THE IMPACT OF PLANT-PRODUCED S-METHYL-METHIONINE ON NUTRITION AND QUORUM SENSING BY *AGROBACTERIUM TUMEFACIENS*

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Ana Lidia Flores-Mireles

January 2012

© 2012 Ana Lidia Flores-Mireles

THE IMPACT OF PLANT-PRODUCED S-METHYL-METHIONINE ON NUTRITION AND QUORUM SENSING BY *AGROBACTERIUM TUMEFACIENS*

Ana Lidia Flores-Mireles, Ph. D.

Cornell University 2012

Agrobacterium tumefaciens infects many higher plants, causing crown gall tumors by transferring and incorporating a DNA fragment (T-DNA) from the tumor-inducible (Ti) plasmid into the chromosome of plant cells. The incorporated T-DNA expresses genes for phytohormone synthases (induces tumor's plant cell proliferation) and opine synthases (causing the production of bacterial nutrients called opines). Opines are utilized by the bacteria during host colonization. Opines are taken up by the bacteria, and induce the transcription of the octopine catabolism operon for opine uptake and catabolism. This operon contains a gene, *msh*, predicted to direct the conversion of S-methylmethionine (SMM, an abundant plant metabolite) and homocysteine into two molecules of methionine. Purified Msh carried out this reaction, verified by ESI-MS/MS, suggesting that SMM is an intermediate in opine catabolism. Purified octopine synthase (Ocs, normally expressed in plant tumors) reductively condensed SMM with pyruvate to produce a novel opine, designated sulfonopine, whose catabolism by the bacteria would generate SMM. Ocs also utilized a large variety of other amino acids as substrates. Tobacco and *Arabidopsis thaliana* plants produced sulfonopine only when colonized by *A. tumefaciens*. Sulfonopine production was corroborated by ESI-MS/MS and NMR. In liquid growth assays, sulfonopine was utilized as sole source of sulfur by *A. tumefaciens* and induced the transcription of the *occ* operon via OccR. This is the first report of an

opine that contains sulfur, an essential element. Moreover, when sulfonopine is catabolized, generating SMM and pyruvate inside of the bacterial cell, Msh uses SMM to synthesize methionine. The methionine thus generated downregulates *msh* together with downstream genes in the *occ* operon, including *traR*. TraR is a master transcriptional regulator, therefore when methionine downregulates its biosynthetic gene, it also represses TraR-dependent regulation of quorum sensing, conjugation and replication of the Ti plasmid of the colonizing bacteria. These data demonstrates that SMM is not only a nutritional source but also serves as a mediator of gene regulation, integrating the nutritional status of the cell with gene expression.

BIOGRAPHICAL SKETCH

Ana Lidia Flores-Mireles was born at 7:00 am on August 26th 1980 in "The Athens of Jalisco", the city of Lagos de Moreno, Jalisco, located in Western México. Early in her life, Ana Lidia found herself profoundly amazed by natural science and chemistry, and often she got in trouble for trying to mix perfumes to improve the fragrance or by doing a lot of experiments to attempt to understand everything around her. Fortunately, she got really supportive parents and a brother that followed her in all the adventures. At the age of 8 she decided that she wanted to become a Marine Biologist and nobody could make her change her mind. Therefore at the age of 17, she moved to La Paz, Baja California Sur, to conquer her most longed dream to be a Marine Biologist at Autonomic University of Baja California Sur (UABCS). She was part of many labs ranging from Microalgae to Marine Mammals. The last two years of her Bachelor degree she actively worked in understand the health of gray whales populations by studying the Major Histocompatibility Complex (MHC) at Dr. Sergio Flores' Laboratory. At that moment she realized that molecular microbiology is essential for genetic manipulation, promptly her to take Dr. Gina Holguin class on Molecular Microbiology. From that moment until Dr. Gina Holguin's last day on Earth. Ana Lidia made a wonderful relationship based on science, but more importantly Dr. Gina Holguin was like a second mother for her. After Ana Lidia got her Bachelor Degree with honors, she was awarded a fellowship from the Mexican Government (CONACyT) to continue towards a Master Degree in the lab of Dr. Gina Holguin at the Northwest Center of Biological Research. The last year of her master degree she came to the lab of Dr. Steve Winans at Cornell University to perform some experiments. She then went back to Mexico to defend her Master thesis on Microbiology of Mangroves. She fall in love with Cornell University and decided to apply to the Department of Microbiology Ph.D. program to continue working with Dr. Winans.

DEDICATORIA

A mis papas Isidro y Lidia, y hermano Josué, las personas más importantes de mi vida, y al resto de mi familia que nunca dejaron de creer en mí. En especial a la persona que me ha apoyado incondicionalmente estos últimos 5 años, y que siempre me ha brindado una sonrisa... Felipe

DEDICATION

To my parents Isidro and Lidia and brother Josué, the most important people in my life, and to the rest of my family that has always believed on me. In special, to the new most important person that has been with me the last 5 years, unconditionally supporting me and always receiving me with a smile… Felipe.

ACKNOWLEDGMENTS

I would like to deeply thank my advisor Dr. Steve Winans for opening the doors of his lab and for his continuous guidance and support ever since I arrived the first time to Cornell during my Master Degree. Although been busy, he always found time to talk to me. He is an enormous source of information, both technical and historical and he has a brilliant mind that remembers every experiment that has been done in the lab. To my committee members, Linda Nicholson and Georg Jander, for their support and suggestions during my PhD. I really enjoyed the conversations with Linda and her class. I will always remember the first time that I met Georg, I got really nervous because he was dressed as a grim ripper for Halloween, however, he was kind and helpful. Especially I would like to thank Dr. Anatol Eberhard for being my chemistry mentor, my friend and for never giving up on me and my work. Thanks to Dr. Carolyn Eberhard for editing my manuscript and for the wonderful friendship. I would also like to thank Esther Angert for being a wonderful person to work with and a nice friend that always has time to talk with me.

I would like to thank the past and current members of the Winans' lab because they are not only wonderful labmates, they are my lab-family, particularly Esther, Uelinton, Gina, Sonny, Nydia, Cathy White and David Kim. I'm grateful to be part of the Department of Microbiology, a nice place to do research and make friends. I want to thank all of the laboratories of this Department for all the help and useful conversations. Especially, to the Helmman lab for sharing the floor, reagents, knowledge, and Ahmed Gaballa, who has been infinitely helpful about technical advices and moral support. My gratitude also goes to members of the Jander and Schroeder labs for helping me with ESI-MS/MS. I would also like to thank Cathy Shappell, Shirley Cramer and Patty Buttler for their continuous support and help, making my life easier. Without them, this work would have not succeeded.

A trillion thanks to all my friends, Veronica, Luis, Yun, Alexis, Jeisa, Giomar,

Abner, Pilar, German, Eli, Juan Camilo, Nelson, la Andrea, Franscico, Andrea, Victor, Karen, Henrique, Gabriela, Daniel, Anthony, Los Putis, Cian, Sophia, Colman and all of the ones that I don't remember right now.

Lastly but not least to my supportive family which I love with all my heart and especially to the great gift that life gave me…. Felipe Santiago.

TABLE OF CONTENTS

ABSTRACT ..... -
BIOGRAPHICAL SKETCH iii
DEDICATION iv
ACKNOWLEDGMENTS v
TABLE OF CONTENTS vii
LIST OF FIGURES xi
LIST OF TABLES xiii
CHAPTER 1 1
1.1 Soil-dwelling bacteria 1
1.2 To Colonize or to disappear: That is the question 4
1.3 The good, the bad and the rhizosphere 5
1.4. *Agrobacterium tumefaciens* mechanism of infection 8
1.4.1 Recognition of the host 12
1.4.2 *Agrobacterium* attachment to the host plants 13
1.4.3 Transport and integration of T-DNA into the plant cell chromosome 14
1.4.4 Function of the oncogenes in the host plant 16
1.4.5 Opine synthesis 16
1.5 Gene regulation in Agrobacterium tumefaciens by opines 20
1.6 Contents of the dissertation 23
1.7 References 24
CHAPTER 2 30
2.1 Abstract 30
2.2 Introduction 30
2.3 Materials and methods 33
DNA manipulation 33
Plasmid construction 34
Chemical synthesis of SMM 35
Overproduction and purification of Msh 35

Msh Activity Assays **36**

Electrospray mass-spectroscopy and Nuclear Magnetic Resonance **37**

2.4 Results **37**

2.5 Discussion **45**

2.6 References **48**

CHAPTER 3 ..... **51**

3.1 Abstract **51**

3.2 Introduction **52**

3.3 Materials and methods **56**

Bacterial strains, plasmids, and oligonucleotides **56**

DNA manipulation **56**

Plasmid construction **57**

Chemical synthesis of SMM **58**

Chemical synthesis of sulfonopine **59**

Mutagenesis of *ocs* and *virD4* by Campbell-type integration **60**

Overproduction and purification of Ocs **61**

Ocs Activity Assays **62**

Enzymatic synthesis of opines **62**

Plant cultivation and inoculation **63**

Electrospray mass-spectroscopy and Nuclear Magnetic Resonance **64**

Utilization of sulfonopine and other sulfur source compounds **64**

Induction of the occ operon by opines **65**

3.4 Results **65**

Octopine synthase utilizes SMM and pyruvate as substrates **65**

Chemical synthesis of sulfonopine **70**

Homogenates and exudates of plants colonized by A. tumefaciens contain

sulfonopine ... **73**

Production of sulfonopine requires host-synthesized SMM ... **75**

Sulfonopine induces transcription of the octopine catabolism operon ... **78**

***A. tumefaciens* can utilize sulfonopine as sole source of sulfur** ... **83**

3.5 Discussion ... **85**

3.6 References ... **88**

3.7 Appendix ... **91**

β-galactosidase and Gus activity ... **107**

Analysis of RNA upstream of *msh in vitro* ... **107**

Synthesis of 1-methyl-7-nitroisatoric anhydride ... **109**

4.4 Results ... **110**

Effect of methionine on the expression of the *occ* operon ... **110**

The intergenic region is enough for the recognition of methionine ... **116**

Methionine interacts directly with the mRNA of the intergenic sequence ... **118**

CHAPTER 5 ... **132**

5.1 Conclusions ... **132**

5.2 Future work directions ... **137**

Appendix 1 ... **139**

A1.1 Abstract ... **139**

A1.2 Introduction ... **140**

A1.3 Results ... **145**

Identification of new CepR regulated promoters. ... **145**

A1.4 Discussion ... **152**

A1.5 Experimental Procedures ... **155**

Bacterial strains and growth conditions ... **155**

Identification of six new CepR regulated promoters. ... **155**

A1.5 References ... **160**

Appendix 2 ... **165**

A2.1 Abstract **165**
A2.2 Introduction **166**
A2.3 Results **171**
RepC overexpression causes strong increases in plasmid copy number. **172**
A2.4 Discussion **177**
A2.4 Experimental Procedures **177**
Bacterial strains and plasmids **177**
DNA manipulations **177**
Construction of flag-tagged RepC **178**
Immunodetection of RepC **178**
A2.5 References **181**
Appendix 3 ..... **184**
A3.1 Rationale and purpose **184**
A3.2 Results and Discussion **184**
A3.3 References **194**

LIST OF FIGURES

Figure 1.1 Factors that influence bacterial communities……. 3
Figure 1.2 Schematic representation of the *A. tumefaciens* genome C58……10
Figure 1.3 Genetic map of the octopine-type Ti plasmid…… 11
Figure 1.4 Plant infeccion process by *A. tumefaciens*……..18
Figure 1.5 Genetic map of the octopine catabolic operon……... 20
Figure 1.6 Gene regulation by octopine…….. 22
Figure 2.1 Genetic organization map of the genes within the *occ* operon……. 32
Figure 2.2 End-to-end sequence similarity between Msh and the MmuM protein of *E. coli*……. 38
Figure 2.3 Msh and MmuM amino acid sequence alignment…….. 39
Figure 2.4 Proposed model for Msh methyltransferase activity……..... 40
Figure 2.5 Authentication of the chemically synthesized SMM by ESI-MS/MS. 40
Figure 2.6 ^{1}H NMR spectrum of SMM in D_20…… 42
Figure 2.7 ESI-MS/MS analysis of the production of methionine by Msh from SMM and homocysteine…… 43
Figure 2.8 ESI-MS/MS detection of methionine in a Msh reaction…… 44
Figure 3.1 Synthesis and utilization of sulfonopine…….. 55
Figure 3.2 Detection of octopine and sulfonopine by ESI-MS/MS in the Ocs enzymatic reaction……. 67
Figure 3.3 Analysis of the chemically synthesized sulfonopine by ESI-MS/MS.. 71
Figure 3.4 ^{1}H NMR spectrum of sulfonopine in D_20……... 72
Figure 3.5 Biological activity of sulfonopine diastereomers……. 72
Figure 3.6 ESI-MS/MS authentication of sulfonopine synthesis in infected tobacco seedlings with different *A. tumefaciens* strains……. 75
Figure 3.7 Authentication of sulfonopine production by ESI-MS?MS…… 77
Figure 3.8 Induction of *occ* operon by octopine-type opines……. 79
Figure 3.9 Induction of *occ* operon of strain KYC16 (R10 *ooxA-uidA*) by octopine-type opines and amino acids…… 80
Figure 3.10 Induction of *occ* operon by octopine-type opines……... 82
Figure 3.11 Utilization of sulfonopine as a sulfur source two strains…… 84
Figure 3.12 The fragmentation pattern of octopine, histopine, canavanopine, and homocysteine…….. 93
Figure 3.13 The fragmentation pattern of octopinic acid, lysopine, methiopine, and heliopine……. 94
Figure 3.14 The fragmentation pattern of homoserinenopine, cystinopine, glycinopine, and homooctopine…….…….. 95

Figure 3.15 The fragmentation pattern of leucinopine, serinopine, alanopine, and asparaginopine…… 96
Figure 3.16 The fragmentation pattern of threonopine, valinopine, and isoleucinopine…… 97
Figure 4.1 Anatomy and transcription regulation of riboswitch……...101
Figure 4.2 Repression of the downstream part of the *occ* operon by methionine… 111
Figure 4.3 Mapping the region involved in the repression of the *occ* operon by methionine…….. 113
Figure 4.4 Repression of the *msh* and downstream genes by methionine…….. 115
Figure 4.5 Mapping the 5'end of the region regulated by methionine…… 117
Figure 4.6 Procedure for analyzing the methionine riboswitch by using SHAPE.. 119
Figure 4.7 Selective 2'-hydroxyl acylation analyzed by primer extension chemistry for resolving mRNA structures…… 119
Figure 4.8 Structure of the PCR product used to generate mRNA……...120
Figure 4.9 Comparison of the mRNA conformational folding in the absence or presence of methionine evaluated by SHAPE analysis…….. 121
Figure 5.1 Synthesis and utilization of sulfonopine……. 135
Figure A2.1 Genetic map of the *repABC* region…….. 171
Figure A2.2 Increased expression of RepC causes elevated copy number……. 175
Figure A3.1 Organization and functions of the catabolic regions…… 185
Figure A3.2 Amino acid similarity and synteny between *ooxBCAD* and genes from *Ochrobactrum intermedium*……187
Figure A3.3 Organization of opine catabolic genes and nearby downstream sequence……... 188
Figure A3.4 OoxC and OoxD amino acids resemble many proteins in the Genbank protein database……..189
Figure A3.5 Coomasie blue stained-PAGE gel of the soluble fraction and purification of the octopine oxidase subunits……... 193

LIST OF TABLES

Table 2.1 Bacterial Strains and plasmid used in this study…….. 34
Table 2.2 Oligonucleotides used in this study…… 35
Table 2.3 Msh catalytic specificity for methyl donor substrates……. 45
Table 3.1 Bacterial Strains and plasmid used in this study……. 56
Table 3.2 Oligonucleotides used in this study……. 58
Table 3.3 Ocs enzymatic utilization of SMM and Arginine…….. 66
Table 3.4 Substrate specificity of octopine synthase for amino acids…… 69
Table 3.5 Substrate specificity of octopine synthase for α-keto acids…… 70
Table 3.6 SMM and sulfonopine content in tobacco tissue and exudate of infected seedlings with *A. tumefaciens* R10 and mutants strains…….. 74
Table 3.7 Sulfonopine contents in floral tissue in *Arabidopsis thaliana* lines. ..…. 76
Table 3.8 Representing fragment peaks of opines analyzed by ESI-MS/MS….. 92
Table 4.1 Strains used in this study……. 104
Table 4.2 Plasmids used in this study……. 105
Table 4.3 Oligonucleotides used in this study…… 106
Table 4.4 Repression ratio of the expression by methionine…… 116
Table A1.1 Induction of *B. cenocepacia* genes by OHL…….. 150
Table A1.2 Activation of CepR-regulated promoters in the heterologous host *E. coli*…….. 151
Table A1.3 Strains used in this study…….. 156
Table A1.4 Plasmids used in this study…….. 157
Table A1.5 Oligonucleotides used in this study……. 158
Table A2.1 Uncontrolled expression of RepC causes lethal runaway replication….176
Table A2.2 Strains and plasmids used in this study……. 179
Table A2.3 Oligonucleotides used in this study……... 180

CHAPTER 1

Introduction

1.1 Soil-dwelling bacteria

Soil is a very heterogeneous environment, varying in conditions such as temperature, moisture, light, and most importantly, nutrients. Therefore, one of the most abundant and diverse inhabitants of the soil, bacteria, are constantly in nutritional stress. Because seeking nutrients is an important task, bacteria have developed an ever growing number of mechanisms to obtain them.

In soil, the particular conditions defining the bacterial surroundings are called habitats, and these can be classified as part of the nonrhizosphere soils or the rhizosphere soils. The bulk of nonrhizosphere soil is oligotrophic, in essence a nutritional desert (Metting, 1985), including numerous microhabitats within which bacterial activity is ephemeral and responsive to fluctuating substrate availability and physical and chemical microenvironmental conditions. On the other hand, rhizosphere soils are rich in nutrients due to the presence of plant roots, which secrete an enormous variety of compounds to the surrounding soils (Lynch, 1987) making them a very biologically active soil zone (Fig. 1.1).

The composition of the root exudates in the rhizosphere varies, depending on growth conditions, rooting medium, stages of plant development and the plant species

(Hale et al., 1978). The presence of the root creates a strong impact on the microbial communities, promoting bacterial biomass formation. This stimulatory effect is due to the release of organic compounds, like carbohydrate and derivatives, which are one of the major sources of carbon and energy for microbial growth and metabolism in the rhizosphere. Therefore the microbial communities in the rhizosphere are highly dynamic and complex, and the microbial activity is more intense compared to non-rhizosphere soils in the vicinity (Bais et al., 2006; Berg and Smalla, 2009; Bolton et al., 1993; Bowen, 1980). For many bacteria, colonization of this nutrient-rich environment can involve microbial interactions that can play important roles in successful bacterial colonization or disappearance from the rhizosphere (Fig. 1.1) (Schmidt, 1979).

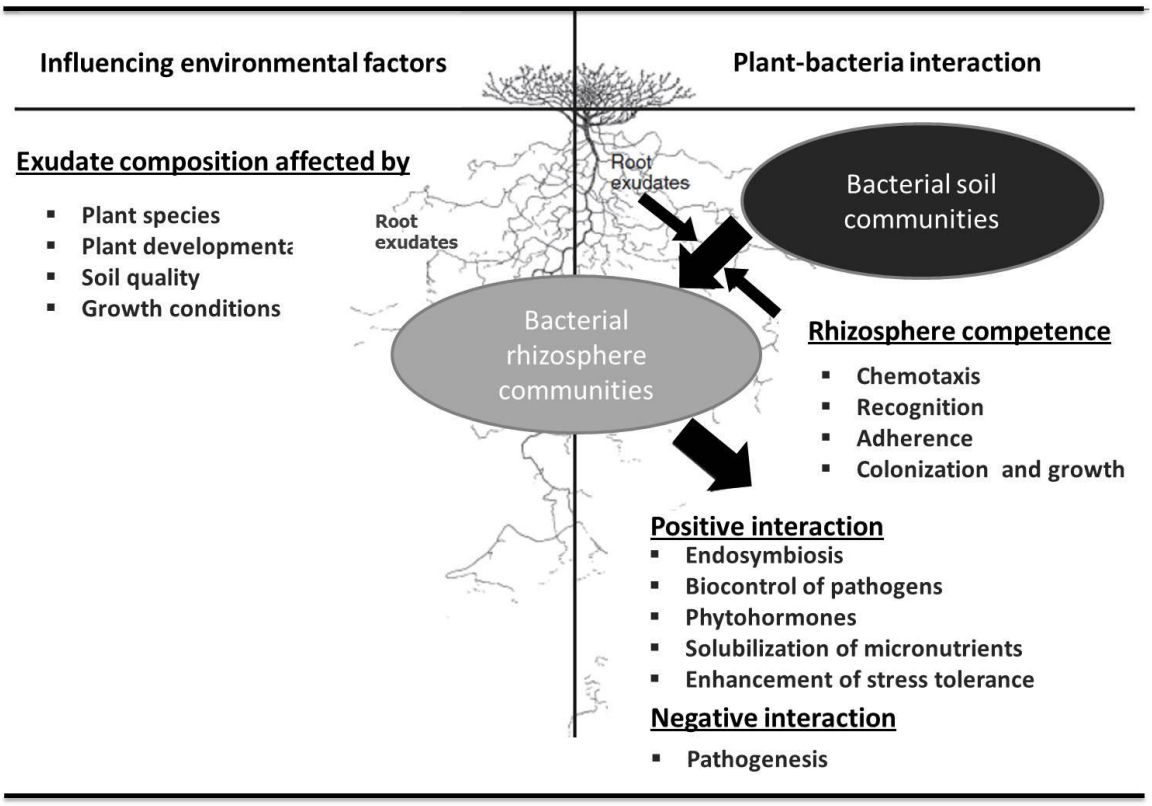


Figure 1.1 Factors that influence bacterial communities from the rhizosphere, and how bacterial communities were selected from soil by their rhizosphere competence. Modified from (Berg and Smalla, 2009).

1.2 To Colonize or to disappear: That is the question

Root exudates, besides being used as nutrients, can also play a role in chemical communication between plants and soil microbes by serving as signal molecules. These molecules are recognized by microbes, which in turn produce other signals that initiate colonization (Bais et al., 2006). Although colonization of plant roots by microorganisms is not well understood, it is a crucial step for their survival. There are several factors that have been implicated in influencing colonization, including the ability of a microorganism to detect, to translocate toward, and to adhere to roots surfaces (Bais et al., 2006; Schmidt, 1979).

Recognition of the plant roots by soil microbes can lead to chemical attraction (chemotaxis) or electrotaxis (Fig. 1.1). Some compounds from root exudates have been described to act as chemoattractants for bacteria, such as sugars, amino acids, organic acids, and aromatic compounds. Following the chemical detection is the actual movement towards the plant roots – the source of the signals (Fig. 1.1); therefore motility is an important trait for microbe competition and enables participation in this cross talk. This motility can be performed by flagella and type IV pili (Bardy et al., 2003; Jarrell and McBride, 2008).

Colonization of root surfaces is often impacted by competition with other microbes (Schmidt, 1979). This competition for nutrients and space can be extreme, therefore microbes have to survive and compete in the rhizosphere until the strongest interaction with the root surface is consolidated (Fig. 1.1). Some microbes can even go further by establishing a long-term interaction with plant root by an attachment process.

Attachment confers nutritional advantages in that bound bacteria are closer to the source of root exudates. Several strategies have been used by microbes to attach to plant cell surfaces, involving fimbri, Type III and Type IV secretion systems, lectins or production of polysaccharides such as lipopolysaccharides (LPS) and extracellular polysaccharide (EPS) (Bolton et al., 1993; Burdman et al., 2000; Lugtenberg et al., 2002; Rodriguez-Navarro et al., 2007; Schmidt, 1979).

1.3 The good, the bad and the rhizosphere

The nutritional richness of the rhizosphere can attract an extremely diverse community of microbes (Jensen et al., 1986). These microbes can mediate both positive and negative interactions with the rhizosphere plants. The positive interactions include symbiotic associations with beneficial microbes, commensalism, mutualism, and protocooperation. On the other hand, negative interactions include association with pathogenic microbes, causing competition, parasitism, and predation, among the various members of the microbial community (Badri and Vivanco, 2009; Bowen, 1980).

Although the production of exudates represents a significant carbon cost to the plant, it can be rewarded by colonization of beneficial bacteria. Many bacteria positively influence plant growth through different mechanisms, including fixation of atmospheric nitrogen and production of metabolites that contribute to the plant by either promoting growth or improving health. The rhizosphere presents a variety of beneficial bacteria such as symbionts and mutualistic rhizobacteria (Bais et al., 2006; Barea et al., 2005).

One of the most important and best-documented mutualistic symbioses is between

plant roots and nitrogen-fixing bacteria (Fitter and Moyersoen, 1996). Nitrogenous compounds are often limiting in soils; therefore bacterial nitrogen fixation is crucial for plants, allowing them to grow in nitrogen deficient habitats. Some nitrogen-fixing bacteria have evolved an intimate interaction with the plant roots by producing root nodules, which the bacteria inhabit as a nitrogen-fixing endosymbionts (Brencic and Winans, 2005; Murray, 2011).

On the other hand, some rhizobacteria have the ability to stimulate plant growth, degrade xenobiotic compounds that may be toxic to plants, and exert biological control over plant pathogenic microorganisms (Raaijmakers et al., 2002), without establishing an intimate interaction with the roots. A group of these bacteria known as plant growth-promoting rhizobacteria (PGPR) can stimulate plant growth by producing phytohormones or by promoting plant nutrition through phosphate solubilization and mineralization of nutrients (Barea et al., 2005; Richardson and Simpson, 2011). Others can create suppressive soils by producing antimicrobial and antifungal metabolites to control plant diseases caused by soil bacteria and fungi. These metabolites range from antibiotics to metal chelators (Bais et al., 2006; Raaijmakers et al., 2002). Lastly, it has also been reported that biofilms produced by the gram positive PGPR *Bacillus subtilis* protect plant roots from gram negative plant pathogens such as *Pseudomonas syringae* by limiting the access to the root surface and by producing an antimicrobial cyclic lipopeptide called surfactin (Bais et al., 2004).

In spite of the fact that the rhizosphere can recruit beneficial soil microorganisms to counteract pathogen assaults by producing antimicrobial metabolites, there are still

some pathogens that can overcome these barriers and cause severe plant diseases (Mendes et al., 2011). Plant pathogens have highly diversified molecular strategies to suppress the host-defense response, prevent pathogen recognition, facilitate nutrient extraction, and promote pathogen reproduction and dissemination. These strategies include production of toxins, phytohormones, detoxification mechanisms, active efflux of antibiotics and enyzmatic antibiotic resistance; in addition, pathogens can produce extracellular proteins, attachment structures, micronutrients uptake systems, and specialized structural layers like lipopolysaccharides (LPS) and exopolysaccharides (EPS) (Duffy et al., 2003; Stravrinides, 2009).

As described before, secretion systems play an important role in pathogenesis by successfully secreting and delivering pathogenic factors for either the extraction of nutrients from their hosts or causing diseases. There are currently six secretion pathways, termed type I to type VI, utilized for a variety of export processes in both pathogenic and non-pathogenic bacteria (Stravrinides, 2009). Many pathogens have successfully managed to use one or more of these secretion systems to secrete or translocate a variety of toxic macromolecules.

