Nitrogen cycling by *Azotobacter vinelandii*

Nitrogen cycling by *A. vinelandii* involved complex processes of nitrogen fixation and nitrate assimilation. In general, the bacterium's nitrogen fixation mechanism was as described previously (section 2.4.2). The mechanism was catalyzed by the nitrogenase complex. Nitrate assimilation involved both nitrate and nitrite reduction. Nitrate was reduced to nitrite by nitrate reductase and nitrite was further reduced by nitrite reductase to ammonium, which serves as nitrogen source for the metabolism of *A. vinelandii* (Payne, 1973). Both nitrogen fixation and nitrate assimilation mechanisms were repressible by nitrate and ammonia (Sorger, 1969). Whilst nitrogen fixation was regulated by the *nif* operon, nitrate assimilatory pathway was regulated by the *nas* operon (Ramos *et al.*, 1993; Wang *et al.*, 2012). In another study by Luque *et al.* (1987), mutant defective of *ntrC* lost nitrate and nitrite reductase simultaneously.

2.6 Genome sequencing of *Azotobacter vinelandii*

The most advanced genomic achievement for *A. vinelandii* is the completion of *A. vinelandii* DJ genome sequencing project (Setubal *et al.*, 1999). Annotation of the genome of *A. vinelandii* DJ revealed massive information on genes that were already known as well as hypothetical ones. The information opens up new avenues to better understand and unveil the biochemical pathways and structures of this bacterium, which has received much interest since 1930s (Lineweaver, 1938) or earlier. The full genome sequence of the bacterium revealed a single circular genome of 5,365,318 bp in size with 65.7% GC content (Setubal *et al.*, 1999). With the completion of the *A. vinelandii* DJ genome sequence, its mechanism of protecting the oxygen-sensitive processes and proteins by high respiration rate was better defined (Setubal *et al.*, 1999). The oxygen-sensitive components include nitrogen fixation proteins, carbon-monoxide dehydrogenase and formate dehydrogenase. In addition, the regulation and production of alginate, a polymer, were elaborated to greater depth. Although much has been achieved through the genome annotation works, there is still a portion of hypothetical genes with unknown functions.

2.7 Post-genomics of *Azotobacter vinelandii*

Advancing from genomics, proteomics is becoming a key tool in systems biology because it provides quantitative and structural information about proteins, which are the major functional determinants of cells (Baginsky *et al.*, 2010). The proteomic technology of two-dimensional gel electrophoresis (2DE) enabled analysis of a bacterium's protein complexes to reveal its global protein expression pattern. By subjecting a bacterium of interest to several different conditions, the global changes in protein expression patterns within the bacterial system can be analyzed. Since

transcriptomic (mRNA) information is not always able to provide true reflection of the adaptation of a microorganism towards their changing environment, proteomics is an invaluable tool to identify the functions or connectivity of specific gene products (Han and Lee, 2006; Hecker *et al.*, 2008; Bumann, 2010; Curreem *et al.*, 2012). The versatility of this differential 2DE approach makes it a powerful means in protein and gene discovery.

The combination of proteomic approach with the full genome information of a bacterium of interest has increased understanding and introduced new insights of many important proteins and their related mechanism. One such example is molybdenum trafficking in the *A. vinelandii* nitrogen fixation mechanism mediated by molybdenum nitrogenase (Hernandez *et al.*, 2009). The genes and proteins involved in molybdenum uptake, homeostasis, storage, regulation and nitrogen cofactor biosynthesis were reviewed. Investigation of molybdenum biochemistry in the bacterium revealed new mechanisms and a novel role for iron-sulfur clusters in the sequestration and delivery of molybdenum (Hernandez *et al.* 2009).

Another example is the regulation of poly-β-hydroxybutyrate (PHB) by an iron-regulatory small RNA ArrF. ArrF is under the negative control of ferric uptake regulatory protein. Deletion of *arrF* gene from the genome of *A. vinelandii* caused overproduction of poly-β-hydroxybutyrate (PHB) (Pyla *et al.*, 2009; Pyla *et al.*, 2010). Gel-based proteomic and real-time RT-PCR analysis revealed a list of other proteins that were affected by *arrF* deletion (Pyla *et al.*, 2010). The proteins were found to express differentially upon the deletion of *arrF* gene. Further investigation of the involvement of these proteins to ArrF may increase understanding of the protein network.