During this introduction we have seen that pathogens can display an array of mechanisms to thrive in the rhizosphere. An interesting fact is that in many plant pathogens, the pathogenic genes are frequently associated with mobile genetic elements such as bacteriophages, transposons, integrative-conjugative elements, or plasmids (Stravrinides, 2009). For example, *Pseudomonas syringae* has a virulent mobile genetic element that contains genes involved in the production of effector proteins and their

delivery system, the type 3 secretion system (T3SS) (Collmer et al., 2000).

Mobile genetic elements are transmissible between bacteria by horizontal/lateral gene transfer, which plays an important role in the combination of virulence gene sets driving the evolution of the plant pathogens and diversification of their strategies (Kado, 2009; Kelly et al., 2009; Stravrinides, 2009). *Agrobacterium tumefaciens* is a fascinating plant pathogen that has a unique virulence mechanism for locating, reaching, and infecting plant cells. This dissertation will focus on the remarkable mechanism of plant infection by *A. tumefaciens*.

1.4. *Agrobacterium tumefaciens* mechanism of infection

Agrobacterium tumefaciens is a ubiquitous soil-borne bacterium that belongs to the Rhizobiaceae family within the α-Proteobacteria group. While some members of the Rhizobiaceae are known to form mutualistic interactions with plants, *Agrobacterium* species are either saprophytic and or plant pathogens (Escobar and Dandekar, 2003). *A. tumefaciens* is a broad-host-range plant pathogen, causing crown gall tumors in more than 90 families of dicotyledonous and some monocots as well (Escobar and Dandekar, 2003; Wood et al., 2001). *A. tumefaciens* can cause major economic losses on grape vines, stone fruit, and nut trees (Shrawat and Lorz, 2006; Zupan et al., 2000).

In 2001, the 5.67-Mb genome of *A. tumefaciens* C58 was sequenced. Its genome consists of four replicons: a circular chromosome, a linear chromosome, and the plasmids pAtC58 and pTiC58 (Fig 2) (Wood et al., 2001). Tumorigenesis requires the Ti plasmid. In fact, introduction of Ti plasmids into non-pathogenic *Agrobacterium* strains is

sufficient to convert them into pathogenic strains. Different types of Ti plasmids of *A. tumefaciens* strains have been described such as octopine, nopaline, and agropine-type plasmids. This classification is based on the low molecular weight amino acid and sugar phosphate derivatives called opines, that the plasmids direct synthesis of, in the host plant nucleus, or utilization of, in the bacteria (Pappas, 2008). In 2000 the octopine-type Ti plasmid sequence was assembled, defining the genes require for host recognition, type IV secretion system (T4SS), and for transferring the oncogenic DNA (T-DNA), as well as genes for opine uptake and catabolism, quorum sensing system and others (Fig. 1.3) (Tzfira and Citovsky, 2008; Zhu et al., 2000). This thesis will focus on the octopine-type Ti plasmid.

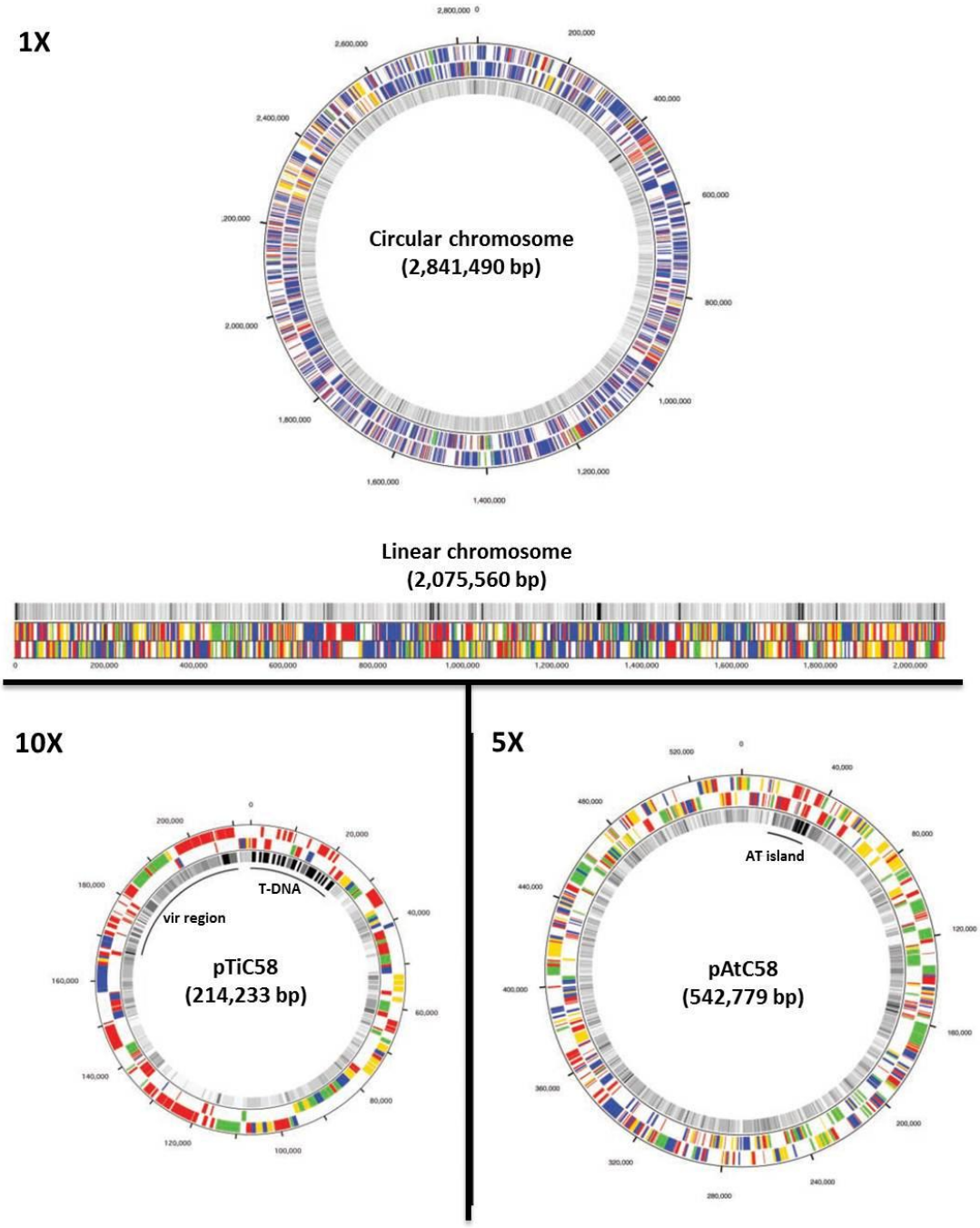


Figure 1.2 Schematic representation of the *A. tumefaciens* genome C58. Chromosomes are drawn to scale with plasmids represented at 5X or 10X magnification, as indicated. The *vir* and T-DNA regions of pTiC58 and the AT island of pAtC58 are indicated. Colors indicate orthology to proteins in the *S. meliloti* replicons: blue, chromosome; green, pSymA; gold, pSymB; red, nonorthologous. The inner circle represents GC content indicated by darker shading. Modified from (Wood et al., 2001).

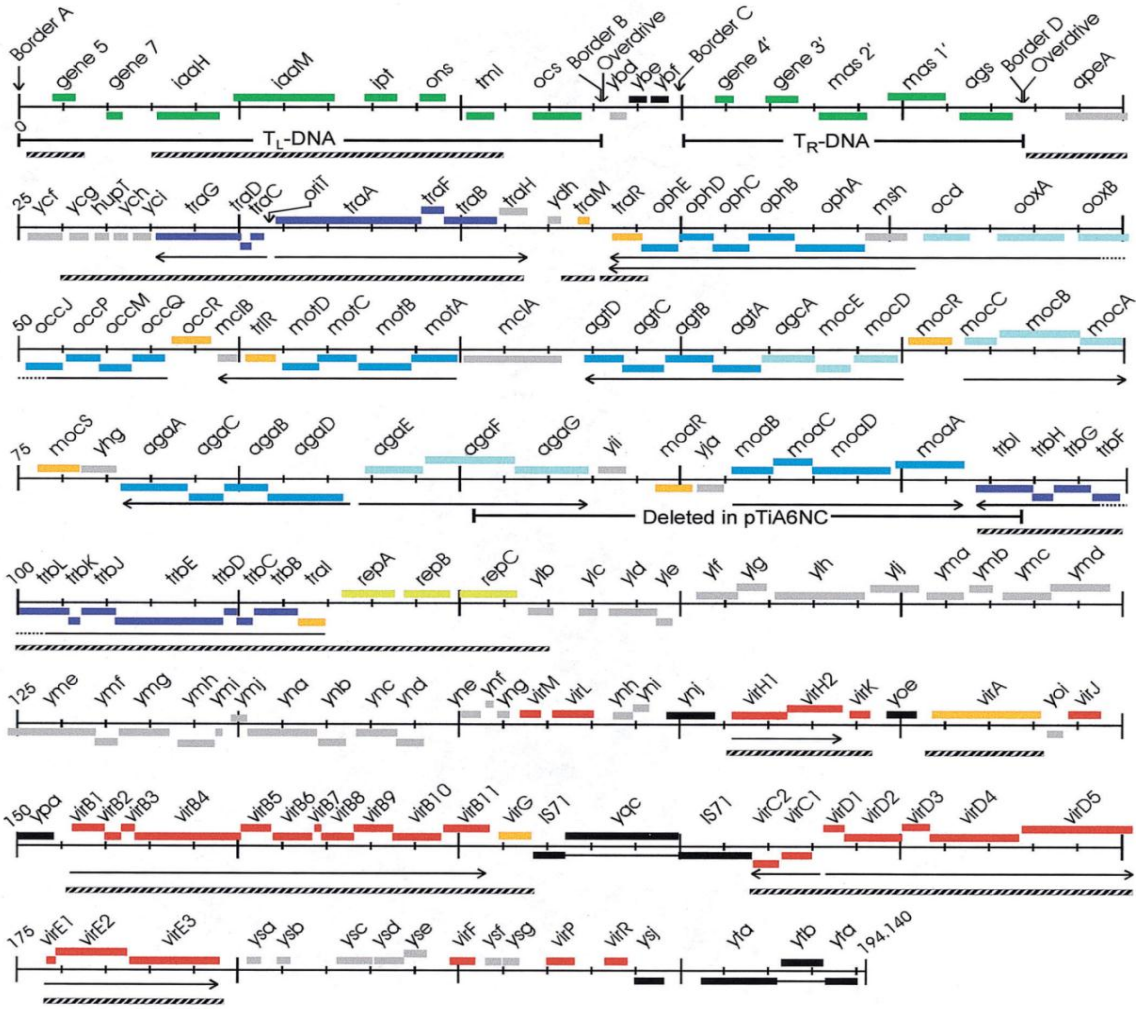


Figure 1.3 Genetic map of the octopine-type Ti plasmid. Genes important for virulence are in red bars, all of which are regulated by VirA and VirG, and many of which are required for T-strand processing and transfer. Green bars are the genes of the T-DNA, which are located on the TL region or TR region and transferred to plant cells. Purple bars indicate genes required for conjugal transfer of the Ti plasmid. Light green bars indicate genes required for vegetative replication and partitioning of the Ti plasmid. Dark blue bars indicate genes encoding ATP binding cassette-type opine permeases, while light blue bars indicate opine catabolic genes. Orange bars indicate regulatory genes. Black bars indicate suspected IS elements, while grey bars indicate ORFs of miscellaneous or unknown function. Modified from (Zhu et al., 2000).

The Ti plasmid makes *A. tumefaciens* pathogenesis a unique and highly specialized process involving interkingdom gene transfer (Escobar and Dandekar, 2003). Initiation of the infection process begins when *A. tumefaciens* detects specific signal molecules from the plant, and uses this information to induce the expression of a set of genes called *vir* genes, located on the Ti plasmid. In consequence a type IV translocation system transfers a set of oncogenic genes (T-DNA) from the Ti plasmid to the plant cell. This T-DNA is integrated into the plant cell chromosomes, a process called Transformation. The products of some of T-DNA genes cause uncontrolled plant cell proliferation, while another transferred gene, *ocs*, is responsible for the synthesis of unusual compounds called opines, that serve as nutrients for *A. tumefaciens*. Catabolism of opines require specific genes that direct opine uptake and catabolism. Most of these genes are located on non-transferred portions of the Ti plasmid (Brencic and Winans, 2005; Zupan et al., 2000). Below I will discuss the molecular mechanism of *A. tumefaciens* that leads to a successful infection, by recognizing plant hosts, attaching to them, transferring, integrating, and expressing T-DNA, and colonizing the plant tumors that it incites.

1.4.1 Recognition of the host

As previously described, plants release chemical compounds that can attract microorganism to the rhizosphere. Phenolic compounds are one of the most ubiquitous groups of secondary metabolites in plants (Bhattacharya et al., 2010). Diverse plant phenolic compounds are released from the plant and they can either attract or repel A. *tumefaciens*. It has been showed that *A. tumefaciens* can positively detect a variety of

phenolic compounds by a two component system VirA/VirG. In addition, VirA can interact with ChvE to also recognize environmental acidity and monosaccharides (Bhattacharya et al., 2010; Brencic et al., 2004). The release of these signals increases when plants are wounded. Activation of this recognition system results in maximal induction of the virulence genes, such as *virA, virB, virC, virD, virE and virG* operons (Escobar and Dandekar, 2003; Gelvin, 2003), initiating the infection process. High activity of the phenolic compound pathway, low pH, and sugars associated with cell wall synthesis are routinely associated with wound repair, making wounds on host plants common sites for *A. tumefaciens* infection (Baron and Zambryski, 1995; Braun, 1952). Non-wounded plants can also be colonized, and can also induce expression of *vir* genes leading to T-DNA transfer. However, this does not lead to plant tumors, indicating that tumorigenesis requires the plant wound response (Brencic et al., 2005). Therefore *A. tumefaciens* has evolved to detect and respond to specific wound-released molecules, an efficient way to look for the portal of entry and initiate infection.

1.4.2 *Agrobacterium* attachment to the host plants

Once the host plant cells are recognized the next step is to attach and initiate the colonization process. The intimate contact between pathogen and host cell is a prerequisite for transformation, therefore motility is required for efficient surface attachment (Brencic and Winans, 2005). The surface of the bacterial cell also plays a crucial role during infection of the plant cells. The specific attachment of *A. tumefaciens* to plant cells is not dependent on the Ti plasmid. Instead, it is facilitated by chromosomally encoded bacterial genes that are involved in the synthesis of cellulose

fibrils that allow the bacterial cells to form large aggregates on wounded plant cells (Matthysse et al., 2005; Pitzschke and Hirt, 2010; Ream, 1989). The bacteria attach only at their cell poles. During the process of attachment several different appendages play an important role, such as adhesins and adhesives (Tomlinson and Fuqua, 2009). Additionally, the host-cell surface also participates in the process of host-pathogen binding. For example, the *Arabidopsis thaliana* At AG17 protein of the plant cell wall is involved in *Agrobacterium*-mediated attachment, as well as several cell wall carbohydrates (Gelvin, 2010). In conclusion, a successful attachment is due to the communication of several components from both the bacterial cell surface and host cell wall.

1.4.3 Transport and integration of T-DNA into the plant cell chromosome

After the activation of the virulence genes and the attachment to the plant cells (Fig. 1.4 Step 1 and 2), the process of transport of the T-DNA and several proteins into the host initiates. In order to transfer the virulence factors, *A. tumefaciens* has to overcome several obstacles, including the bacterial inner membrane, the periplasm/peptidoglycan wall, the outer membrane, as well as the plant host cell wall and membrane (McCullen and Binns, 2006). *A. tumefaciens* has solved these obstacles by orchestrating the assembly of a specialized transporter complex (T4SS) that will ultimately allow passage of the T-DNA and proteins. The octopine-type Ti plasmid has two fragments of DNA that are transferred to the nuclei of the host plants (Zhu et al., 2000).

Once the expression of the *vir* genes is induced (Fig. 1.4 Step 3), two processes

have to happen: the construction of the T4SS and the T-DNA synthesis. For the formation of the transporter complex 12 proteins are needed, VirB1-11 and VirD4 (Christie, 2004). For the synthesis and processing of the T-DNA (T-strand), several proteins are required such as VirD1, and VirD2. VirD1 and VirD2 work co-operatively to nick a single strand on the left and right borders of each T-DNA. The DNA between these nicks is converted to a single-strand linear form (called a T-strand) by rolling circle DNA synthesis (Fig 1.4 Step 4). VirD2 nicks and covalently binds to 5'end of the T-strand. After the VirD2-T-DNA complex is formed, then it is ready to be transported through the T4SS into the plant cell. This process is done by the recognition of the VirD2 by VirD4 from T4SS (Fig, 1.4 Step 5). Simultaneously, VirE2, VirE3, and VirF are transported into the plant cell. Once translocated, the VirD2-T-DNA is coated by VirE2, which is a single-stranded DNA binding protein, and protects it from nuclease activity inside of the plant (Fig. 1.4 Step 6). VirD2 contains a nuclear localization signal (NLS), acting as a pilot protein which directs the complex inside the plant cells, and leads the T-DNA through the nuclear pore. VirE2 also contains a NLS that will help direct the VirD2-T-DNA to the plant nucleus (Christie, 2004; Citovsky et al., 2007; Escobar and Dandekar, 2003; McCullen and Binns, 2006; Pitzschke and Hirt, 2010; Tzfira and Citovsky, 2008).

The trafficking of the VirD2-T-DNA-VirE2 complex (T-complex) from the cytoplasm to the nuclear membrane is carried out by a dynein-like motor protein that interacts with the NLS from VirE2 (Fig. 1.4 Step 7). Once in the nuclear membrane the NSL of VirD2 is recognized by importins α and VirE2 interacts with VIP proteins (VirE2-interacting proteins; Step 8). These interactions with host proteins facilitate the

nuclear transport of the T-complex. Once inside of the nucleus, the T-complex has to be uncoated to allow the proper integration into the host chromatin. VirF is believed to mediate degradation of the T-DNA complex by coordinating the host proteasome machinery, thus facilitating the release of the T-DNA (Fig. 1.4 Step 9 and 10) (Pitzschke and Hirt, 2010; Tzfira and Citovsky, 2008).

1.4.4 Function of the oncogenes in the host plant

As previously explained, the octopine-type Ti plasmid transfers two T-strands, T_L-DNA and T_R-DNA, that together encode for 13 proteins. Collectively, these genes direct tumor development and opine synthesis.

Uncontrolled proliferation of the transformed plant cell requires the overproduction of phytohormones (Fig. 1.4 Step 11). The genes responsible for the production of the phytohormones auxins and cytokinins are located on the T_L-DNA. These genes encode for several enzymes including iaaM, iaaH, and ipt. The iaaM and iaaH proteins are required for auxin biosynthesis from tryptophan via indole-acetamide. For the production of cytokinin, *ipt* encodes for an isopentenyltransferase, which condenses isopentenyl pyrophosphate and AMP into cytokinin. Auxin and cytokinin concentrations increase in the transformed plant cells far beyond those of healthy host tissues, a mechanism that is necessary for the development of the crown galls (Gnanamanickam, 2006; Tzfira and Citovsky, 2008; Zhu et al., 2000).

1.4.5 Opine synthesis

On the transferred T-DNA, there is a set of genes that are involved in the

production of *A. tumefaciens* nutrients called opines. Opines have important roles in the epidemiology of crown gall and ecology of *A. tumefaciens*, as they provide a source of nutrients, and also provide a signal that leads to the dissemination of Ti plasmid (Gnanamanickam, 2006; Savka et al., 2002; Zhu et al., 2000).

More than 20 different opines have been identified in crown galls (Dessaux et al., 1993); they are carbon compounds of low molecular weight based on amino acids or sugar-phosphate derivatives. They serve as carbon and/or nitrogen sources, and recently as a sulfur-containing nutrient, described in this dissertation. In the octopine-type T-DNAs, there are four opine synthesis genes, including *ocs*, *ags*, and *mas1*, and *mas2*, which catalyze the production of octopine, agropine, and mannopine, respectively. For the purpose of this thesis, I am going to focus in the *ocs g*ene. This gene encodes octopine synthase, which reductively condenses pyruvate with either arginine, lysine, histidine, glutamine, or ornithine to produce octopine, lysopine, histopine, heliopine, or octopinic acid, respectively, all of which can be detected in crown gall tumors (Fig. 1.4 Step 12) (Chilton et al., 2001; Gnanamanickam, 2006; Spanik et al., 1998; Zhu et al., 2000)

As discussed previously, *A. tumefaciens* is attracted to the rhizosphere by different small molecules, and it has been also demonstrated that *A. tumefaciens* can be chemoattracted to opines, suggesting that *A. tumefaciens* detects and takes advantage of nutrient-rich environments (Kim and Farrand, 1998).

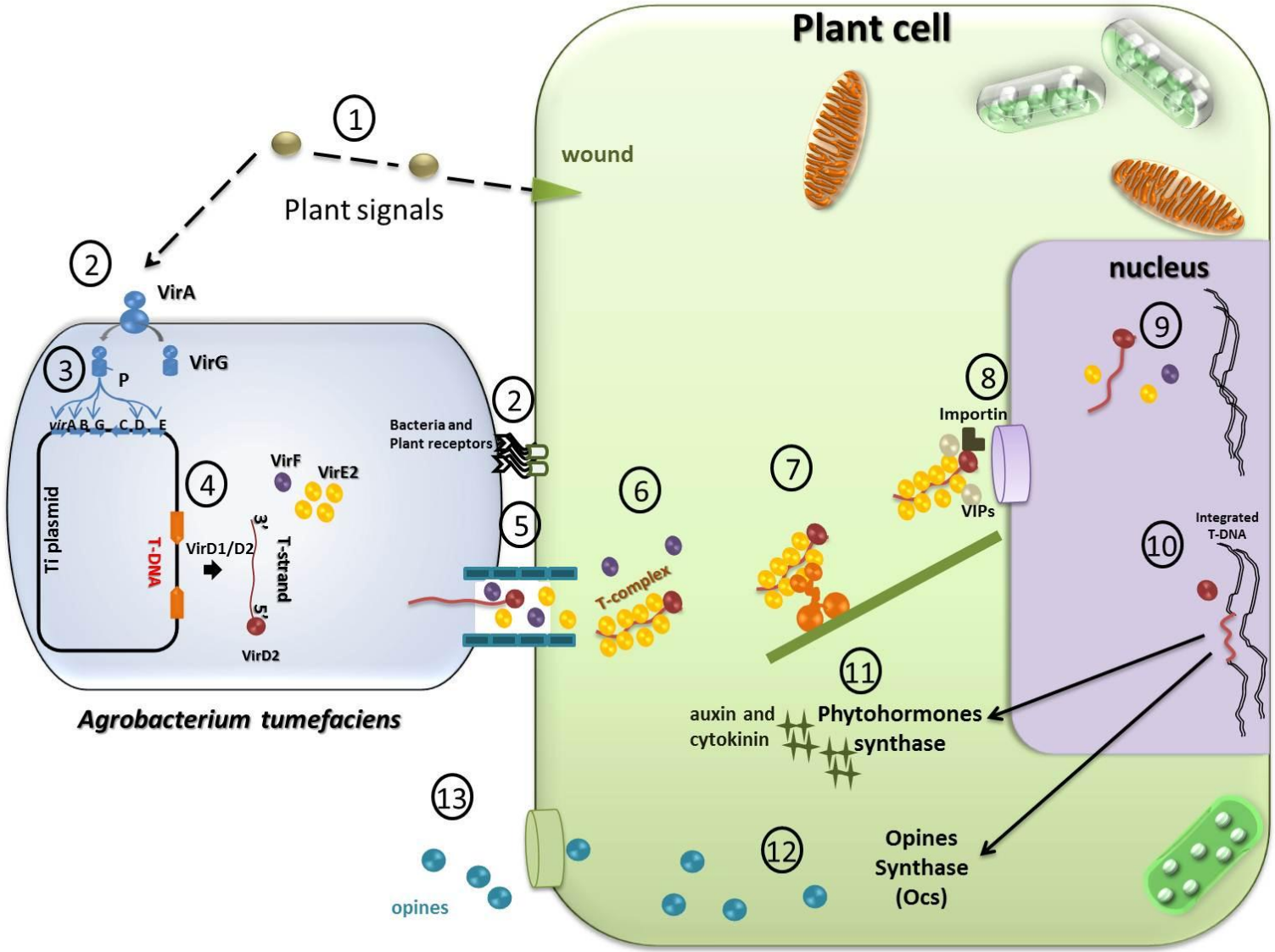

Plant cell
wound
1
Plant signals
2
VirA
3
P
VirG
virA B G C D E
Ti plasmid
T-DNA
4
VirD1/D2
VirF
VirE2
3'
T-strand
5'
VirD2
Bacteria and
Plant receptors
2
5
6
T-complex
7
8
Importin
VIPs
nucleus
9
10
Integrated
T-DNA
11
auxin and
cytokinin
Phytohormones
synthase
12
Opines
Synthase
(Ocs)
13
opines
Agrobacterium tumefaciens

Figure 1.4 Plant infeccion process by *A. tumefaciens*. The infection process comprises 12 major steps, beginning with the recognition of the released signals from the wounded host plant cell (Step1), and the sensing of plant signals by VirA/VirG and polar attachment (Step 2). Consequently VirG-P activates the vir genes (Step 3), and a copy of the T-DNA is produce by VirD1/VirD2 protein complex (Step 4) and delivered to the plant cell as a VirD2-T-DNA complex (T-complex) through the T4SS (Step 5). Following the association of VirE2 with T-complex (Step 6), dynein like motor ptotein recognizes VirE2 proteins and the T-complex is tranported to the nuclear pore (Step 7). Importin proteins recognize VirD2 and actively transports the complex into the nucleus (Step 8). Once inside of the nucleus, the T-DNA is uncoated by proteosome activity coordinated by VirF (Step 9), and the uncoated T-DNA is integrated into the plant cell chromosome (Step 10). Genes on the T-DNA encode for production of phytohormones for tumor generation (Step 11) and for production of opines for *A. tumefaciens* nutrition (Step 12). Realease of opines from the infected plant possibly by using plant permeases (Step 13).

1.5 Gene regulation in Agrobacterium tumefaciens by opines

After the synthesis and exudation of opines from the grown gall tumor, opines are utilized as nutrients by *A. tumefaciens*. In addition to being a nutritional source, opines can also serve as a signal molecules, regulating the transcription of genes involved in the uptake and catabolic pathway (Habeeb et al., 1991). It is known that *A. tumefaciens* contains one or more operons of the Ti-plasmid-encoded genes that are required for the internalization and utilization genes (Zhu et al., 2000). Among these, the octopine catabolism operon (occ) is required for the uptake and utilization of octopine-type opines (Fig. 1.5).

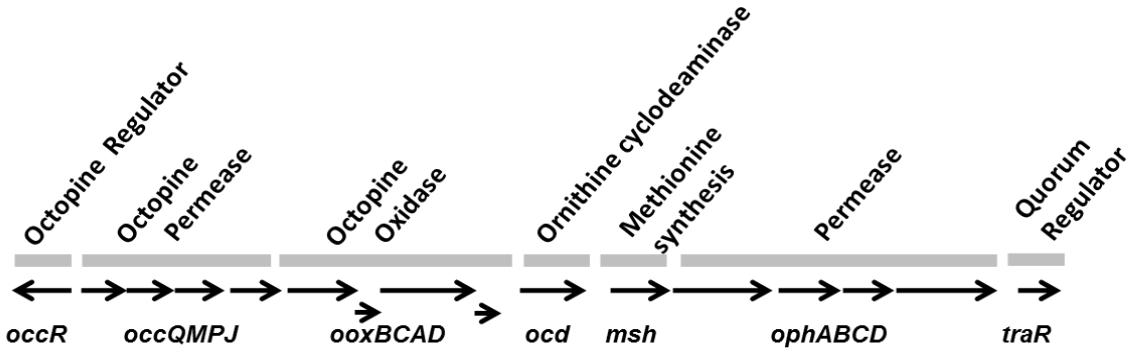


Figure 1.5 Genetic map of the octopine catabolic operon. The *occQMPJ* genes direct the uptake of octopine and probably related opines, while *ooxBCAD* direct the oxidative hydrolysis of octopine-type opines, yielding pyruvate and the corresponding amino acid. The *ocd* gene directs the conversion of ornithine to proline. The *msh* resembles *S*-methylmethionine-homocysteine methyltransferase. The remaining six genes constitute an ABC-type permease for an unknown substrate, and the *traR* gene, which directs quorum-dependent Ti plasmid conjugation and vegetative replication. OccR positively regulates the operon in the presence of octopine-type opines.

The function of the genes in the *occ* operon had been characterized by the utilization of octopine, an opine based on arginine and pyruvate. This operon contains 15 genes, and is regulated by the product of the divergent *occR* gene in response to octopine (Fig. 1.5) (Habeeb et al., 1991; Klapwijk et al., 1978; von Lintig et al., 1991; Zanker et al., 1992). The first four genes in the operon, *occQ,-M,-P* and *–J,* encode for an ABC-type octopine uptake system (Valdivia et al., 1991). The *ooxB, C, A* and *D* genes carry out the oxidative hydrolysis of the opine, generating pyruvate and the corresponding amino acid (Zanker et al., 1994). The *ocd* gene directs the conversion of ornithine (produced from metabolism of arginine) to proline (Schindler et al., 1989). Downstream of *ocd* are six additional genes. The first is *msh,* a gene with predicted methyl transferases activity but without a known in vivo function. I show that Msh is a S-methyl transferase, producing methionine from S-methylmethionine and homocysteine (see Chapter Two of this thesis). The next four genes, *ophABC* and -D, encode for a second ABC-type uptake system for an unknown substrate, presumably an opine. The last gene in the operon is *traR.* TraR is a LuxR-type quorum sensing transcription factor that regulates Ti plasmid-encoded operons that direct conjugal transfer and vegetative replication of the Ti plasmid (Cho and Winans, 2007; Fuqua and Winans, 1996). Quorum sensing and conjugation therefore occur only within or near crown gall tumors, as these tumors are the only natural source of these opines in the rhizosphere (Fig. 1.6) (White and Winans, 2008).

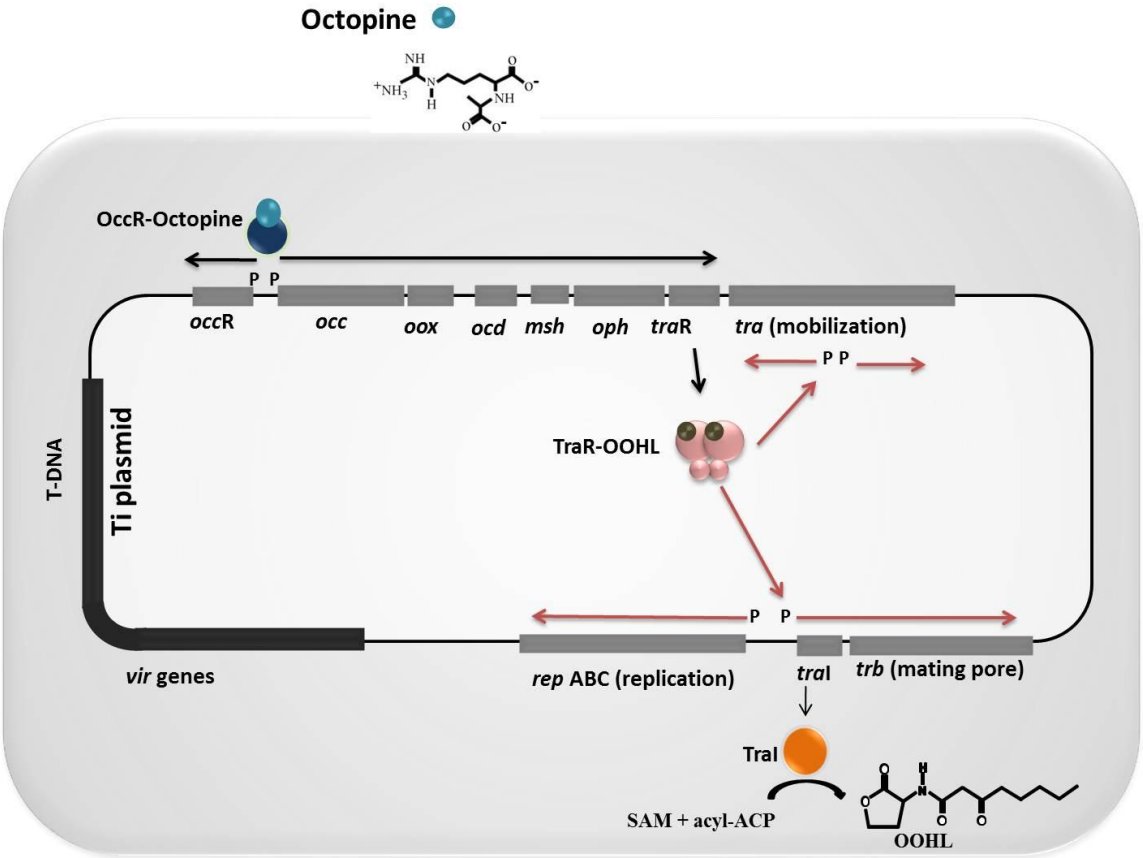


Figure 1.6 Gene regulation by octopine. Octopine not only can be used as a nitrogen and carbon source but also as a regulatory signal molecule. It activates genes involved in its catabolism located in the octopine catabolism operon (*occ* operon). In addition it activates *traR* gene, which encodes for a master regulator of genes in the Ti plasmid. TraR positively regulates the conjugation and replication of the Ti plasmid as well as quorum sensing system in *A. tumefaciens*.

1.6 Contents of the dissertation

This thesis is focused on the role of S-methylmethionine (SMM) in nutrition and gene regulation of *A. tumefaciens*. SMM is a compound produced by plants as a stored form of methionine. *A. tumefaciens* uses SMM to produce methionine, and to regulate the conjugation of Ti plasmid and cell-cell signaling in bacteria colonizing plant tumors. However, how *Agrobacterium* takes in SMM from the plant was not known. In this thesis, I show that the *ocs* gene, found on the T-DNA can direct the reductive condensation of SMM and pyruvate, generating a new opine, that I designate sulfonopine. Sulfonopine can be catabolized by A*grobacterium* and used as a source of methionine and sulfur and as a regulatory signal. The methionine thus generated may regulate the downstream half of the *occ* via a riboswitch (Chapter 4).

In Chapter 2, I described the function of the uncharacterized gene, *msh*, located in the *occ* operon. Msh had a limited amino acid similarity to proteins involved in methionine synthesis. By reducing the gap penalty of the BLAST algorithm, I was able to identify a homolog with end-to-end amino acid sequence similarity to an *E. coli* gene called *mmuM*. Similarly to its homolog, Msh is a methyl transferase that uses SMM and homocysteine to produce two molecules of methionine. The most intriguing finding was not that Msh is involved in methionine production but the utilization of SMM. As I mention above, SMM is a plant produced molecule. The main question is how SMM was taken up from the plant tissue. Since *msh* is located in the *occ* operon, I hypothesized that it played a role in opine catabolism. Therefore, I hypothesized that SMM is used by Ocs together with pyruvate to produce a new sulfur-containing opine, sulfonopine.

In Chapter 3, I confirmed the prediction that Ocs can utilize SMM and pyruvate to produce sulfonopine. This is the first described opine that provides sulfur to *A. tumefaciens*. Here I described the production and release of sulfonopine in infected plant tissues, as well as its utilization by *A. tumefaciens*. Sulfonopine also serves as a signal molecule for induction of the *occ* operon.

Once inside of the bacterial cell, sulfonopine is catabolized generating SMM and pyruvate. Then SMM is used by Msh to synthesize methionine. I found that the methionine thus generated downregulates *msh* together with downstream genes in the *occ* operon, including *traR* (Fig 1.5). Therefore, when methionine downregulates its biosynthetic gene, it also represses TraR-dependent regulation of quorum sensing, conjugation and replication of the Ti plasmid (Fig. 1.6). In Chapter 4, I present evidence that methionine represses the downstream region of the *occ* operon, possibly via a riboswitch.

1.7 References

Badri, D.V., and Vivanco, J.M. (2009). Regulation and function of root exudates. **Plant, Cell & Environment** *32,* 666-681.

Bais, H.P., Fall, R., and Vivanco, J.M. (2004). Biocontrol of *Bacillus subtilis* against infection of *Arabidopsis* roots by *Pseudomonas syringae* is facilitated by biofilm formation and surfactin production. **Plant Physiology** *134*, 307-319.

Bais, H.P., Weir, T.L., Perry, L.G., Gilroy, S., and Vivanco, J.M. (2006). The role of root exudates in rhizosphere interactions with plants and other organisms. **Annual Review of Plant Biology** *57,* 233-266.

Bardy, S.L., Ng, S.Y., and Jarrell, K.F. (2003). Prokaryotic Motility Structures.

Microbiology *149*, 295-304.

Barea, J.M., Pozo, M.J., Azcon, R., and Azcon-Aguilar, C. (2005). Microbial co-operation in the rhizosphere. **Journal of Experimental Botany** *56*, 1761-1778.

Baron, C., and Zambryski, P.C. (1995). The plant response in pathogenesis, symbiosis, and wounding: variations on a common theme? **Annual Review of Genetics** *29*, 107-129.

Berg, G., and Smalla, K. (2009). Plant species and soil type cooperatively shape the structure and function of microbial communities in the rhizosphere. **FEMS Microbiology Ecology** *68*, 1-13.

Bhattacharya, A., Sood, P., and Citovsky, V. (2010). The roles of plant phenolics in defence and communication during *Agrobacterium* and *Rhizobium* infection. **Molecular Plant Pathology** *11*, 705-719.

Bolton, H., Fredrickson, J.K., and Elliott, L.F. (1993). Microbial ecology of the rizhosphere. **In Soil Microbial Ecology**, B. Metting, ed. (New York: Marcell Dekker, Inc), pp. 27-63.

Bowen, G.D. (1980). Misconceptions, concepts and approaches in rhizosphere biology. In **Contemporary Microbial Ecology**, D.C. Ellwood, J.W. Hedger, M.J. Latham, M.J. Lynch, and J.H. Slater, eds. (New york: Academic Press).

Braun, A.C. (1952). Conditioning of the host cell as a factor in the transformation process in crown gall. **Growth** *16*, 65-74.

Brencic, A., Angert, E.R., and Winans, S.C. (2005). Unwounded plants elicit *Agrobacterium* vir gene induction and T-DNA transfer: transformed plant cells produce opines yet are tumour free. **Molecular Microbiology** *57*, 1522-1531.

Brencic, A., and Winans, S.C. (2005). Detection of and response to signals involved in host-microbe interactions by plant-associated bacteria. **Microbiology and Molecular Biology Reviews: MMBR** *69*, 155-194.

Brencic, A., Xia, Q., and Winans, S.C. (2004). VirA of *Agrobacterium tumefaciens* is an intradimer transphosphorylase and can actively block vir gene expression in the absence of phenolic signals. **Molecular Microbiology** *52*, 1349-1362.

Burdman, S., Okon, Y., and Jurkevitch, E. (2000). Surface characteristics of *Azospirillum brasilense* in relation to cell aggregation and attachment to plant roots. **Crit Rev Microbiol** *26*, 91-110.

Chilton, W.S., Petit, A., Chilton, M.D., and Dessaux, Y. (2001). Structure and characterization of the crown gall opines heliopine, vitopine and rideopine.

Phytochemistry *58*, 137-142.

Cho, H., and Winans, S.C. (2007). TraA, TraC and TraD autorepress two divergent quorum-regulated promoters near the transfer origin of the Ti plasmid of *Agrobacterium tumefaciens.* **Molecular Microbiology** *63*, 1769-1782.

Christie, P.J. (2004). Type IV secretion: the *Agrobacterium* VirB/D4 and related conjugation systems. **Biochimica et Biophysica Acta** *1694*, 219-234.

Citovsky, V., Kozlovsky, S.V., Lacroix, B., Zaltsman, A., Dafny-Yelin, M., Vyas, S., Tovkach, A., and Tzfira, T. (2007). Biological systems of the host cell involved in *Agrobacterium* infection. **Cellular Microbiology** *9*, 9-20.

Collmer, A., Badel, J.L., Charkowski, A.O., Deng, W.L., Fouts, D.E., Ramos, A.R., Rehm, A.H., Anderson, D.M., Schneewind, O., van Dijk, K., *et al.* (2000). *Pseudomonas syringae* Hrp type III secretion system and effector proteins. **Proceedings of the National Academy of Sciences of the United States of America** *97*, 8770-8777.

Dessaux, Y., Petit, A., and Tempe, J. (1993). Chemistry and biochemistry of opines, chemical mediators of parasitism. **Phytochemistry** *34*, 31-38.

Duffy, B., Schouten, A., and Raaijmakers, J.M. (2003). Pathogen self-defense: mechanisms to counteract microbial antagonism. **Annual Review of Phytopathology** *41*, 501-538.

Escobar, M.A., and Dandekar, A.M. (2003). *Agrobacterium tumefaciens* as an agent of disease. **Trends Plant Sci** *8*, 380-386.

Fitter, A.H., and Moyersoen, B. (1996). Evolutionary trends in root-microbe symbioses. **Philos T Roy Soc B** *351*, 1367-1375.

Fuqua, C., and Winans, S.C. (1996). Localization of OccR-activated and TraR-activated promoters that express two ABC-type permeases and the traR gene of Ti plasmid pTiR10. **Molecular Microbiology** *20*, 1199-1210.

Gelvin, S.B. (2003). *Agrobacterium*-mediated plant transformation: the biology behind the "gene-jockeying" tool. **Microbiology and Molecular Biology Reviews: MMBR** *67*, 16-37, table of contents.

Gelvin, S.B. (2010). Plant proteins involved in *Agrobacterium*-mediated genetic transformation. **Annual Review of Phytopathology** *48*, 45-68.

Gnanamanickam, S.S. (2006). Plant-associated bacteria (The Netherlands: **Springer**).

Habeeb, L.F., Wang, L., and Winans, S.C. (1991). Transcription of the octopine catabolism operon of the *Agrobacterium* tumor-inducing plasmid ptia6 is activated by a

Lysr-type regulatory protein. **Mol Plant Microbe In** *4*, 379-385.

Hale, H.G., Moore, L.D., and Griffin, C.J. (1978). Root exudates and exudation. In interactions between non-pathogenic soil microorganisms and plants, Y.R. Dommergues, and S.V. Krupa, eds. (New York: **Elsevier Scientific Publishing**), pp. 163-203.

Jarrell, K.F., and McBride, M.J. (2008). The surprisingly diverse ways that prokaryotes move. **Nat Rev Microbiol** *6*, 466-476.

Jensen, V., Kjoller, A., and Sorensen, L.H. (1986). Microbial communities in soil (England: Elsevier Applied Science Publisher LTD).

Kado, C.I. (2009). Horizontal gene transfer: sustaining pathogenicity and optimizing host-pathogen interactions. **Molecular Plant Pathology** *10*, 143-150.

Kelly, B.G., Vespermann, A., and Bolton, D.J. (2009). Gene transfer events and their occurrence in selected environments. **Food Chem Toxicol** *47*, 978-983.

Kim, H., and Farrand, S.K. (1998). Opine catabolic loci from *Agrobacterium* plasmids confer chemotaxis to their cognate substrates. **Mol Plant Microbe In** *11*, 131-143.

Klapwijk, P.M., Scheulderman, T., and Schilperoort, R.A. (1978). Coordinated regulation of octopine degradation and conjugative transfer of Ti plasmids in *Agrobacterium tumefaciens* - Evidence for a common regulatory gene and separate operons. **Journal of Bacteriology** *136*, 775-785.

Lugtenberg, B.J., Chin, A.W.T.F., and Bloemberg, G.V. (2002). Microbe-plant interactions: principles and mechanisms. **Antonie van Leeuwenhoek** *81*, 373-383.

Lynch, M.J. (1987). The Rhizosphere (Chichester, UK: **Wyle Interscience**).

Matthysse, A.G., Marry, M., Krall, L., Kaye, M., Ramey, B.E., Fuqua, C., and White, A.R. (2005). The effect of cellulose overproduction on binding and biofilm formation on roots by *Agrobacterium tumefaciens*. **Molecular Plant-Microbe Interactions: MPMI** *18*, 1002-1010.

McCullen, C.A., and Binns, A.N. (2006). *Agrobacterium tumefaciens* and plant cell interactions and activities required for interkingdom macromolecular transfer. **Annual Review of Cell and Developmental Biology** *22*, 101-127.

Mendes, R., Kruijt, M., de Bruijn, I., Dekkers, E., van der Voort, M., Schneider, J.H., Piceno, Y.M., DeSantis, T.Z., Andersen, G.L., Bakker, P.A., *et al.* (2011). Deciphering the rhizosphere microbiome for disease-suppressive bacteria. **Science** *332*, 1097-1100.

Murray, J.D. (2011). Invasion by invitation: rhizobial infection in legumes. **Molecular Plant-Microbe Interactions: MPMI** *24*, 631-639.

Pappas, K.M. (2008). Cell-cell signaling and the *Agrobacterium tumefaciens* Ti plasmid copy number fluctuations. **Plasmid** *60*, 89-107.

Pitzschke, A., and Hirt, H. (2010). New insights into an old story: *Agrobacterium*-induced tumour formation in plants by plant transformation. **EMBO J** *29*, 1021-1032.

Raaijmakers, J.M., Vlami, M., and de Souza, J.T. (2002). Antibiotic production by bacterial biocontrol agents. **Antonie van Leeuwenhoek** *81*, 537-547.

Ream, W. (1989). *Agrobacterium tumefaciens* and Interkingdom Genetic Exchange. **Annual Review of Phytopathology** *27*, 583-618.

Richardson, A.E., and Simpson, R.J. (2011). Soil microorganisms mediating phosphorus availability update on microbial phosphorus. **Plant Physiology** *156*, 989-996.

Rodriguez-Navarro, D.N., Dardanelli, M.S., and Ruiz-Sainz, J.E. (2007). Attachment of bacteria to the roots of higher plants. **FEMS Microbiology Letters** *272*, 127-136.

Savka, M.A., Dessaux, Y., Oger, P., and Rossbach, S. (2002). Engineering bacterial competitiveness and persistence in the phytosphere. **Molecular Plant-Microbe Interactions: MPMI** *15*, 866-874.

Schindler, U., Sans, N., and Schroder, J. (1989). Ornithine cyclodeaminase from octopine Ti plasmid Ach5 - Identification, DNA-sequence, enzyme properties, and comparison with gene and enzyme from nopaline Ti plasmid-C58. **Journal of Bacteriology** *171*, 847-854.

Schmidt, E.L. (1979). Initiation of plant root-microbe interactions. **Annu Rev Microbiol** *33*, 355-376.

Shrawat, A.K., and Lorz, H. (2006). *Agrobacterium*-mediated transformation of cereals: a promising approach crossing barriers. **Plant Biotechnology Journal** *4*, 575-603.

Spanik, H.P., Kondorosi, A., and Hooykaas, P.J. (1998). The Rhizobiaceae (The Netherlands: **Kluwer Academic Publishers**).

Stravrinides, J. (2009). Origin and evolution of phytopathogenic bacteria. In plant pathogenic bacteria: Genomics and Molecular Biology, R.W. Jackson, ed. (UK: **Caister Academic Press**).

Tomlinson, A.D., and Fuqua, C. (2009). Mechanisms and regulation of polar surface attachment in *Agrobacterium tumefaciens*. **Current Opinion in Microbiology** *12*, 708-714.

Tzfira, T., and Citovsky, V. (2008). *Agrobacterium* (NY: **Springer**).

Valdivia, R.H., Lu, W., and Winans, S.C. (1991). Characterization of a putative periplasmic transport-system for octopine accumulation encoded by *Agrobacterium tumefaciens* Ti plasmid pTia6. **Journal of Bacteriology** *173*, 6398-6405.

von Lintig, J., Zanker, H., and Schroder, J. (1991). Positive regulators of opine-inducible promoters in the nopaline and octopine catabolism regions of Ti plasmids. **Mol Plant Microbe In** *4*, 370-378.

White, C.E., and Winans, S.C. (2008). The cell-cell communication system in *Agrobacterium tumefaciens*. In *Agrobacterium*: From Biology to Biotechnology, T. Tzfira, and V. Citovsky, eds. (**Springer**), pp. 593-622.

Wood, D.W., Setubal, J.C., Kaul, R., Monks, D.E., Kitajima, J.P., Okura, V.K., Zhou, Y., Chen, L., Wood, G.E., Almeida, N.F., Jr., *et al.* (2001). The genome of the natural genetic engineer *Agrobacterium tumefaciens* C58. **Science** *294*, 2317-2323.

Zanker, H., Lurz, G., Langridge, U., Langridge, P., Kreusch, D., and Schroder, J. (1994). Octopine and nopaline oxidases from Ti plasmids of *Agrobacterium tumefaciens* - Molecular analysis, relationship, and functional-characterization. **Journal of Bacteriology** *176*, 4511-4517.

Zanker, H., Vonlintig, J., and Schroder, J. (1992). Opine transport genes in the octopine (Occ) and nopaline (Noc) catabolic regions in Ti plasmids of *Agrobacterium tumefaciens*. **Journal of Bacteriology** *174*, 841-849.

Zhu, J., Oger, P.M., Schrammeijer, B., Hooykaas, P.J., Farrand, S.K., and Winans, S.C. (2000). The bases of crown gall tumorigenesis. **Journal of Bacteriology** *182*, 3885-3895.

Zupan, J., Muth, T.R., Draper, O., and Zambryski, P. (2000). The transfer of DNA from *Agrobacterium tumefaciens* into plants: a feast of fundamental insights. **Plant J** *23*, 11-28.

CHAPTER 2

A gene in the octopine catabolism operon can utilize S-methylmethionine

2.1 Abstract

Agrobacterium tumefaciens incites plant tumors that produce nutrients called opines, which are utilized by the bacteria during host colonization. Some strains encode an operon required for the catabolism of the opine octopine. This operon contains a gene of unknown function, *msh*, that has a limited similarity to a family of proteins involved in methionine synthesis. Here I show that Msh has an end-to-end amino acid similarity to MmuM, an *E. coli* protein. MmuM is a methyltransferase that directs the conversion of S-methylmethionine (SMM) and homocysteine to produce two molecules of methionine. I confirmed that Msh performed the same enzymatic reaction as its homolog. In addition, I showed that Msh preferentially used SMM over other known methyl donors. In light of these results, I speculate about the role of Msh in plant colonization.

2.2 Introduction

The *msh* gene has been puzzling ever since its discovery (Fuqua & Winans 1996). It is located near the middle of the octopine catabolism (*occ*) operon, suggesting a function in opine catabolism (Fig. 2.1). As discussed in Chapter 1, the *occ* operon is located in the Ti plasmid, and the genes within this operon have been intensively studied. In 1977, two research groups showed that opines induce specific genes required for their

transport and catabolism (Klapwijk et al 1977, Montoya et al 1977). The genes in the *occ* operon are involved in the internalization and utilization of opines, and that transcription of these catabolic genes is induced by cognate opines.

The localization of the *occ* operon in the Ti plasmid was identified by a construction of a cosmid library. The cosmid pVK261 containing 17.9 kb of the Ti plasmid DNA had the ability to catabolize octopine, and the catabolic enzymes were inducible by octopine in the host lacking the Ti plasmid (Knauf & Nester 1982, Stachel et al 1985). Habeeb and collaborators narrowed down the size of the *occ* operon, finding that octopine regulates the transcription of a ~14 Kb operon by using *lacZ* fusions to map the *occ* operon. In addition, it was found that in order to activate the transcription of the operon, octopine is required to bind to OccR, a transcription regulator (Habeeb et al 1991, Wang & Winans 1995).

Valdivia and collaborators defined four open reading frames (ORF), *occQ*, *occM*, *occP* and *occJ*, which were involved in the internalization of octopine. By sequence analysis, these genes showed similarity to proteins involved in active transport, and were especially similar to genes that direct the uptake of histidine, arginine, ornithine, and lysine. These four genes were expressed in *E. coli* and by adding radiolabelled octopine, it was shown that these proteins were permeases responsible for octopine uptake (Valdivia et al 1991). During a screening for octopine-regulated genes, *tra*R gene was identified and located as the last gene of the operon. TraR is a quorum sensing transcriptional regulator, which is important for cell-cell bacterial communication (Fuqua & Winans 1994). The Schroder group identified two genes, *ooxA* and *ooxB*, involved in

the octopine catabolism downstream of the octopine permeases (Fig. 2.1) (Zanker et al 1994). Additionally, downstream of the octopine catabolic genes, *ocd* was found, which encodes for an ornithine cyclodeaminase that converts L-ornithine to L-proline (Schindler et al 1989).

msh was discovered during a mapping of TraR-inducible promoters in the *occ* operon. Fuqua and collaborators found a putative promoter located upstream of an uncharacterized region of the Ti plasmid between *ocd* and *traR* genes (Fig. 2.1). The first ORF downstream of the putative promoter, *msh*, encodes a 316 amino acid protein. The predicted protein was noted to have a limited similarity to a family of proteins involved in methionine biosynthesis, hence the mnemonic methionine synthase homolog. (Fuqua & Winans 1996).

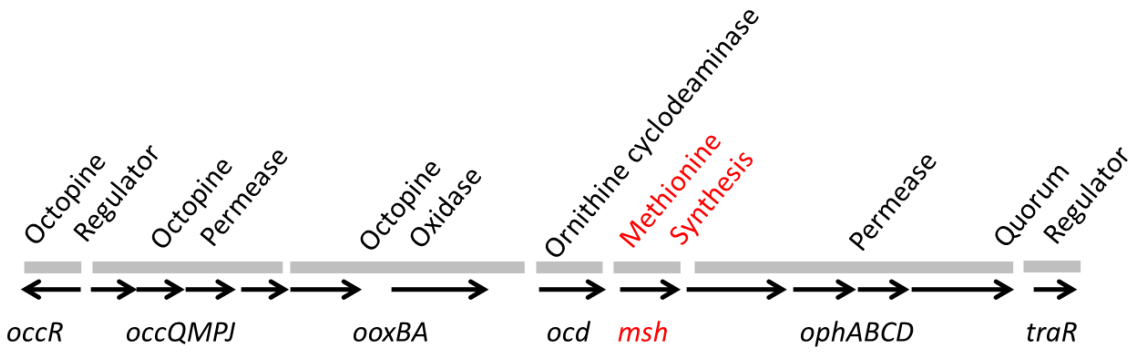


Figure 2.1 Genetic organization map of the genes within the *occ* operon.

In this study, I describe that Msh is a methyltransferase, which produces two methionines by transferring a methyl group from SMM to homocysteine. In addition Msh preferentially uses SMM as methyl donor, a non-proteinogenic amino acid found in a wide variety of plants.