Lery *et al.* (2010) used the genomic information of *A. vinelandii* and

Gluconacetobacter diazotrophicus to produce protein structure information of nitrogenase conformational protection through FeSII-nitrogenase interactions. The research group combined approaches of bioinformatics analysis, comparative protein modeling, protein docking and molecular dynamics to elucidate the molecular mechanisms and structural features of FeSII-nitrogenase interaction.

2.8 Two-dimensional gel electrophoresis

Two-dimensional gel electrophoresis (2DE) is a gel-based protein separation technique. 2DE was developed in the 1970s for large-scale protein separation (Klose, 1975; O'Farrell, 1975). This technique has hence remained as one of the preferred methods to separate crude protein samples in order to detect differentiation in levels of abundance as well as patterns. The electrophoresis method has the capacity to resolve thousands of protein in one electrophoresis trial. In general, the technique of 2DE involves two electrophoresis steps. The first step or 1^{st} dimension electrophoresis involves separation of proteins according to their charges by isoelectric focusing (IEF), while the second step or 2^{nd} dimension electrophoresis involves separation of the isoelectric-focused proteins according to their molecular weight by the sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) (Gorg *et al.*, 2004). The method hence produced information on molecular weight and isoelectric point (pI) of each protein, as well as the quantity of proteins in a protein complex (Gorg *et al.*, 2004; Wittmann-Liebold *et al.*, 2006).

The 2DE protein spots can be visualized through staining with Coomassie Blue, silver stain or SYPRO (ruby, red, orange) (Westermeier and Marouga, 2005; Miller *et al.*, 2006). Coomassie blue and silver stains are the more common staining methods. Silver stain is the more sensitive staining method able to detect as low as

0.1 ng of protein. However, it has the potential to interfere with subsequent mass spectrometry (MS) analysis. In addition, silver stain does not produce linear relationship between spot intensity and protein expression for accurate estimation of spot volumes in relation to protein expression. In contrast, Coomassie blue stain produces linear relationship between the protein spot intensity and its expression level, thus more frequently employed to stain and visualize the 2DE protein spots (Gorg *et al.*, 2004; Wittmann-Liebold *et al.*, 2006). 2DE protein spot of interest which shows differentiation of up- or down-regulation is excised from the 2DE gel, digested with trypsin and analyzed by mass spectrometry to determine its identity (Rose *et al.*, 2004).

2.9 Peptide analysis by mass spectrometry

Mass spectrometry (MS) is employed for protein identification (Lambert *et al.*, 2005; Domon and Aebersold, 2006). Protein spot is first excised from the 2DE gel and digested with trypsin. The resulting peptides are then analyzed by MS.

The mass spectrometer consists of an ion source, mass analyzer and ion detection system. Analysis of proteins by MS occurs in three major steps, specifically protein ionization and generation of gas-phase ions, separation of ions according to their mass to charge ratio and the detection of ions (Mann *et al.*, 2001). There are two main ionization sources and four major mass analyzers used for protein identification and characterization (Mann *et al.*, 2001). The ionization sources are the matrix assisted laser desorption/ionization (MALDI) and electrospray ionization (ESI). The mass analyzers include time-of-flight (TOF), ion trap, quadrupole and fourier transform ion cyclotron (FTIC). The ionization source can be combined with different types of mass analyzer depending on the specific application, therefore

provides a large variety of specialized mass spectrometers for protein identification and characterization (Domon and Aebersold, 2006). Simple mass spectrometers such as MALDI-TOF are used only for measurement of mass, whereas tandem mass spectrometers are used for amino acid sequence determination (Dubey and Grover, 2001; Mann *et al.*, 2001). MALDI measures the mass of peptides derived from a trypsinized protein thus generates a list of experimental peptide masses, often referred to as peptide mass fingerprints (Karas and Hillenkamp, 1988; Medzihradszky *et al.*, 2000). These peptide mass fingerprints are then correlated with the peptide fingerprints of known proteins in the protein database using search engines such as Mascot and Sequest.