2.3 Materials and methods

Bacterial strains, plasmids, and oligonucleotides

The bacterial strains-plasmids and oligonucleotides used in this study are listed in Table 2.1 and 2.2, respectively. *E. coli* strains were grown at 37°C in Luria Broth (LB) and Terrific Broth (TB) as complex media (Sambrook et al 1989). L-homocysteine and L-methionine were purchased from Sigma-Aldrich. L-SMM was produced by chemical synthesis (described below).

DNA manipulation

Recombinant DNA techniques were performed using established procedures (Sambrook et al 1989). PCR amplification of genes was done using Platinum *Taq* DNA Polymerase High Fidelity (Invitrogen). Plasmid DNA was isolated using QIAprep spin miniprep kits (Qiagen). DNA fragments generated by PCR or restriction digestion were gel purified using QIAquick Gel Extraction Kit (Qiagen). Restriction enzymes and other DNA modification reagents were purchased from New England Biolabs and used according to the methods described by the manufacturers. Plasmid DNA was introduced into *E. coli* by electroporation.

Table 2.1 Bacterial Strains and plasmid used in this study

Strains or plasmids	Relevant features	References
Strains		
BL21/DE3	*E. coli* P_{lac}-gene 1 of bacteriophage T7	(Studier et al 1990)
R10	*A. tumefaciens* R10	(Dessaux et al 1989)
Plasmids		
pMCSG19	P_{T7}-MBP-TVMV-his_{6}-TEV, Amp^{R}	(Donnelly et al 2006)
PT7-groE	P_{T7}-*gro*ESL, ColE1; Cm^{R}	(Yasukawa et al 1995)
pRK1037	P_{L}-tetO-TVMV protease, Km^{R}	(Donnelly et al 2006)
pAFM11	*msh* cloned into pMCSG19, Amp^{R}	This study

Plasmid construction

The *msh* gene was subcloned by PCR amplification using primers ALFM21 and ALFM27 inserted into pMCSG19 using ligation-independent cloning (LIC) (Donnelly et al 2006) (Table 2.2), resulting in pAFM11, which encodes an MBP-His_6-Msh fusion. Between MBP and His_6- tag, there is a recognition site for TVMV protease. pAFM11 and pT7-*groESL* were electroporated into *E. coli* BL21/DE3(pRK1037) (pRK1037 encodes TVMV protease), so the MBP portion of the tripartite fusion was removed immediately after protein synthesis. Plasmid pT7-*groESL* was provided to enhance accumulation of soluble protein (Chai & Winans 2009, Yasukawa et al 1995).

Table 2.2 Oligonucleotides used in this study

Oligonucleotide Name	DNA Sequence
ALFM21	5'-CGCGGATCCATGTCATCGAAAGTC-3'
ALFM27	5'-ATA GTTTAGCGGCCGCTC AGGCTGCGG C-3'

Chemical synthesis of SMM

L-SMM was synthesized using 1.5 g of L-methionine or D -methionine dissolved in 16 ml of 89% formic acid and 5 ml of acetic acid and combined with 2.5 ml of methyl iodide (Toennies & Kolb 1945). The reaction was incubated for 3 days at room temperature in the dark, and then evaporated to a syrup. Methanol (10 ml) was added to obtain granular particles, which were filtered and washed with methanol and acetone, and dissolved in 8 ml of warm 50% ethanol, and 25 ml of 100% ethanol was then added to allow crystallization in the dark. Crystals were filtered, washed and dried with acetone. Purity of L-SMM was determined using ESI MS/MS (Fig. 2.5) and NMR (Fig. 2.6).

Overproduction and purification of Msh

To overproduce Msh, *E. coli* strain BL21/DE3(pAFM11)(pRK1037)(pT7-*groESL*) was cultured at 37°C in 1 L of TB containing 400 µg/ml of ampicillin, 400 µg/ml of kanamycin, and 35 µg/ml of chloramphenicol until OD_{600} 0.6 was reached. The culture was cooled on ice to 28°C and overexpression of Msh was induced by adding 0.3 mM IPTG. Incubation was continued at 28°C for 5 additional hours. Cells were concentrated by centrifugation for 10 min at 4°C. The pellet was suspended in lysis

buffer (20 mM sodium phosphate buffer (pH 7.4), 200 mM NaCl, 20% glycerol, and 10 mM imidazole) and disrupted by passage through a French pressure cell (20,000 psi). The lysate was cleared by ultracentrifugation (10,000 x g at 4°C for 30 min). The supernatant was applied to Ni Sepharose^{TM} 6 Fast Flow (GE Lifescience) chromatography resin. The column was washed extensively using lysis buffer, and Msh was eluted using lysis buffer supplemented with 250 mM imidazole. Fractions containing Msh were combined and concentrated by using an Amicon Ultra cell with YM-30 filter membrane (30,000 MWCO; Millipore, Eschborn, Germany). During concentration the buffer was changed to 20 mM sodium phosphate buffer (pH 7.4), 200 mM NaCl, 20% glycerol and 1 mM DTT. Msh was further purified by gel filtration chromatography using a Superdex 200 column (GE Lifescience). Peak fractions were pooled and concentrated as described above, then Msh dialyzed using a storage buffer containing 20 mM Tris (pH 7.9), 50% glycerol and 1 mM DTT. Protein purity was analyzed using 12% SDS-PAGE gels and visualized by Coomassie staining.

Msh Activity Assays

Msh enzymatic reactions were carried out by using 0.61 µM His_6-Msh in 20 mM Tris buffer (pH 7.9), 10 mM homocysteine, and different concentrations of a variety of methyl donors such as SMM, SAM, methylcobalamine, betaine and, dimethylglycine. The reactions were stopped by adding an equal volume of 3:1 (v/v) methanol: 1% formic acid in water and centrifuged for 30 min to precipitate Msh. The supernatants were analyzed for methionine production by ESI MS/MS, and quantify using a standard curve. Enzyme kinetics were calculated using the initial velocities for various concentrations of

methyl donor. Initial velocities obtained from Msh enzymatic reactions were used to calculate the kinetics constants by using nonlinear least square analysis of the data fitted to the Michaelis-Menten and Lineweaver-Burk rate equations using SIGMA PLOT 9.0 (Systat Software, Ekrath, Germany) and the enzyme kinetic module 2.0.

Electrospray mass-spectroscopy and Nuclear Magnetic Resonance

Mass spectroscopy analysis was carried out using a Micromass Quattro II tandem MS operated in positive ion electrospray mode. Samples were injected directly using a syringe pump at a rate of 4 µl per min. Data acquisition and processing for the MS scans were controlled by the MassLynx software (Waters Corporation, Milford, MA). Twenty or more scans were averaged for each sample. When possible, spectra of all compounds were compared to those of commercial preparations of the same or similar compounds. NMR analysis was done by using a JEOL ECX-400 NMR spectrometer (Peabody, MA).

2.4 Results

The predicted Msh protein was previously found to have a limited sequence similarity to a family of proteins involved in methionine biosynthesis. However, by simply decreasing the gap penalty of the BLAST algorithm (Altschul et al 1990), I was able to detect an end-to-end sequence similarity between Msh and the MmuM protein of *E. coli* (Fig. 2.2 and 2.3).

A

BLASTP Parameters:

Expect Threshold: 10
Word size: 3
Max matches in a query range: 0
Matrix: BLOSUM62
Gap Costs: existence 10, extension 1*
Compositional Adjustments: Conditional compositional score matrix adjustment

*Default values are: existence 11, extension 1.

B

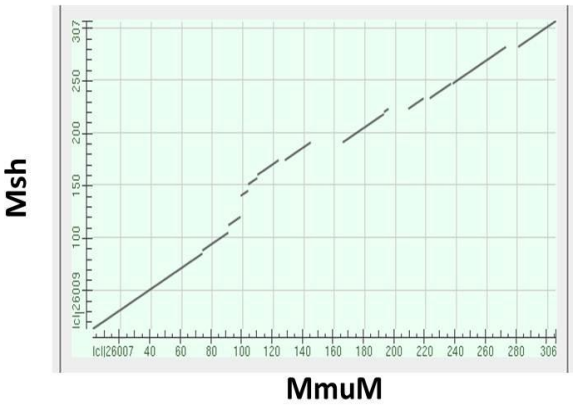


Figure 2.2 End-to-end sequence similarity between Msh and the MmuM protein of *E. coli*. A) BLAST parameters. B) Scatter plot of the amino acid comparison between Msh and MmuM.

The *mmu*M was found when studying a plant gene, *smt*A, which encodes for a selenocysteine methyltransferase involved in selenium tolerance (Neuhierl et al 1999). By sequence comparison with entries in the databases, it was found that an *E. coli* gene, *mmu*M, shared a 40% amino acid sequence identity. Likewise, *mmu*M showed low sequence similarity to the amino-terminal of MetH (methyltransferase) in *E. coli* and to a human betaine:homocysteine methyltransferase. MmuM enzymatic activity was elucidated, finding that it transfers a methyl group from S-methylmethionine (SMM) to homocysteine (Hcy) to yield two molecules of methionine (Met), and can also use selenohomocysteine in place of Hcy, hence MmuM stands for S-methylmethionine utilization,(Neuhierl et al 1999, Thanbichler et al 1999).

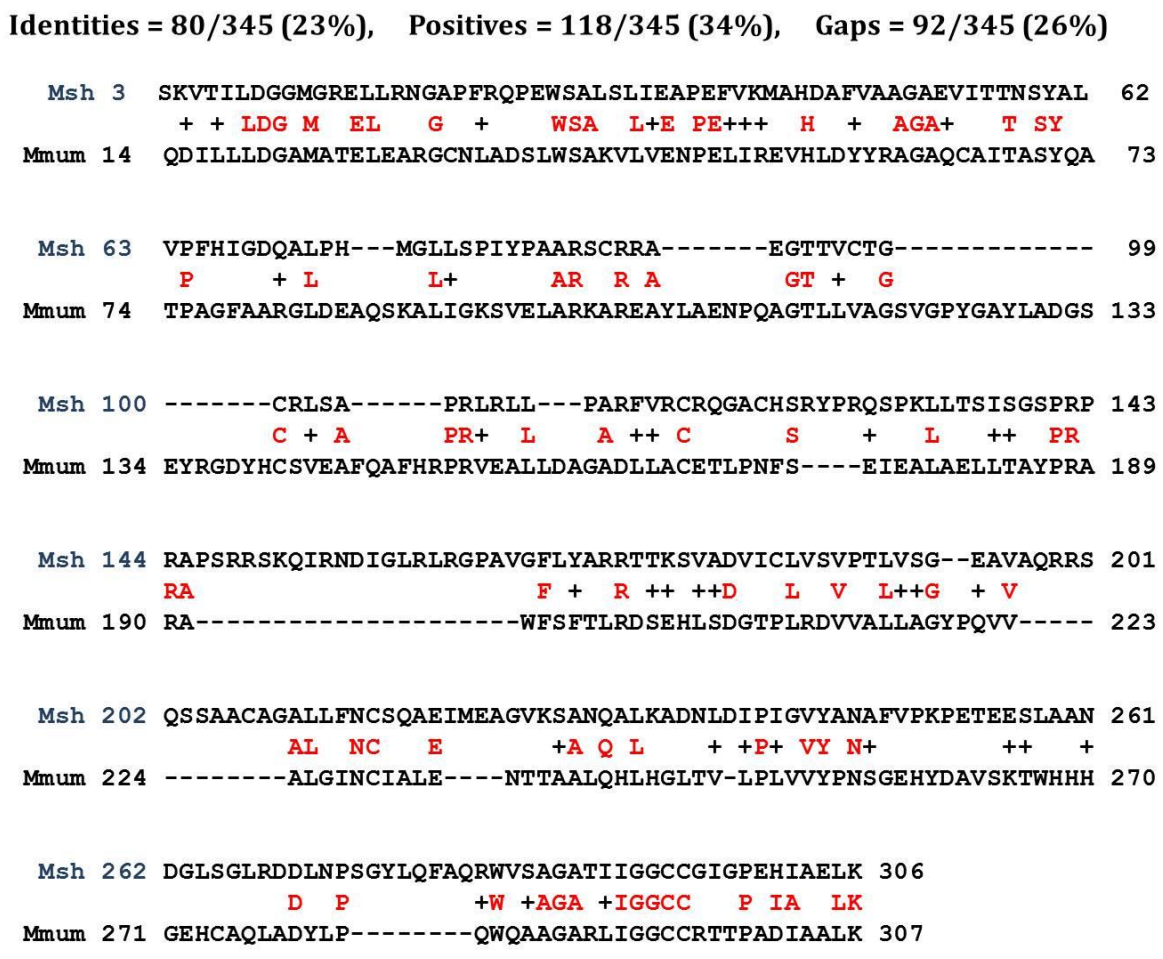


Figure 2.3. Msh and MmuM amino acid sequence alignment. Conserved amino acids are in red.

In order to determine whether Msh could convert SMM and Hcy to two equivalents of Met (Fig. 2.4). First, L-SMM was synthesized and its purity was analyzed by ESI-MS/MS (Fig. 2.5) and NMR (Fig. 2.6), this was done by collaborating with Dr. Anatol Eberhard. ESI-MS/MS analysis showed that SMM has a unique fingerprint fragment with a molecular weight of 102.

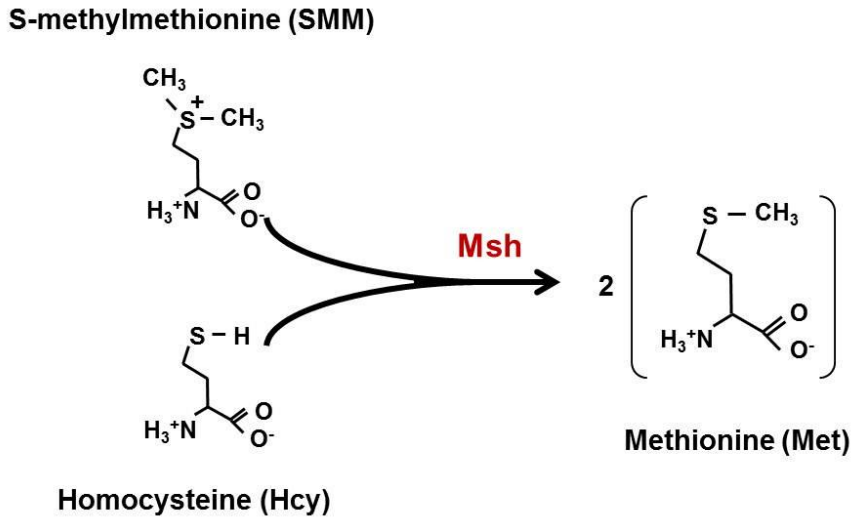


Figure 2.4 Proposed model for Msh methyltransferase activity. Msh uses SMM and homocysteine to produce two molecules of methionine.

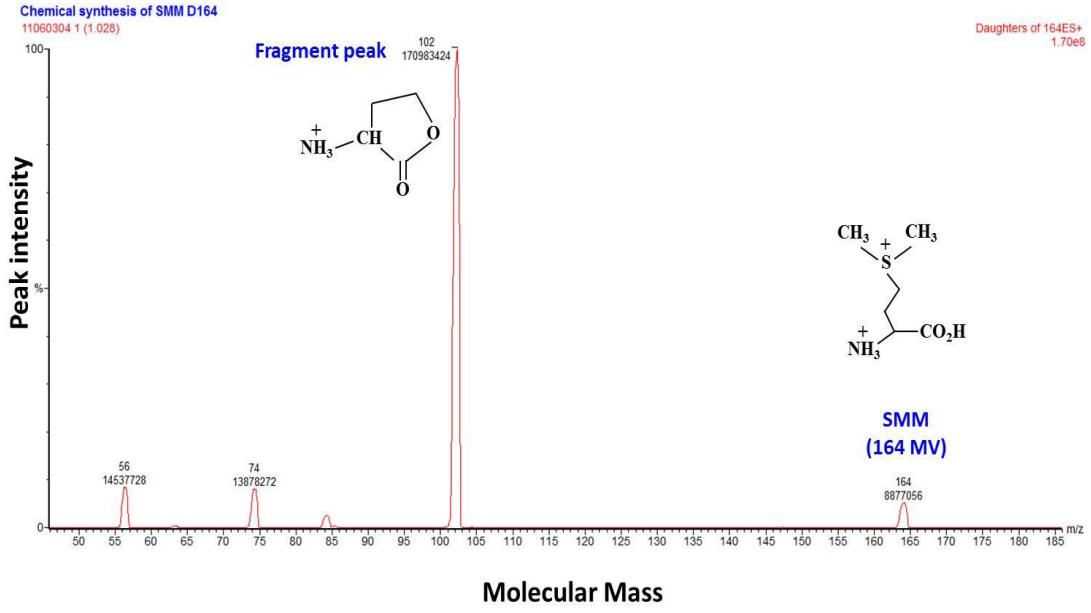


Figure 2.5 Authentication of the chemically synthesized SMM by ESI-MS/MS. The authentication of SMM was done by fragmenting it by adjusting the collision energy of the spectrometer Micromass Quattro to 11. It was found a finger print peak, which has a molecular weight of 102. x-axis is the molecular weight and y-axis is the peak intensity.

Subsequently, His_6-Msh was purified, however, early attempts resulted in low yield, prompting us to also co-overexpress GROESL, which resulted in better yield. I purified recombinant His_6-Msh and enzymatic reactions were performed by using homocysteine and SMM as substrates. Samples were taken at different time points during the enzymatic reaction and the substrates and products were analyzed by electrospray tandem mass spectrometry (ESI-MS/MS). I found that indeed Msh can utilize SMM and Hcy as substrates to produce methionine (Fig. 2.7 and Fig. 2.8).

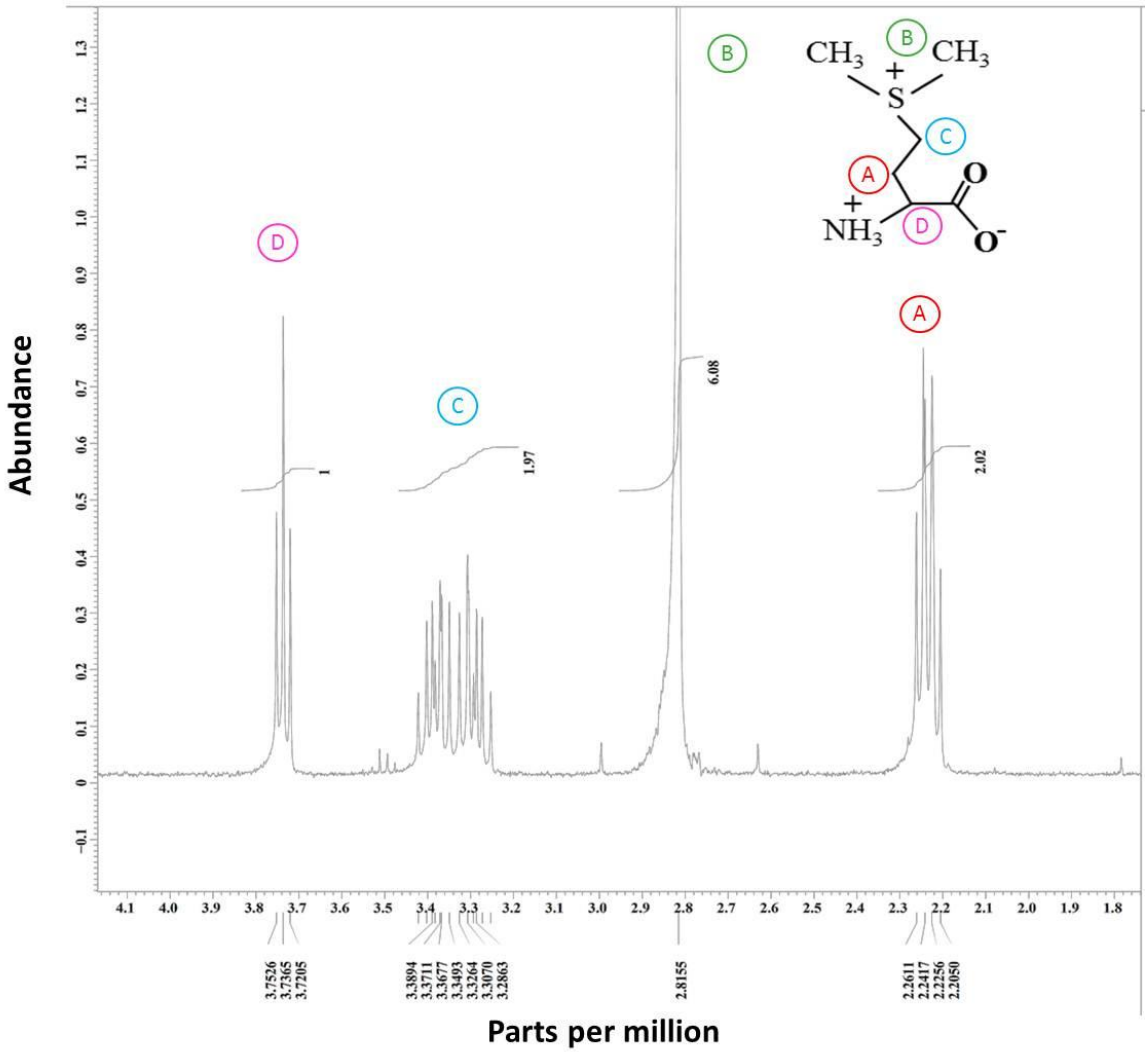


Figure 2.6 ^{1}H NMR spectrum of SMM in D_20. x-axis parts per Million. y-axis abundance. The letters indicated the position of the hydrogens in the SMM molecule. NMR analysis was done by using JEOL ECX-400 NMR spectrometer.

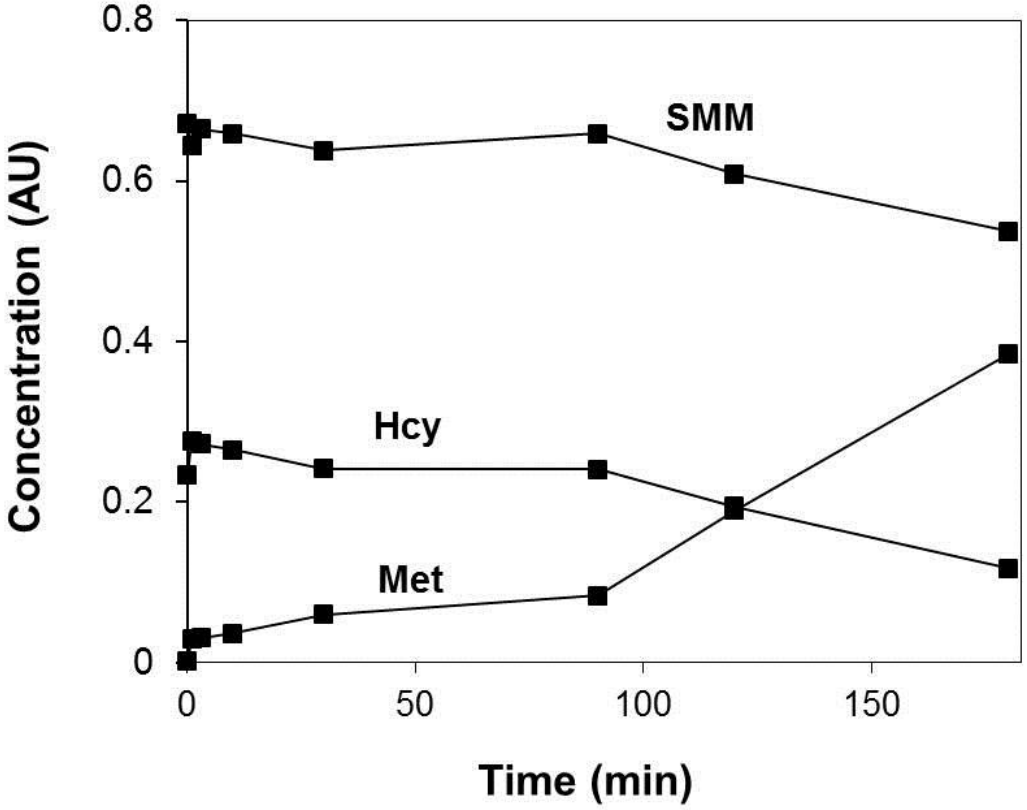


Figure 2.7 ESI-MS/MS analysis of the production of methionine by Msh from SMM and homocysteine. x-axis is the time of the reaction in min. y-axis is the relative concentration of the substrates and product which were normalized against the Tris buffer concentration.

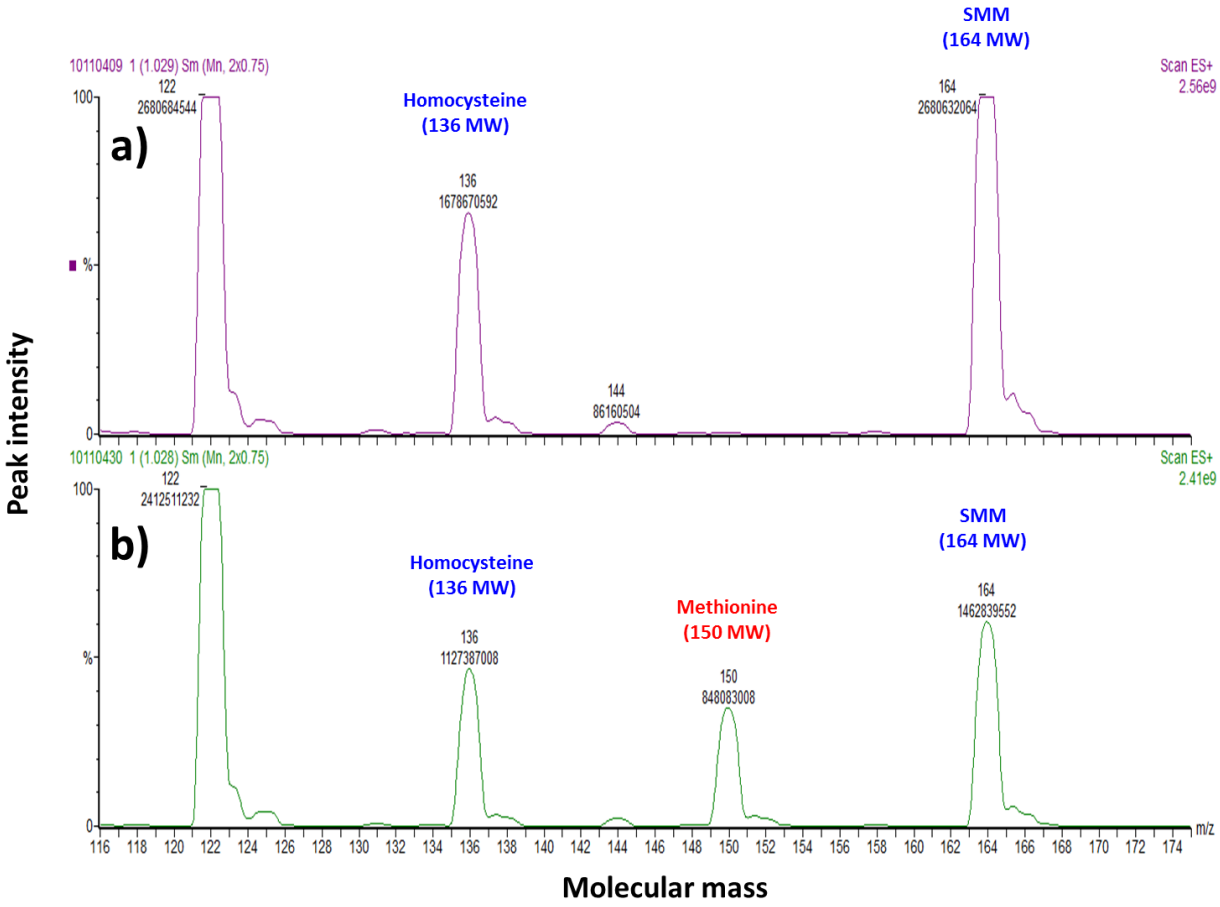


Figure 2.8 ESI-MS/MS detection of methionine in a Msh reaction. a) Time 0 and b) 1 hour. x-axis molecular weight and y-axis peak intensity. Peak 122 refers to Tris buffer from the enzymatic reaction.

Msh was able to use SMM as a methyl donor, however, the spectrum of substrates that can be used by Msh was unknown. To investigate Msh substrate specificity, enzymatic reactions were performed using different compounds known to be methyl donors such as *S*-adenosyl-methionine (SAM), dimethyl glycine, betaine and methylcobalamine (Ranocha et al 2000, Szegedi et al 2008). I found that the catalytic

specificity of Msh with SMM was 39-fold greater than with methylcobalamin, 272-fold greater than dimethyl glycine, and 775-fold greater than SAM, indicating that SMM is the preferred substrate for Msh (Table 2.3). These properties are similar to those of MmuM (Neuhierl et al., 1999; Thanbichler et al., 1999), suggesting that these two enzymes are functionally similar.

Table 2.3. Msh catalytic specificity for methyl donor substrates

Substrate	**k_m (mM)**	**k_{cat} (min^{-1})**	**Enzymatic efficiency (k_{cat}/k_m)**	**Relative activity (%)**
SMM	3.7 ± 2	114	31	(100)
Methylcobalamin	17 ± 8	14	0.8	2.7
Dimethylglycine	21 ± 8	2.4	0.114	0.4
SAM	18 ± 4	0.7	0.04	0.12
Betaine	>1.3 ± 0.6[α]	>0.04 [α]	<0.03 [α]	<0.9 [α]

[α] Detection limit.

Mean ± SD of n=3 enzymatic reactions.

2.5 Discussion

In *E. coli*, there are three known pathways for methionine biosynthesis; all involve methylation of homocysteine, but differ in the source of the methyl group. The MetE protein transfers the methyl group directly from *N*5-methyl-tetrahydrofolate (Gonzalez et al 1996). MetH, transfers the methyl group from N5-methyl-tetrahydrofolate to a cobalamine coenzyme and from there to homocysteine (Frasca et al 1988). The third

methionine synthase, MmuM transfers the methyl group from SMM, converting it to a second molecule of methionine (Thanbichler et al 1999).

The Msh protein of *A. tumefaciens* has limited sequence and structural homology with MetE and MetH, restricted to their N-terminal homocysteine binding domains (Evans et al 2004). In contrast, I found that Msh has end-to-end sequence similarity with MmuM of *E. coli*, which utilizes SMM to make methionine.

Msh clearly showed selectivity for SMM (Table 2.3). This noticeable selectivity for SMM has been also reported in other SMM-homocysteine methyltransferases such as MmuM found in *E. coli* (Thanbichler et al 1999), AtHMT-1 and 2 in *Arabidopsis* (Ranocha et al 2000), Mht1p in *S. cerevisiae* (Thomas et al 2000), and BHMT in human liver (Szegedi et al 2008).

SMM is a ubiquitous metabolite in many or possibly all plants, though it is not essential (Lee et al 2008). The methylation of methionine creates a strong positive charge on the sulfur atom. As methionine is a zwitterion with a rather hydrophobic side chain, its solubility in water is limited. The sulfonium ion of SMM causes it to be far more soluble in water than methionine. SMM is therefore the transported form of methionine, and is concentrated in floral tissues (Ranocha et al 2001).

The overall results showed that Msh has a strong specificity for SMM as a methyl donor substrate to produce methionine. This finding is intriguing since SMM is a plant produced molecule, therefore a big question is how *Agrobacterium tumefaciens* can take up this plant produced molecule to generate methionine. Moreover, why would this gene

be present in the *occ* operon? How a gene in the *occ* operon is involved in SMM utilization? This localization in the octopine catabolism operon suggests that *msh* might have a role in utilization of opines, therefore I hypothesized that SMM is an intermediate in opine catabolism.

By reviewing the *A. tumefaciens* plant infection process (see Chapter 1, Figure 1.4), it is known that the T-DNA encoded octopine synthase (Ocs) , which is expressed only in tumor cells, uses many basic amino acids as substrates, including arginine, lysine, ornithine, and histidine, as well as methionine and glutamine together with pyruvate to produce the corresponding opines. SMM is positively charged molecule and it has an amino acid moiety, therefore, the utilization of SMM together with pyruvate to synthesize the corresponding opine seems entirely plausible. If this hypothesis is true, the new opine, that I termed sulfonopine, would be the first described opine that contains sulfur, an essential bacterial nutrient.

The production *in vitro* and *in vivo* and release of sulfonopine as well as its utilization as a sulfur-source and as a signal molecule for regulation of the octopine catabolism operon is going to be discussed in the Chapter 3.

2.6 References

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. **Journal of Molecular Biology** 215: 403-10

Chai Y, Winans SC. 2009. The chaperone GroESL enhances the accumulation of soluble, active TraR protein, a quorum-sensing transcription factor from *Agrobacterium tumefaciens*. **Journal of Bacteriology** 191: 3706-11

Dessaux Y, Petit A, Ellis JG, Legrain C, Demarez M, et al. 1989. Ti Plasmid-Controlled Chromosome Transfer in *Agrobacterium tumefaciens*. **Journal of Bacteriology** 171: 6363-66

Donnelly MI, Zhou M, Millard CS, Clancy S, Stols L, et al. 2006. An expression vector tailored for large-scale, high-throughput purification of recombinant proteins. **Protein Expression and Purification** 47: 446-54

Evans JC, Huddler DP, Hilgers MT, Romanchuk G, Matthews RG, Ludwig ML. 2004. Structures of the N-terminal modules imply large domain motions during catalysis by methionine synthase. **Proc Natl Acad Sci U S A** 101: 3729-36

Frasca V, Banerjee RV, Dunham WR, Sands RH, Matthews RG. 1988. Cobalamin-dependent methionine synthase from *Escherichia coli* B: electron paramagnetic resonance spectra of the inactive form and the active methylated form of the enzyme. **Biochemistry** 27: 8458-65

Fuqua C, Winans SC. 1996. Localization of OccR-activated and TraR-activated promoters that express two ABC-type permeases and the traR gene of Ti plasmid pTiR10. **Mol Microbiol** 20: 1199-210

Fuqua WC, Winans SC. 1994. A LuxR-LuxI type regulatory system activates *Agrobacterium* Ti plasmid conjugal transfer in the presence of a plant tumor metabolite. **Journal of bacteriology** 176: 2796-806

Gonzalez JC, Peariso K, Penner-Hahn JE, Matthews RG. 1996. Cobalamin-independent methionine synthase from *Escherichia coli*: a zinc metalloenzyme. **Biochemistry** 35: 12228-34

Habeeb LF, Wang L, Winans SC. 1991. Transcription of the octopine catabolism operon of the *Agrobacterium* tumor-inducing plasmid pTiA6 is activated by a LysR-type regulatory protein. **Molecular plant-microbe interactions***: MPMI* 4: 379-85

Klapwijk PM, Oudshoorn M, Schilperoort RA. 1977. Inducible permease involved in the uptake of octopine, lysopine and octopinic acid by *Agrobacterium tumefaciens* strains carrying virulence-associated plasmids. **Microbiology** 102: 1-11

Knauf VC, Nester EW. 1982. Wide host range cloning vectors: a cosmid clone bank of an *Agrobacterium* Ti plasmid. **Plasmid** 8: 45-54

Lee M, Huang T, Toro-Ramos T, Fraga M, Last RL, Jander G. 2008. Reduced activity of *Arabidopsis thaliana* HMT2, a methionine biosynthetic enzyme, increases seed methionine content. **The Plant Journal: for cell and molecular biology** 54: 310-20

Montoya AL, Chilton MD, Gordon MP, Sciaky D, Nester EW. 1977. Octopine and nopaline metabolism in *Agrobacterium tumefaciens* and crown gall tumor cells: role of plasmid genes. **Journal of Bacteriology** 129: 101-7

Neuhierl B, Thanbichler M, Lottspeich F, Bock A. 1999. A family of S-methylmethionine-dependent thiol/selenol methyltransferases. Role in selenium tolerance and evolutionary relation. **The Journal of Biological Chemistry** 274: 5407-14

Ranocha P, Bourgis F, Ziemak MJ, Rhodes D, Gage DA, Hanson AD. 2000. Characterization and functional expression of cDNAs encoding methionine-sensitive and -insensitive homocysteine S-methyltransferases from *Arabidopsis*. **The Journal of biological chemistry** 275: 15962-8

Ranocha P, McNeil SD, Ziemak MJ, Li C, Tarczynski MC, Hanson AD. 2001. The S-methylmethionine cycle in angiosperms: ubiquity, antiquity and activity. **The Plant Journal: for cell and molecular biology** 25: 575-84

Sambrook J, Fritsch EF, Maniatis T. 1989. *Molecular cloning: a laboratory manual*. **Cold Spring Harbor**, NY Cold Spring Harbor Laboratory Press.

Schindler U, Sans N, Schroder J. 1989. Ornithine cyclodeaminase from octopine Ti plasmid Ach5: identification, DNA sequence, enzyme properties, and comparison with gene and enzyme from nopaline Ti plasmid C58. **Journal of Bacteriology** 171: 847-54

Stachel SE, An G, Flores C, Nester EW. 1985. A Tn3 *lac*Z transposon for the random generation of beta-galactosidase gene fusions: application to the analysis of gene expression in *Agrobacterium*. **The EMBO journal** 4: 891-8

Studier FW, Rosenberg AH, Dunn JJ, Dubendorff JW. 1990. Use of T7 RNA-Polymerase to direct expression of cloned genes. **Methods in Enzymology** 185: 60-89

Szegedi SS, Castro CC, Koutmos M, Garrow TA. 2008. Betaine-homocysteine S-methyltransferase-2 is an S-methylmethionine-homocysteine methyltransferase. **Journal of Biological Chemistry** 283: 8939-45

Thanbichler M, Neuhierl B, Bock A. 1999. S-methylmethionine metabolism in *Escherichia coli*. **Journal of Bacteriology** 181: 662-5

Thomas D, Becker A, Surdin-Kerjan Y. 2000. Reverse methionine biosynthesis from S-

adenosylmethionine in eukaryotic cells. **The Journal of Biological Chemistry** 275: 40718-24

Toennies G, Kolb JJ. 1945. Methionine studies VII. Sulfonium derivatives. **J. Am. Chem. Soc.** 67: 849-51

Valdivia RH, Wang L, Winans SC. 1991. Characterization of a putative periplasmic transport system for octopine accumulation encoded by *Agrobacterium tumefaciens* Ti plasmid pTiA6. **Journal of Bacteriology** 173: 6398-405

Wang L, Winans SC. 1995. The sixty nucleotide OccR operator contains a subsite essential and sufficient for OccR binding and a second subsite required for ligand-responsive DNA bending. **Journal of Molecular Biology** 253: 691-702

Yasukawa T, Kanei-Ishii C, Maekawa T, Fujimoto J, Yamamoto T, Ishii S. 1995. Increase of solubility of foreign proteins in *Escherichia coli* by coproduction of the bacterial thioredoxin. **The Journal of Biological Chemistry** 270: 25328-31

Zanker H, Lurz G, Langridge U, Langridge P, Kreusch D, Schroder J. 1994. Octopine and nopaline oxidases from Ti plasmids of *Agrobacterium tumefaciens*: molecular analysis, relationship, and functional characterization. **Journal of Bacteriology** 176: 4511-7

CHAPTER 3

Sulfur acquisition by *A. tumefaciens* via sulfonopine, a novel sulfur-containing opine produced in the infected plants

3.1 Abstract

Agrobacterium tumefaciens induces tumors in infected plants by introducing a piece of DNA (T-DNA) into the plant cell chromosome. T-DNA encodes for an enzyme, Octopine synthase (Ocs), that produces specialized nutrients, opines, that are rich in nitrogen and carbon. This strategy is a great advantage for survival of the bacterium in oligotrophic soils. However sulfur, one of the essential elements for life, is present in limiting amounts in the environment, and how *A. tumefaciens* manages sulfur nutritional stress was not understood. In this study I demonstrate that Ocs utilizes S-methylmethionine, an abundant plant produced compound, together with pyruvate to produce a novel sulfur containing opine, sulfonopine. This is the first described opine that is used as a sulfur-source by *A. tumefaciens*. I also characterized the molecular mechanism of sulfonopine synthesis and release from infected plants, as well as how sulfonopine serves as a nutrient and signal molecule for gene regulation once taken up by *A. tumefaciens*.

3.2 Introduction

Agrobacterium tumefaciens provides fascinating examples of the chemical ecology that underlies the interactions between plant-associated bacteria and their hosts (Brencic & Winans 2005, Ranocha et al 2001, Zhu et al 2000). This bacterium uses a conjugation-like machinery to transfer discrete fragments of oncogenic DNA from the tumor-inducing (Ti) plasmid to the nuclei of host plants (Alvarez-Martinez & Christie 2009, Gelvin 2009, Tzfira & Citovsky 2006). These DNA fragments are integrated into the host genome, and genes encoded within them are expressed in the plant nuclei using promoters that have evolved to function in plants even though these genes are not of plant origin. Several transferred genes direct the production of phytohormones such as auxin and cytokinin, which can lead to uncontrolled cell proliferation, producing a crown gall tumor (Kamada-Nobusada & Sakakibara 2009, Tzfira & Citovsky 2006). Other transferred genes direct the production of novel compounds called opines, which serve as nutrients for the colonizing bacteria (Savka et al 2002, White & Winans 2007). Bacteria catabolize opines via specialized permeases and catabolic enzymes that are encoded on non-transferred portions of the Ti plasmid (Zhu et al 2000).

Different strains of *Agrobacterium* collectively can direct plant tumors to synthesize a bewildering array of opines (Dessaux et al 1998). First, different strains transfer diverse types of opine biosynthetic genes. Second, each strain generally transfers more than one opine biosynthetic gene. Third, particular opine biosynthetic enzymes can sometimes utilize a variety of substrates. For example, *A. tumefaciens* strains that carry a so-called octopine type Ti plasmid, such as pTiA6, pTiB6, pTi15955, pTiR10, or

pTiAch5, transfer the opine biosynthetic genes *ocs*, *ags*, *mas1*, and *mas2*. Ocs can reductively conjugate pyruvate with several different amino acids, including arginine, lysine, ornithine, histidine, methionine, and glutamine, synthesizing the corresponding opines, all of which provide sources of carbon and nitrogen (Biemann et al 1960, Menage & Morel 1964, Menage & Morel 1965). The Mas1 and Mas2 enzymes conjugate mannose with glutamine, forming mannopine, which can be lactonized by Ags to form agropine or can spontaneously lactamize to form agropinic acid (Dessaux et al 1986, Dessaux et al 1988, Hong et al 1997). *A. tumefaciens* strains carrying a different Ti plasmid, pTiC58, direct plants to produce a completely different set of opines, including nopaline, which resembles octopine, but is synthesized using α-ketoglutarate in place of pyruvate (Goldmann et al 1969), and agrocinopines, which are disaccharides linked by phosphodiester bonds (Ryder et al 1984). Agrocinopines therefore provide sources of carbohydrate and phosphate but not nitrogen.

Opine utilization genes are generally located on non-transferred portions of the Ti plasmids, and are transcriptionally induced by the opines whose catabolism they direct (Dessaux et al 1998). Opine catabolic operons invariably include ABC-type uptake systems dedicated to importing a particular opine. These operons also contain at least one gene that directs the first step in the catabolism of the opine. Other steps in its catabolism may be encoded in the same operon or elsewhere in the genome. Some opine catabolic operons also include a *traR* gene, which encodes a LuxR-type transcription factor that directs the transcription of Ti plasmid vegetative replication and conjugative transfer genes (Fuqua & Winans 1996a, Piper et al 1999). TraR functions only in the presence of 3-oxooctanoylhomoserine lactone (OOHL), whose synthesis is directed by

TraI, also encoded on the Ti plasmids (Fuqua and Winans, 1994; Piper et al., 1993). The result is that some opines cause expression of TraR, which leads to quorum-dependent conjugation and increased rates of Ti plasmid vegetative replication.

Many of the ideas described above are exemplified by the octopine catabolism *(occ)* operon of octopine-type Ti plasmids. This operon contains 16 genes, and is activated by the product of the divergent *occR* gene in response to octopine and similar opines (Habeeb et al 1991, von Lintig et al 1991, Wang et al 1992). Opines are not only important as nutrients but also as signal molecules since quorum sensing and conjugation occur only within or near colonized plants, as they are the only natural source of these opines in the rhizosphere (Fuqua & Winans 1996b).

In chapter 2, I described that *msh*, a gene from the octopine catabolism operon, can utilize S-methylmethionine (SMM) and homocysteine to produce methionine. However, SMM is a plant produced molecule. The location of *msh* in the *occ* operon indicates that it should be related to opine utilization. Reviewing the *A. tumefaciens* ecology, it is known that T-DNA encoded gene, *ocs*, uses positively charged amino acid and pyruvate to produce the corresponding opines. Therefore I hypothesize that Ocs can take SMM to produce a new sulfur-containing opine, sulfonopine. In my proposed model Ocs produces sulfonopine in infected plants, then is released and catabolized by *A. tumefaciens* colonizing bacteria (Fig. 3.1). The catabolized sulfonopine will generate SMM and pyruvate. Now SMM is available inside of the bacterial cell to be used by Msh to produce methionine (Fig. 3.1).

In this chapter I show that Ocs directs the synthesis of a novel, sulfur-containing opine derived from S-methylmethionine, a non-proteinogenic amino acid found in a wide variety of plants. I also characterized the molecular mechanism of sulfonopine synthesis and release from infected plants, as well as how sulfonopine serves as a nutrient and signal molecule for gene regulation once taken up by *A. tumefaciens*. This is a new opine that provides the bacteria with sulfur, an essential nutrient.

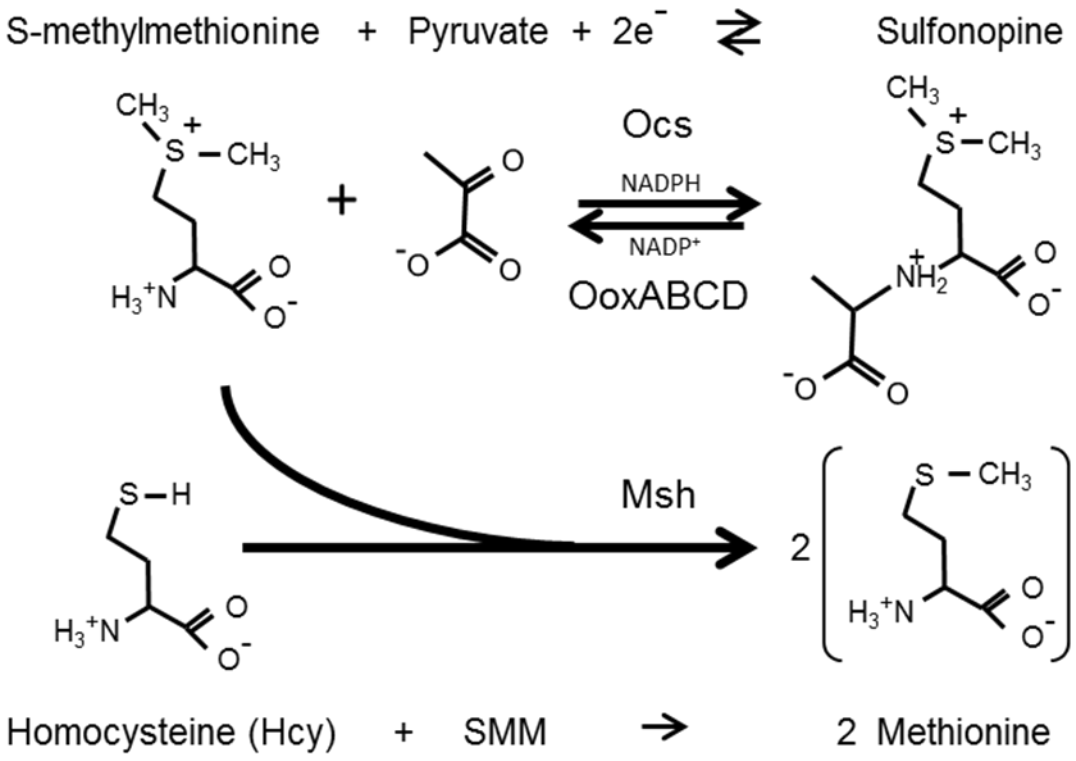


Figure 3.1 Synthesis and utilization of sulfonopine. Proposed synthesis of sulfonopine in plant cells by Ocs and its metabolism in bacterial cells by OoxABCD (top). SMM, derived from sulfonopine, is demethylated by Msh, yielding methionine (bottom).

3.3 Materials and methods

Bacterial strains, plasmids, and oligonucleotides

The bacterial strains/plasmids and oligonucleotides used in this study are listed in Table 3.1 and 3.2, respectively. Luria Broth (LB) and Terrific Broth (TB) were used as complex media while AT and AB media were used as defined media (Cangelosi et al 1991, Tartoff & Hobbs 1987). A modified form of AB medium lacking sulfur was prepared to test alternative sulfur sources. *E. coli* strains were grown at 37°C and *Agrobacterium tumefaciens* strains at 27°C. Octopine, amino acids, NADPH, NADH, and α-keto acids were purchased from Sigma-Aldrich. Sulfonopine was produced either by chemical synthesis or enzymatic synthesis and other opines were enzymatically synthesized (described below).

DNA manipulation

Recombinant DNA techniques were performed using established procedures (Sambrook et al 1989). PCR amplification of genes was done using Platinum *Taq* DNA Polymerase High Fidelity (Invitrogen). Plasmid DNA was isolated using QIAprep spin miniprep kits (Qiagen). DNA fragments generated by PCR or restriction digestion were gel purified using QIAquick Gel Extraction Kit (Qiagen). Restriction enzymes and other DNA modification reagents were purchased from New England Biolabs and used according to the methods described by the manufacturers. Plasmid DNA was introduced into *E. coli* and *A. tumefaciens* by electroporation (Cangelosi et al 1991). Plasmids and bacterial strains are listed in Table 3.1, and primers are described in Table 3.2.

Table 3.1. Bacterial Strains and plasmid used in this study.

Strains or plasmids	Relevant features	References
Strains		
BL21/DE3	*E. coli* P_{lac}-gene 1 of bacteriophage T7	(Studier et al., 1990)
S17-1/λpir	RK2, *tra* regulon, *pir*, host for *pir*-dependent plasmid	(Simon et al., 1983)
R10	*A. tumefaciens* R10	(S. K. Farrand)
KYC55	*A. tumefaciens* R10, Ti plasmid less	(Cho et al. 1997)
KYC16	R10 (*ooxA*::Tn5*gus*A7), Km^R	(Cho et al. 1996)
ALFM20	R10::pAFM110, *ocs*-lacZ, Km^R	This study
ALFM21	R10::pAFM111, *virD4*-lacZ, Km^R	This study
Plasmids		
pVIK107	*lacZY* for translational fusions, Km^R	(Kalogeraki and Winans, 1996)
pMCSG19	P_{T7}-MBP-TVMV-his_{6}-TEV, Amp^R	(Donnelly et al. 2005)
PT7-groE	P_{T7}-*groESL*, ColE1; Cm^R	(Yasukawa et al., 1995)
pRK1037	P_L-tetO-TVMV protease, Km^R	(Donnelly et al. 2005)
pAFM04	*ocs* cloned into pMCSG19, Amp^R	This study
pAFM110	pVIK107, *ocs* internal fragment, in-frame translational fusion	This study
pAFM111	pVIK107, *virD4* internal fragment, in-frame translational fusion	This study

Plasmid construction

The *ocs* gene was PCR amplified by using primers ALFM28 and ALFM29, and inserted into pMCSG19 using ligation-independent cloning (LIC) (Donnelly et al 2006), resulting in pAFM04, which encodes MBP-His_6-Ocs fusion. Between MBP and His_6-tag, there is a recognition site for TVMV protease. pAFM04 and pT7-*groESL* were electroporated into *E. coli* BL21/DE3(pRK1037) (pRK1037 encodes TVMV protease),

so the MBP portion of the tripartite fusion was removed immediately after protein synthesis. Plasmid pT7-*groESL* was provided to enhance accumulation of soluble protein (Chai & Winans 2009).

Table 3.2. Oligonucleotides used in this study.

Oligonucleotide Name	DNA Sequence
ALFM28	5'-TACTTCCAATCCAATGCAATGGCTAAAGTGGCA-3'
ALFM29	5'-TTATCCACTAATTCAAACTCCATTGAG-3'
ALFM218	5'-GGGGTACCTCGGTCCTCGTAGCATTGCCC-3'
ALFM219	5'-CGCGGATCCTCTTGGAGTTTCGATATCAGC-3'
ALFM220	5'-GGGGTACCATTGGCGAAATGCAGCATGCT-3'
ALFM221	5'-CGCGGATCCTAGGCTTTCCTCCGCGAGTTG-3'

Chemical synthesis of SMM

L-SMM and D-SMM were synthesized using 1.5 g of L-methionine or D -methionine dissolved in 16 ml of 89% formic acid and 5 ml of acetic acid and combined with 2.5 ml of methyl iodide (Toennies & Kolb 1945). The reaction was incubated for 3 days at room temperature in the dark, and then evaporated to a syrup. Methanol (10 ml) was added to obtain granular particles, which were filtered and washed with methanol and acetone, and dissolved in 8 ml of warm 50% ethanol, and 25 ml of 100% ethanol was

then added to allow crystallization in the dark. Crystals were filtered, washed and dried with acetone. Purity of L-SMM and D-SMM was determined using ESI MS/MS and NMR (Chapter 2, Figures 2.3 and 2.4).

Chemical synthesis of sulfonopine

The chemical synthesis of sulfonopine was done in two steps: 1) reductive condensation of methionine with pyruvate to yield methiopine, and 2) methylation of methiopine to yield sulfonopine. In the first step, 2 g of methionine methyl ester-HCl (Sigma-Aldrich) was combined with 10 ml of water and 1 M NaOH was added to reach pH 10.0 (~10 ml). The free base was extracted using chloroform and dried using Na_2SO_4. A solution of 35 ml of chloroform containing 1.5 g of methyl pyruvate and 4 g of sodium triacetoxyborohydride was added and incubated overnight under slight N_2 pressure at room temperature. Ten ml of $NaHCO_3$ solution (pH ~7.9), was then added and mixed until bubble formation ceased. The mixture was transferred to a separatory funnel, 20 ml of 1M NaOH was added, and extracted three times using 50 ml of CH_2Cl_2. The organic phases were combined and dried using MgSO4. The dried extract was dissolved in 3 ml of MeCN and then 2 ml of water was added. After centrifugation, the supernatant was subjected to preparative HPLC using a 1x25 cm C-18 column (Phenomenex) and eluted with an acetonitrile-water gradient. Testing with ESI-MS/MS showed two major peaks to correspond to the two expected diastereomers of methiopine dimethyl ester. The pooled fractions of each diastereomer were evaporated to dryness, redissolved in 10 ml of 1M HCl, hydrolyzed by heating for 15 hours using an autoclave (120°C), and dried *in vacuo* to remove the HCl. The purity of each diastereomer was checked by ESI-MS/MS.