2.10 Real-time polymerase chain reaction

Real-time polymerase chain reaction (RT-PCR) is an advanced technology of the conventional polymerase chain reaction (PCR). The technology enables continuous monitoring of the amplicon during the course of a RT-PCR reaction while conventional PCR enables only end-point analysis. The continuous monitoring of RT-PCR amplicon is mediated by fluorescence dyes or probes added into the RT-PCR mixture. During RT-PCR reaction, fluorescence signal is emitted, captured and correlated in proportion to the amount of RT-PCR amplicon generated. A threshold value of the amplification cycle is registered, which indicates the generation of a specific amount of DNA amplicon (Kubista *et al.*, 2006). By assuming a certain amplification efficiency that is almost the doubling of the number of DNA amplicon per amplification cycle, the initial amount of DNA molecules in the reaction can be calculated.

RT-PCR technology has been applied in various research disciplines, most

commonly the medical science. Later, the technology was extended to food, industrial, agriculture and microbial studies. Common applications include single nucleotide polymorphism (SNP) analysis, pathogen detection, gene expression analysis and analysis of chromosome aberrations. Although the advancement in bioscience has generated comprehensive genomic information (i.e. through the genome sequencing projects), a major portion of the genome was still undefined. The unknown gene or protein was referred to as hypothetical gene/ protein. The still unexplored information could represent significant role in the functional gene/ protein networks. RT-PCR technology has become a popular means for detecting the dynamics of gene expression in plant-microbe associations (Deepak *et al.*, 2007) due to its high sensitivity and specificity.

2.11 Root-microbe interaction

In the context of soil, rhizosphere is the layer of soil that is in contact with plant root. Therefore, the bacteria that reside in the rhizosphere are defined as the rhizosphere bacteria. Due to their close proximity, these bacteria pose direct effects to the plant. The rhizosphere bacteria can improve the uptake of nutrient by plant and produce plant-growth-promoting compounds. In a natural root ecosystem of mixed microorganism, the rhizosphere bacteria protect plant root surface from colonization by pathogenic microbes through direct competitive effects and production of antimicrobial agents (Maunuksela, 2011).

The provision of atmospheric ammonium-N to plant is one of the direct plant-growth-promoting effects of rhizosphere bacteria (Glick, 1995). The bacteria could also promote plant growth directly through the production of plant hormones such as

auxins (Tien *et al.*, 1979), gibberellins (Gutiérrez-Mañero *et al.*, 2001) and ethylene (Lynch and Whipps, 1990).

Upon detecting an approaching microorganism in the root zone, the plant root secretes signal molecules such as flavonoids and flavones for protection against the invasion (Parmar and Dufresne, 2011). Therefore, it is important for a plant-growth-promoting bacterium to be rhizospheric competent (Karuppiah and Rajaram, 2011) before it can exert its beneficial effect to the plant effectively. The bacterium must be able to colonize and multiply in the plant rhizosphere (Karuppiah and Rajaram, 2011; Maunuksela, 2011). According to Benizri *et al.* (2001), bacterial colonization and survival at the root rhizosphere were influenced by various biotic and abiotic factors.

A classic root-microbe interaction is the mutualistic symbiosis between the rhizobia and legumes (Salavati *et al.*, 2013). The root nodules are their symbiosis machinery mediated by the Nod factors (Perret *et al.*, 2000). In this interaction model, the rhizobia reside in the root nodules. The root nodules provide an anaerobic environment for the rhizobia to fix N_2 which was then supplied to the hosting plant. In exchange, the host legume plant provides the rhizobia with carbohydrates. Through bi-direction quorum-sensing (signal transductions) mechanisms between the rhizobia and legumes, the rhizobia were first attracted to the legume rhizosphere by the plant-secreted flavovoids (Schlaman *et al.*, 1998; Broughton *et al.*, 2000; Perret *et al.*, 2000). The bacteria then proceeded to infect the root hairs (Perret *et al.*, 2000) before penetrate the root hair tissues (Broughton *et al.*, 2000; Geurts and Bisseling, 2002). Provoked by the bacterial intrusion, the root cortical cells differentiated to form nodule (Gage, 2004; Oldroyd and Downie, 2004).