Conversion of the methiopine diastereomers to those of sulfonopine was carried out by combining 3 mmoles of free methiopine, 15 ml of formic acid and 5 ml of acetic acid, and adding 1.25 ml of methyl iodide. The reactions were incubated for 3 days at room temperature in the dark, and then evaporated to syrups. Methanol (10 ml) was added to obtain granular particles. Particles were filtered and washed with methanol and acetone, and dissolved in 8 ml of warm 50% ethanol, and 25 ml of 100% ethanol was added. The solutions were incubated in the dark to allow formation of sulfonopine crystals. These were filtered, washed, and dried with acetone. The purity of the sulfonopine diastereomers was checked by ESI-MS/MS (Fig. 3.1) and NMR (Fig. 3.2). Only one of the diastereomers was found to have biological activity (Fig. 4.1).

Mutagenesis of *ocs* and *virD4* by Campbell-type integration

Plasmids pAFM110 and pAFM111 are suicide plasmids containing 0.5 kb internal fragments of *ocs* and *virD4*, respectively. These fragments were PCR amplified using ALFM218 and ALFM219 (for *ocs*) and ALFM220 and ALFM221 (for *virD4*), and inserted between the *Kpn*I and *Bam*HI sites of the suicide plasmid pVIK107 (Kalogeraki & Winans 1997). The resulting plasmids, pAFM110 and pAFM111, respectively, were transferred into strain R10 by conjugation (Kalogeraki & Winans 1997), and transconjugants were selected using AB minimal agar plates containing 200 µg/ml of kanamycin. Since these plasmids cannot replicate in *A. tumefaciens*, they conferred antibiotic resistance by Campbell-type integration into *ocs* and *vir*D4, respectively, creating null mutations in each gene. Campbell-type integration was confirmed by PCR amplification.

Overproduction and purification of Ocs

To overproduce *Ocs*, *E. coli* strain BL21/DE3(pAFM04)(pRK1037)(pT7-*groESL*) was cultured at 37°C in 1 L of TB containing 400 µg/ml of ampicillin, 400 µg/ml of kanamycin, and 35 µg/ml of chloramphenicol until OD_{600} 0.6 was reached. The culture was cooled on ice to 28°C and overexpression of Ocs was induced by adding 0.3 mM IPTG. Incubation was continued at 28°C for 5 additional hours. Cells were concentrated by centrifugation for 10 min at 4°C. The pellet was suspended in lysis buffer (20 mM sodium phosphate buffer (pH 7.4), 200 mM NaCl, 20% glycerol, and 10 mM imidazole) and disrupted by passage through a French pressure cell (20,000 psi). The lysate was cleared by ultracentrifugation (10,000 x g at 4°C for 30 min). The supernatant was applied to Ni Sepharose™ 6 Fast Flow (GE Lifescience) chromatography resin. The column was washed extensively using lysis buffer, and Ocs was eluted using lysis buffer supplemented with 250 mM imidazole. Fractions containing Ocs were combined and concentrated by using an Amicon Ultra cell with YM-30 filter membrane (30,000 MWCO; Millipore, Eschborn, Germany). During concentration the buffer was changed to 20 mM sodium phosphate buffer (pH 7.4), 200 mM NaCl, 20% glycerol and 1 mM DTT. Ocs was further purified by gel filtration chromatography using a Superdex 200 column (GE Lifescience). Peak fractions were pooled and concentrated as described above and dialyzed using a buffer containing 150 mM sodium phosphate buffer (pH 6.6), 50% glycerol, and 1 mM DTT. Protein purity was analyzed using 12% SDS-PAGE gels and visualized by Coomassie staining.

Ocs Activity Assays

Enzymatic assays were performed by using 10 mM sodium pyruvate, 0.3 mM NADPH and different concentrations of amino acids in 150 mM PIPES (pH 6.6) and 0.08 µM of His_6-Ocs. The reaction was started by the addition of the amino acid to be tested and followed by measuring the oxidation of NADPH spectrophotometrically at 340 nm at room temperature by using a Synergy™ HT multi-detection microplate reader (Biotek Intruments). The concentrations of NADPH in solution were calculated on the basis of the molar absorption coefficient 6200 M^{-1} cm^{-1} at 340 nm. Production of opines in the enzymatic mixture was confirmed by ESI MS/MS. Enzyme kinetics were calculated using the initial velocities. Initial velocities obtained from Ocs enzymatic reactions were used to calculate the kinetics constants by using nonlinear least square analysis of the data fitted to the Michaelis-Menten and Lineweaver-Burk rate equations using SIGMA PLOT 9.0 (Systat Software, Ekrath, Germany) and the enzyme kinetic module 2.0.

Enzymatic synthesis of opines

Octopine-type opines were produced using purified Ocs and glucose dehydrogenase, which regenerates NADPH from $NADP^+$ at the expense of glucose (Weckbecker & Hummel 2005). Enzymatic assays were performed using 150 mM sodium pyruvate, 90 mM of the amino acid substrate, 10 mM NADPH, 150 mM glucose, 150 mM sodium phosphate (pH 6.6), 0.08 µM of His_6-Ocs and 2 units of glucose dehydrogenase. Reactions were incubated at 28°C for 24 hours. Enzymes were removed by methanol precipitation and filtration (Amicon Ultra filters, Millipore; cut-off 3,000). The concentrations of reactants and products were determined using ESI MS/MS.

Plant cultivation and inoculation

Arabidopsis thaliana wild-type Col-O, and mutants *hmt2-2* and *mmt-2* (Lee et al 2008) were obtained as seeds from G. Jander (Boyce Thompson Institute). *Nicotiana tabacum* seeds were obtained from A. Colmer (Cornell University). Seeds were surface-sterilized by soaking in 50% bleach and 0.1% SDS for 10 min, followed by extensive washing in sterile water. Seeds were transferred to sterile water in Petri dishes and incubated for 48 h at 4°C in the dark. *A. thaliana* sterilized seeds were transferred to trays containing sterilized soil. Plants were maintained at room temperature with exposure to natural and artificial lighting.

Prior to plant infection, *A. tumefaciens* strains were cultured overnight in AT medium. Cells were then washed with sterile water and suspended to OD_{600} of 0.5 in a solution containing 5% sucrose and 0.005% of Triton-X100. *A. thaliana* plants cultivated in soil were infected by the floral dip method (Zhang et al 2006). Plants were then cultivated for 2 weeks. Floral bolts were collected and homogenized using a Potter-Elvehjem homogenizer. The homogenate was centrifuged for 5 min at 10,000 x g at 4°C and the supernatant was evaporated to dryness under a vacuum. The resulting solid was resuspended in 3:1 (v/v) methanol: 1% formic acid in water and centrifuged. The clear supernatant was analyzed by ESI MS/MS.

Approximately 200 *N. tabacum* seeds were germinated in water supplemented with MS salts (Murashige 1962). When the seedlings were 1-2 cm in length (approximately 2 weeks), they were submerged in the bacterial suspensions for 5 min, washed gently in 5% sucrose, and transferred to MS salts medium. Three days later, the

infected seedlings were extensively washed with medium containing 300 µg per ml of carbenicillin, and incubated for 1 day in medium containing carbenicillin, then transferred to water and incubated for 2 weeks. Seedlings and water were collected for ESI MS/MS analysis. Seedling extracts were obtained following the same protocol as for *A. thaliana* described above. Exudates were obtained by evaporating the water to dryness and resuspending in 3:1 (v/v) methanol: 1% formic acid in water for further ESI MS/MS analysis.

Electrospray mass-spectroscopy and Nuclear Magnetic Resonance

Mass spectroscopy analysis was carried out using a Micromass Quattro II tandem MS operated in positive ion electrospray mode. Samples were injected directly using a syringe pump at a rate of 4 µl per min. Data acquisition and processing for the MS scans were controlled by the MassLynx software (Waters Corporation, Milford, MA). Twenty or more scans were averaged for each sample. When possible, spectra of all compounds were compared to those of commercial preparations of the same or similar compounds. NMR analysis was done by using a JEOL ECX-400 NMR spectrometer (Peabody, MA).

Utilization of sulfonopine and other sulfur source compounds

Sulfonopine, SMM, and methionine were tested as sources of sulfur by culturing strains in modified AB broth lacking sulfur and supplemented with 1 mM of the tested compound as a sole sulfur source. Cultures were incubated at 28°C with vigorous aeration. Experiments were performed in triplicate using three different isolates for each strain.

Induction of the *occ* operon by opines

Opines were tested for induction of the *occ* operon using strain KYC16 (*ooxA-uidA*) cultured in AT medium containing 100 µg/ml of kanamycin at 27°C. β-glucuronidase specific activity was measured as described previously (Gallagher 1992).

To determine whether OccR is responsible for induction, KYC1203(pKY148) and KYC1203(pKY148)(pRJM101) were cultured in AT medium supplemented with 1 mM of opines under vigorous aeration at 27°C to an OD_{600} of 0.3-0.4, and assayed for β-galactosidase specific activity (Miller 1972).

3.4 Results

Octopine synthase utilizes SMM and pyruvate as substrates

The model proposed in the Introduction predicts that SMM is derived from a hypothetical opine, sulfonopine, which is synthesized by the Ocs protein that is expressed within crown gall tumors. To determine whether Ocs can synthesize sulfonopine, I constructed a gene encoding an MBP-His_6-Ocs fusion protein that contains a cleavage site for TVMV protease between MBP and the His_6 tag. This fusion was co-expressed with TVMV protease for *in vivo* removal of the MBP portion of the fusion (Donnelly et al 2006). An additional plasmid was added to overexpress the chaperone proteins GroESL, as this caused a significant increase in His_6-Ocs solubility. His_6-Ocs protein was purified to apparent homogeneity by sequential IMAC and gel filtration chromatography.

Enzymatic reactions were prepared containing Ocs, pyruvate, NADPH, and either SMM or arginine (Arg), and the reaction was monitored by using UV spectroscopy to detect the conversion of NADPH to $NADP^+$. I found that both amino acids stimulated NADPH oxidation (Table 3.3).

Table 3.3. Ocs enzymatic utilization of SMM and Arginine

Amino Acid	**α-keto acid**	**k_m (mM)**	**k_{cat} (min^{-1})**	**Enzymatic efficiency (k_{cat}/k_m)**	**Relative activity (%)**
L-SMM	Pyruvate	0.31 ± 0.04	49	160	76
L-Arginine	Pyruvate	3.5 ± 0.3	720	210	(100)

Mean ± SD of n=3 enzymatic reactions.

[α] Detection limit

The enzymatic reactions were analyzed by ESI-MS/MS to verify the production of octopine and sulfonopine. As a negative control a reaction containing pyruvate, NADPH and Ocs but not amino acid was used. I confirmed that reactions containing arginine produced octopine, which have molecular weights of 175 and 247 respectively, while the reaction containing SMM produced sulfonopine, which has molecular weights of 164 and 236 respectively (Fig. 3.2). The enzymatic efficiency (k_{cat}/k_m) for SMM was similar to that for arginine, indicating that Ocs utilized SMM with high efficiency (Table 3.3).

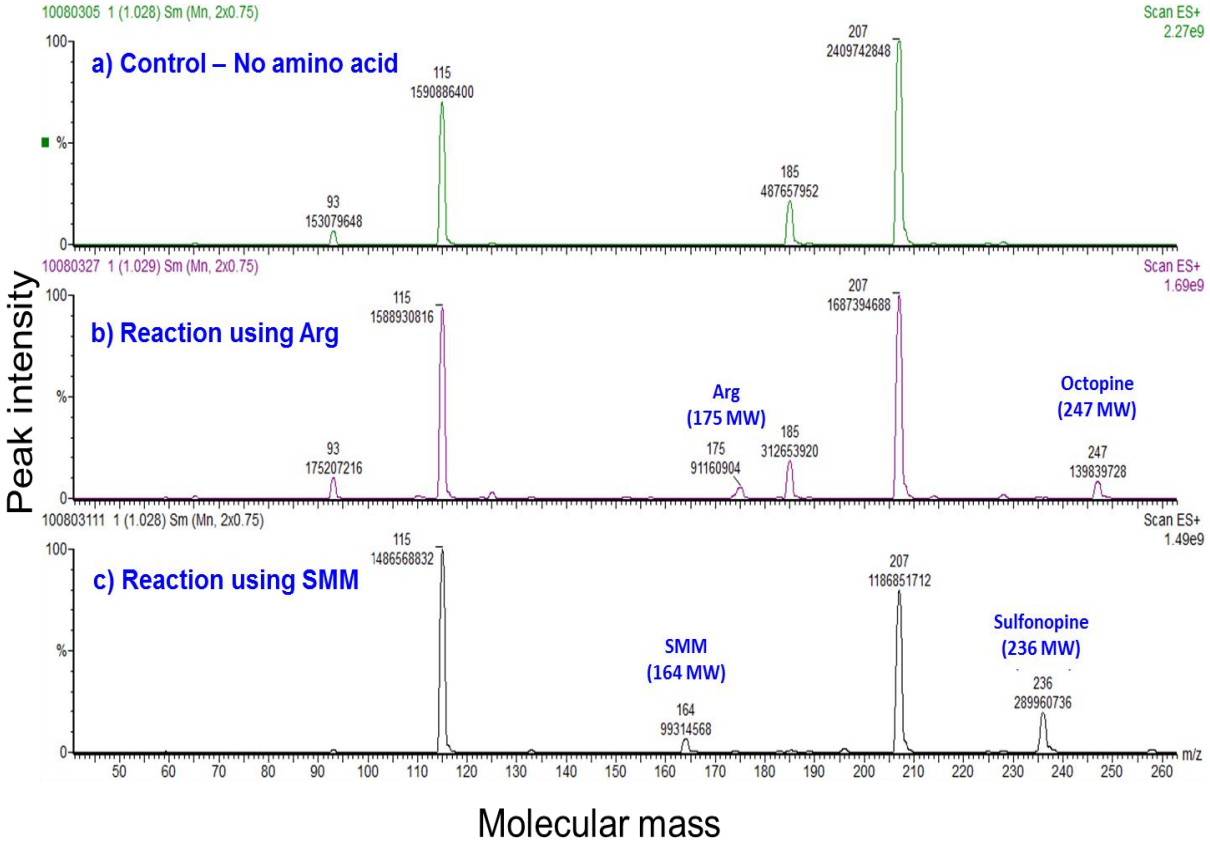


Figure 3.2 Detection of octopine and sulfonopine by ESI-MS/MS in the Ocs enzymatic reaction. a) Control enzymatic reaction- no amino acid. b) Reaction with arginine. c) Reaction with SMM. X-axis molecular weight and y-axis peak intensity. Peaks 93, 115 and 207 are compounds from the enzymatic reaction.

Previous reports showed that Ocs can utilized positively charged amino acids such as arginine, lysine, ornithine, histidine, as well as methionine and glutamine (Hack & Kemp 1977, Hack & Kemp 1980, Otten et al 1977). To better understand Ocs substrate specificity, I compared all 20 proteinogenic amino acids as well as several nonproteinogenic amino acids that are commonly found in plants (Bell 2003). I also wanted to know the ability of Ocs to use other α-keto acids in place of pyruvate. Additionally, the utilization of NADPH and NADH was compared. I found that Ocs can utilized SMM along with lysine, ornithine, histidine, methionine, and glutamine, confirming previous reports. In addition, Ocs can also utilize canavanine, homocysteine, homoserine, cysteine, glycine, and homoarginine, as well as others at reduced efficiencies (Table 3.4). Of these, histidine and canavanine were utilized even more efficiently than arginine. The enzyme did not detectably utilize aromatic or acidic amino acids, nor did it utilize proline, carnosine, or any of three tested D-amino acids.

The enzyme utilized α-ketobutyrate, oxaloacetate, and glyoxylate in place of pyruvate, although significantly less efficiently (Table 3.5). It did not detectably utilize α-ketoglutarate as a substrate. I confirmed earlier reports that Ocs utilizes NADPH slightly more efficiently than NADH (Otten et al 1977).

Table 3.4. Substrate specificity of octopine synthase for amino acids.

Amino Acid	α-keto acid	k_m (mM)	k_{cat} (min^{-1})	Enzymatic efficiency (k_{cat}/k_m)	Relative activity (%)
L-SMM	Pyruvate	0.31 ± 0.04	49	160	76
L-Arginine	Pyruvate	3.5 ± 0.3	720	210	(100)
L-Histidine	Pyruvate	2.0 ± 0.21	1200	600	286
L-Canavanine	Pyruvate	2.8 ± 0.5	1000	360	171
L-Homocysteine	Pyruvate	8.9 ± 1.1	1600	180	86
L-Ornithine	Pyruvate	3.0 ± 0.4	540	180	86
L-Lysine	Pyruvate	0.74 ± 0.08	130	178	85
L-Methionine	Pyruvate	27 ± 6	4700	170	81
L-Glutamine	Pyruvate	5.6 ± 0.6	780	140	67
L-Homoserine	Pyruvate	15 ± 3	1800	120	57
L-Cysteine	Pyruvate	11 ± 3	670	61	29
L-Glycine	Pyruvate	30 ± 6	1300	43	20
L-Homoarginine	Pyruvate	11 ± 2	410	37	18
L-Leucine	Pyruvate	11 ± 2	410	37	18
L-Serine	Pyruvate	12 ± 2	220	18	9
L-Alanine	Pyruvate	47 ± 14	340	7.2	3.4
L-Asparagine	Pyruvate	41 ± 13	220	5.4	2.6
L-Threonine	Pyruvate	976 ± 30	4600	4.7	2.2
L-Valine	Pyruvate	36 ± 5	170	4.7	2.2
L-Isoleucine	Pyruvate	70 ± 30	180	2.6	1.2
L-Glutamic acid	Pyruvate	$>1.5 \times 10^6$	$>6.9 \times 10^4$	<0.6	<0.09
L-Aspartic acid	Pyruvate	$>1.5 \times 10^6$	$>6.9 \times 10^4$	<0.6	<0.09
L-Tryptophan	Pyruvate	$>1.5 \times 10^6$	$>6.9 \times 10^4$	<0.6	<0.09
L-Tyrosine	Pyruvate	$>1.5 \times 10^6$	$>6.9 \times 10^4$	<0.6	<0.09
L-Phenylalanine	Pyruvate	$>1.5 \times 10^6$	$>6.9 \times 10^4$	<0.6	<0.09
L-Proline	Pyruvate	$>1.5 \times 10^6$	$>6.9 \times 10^4$	<0.6	<0.09
L-Carnosine	Pyruvate	$>1.5 \times 10^6$	$>6.9 \times 10^4$	<0.6	<0.09
D-Arginine	Pyruvate	$>1.5 \times 10^6$	$>6.9 \times 10^4$	<0.6	<0.09
D-Methionine	Pyruvate	$>1.5 \times 10^6$	$>6.9 \times 10^4$	<0.6	<0.09
D-SMM	Pyruvate	$>1.5 \times 10^6$	$>6.9 \times 10^4$	<0.6	<0.09

Mean ± SD of n=3 enzymatic reactions.
[α] Detection limit

Table 3.5. Substrate specificity of octopine synthase for α-keto acids

Amino Acid	α-keto acid	km (mM)	kcat (min-1)	Enzymatic efficiency (kcat/km)	Relative activity (%)
Arginine	Pyruvate	0.44 ± 0.07	561	1275	(100)
Arginine	α-ketobutyrate	6 ± 1	573	96	7.5
Arginine	Oxaloacetate	28 ± 4.5	1073	38	3
Arginine	Glyoxylate	22 ± 5	354	16	1.3
Arginine	α-ketoglutarate	>1.5 e^{+6} [α]	>6.9 e^{+4} [α]	<0.6 [α]	<0.09 [α]

Mean ± SD of n=3 enzymatic reactions.
[α] Detection limit

Chemical synthesis of sulfonopine

After this finding, in collaboration with Dr. Anatol Eberhard I chemically synthesized sulfonopine to further confirm the authenticity of the enzymatically produced compound. From the chemical synthesis of sulfonopine, two diastereomers were generated. These diastereomers were separated by HPLC, and checked by ESI-MS/MS (Fig. 3.3) and NMR (Fig. 3.4). Chemically synthesized sulfonopine has a molecular weight (MW) of 247 and when fragmented it generated to fingerprint peaks at 174 and 128 MW (Fig. 3.3). Subsequently, the sulfonopine diastereomers were tested for biological activity. For this experiment, I used the strain KYC16, which harbors an octopine-inducible *oox*A-uidA fusion. It is known that octopine induces the octopine catabolism operon. Therefore as a positive control I used octopine and as a negative control I used SMM. I found that only one of the diastereomers was able to induce the fusion as octopine did (Fig. 3.5). I designated sulfonopine as (N-(R-1-carboxyethyl)-S-methyl-S-methionine).

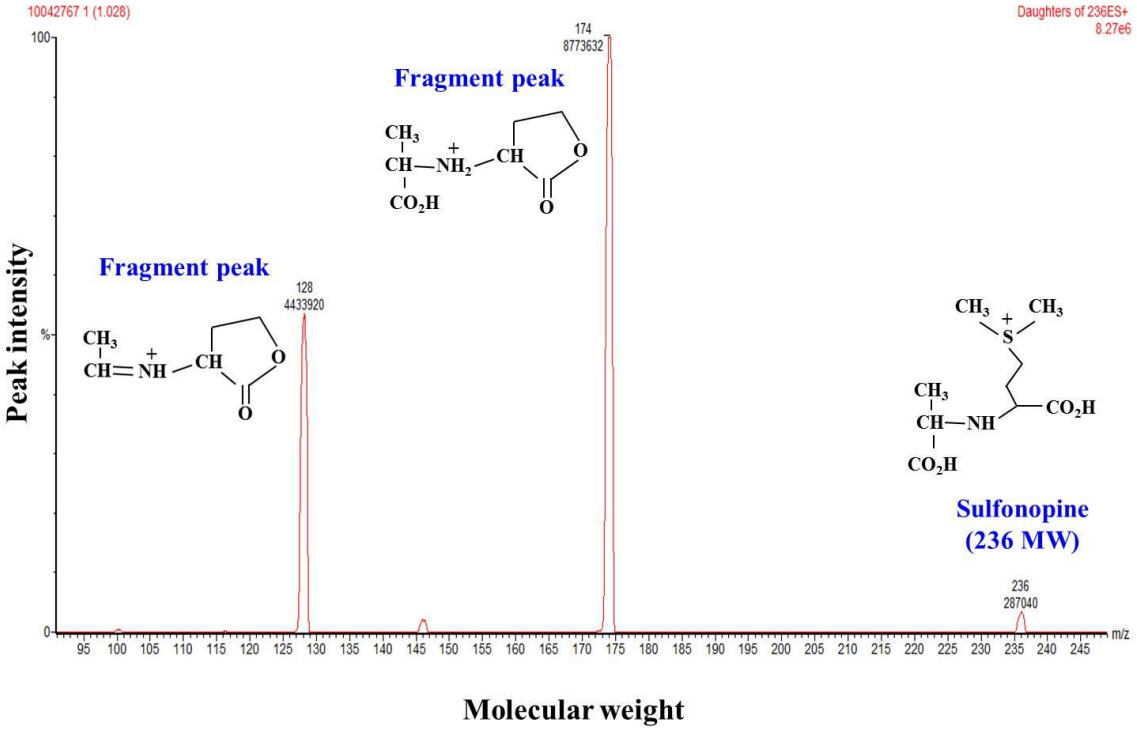


Figure 3.3 Analysis of the chemically synthesized sulfonopine by ESI-MS/MS. Sulfonopine was fragmented by adjusting the collision energy of the spectrometer Micromass Quattro to 11. Sulfonopine generated two fingerprint fragments with a molecular weight of 174 and 128. The x-axis shows the molecular weight and y-axis is the peak intensity.

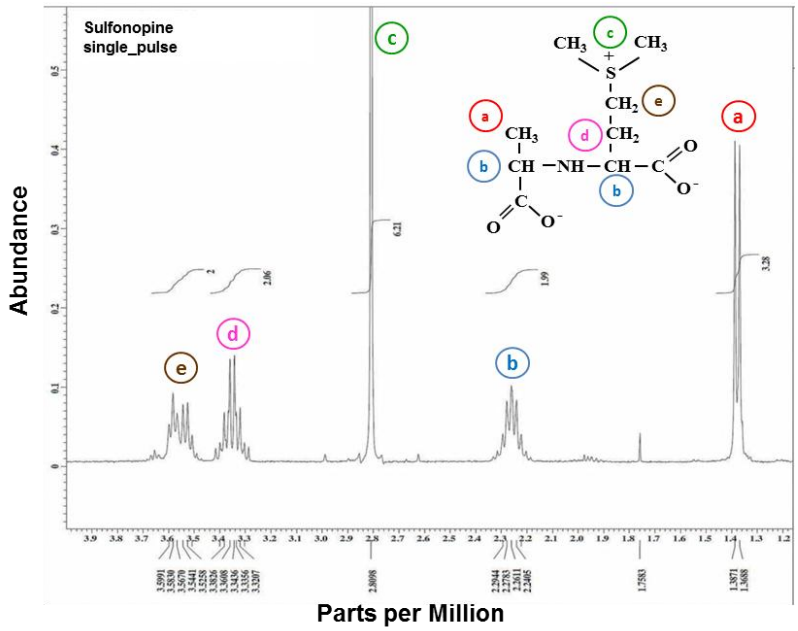


Figure 3.4 ^{1}H NMR spectrum of sulfonopine in D_20. x-axis parts per Million. y-axis abundance. NMR analysis was done by using JEOL ECX-400 NMR spectrometer.

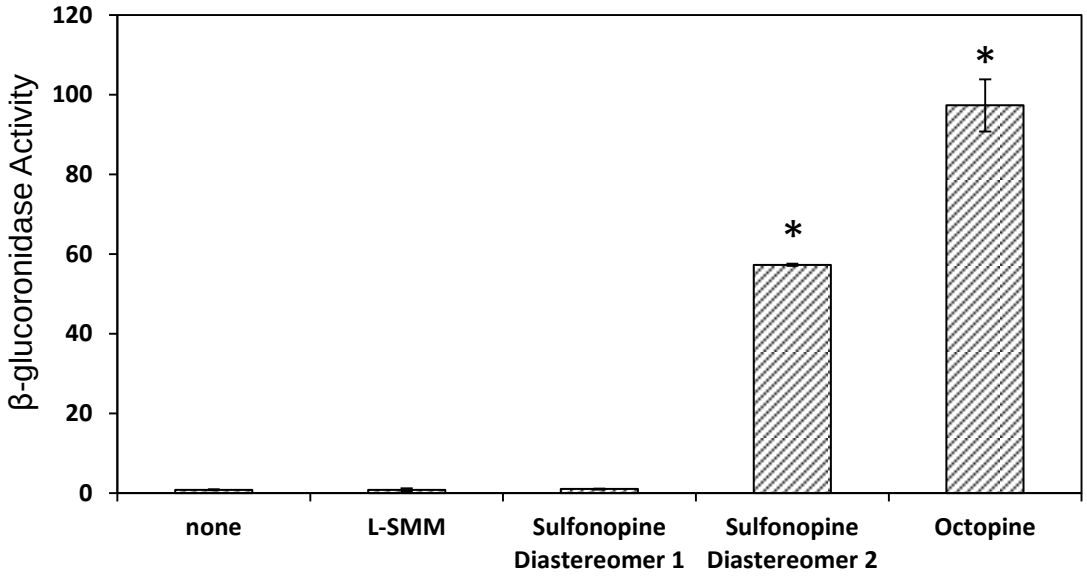


Figure 3.5 Biological activity of sulfonopine diastereomers. B-glucoronidase specific activity of *ooxA-uidA* strain KYC16 cultured in the presence of L-SMM, sulfonopine diastereomers and octopine. Mean ± SD of n=3. *P<0.001, *t* test relative to none sample.

Homogenates and exudates of plants colonized by *A. tumefaciens* contain sulfonopine

I next wanted to determine whether plants can synthesize and exude sulfonopine, and whether transfer and expression of the *ocs* gene is required. To address this, tobacco seedlings were cultured hydroponically and inoculated using four different strains: R10 (wild type), KYC55 (which lacks a Ti plasmid), R10 *ΔvirD4* (which cannot transfer T-DNA to plant cells), and R10 *Δocs* (which fails to produce octopine-type opines). After three days of cocultivation, the bacteria were killed by addition of carbenicillin and the seedlings were suspended in water. Two weeks later the water and seedling tissues were assayed for SMM and sulfonopine content by ESI-MS/MS.

As expected, SMM concentrations were similar in all seedlings. Seedlings infected with strain R10 synthesized sulfonopine, which was readily detected in homogenized tissues and in seedling exudates (Table 3.6). Sulfonopine was not detected in mock-infected seedlings, nor in seedlings infected with a strain lacking a Ti plasmid, nor in seedlings infected with a strain lacking *virD4* or *ocs* (Table 3.6, Fig. 3.6). These results show that Ocs can direct sulfonopine production in infected plants, and that this opine is released to the rhizosphere, where it could be consumed by Agrobacteria colonizing bacteria.

Table 3.6. SMM and sulfonopine content in tobacco tissue and exudate of infected seedlings with *A. tumefaciens* R10 and mutants strains.

	R10 (wild type)	KYC55 (no Ti plasmid)	Δ*vir*D4 (cannot transfer T-DNA)	Δ*ocs*	Mock infected
SMM in Homogenate[a]	21 ± 16[c]	18 ± 14	31 ± 25	26 ±19	40 ± 25
Sulfonopine in Homogenate[a]	5 ± 0.6	ND[d]	ND	ND	ND
Sulfonopine in exudate[b]	2 ± 0.7	ND	ND	ND	ND

a: nmol mg^{-1} of plant tissue
b: nmol ml^{-1}
c: Mean ± SD of triplicate samples.
d: ND. Not detected

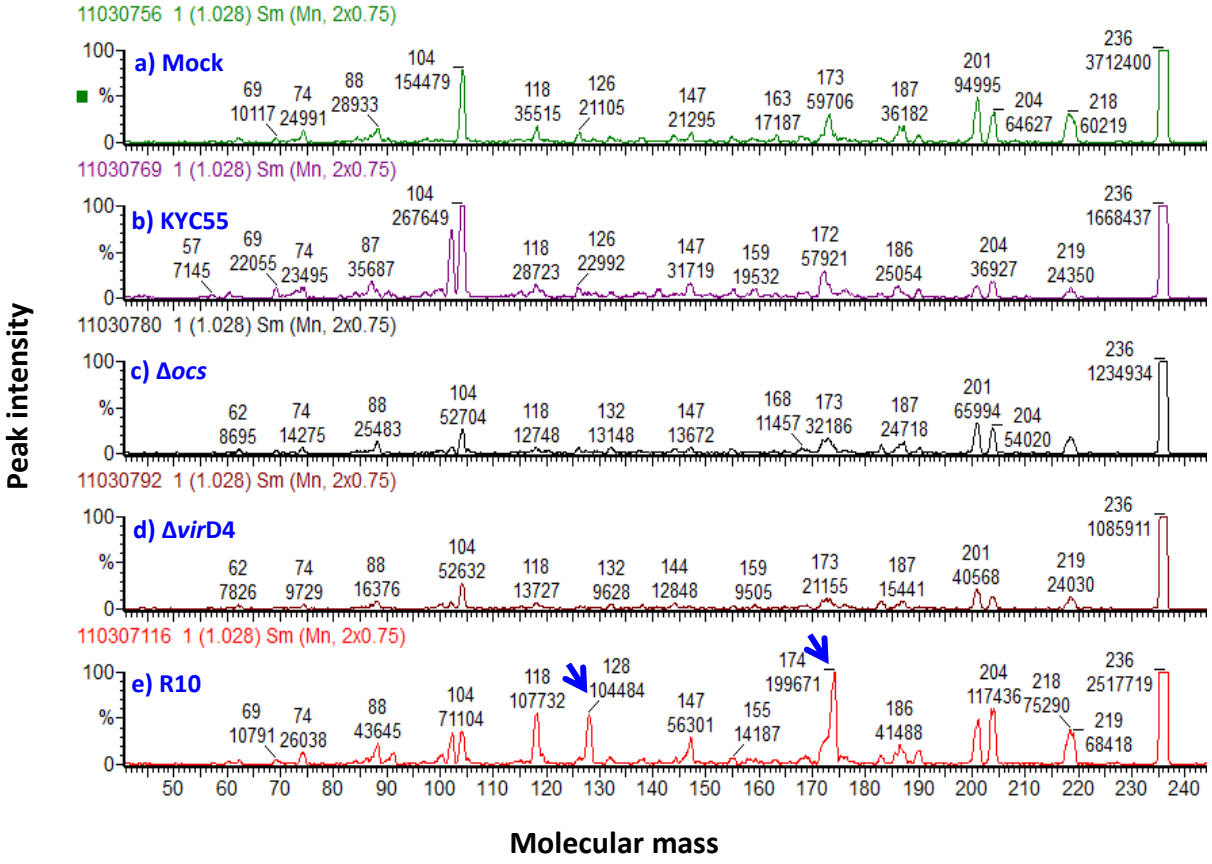


Figure 3.6 ESI-MS/MS authentication of sulfonopine synthesis in infected tobacco seedlings with different *A. tumefaciens* strains. The authentication of sulfonopine was done by fragmenting it by adjusting the collision energy of the spectrometer Micromass Quattro to 11. a) mock infection; b) KYC55; c) Δ *ocs*; d) Δ*virD4*; e) R10. Unique sulfonopine fragments are indicated with the arrows.

Production of sulfonopine requires host-synthesized SMM

I wanted to know whether the production of sulfonopine by plants is dependent on SMM. I therefore used *Arabidopsis thaliana* wild type (Col-0), and two mutant lines, one of which (*hmt*2-2) overexpresses SMM, while the other (*mmt*) does not synthesize any SMM (Lee et al 2008). Sulfonopine was produced by infected Col-0 and was

produced at higher levels by infected *hmt*2-2 plants, but was not produced by infected *mmt* plants (Table 3.7, Fig. 3.7). The concentration of SMM therefore influences the production and accumulation of sulfonopine, consistent with SMM being a direct substrate for Ocs.

Table 3.7. Sulfonopine contents in floral tissue in *Arabidopsis thaliana* lines.

	Infection	*Arabidopsis* strain		
		Wild type[a]	*mmt*[b]	*hmt2-2*[c]
Sulfonopine	-	ND[d]	ND	ND
Sulfonopine	+	7 ± 1[e, f]	ND	140 ± 90

a: Wild type (Columbia ecotype).
b: does not produce SMM.
c: SMM hyperaccumulation.
d: Not detected
e: (nmol mg^{-1})
f: Mean ± SD of triplicate samples.

The production of sulfonopine in an enzymatic reaction, in infected *A. thaliana* hmt2-2, in homogenates, and in exudates from tobacco seedlings was confirmed by comparing their fragmentation pattern with that of the chemically synthesized sulfonopine (Fig.3.7). Sulfonopine produced in the different experiments showed the characteristic peaks 174 and 128.

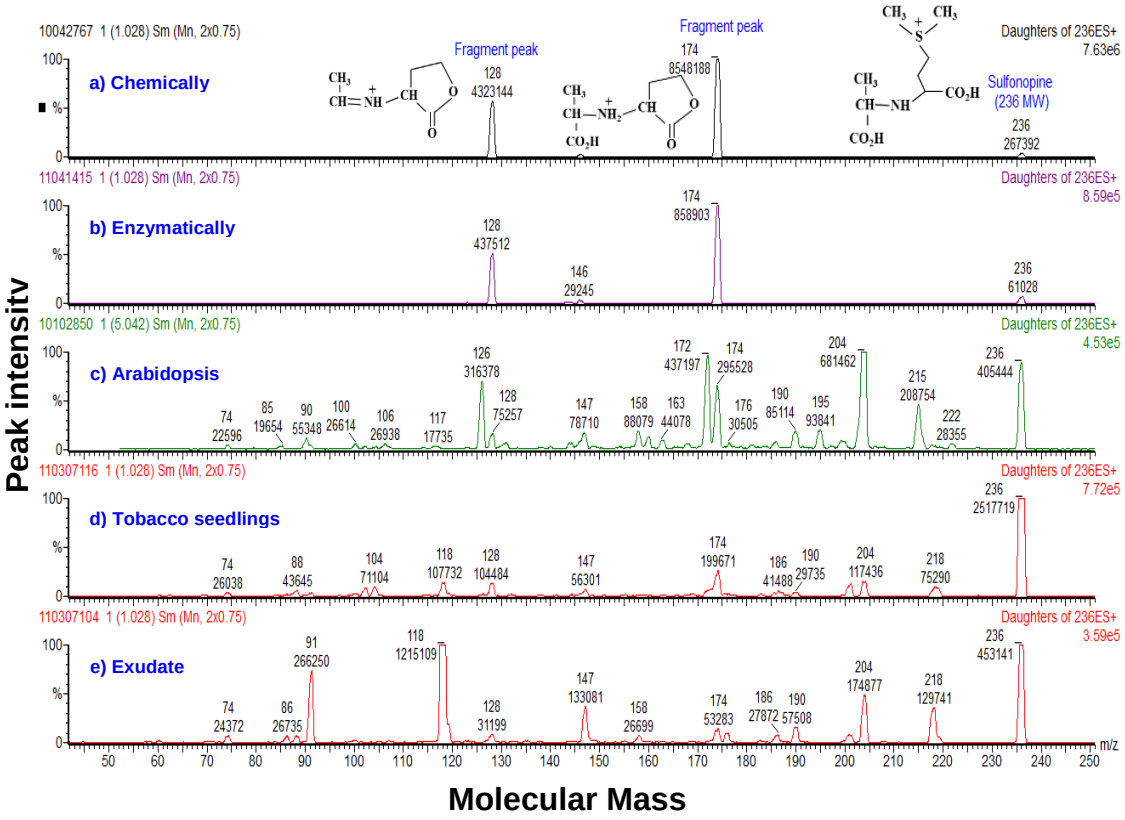


Figure 3.7 Authentication of sulfonopine production by ESI-MS/MS: The authentication of sulfonopine was done by fragmenting it by adjusting the collision energy of the spectrometer Micromass Quattro to 11. a) chemically synthesized, b) enzymatically synthesized, c) *Arabidopsis* tissue, d) tobacco seedling tissue and e) in tobacco seedling exudates. The authentication of sulfonopine was done by fragmenting it and comparing the fragments with the authentic compound chemically synthesized. Sulfonopine has two fingerprint fragments, which have a molecular weight of 174 and 128. x-axis molecular weight. y-axis peak intensity.

Sulfonopine induces transcription of the octopine catabolism operon

It is known that octopine induces the transcription of the *occ* operon via the LysR-type regulator OccR (Habeeb et al 1991). To investigate whether the octopine-type opines found in exudates and homogenates from infected-tobacco seedlings were able to induce transcription of the *occ* operon, I used the strain KYC16, which harbors an octopine-inducible *ooxA-uidA* fusion made by transposition of Tn5*gus*A7 (Cho et al 1996). KYC16 was cultured in the presence of exudates or homogenates from the infected-tobacco seedlings. β-glucoronidase activity was elevated by exudates and homogenates of tobacco seedlings infected with *A. tumefaciens* R10 (Fig. 3.8 a-b).

As described above, the regulation of this operon is based only on octopine experiments. To investigate whether sulfonopine and other opines can induce the expression of the *occ* operon, I enzymatically prepared a large variety of octopine-type opines using a variety of amino acids. I tested eleven of these opines for induction of the *ooxA-uidA* fusion. I used a commercial source of octopine as a positive control and the corresponding amino acids as negative controls.

Each of the opines activated the fusion at different levels, sulfonopine together with the commercial octopine, octopinic acid, lysopine, canavanopine, and homooctopine induced better the operon while heliopine, histopine, cystinopine and glycinopine induced poorly and each of the other opines (Fig. 3.9a). The corresponding amino acids did not induce the operon (Fig. 3.9 b).

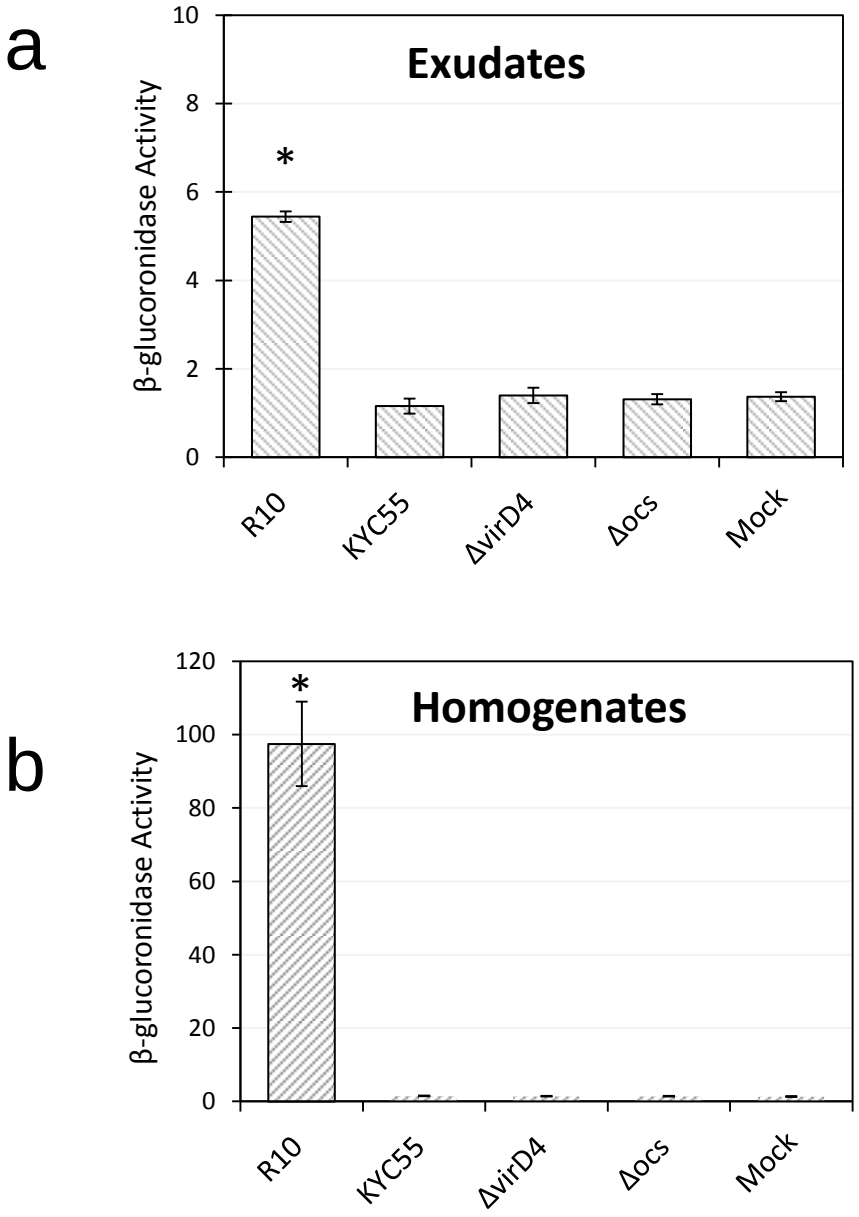


Figure 3.8 Induction of *occ* operon by octopine-type opines. (a-b) β-glucoronidase specific activity of *ooxA-uidA* strain KYC16 cultured in the presence of (a) exudates or (b) tissue homogenates from infected tobacco seedlings. Mean ± SD of n=3. *P<0.001, *t* test relative to mock sample.

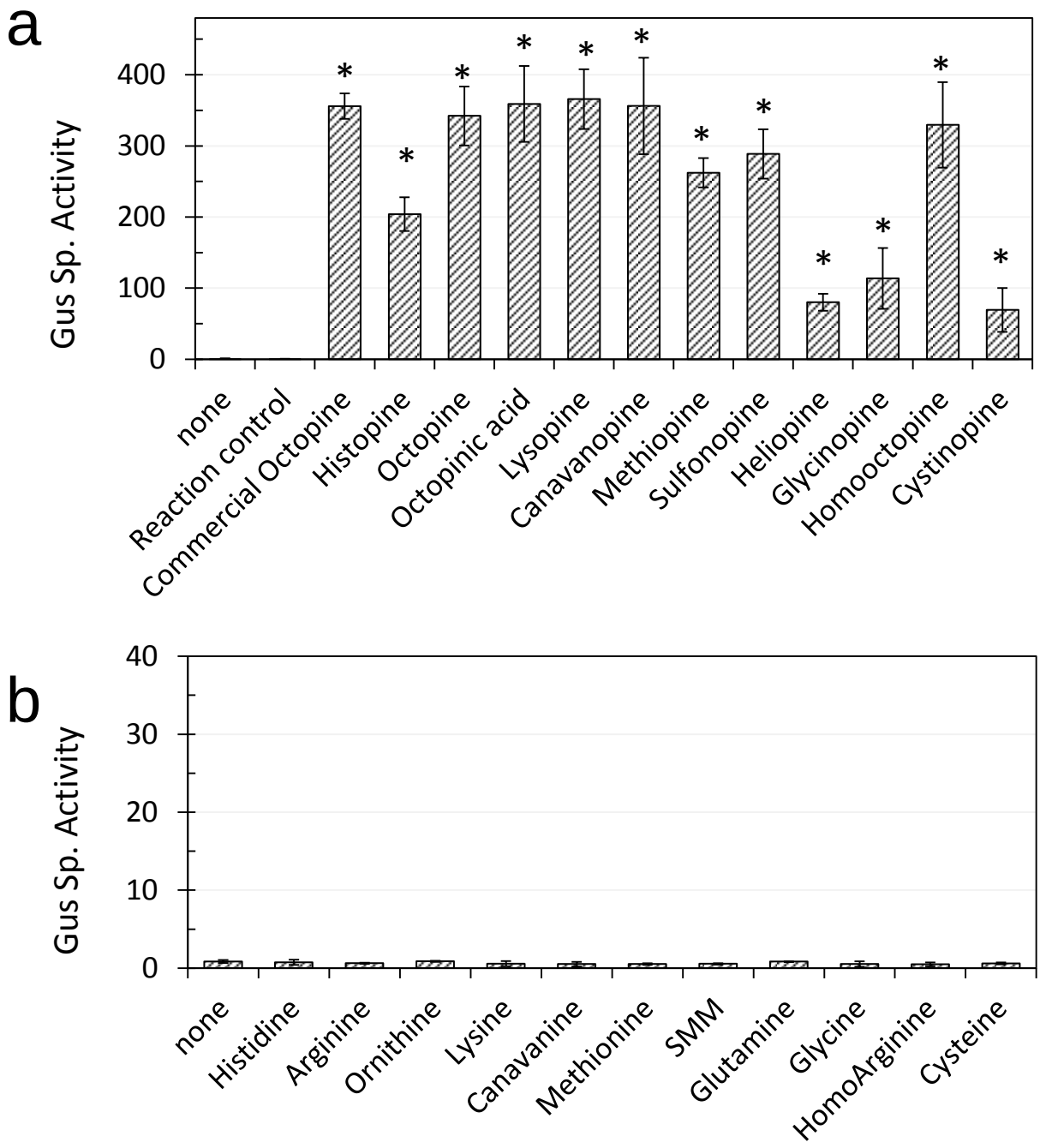


Figure 3.9 Induction of *occ* operon of strain KYC16 (R10 *ooxA-uidA*) by octopine-type opines and amino acids. a) Enzymatically produced opines induce expression of an *ooxA-uidA* fusion on the Ti plasmid. "Reaction control" refers to an enzymatic reaction mixture lacking any amino acid. "Commercial octopine" was purchased from Aldrich. b) amino acids. Mean ± SD of n=3. *P<0.001, *t* test relative to none sample.

Previous reports showed that octopine binds to OccR, activating the transcription of the permeases and catabolic genes for octopine utilization. However, it is not known if OccR is the only regulatory protein of this operon that responds to the presence of octopine. I wanted to examine whether OccR is required for induction by sulfonopine and other opines, as a positive control in this experiment I used octopine. I cultured two congenic strains in the presence or absence of five different opines. Strains KYC1203(pKY148) (pRJM101) and KYC1203(pKY148) have a plasmid-borne *occQ-lacZ* fusion, and the former strain expresses OccR, while the latter strain does not.

I found that the strain that expresses OccR was induced in the presence of the five opines. However, the level of activation was different among them, octopine and sulfonopine induced better the *occ* operon than methiopine, histopine and heliopine (Fig. 3.10a). The strain lacking OccR was not induced (Fig. 3.10b), indicating that OccR is the only regulatory protein of this operon and it is required for detection of these opines.

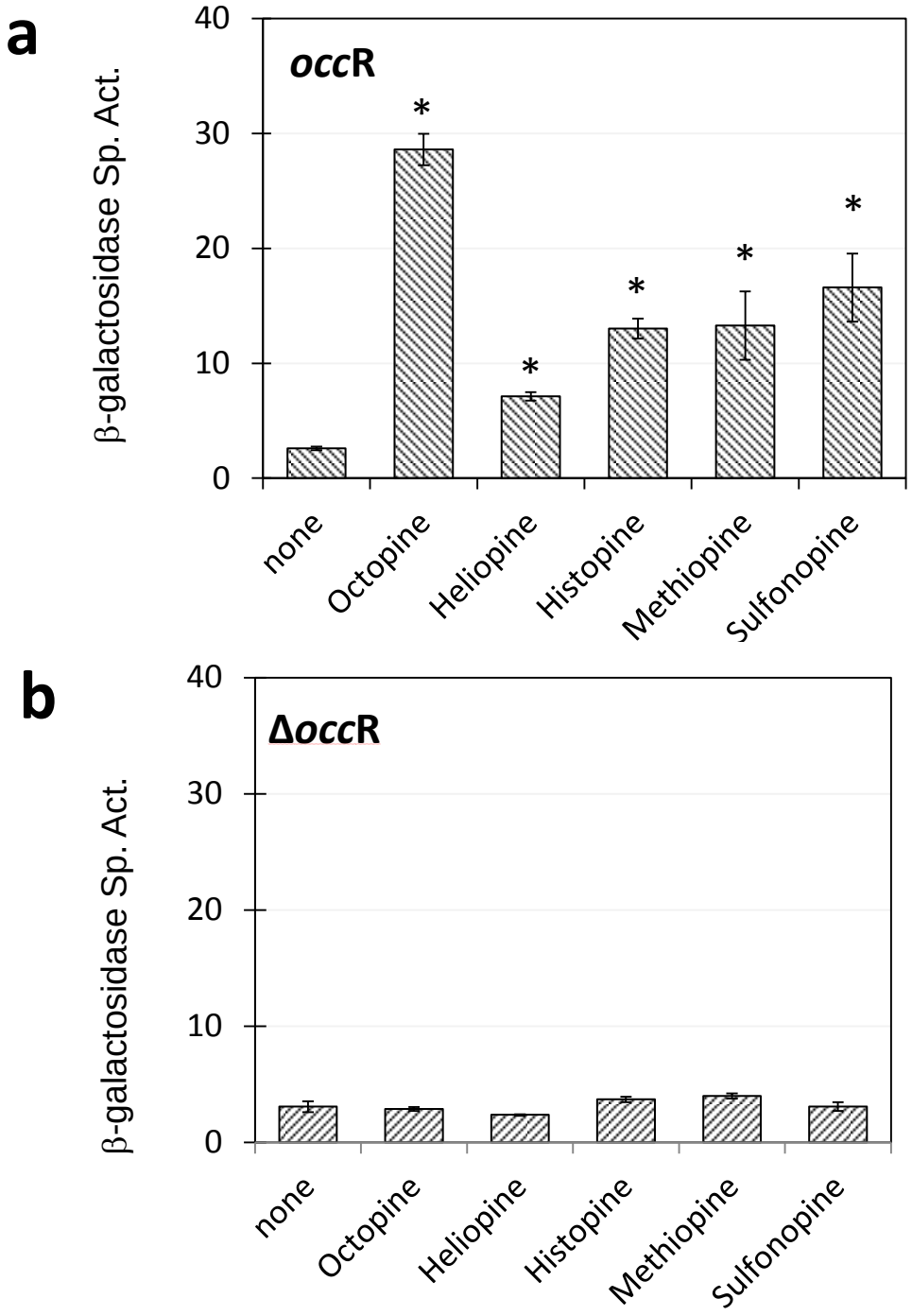


Figure 3.10 Induction of *occ* operon by octopine-type opines. β-galactosidase activity of *A. tumefaciens* strains a) KYC1203(pKYC148)(pRM101) and b) KYC1203 (pKYC148) (Δ *occ*R) cultured in the presence of different opines. pKYC148 contains an *occQ-lacZ* translational fusion. Mean ± SD of n=3. *P<0.001, *t* test relative to none sample.

A. tumefaciens can utilize sulfonopine as sole source of sulfur

Overall all the previous results showed that sulfonopine indeed induced the octopine catabolism operon by interacting with OccR, showing that sulfonopine acts as a signal molecule inducing gene expression in *A. tumefaciens.* However, I still did not know if sulfonopine can be used as a sulfur source. Therefore the next experiment was to determine whether *A. tumefaciens* is able to utilize sulfonopine and whether the octopine catabolism operon is required for its utilization.

For this experiment, I used two *A. tumefaciens* strains, R10 and KYC55 (Ti-plasmid less). The strains were cultured in modified AB minimal lacking sulfur supplemented with 1 mM of sulfonopine, SMM, or methionine as sole sulfur sources. Sulfonopine was synthesized enzymatically, therefore I used a control reaction negative control "mock" which refers to a control reaction conducted in the absence of any amino acid. An additional control was used referred as "buffer", in this control the strains were growing in the medium lacking any additions.

I found that the strain R10 could utilize all three compounds, sulfonopine, methionine, and SMM. Unexpectedly, strain KYC55 also grew at the expense of any of these three compounds. This indicates that the *occ* operon is dispensable for the use of sulfonopine as a sulfur source (Fig. 3.11a-b). Also, the growth of either the strain R10 or KYC55 was slower in the presence SMM compared with sulfonopine or methionine.