Although unable to colonize root tissue like the symbiotic root-nodulating bacteria, some free-living bacteria are able to form looser interaction with root. These

bacteria adhere to the surface of root in an associative relationship. Several nitrogen-fixing root-associative bacteria have been identified, including *Azospirillum*, *Enterobacter*, *Klebsiella* and *Pseudomonas* (Haahtela *et al.*, 1988). *Azospirillum* was also found to produce auxin in root rhizosphere, stimulate rooting and enhance plant growth (Bloemberg and Lugtenberg, 2001). Another root-associative bacterium, *Acetobacter diazotrophicus*, provides up to 80% of the nitrogen required by sugarcane through biological nitrogen fixation (Boddey *et al.*, 1991).

In this study, comparative profiling of the global proteome of *A. vinelandii* ATCC 12837 revealed a hypothetical protein Avin_16040 which is upregulated during root adhesion (attachment to root surface). The behaviour of the protein indicated its potential importance based on the fact that bacterial colonization (adhesion) of plant root is a pre-requisite for an effective root-microbe interaction. Ultimately, a deletion mutant devoid of the gene sequence of Avin_16040 was generated and analyzed.

2.12 Bacterial S-layer protein

The bacterial surface layers (S-layers) are glycoproteins which were present as the outermost structures on the cell envelope of many Gram-positive and Gram-negative bacteria. They can exist as uniform nanolattices of oblique (p1, p2), square (p4), or hexagonal (p3, p6) symmetry (Sára and Sleytr, 2000; Sleytr *et al.*, 2007). The Gram-negative bacterial strains such as *A. vinelandii* (Bingle *et al.*, 1984; 1986; 1987a; 1987b), *Aeromonas salmonicida* (Kay *et al.*, 1981; 1984; Stewart *et al.*, 1986), *Aeromonas hydrophila* (Al-Karadaghi *et al.*, 1988; Dooley *et al.*, 1988; 1989; Dooley and Trust, 1988; Murray *et al.*, 1988), *Comamonas acidovorans* (Chalcroft *et al.*, 1986; Gerbi-Reiger *et al.*, 1988) and *Pseudomonas* (Austin *et al.*, 1990) were

reported to produce S-layers of tetragonal (p4) symmetry.

Although no precise function yet has been attributed to the S-layers, these bacterial structures were speculated to possibly protect the prokaryotic bacteria from dessication and environmental stress (Ristl *et al.*, 2011). Besides, the S-layers were found to contribute to the surface charge of bacterial cell, facilitate the adherence of bacterial cells to external biotic/ abiotic surfaces, influence the formation of biofilm, and involved in bacterial pathogenesis (Schneitz *et al.*, 1993; Lee *et al.*, 2006; Fletcher *et al.*, 2007; Ristl *et al.*, 2011). In the symbiotic legume-rhizobia interaction, the bacterial surface proteins were found to be involved in rhizobia attachment to root hairs during the initial step of the symbiosis (Peters and Verma, 1990; Deakin and Broughton, 2009).

CHAPTER 3

COMPARATIVE PROTEOMICS OF *Azotobacter vinelandii* ATCC 12837 IN ASSOCIATION AND NON-ASSOCIATION WITH *Oryza sativa* MR 219 ROOT UNDER NITROGEN-ENRICHED AND NITROGEN-FREE ENVIRONMENTS

3.1 Introduction

Proteomics has become an indispensable approach for large-scale protein analysis in functional genomics (Han and Lee, 2006). Comparative proteomics using 2DE/MS-based approach has for the recent decades been widely applied to understand the complex biological systems of human, animals, eukaryotes and prokaryotes. It has been proven useful in the search for novel proteins and novel protein-coding genes (Kovarova *et al.*, 2002; Dai *et al.*, 2006; Ummanni *et al.*, 2011). Over the years, the comparative proteomics approach has revealed the workings of various biological mechanisms, uncovered novel proteins with beneficial or detrimental effects. The discoveries from the comparative proteomic attempts brought about research works to solve relevant biological issues in addition to providing new insights and understanding of the biological relationships.

In accordance with the long history of applying proteomics in the medical investigations, adaptation of comparative proteomic methodologies (i.e. 2DE/MS and 2DE MS/MS) to study the plant-microbe interaction has become popular in recent years. These studies include the plant-microbe symbiosis (Natera *et al.*, 2000; Wienkoop and Saalbach, 2003; Bestel-Corre *et al.*, 2004; Hauberg-Lotte *et al.*, 2012), pathogenesis (Kav *et al.*, 2007) and heavy metal remediation (Farinati *et al.*,