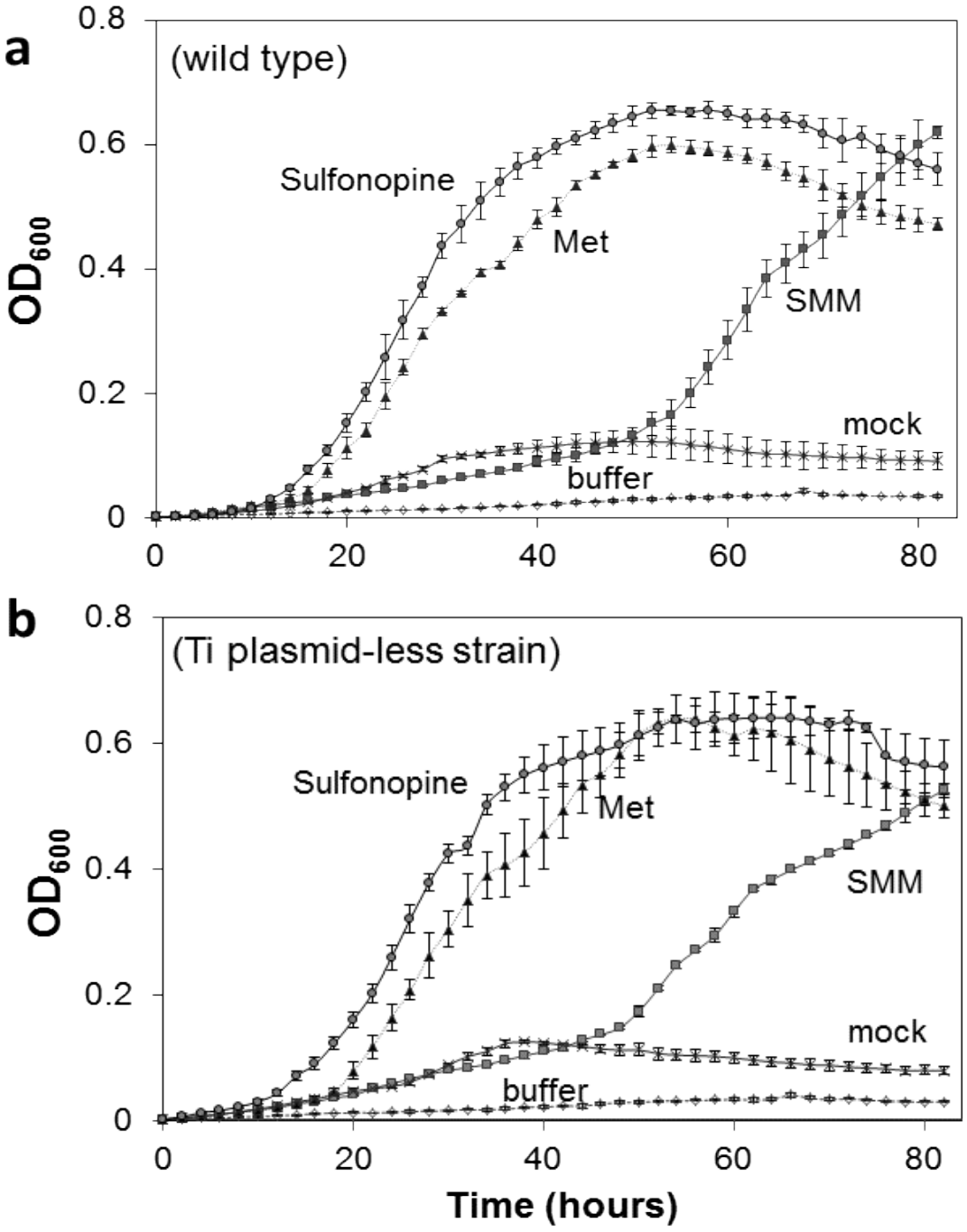


Figure 3.11 Utilization of sulfonopine as a sulfur source by two strains: a) R10 (wild type) and b) KYC55 (lacking a Ti plasmid). Strains were cultured in modified AB minimal lacking sulfur supplemented with 1mM of the tested compounds. Sulfonopine was synthesized enzymatically, and "mock" refers to a control reaction conducted in the absence of any amino acid. "Buffer" refers to growth medium lacking any additions. Mean ± SD of n=3.

3.5 Discussion

As I described in Chapter 2, SMM is a ubiquitous metabolite in many or possibly all plants, though it is not essential in at least one plant (Lee et al., 2008). The plant uses SMM as a transported form of methionine, and is concentrated in floral tissues (Ranocha et al., 2001). It should perhaps not be surprising that SMM could be a substrate for Ocs, as Ocs can utilize a rather broad variety of amino acids, especially the positively charged amino acids arginine, lysine, ornithine and histidine, as well as the non-charged amino acids glutamine and methionine (Bates et al., 1984; Biemann et al., 1960; Chilton et al., 2001; Hack and Kemp, 1977; Menage and Morel, 1964).

Lysopine (a condensation product of lysine and pyruvate) was the first discovered opine (Biemann et al., 1960; Seitz, 1964), followed by similar opines derived from either arginine, ornithine, histidine, glutamine, or methionine (Firmin et al., 1985; Hack and Kemp, 1977; Menage and Morel, 1964). Two research groups purified small amounts of Ocs directly from crown gall tumors (Hack and Kemp, 1980; Otten et al., 1977), and both studies found that this enzyme preferentially utilizes basic amino acids. In the present study, I produced a highly purified recombinant Ocs, which was stable and provided highly reproducible kinetic data for different substrates. This enzyme utilizes a surprising variety of proteogenic and nonproteogenic substrates, and several different alpha-keto acids, as well as NADH or NADPH. Many of the nonproteinogenic substrates are found in plants, including canavanine, homocysteine, homoserine, homoarginine, and selenomethionine (Bell, 2003). From an evolutionary perspective, it would seem advantageous for *A. tumefaciens* to utilize such a wide array of compounds as substrates,

using either NADH or NADPH, and to derivatize them in such a way that they are unavailable to most other bacteria or presumably to plants.

Production of sulfonopine by tobacco required transfer of *ocs* from the bacterium into plant cells, as a bacterial *vir* mutant or *ocs* mutant blocked sulfonopine production. I do not know how sulfonopine or other opines are released from plant cells. An older report suggested that the *ons* gene facilitated opine export (Messens et al., 1985), but Ons is a member of the RolB-RolC family, whose members regulate the abundance or activity of phytohormones (Bulgakov, 2008). Therefore, any effect of Ons on opine export would probably be indirect. The export of opines from transformed plant cells is therefore an unanswered question. However, the roots of many nontransformed plants exude surprisingly high levels of amino acids and carbohydrates (Phillips et al., 2004), suggesting that opine secretion may not require T-DNA encoded proteins.

Given the abundance of SMM in plants and the fact that it contains sulfur, a potentially scarce nutrient, it should not be surprising that at least some strains of *A. tumefaciens* can utilize it to form a sulfur-containing opine. I have not tested whether nopaline-type Ti plasmids can direct plant tumors to make a similar opine, but it seems quite plausible that nopaline synthase could conjugate SMM with α-ketoglutarate (rather than pyruvate) to create a compound quite similar to sulfonopine. I have evidence that expression of *msh* is downregulated by methionine (Chapter 4), which underscores the importance of sulfonopine as a source of methionine and by extension, as a source of sulfur.

Opines have long been known to induce genes required for their uptake and catabolism (Klapwijk et al., 1978; Montoya et al., 1977; von Lintig et al., 1991). Octopine has been reported to activate the transcription of the opine catabolic operon by binding to OccR, which is the transcriptional regulator of this operon (Habeeb et al., 1991). In the present study, I found that all tested opines can activate the transcription of the *occ* operon and that for all five opines tested, OccR was required for induction. These findings suggest that OccR detects primarily the amino acid backbone and pyruvate moieties of these opines, and has little or no specificity for their amino acid side chains.

The finding that the utilization of sulfonopine did not require the Ti plasmid was somewhat counterintuitive. Apparently the uptake and utilization of these compounds can be directed by other genes of the 5.6 megabase genome of this organism. Previous reports have shown that the first four genes of the *occ* operon direct octopine uptake into *E. coli* (Valdivia et al., 1991), but are not essential for its uptake into *A. tumefaciens* (Cho et al., 1996). Evidently, octopine uptake is encoded redundantly. The present study indicates that this is also true of sulfonopine uptake. However, utilization of octopine as a carbon or nitrogen source requires at least some genes of the *occ* operon (Cho et al., 1996). My observations that sulfonopine can provide sulfur in the absence of *occ* indicates that its catabolism is encoded redundantly.

One of the opines described in this study, methiopine, was previously described as a pseudo-opine, in that it was found to be produced by crown gall tumors but not imported by *A. tumefaciens (*Firmin et al., 1985). My finding that methiopine induces expression of the *occ* operon suggests that the earlier report should be re-evaluated.

Since sulfonopine is able to activate the *occ* operon, which includes TraR, this opine can also play a role in quorum sensing and conjugation of the Ti plasmid. In addition, sulfonopine is the first reported opine that can provide sulfur. Overall, these studies establish that SMM, a plant produced molecule, plays an important role in the nutrition and gene regulation in *A. tumefaciens*.

3.6 References

Alvarez-Martinez CE, Christie PJ. 2009. Biological diversity of prokaryotic type IV secretion systems. **Microbiol Mol Biol Rev** 73: 775-808

Bell EA. 2003. Nonprotein amino acids of plants: significance in medicine, nutrition, and agriculture. **Journal of Agricultural and Food Chemistry** 51: 2854-65

Biemann K, Lioret C, Asselineau J, Lederer E, Polonsky J. 1960. On the structure of lysopine, a new amino acid isolated from crown gall tissue. **Biochimica et Biophysica Acta** 40: 369-70

Brencic A, Winans SC. 2005. Detection of and response to signals involved in host-microbe interactions by plant-associated bacteria. **Microbiol Mol Biol Rev** 69: 155-94

Cangelosi GA, Best EA, Martinetti G, Nester EW. 1991. Genetic analysis of *Agrobacterium*. **Methods in Enzymology** 204: 384-97

Chai Y, Winans SC. 2009. The chaperone GroESL enhances the accumulation of soluble, active TraR protein, a quorum-sensing transcription factor from *Agrobacterium tumefaciens*. **Journal of Bacteriology** 191: 3706-11

Cho K, Fuqua C, Martin BS, Winans SC. 1996. Identification of *Agrobacterium tumefaciens* genes that direct the complete catabolism of octopine. **Journal of Bacteriology** 178: 1872-80

Dessaux Y, Guyon P, Farrand SK, Petit A, Tempe J. 1986. *Agrobacterium* Ti and Ri plasmids specify enzymic lactonization of mannopine to agropine. **J Gen Microbiol** 132: 2549-59

Dessaux Y, Guyon P, Petit A, Tempe J, Demarez M, et al. 1988. Opine utilization by *Agrobacterium* spp.: octopine-type Ti plasmids encode two pathways for mannopinic acid degradation. **Journal of Bacteriology** 170: 2939-46

Dessaux Y, Petit A, Farrand SK, Murphy PJ. 1998. Opines and opine-like molecules involved in plant-*Rhizobiaceae* interactions **In *The Rhizobiaceae***, ed. HP Spaink, A Kondorosk, PJ Hooykaas, pp. 173-97. Dordrecht: Kluwer Academic Publishers

Donnelly MI, Zhou M, Millard CS, Clancy S, Stols L, et al. 2006. An expression vector tailored for large-scale, high-throughput purification of recombinant proteins. **Protein Expression and Purification** 47: 446-54

Fuqua C, Winans SC. 1996a. Conserved cis-acting promoter elements are required for density-dependent transcription of *Agrobacterium tumefaciens* conjugal transfer genes. **Journal of Bacteriology** 178: 435-40

Fuqua C, Winans SC. 1996b. Localization of OccR-activated and TraR-activated promoters that express two ABC-type permeases and the traR gene of Ti plasmid pTiR10. **Mol Microbiol** 20: 1199-210

Gallagher SR. 1992. *Gus Protocols: Using the Gus gene as a reporter of gene expression*. San Diego: Academic Press.

Gelvin SB. 2009. *Agrobacterium* in the genomics age. **Plant Physiol** 150: 1665-76

Goldmann A, Thomas DW, Morel G. 1969. Sur Ia structure de la nopaline metabolite anormal de certaines tumeurs de crown-gall. **C. R. Acad. Sci. Paris** 268: 852-54.

Habeeb LF, Wang L, Winans SC. 1991. Transcription of the octopine catabolism operon of the *Agrobacterium* tumor-inducing plasmid pTiA6 is activated by a LysR-type regulatory protein. **Mol Plant Microbe Interact** 4: 379-85

Hack E, Kemp JD. 1977. Comparison of octopine, histopine, lysopine, and octopinic acid synthesizing activities in sunflower crown gall tissues. **Biochemical and Biophysical Research Communications** 78: 785-91

Hack E, Kemp JD. 1980. Purification and characterization of the crown gall-specific enzyme, octopine synthase. **Plant Physiol** 65: 949-55

Hong SB, Hwang I, Dessaux Y, Guyon P, Kim KS, Farrand SK. 1997. A T-DNA gene required for agropine biosynthesis by transformed plants is functionally and evolutionarily related to a Ti plasmid gene required for catabolism of agropine by *Agrobacterium* strains. **Journal of Bacteriology** 179: 4831-40

Kalogeraki VS, Winans SC. 1997. Suicide plasmids containing promoterless reporter genes can simultaneously disrupt and create fusions to target genes of diverse bacteria. **Gene** 188: 69-75

Kamada-Nobusada T, Sakakibara H. 2009. Molecular basis for cytokinin biosynthesis. **Phytochemistry** 70: 444-9

Lee M, Huang T, Toro-Ramos T, Fraga M, Last RL, Jander G. 2008. Reduced activity of *Arabidopsis thaliana* HMT2, a methionine biosynthetic enzyme, increases seed methionine content. **Plant J** 54: 310-20

Menage A, Morel G. 1964. Sur la presence d'octopine dans les tissus de crown-gall. **C. R. Acad. Sci. Paris** 259: 4795-96

Menage A, Morel G. 1965. Sur la presence d'un acide amine nouveau dans les tissus de crown-gall. **C. R. Acad. Sci. Paris** 261: 2001-02

Miller JH. 1972. *Experiments in Molecular Genetics*. pp. 448. **Cold Spring Harbor**: Cold Spring Harbor Laboratoey Press.

Murashige T, and Skoog, F. . 1962. A revised medium for rapid growth and bioassays with tobacco tissue cultures. **Physiol Plant** 15: 473-97

Otten LA, Vreugdenhil D, Schilperoort RA. 1977. Properties of D(+)-lysopine dehydrogenase from crown gall tumour tissue. **Biochimica et Biophysica Acta** 485: 268-77

Piper KR, Beck Von Bodman S, Hwang I, Farrand SK. 1999. Hierarchical gene regulatory systems arising from fortuitous gene associations: controlling quorum sensing by the opine regulon in *Agrobacterium*. **Mol Microbiol** 32: 1077-89

Ranocha P, McNeil SD, Ziemak MJ, Li C, Tarczynski MC, Hanson AD. 2001. The S-methylmethionine cycle in angiosperms: ubiquity, antiquity and activity. **Plant J** 25: 575-84

Ryder MH, Tate ME, Jones GP. 1984 Agrocinopine A, a tumor-inducing plasmid-coded enzyme product, is a phosphodiester of sucrose and L-arabinose. **J Biol Chem**. 259: 9704-10

Sambrook J, Fritsch EF, Maniatis T. 1989. *Molecular cloning: a laboratory manual.* . **Cold Spring Harbor**, NY Cold Spring Harbor Laboratory Press.

Savka MA, Dessaux Y, Oger P, Rossbach S. 2002. Engineering bacterial competitiveness and persistence in the phytosphere. **Mol Plant Microbe Interact** 15: 866-74

Tartoff KD, Hobbs CA. 1987. Improved media for growing plasmid and cosmid clones. **Bethesda Res. Lab. Focus** 9: 12

Toennies G, Kolb JJ. 1945. Methionine studies VII. Sulfonium derivatives**. J. Am. Chem. Soc**. 67: 849-51

Tzfira T, Citovsky V. 2006. *Agrobacterium*-mediated genetic transformation of plants: biology and biotechnology. **Curr Opin Biotechnol** 17: 147-54

von Lintig J, Zanker H, Schroder J. 1991. Positive regulators of opine-inducible promoters in the nopaline and octopine catabolism regions of Ti plasmids. **Mol Plant Microbe Interact** 4: 370-8

Wang L, Helmann JD, Winans SC. 1992. The *A. tumefaciens* transcriptional activator OccR causes a bend at a target promoter, which is partially relaxed by a plant tumor metabolite. **Cell** 69: 659-67

Weckbecker A, Hummel W. 2005. Glucose Dehydrogenase for the Regeneration fo NADPH and NADH In **Microbial Enzymes and Biotransformations**, ed. JL Barredo: SpringerLink

White CE, Winans SC. 2007. Cell-cell communication in the plant pathogen *Agrobacterium tumefaciens*. **Philos Trans R Soc Lond B Biol Sci** 362: 1135-48

Zhang X, Henriques R, Lin SS, Niu QW, Chua NH. 2006. *Agrobacterium*-mediated transformation of *Arabidopsis thaliana* using the floral dip method. **Nature Protocols** 1: 641-6

Zhu J, Oger PM, Schrammeijer B, Hooykaas PJ, Farrand SK, Winans SC. 2000. The bases of crown gall tumorigenesis. **Journal of Bacteriology** 182: 3885-95

3.7 Appendix

In addition to sulfonopine, many other opines were enzymatically produced in this study using proteogenic and non-proteogenic amino acids. Here we present a table and the ESI-MS/MS spectra showing the unique fragment peaks of each opine. The opines were listed according to their preference by Ocs (Table 3.8). The biggest fragment peak is indicated in red, while the second biggest is in black bold and the third major peak in blue bold. The fragmentation pattern of opines are presented from Figure 3.12 to 3.16. The peaks were verified by fragmenting and looking for the parent peak, at different collision energies. The collision settings are specific for the mass spectrometer Micromass Quattro II tandem MS operated in positive ion electrospray mode.

Table 3.8 Representing fragment peaks of opines analyzed by ESI-MS/MS.

Opine	Parent amino acid	Molecular Weight	Fragments (MW)	Collision settings
Sulfonopine	L-SMM	236	**174** and **128**	15
Octopine	L-Arginine	247	230, **188**, **175**, **158**, **142**, **130** and 60	15
Histopine	L-Histidine	228	**182**, **138**, 136, **115** and 95	15
Canavanopine	L-Canavanine	249	203, **160**, **128**, **118** and **76**	15
Homocystinopine	L-Homocysteine	208	174, **162**, **116** and **93**	11
Octopinic acid	L-Ornithine	205	**188**, **142**, **116** and 70	11
Lysinopine	L-Lysine	219	**202**, 184, 173, **156**, **130** and **84**	15
Methiopine	L-Methionine	222	**176**, **133**, **128** and 102	11
Heliopine	L-Glutamine	219	202, 184, **173**, **156**, **130** and **84**	15
Homoseropine	L-Homoserine	192	**146** and **116**	11
Cystinopine	L-Cystinopine	194	**102**	11
Glycinopine	L-Glycine	148	**116** and **102**	11
Homooctopine	L-Homoarginine	261	**199**, **189**, and **144**	11
Leucinopine	L-Leucine	204	**158**, **112** and **56**	15
Serinopine	L-Serine	178	**132**, **88** and **86**	15
Alanopine	L-Alanine	162	**116**, **88** and **70**	15
Asparaginopine	L-Asparagine	205	**159**, 146, 142, **115** and 57	15
Threonopine	L-Threonine	192	**146**, 128, 102 and **100**	15
Valinopine	L-Valine	190	**144** and **98**	15
Isoleucinopine	L-Isoleucine	204	**158**, **112** and **56**	15

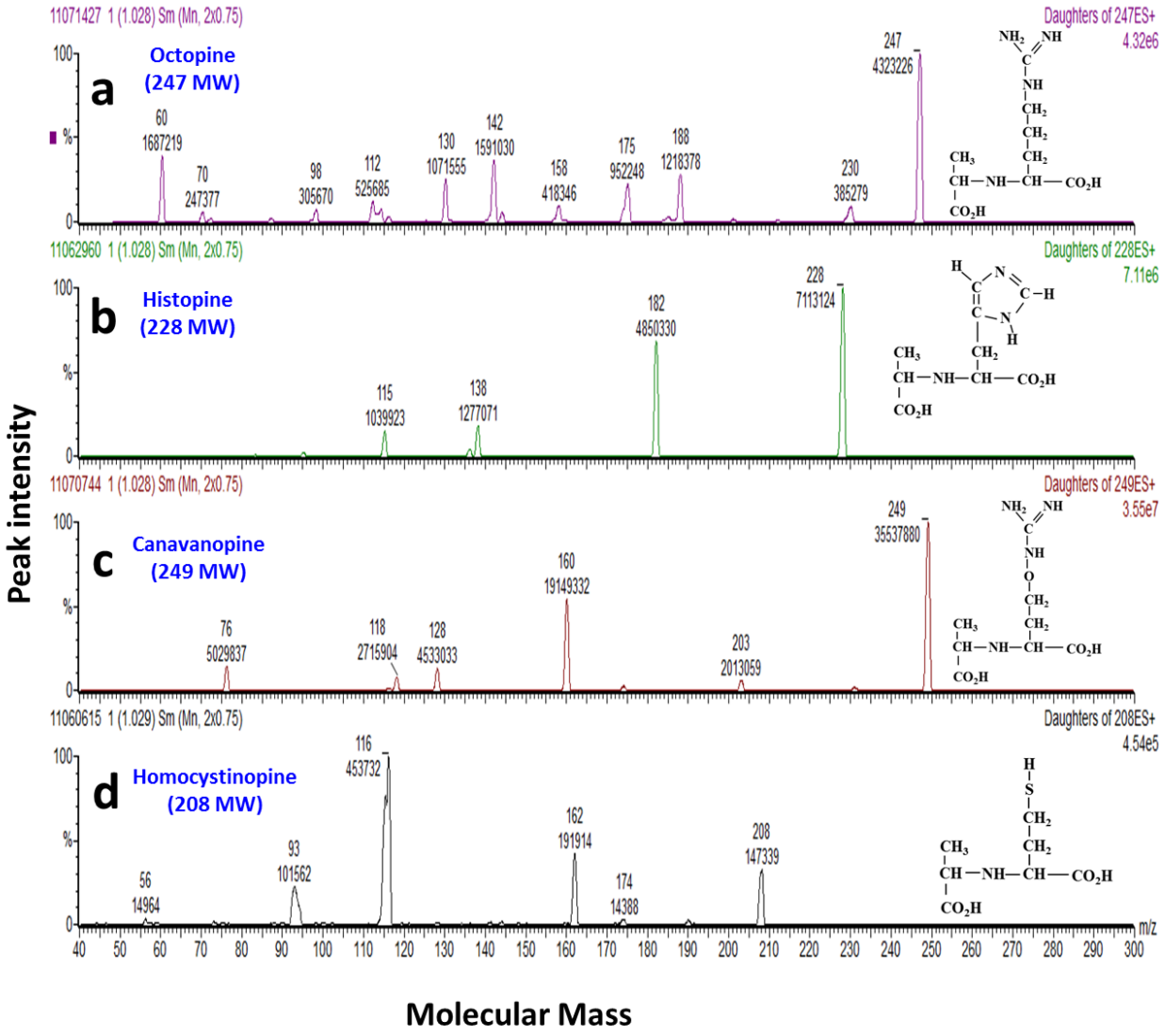


Figure 3.12 ESI-MS/MS analysis of the fragmentation pattern of the enzymatically synthesized opines. a) Octopine; b) histopine; c) canavanopine; and d) homocystinopine. x-axis molecular weight. y-axis peak intensity. Opines were fragmented by adjusting the collision energy of the spectrometer Micromass Quattro to 11 or 15 (See Table 3.8).

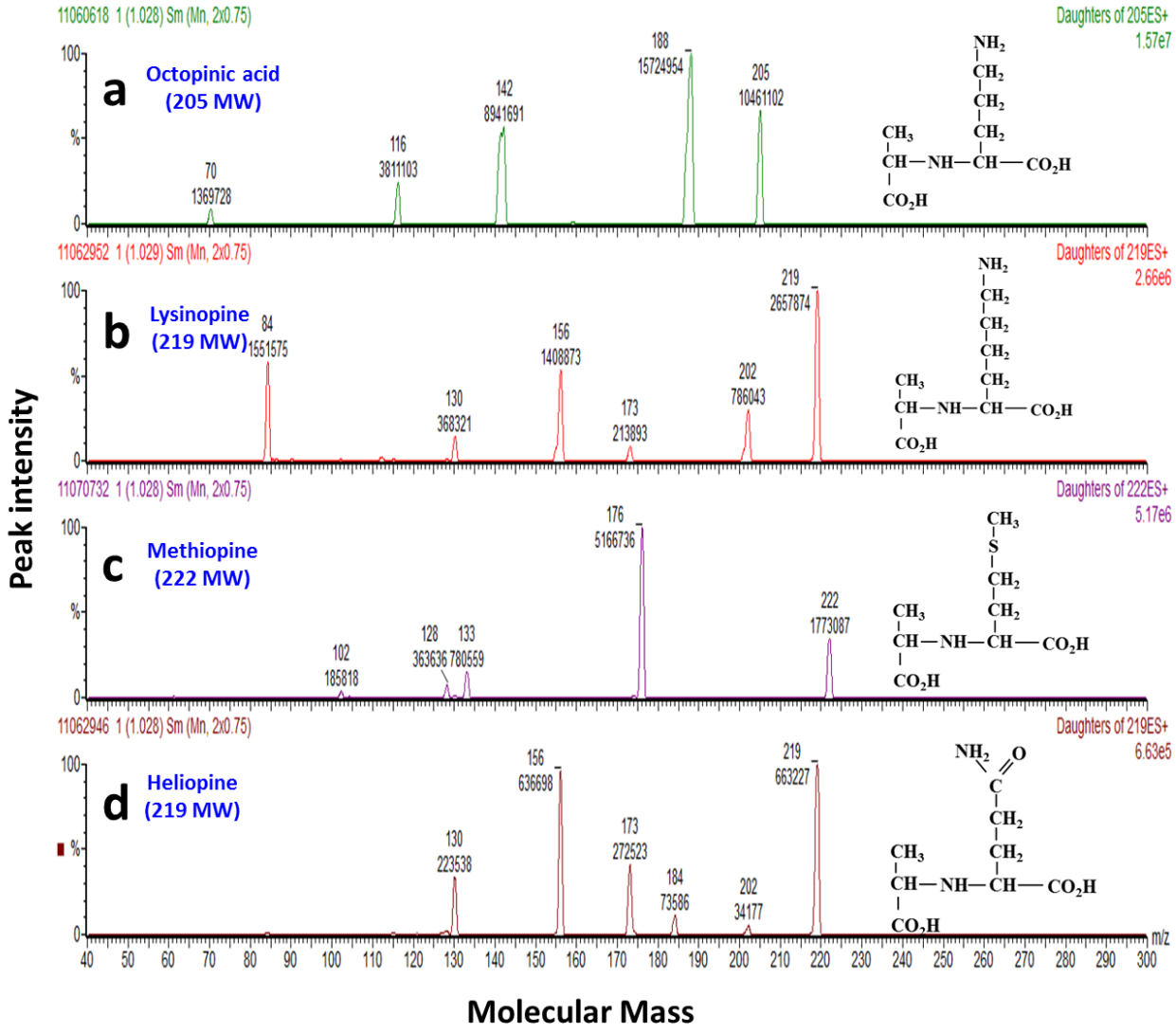


Figure 3.13 ESI-MS/MS analysis of the fragmentation pattern of the enzymatically synthesized opines. a) Octopinic acid; b) lysopine; c) methiopine; and d) heliopine. x-axis molecular weight. y-axis peak intensity. Opines were fragmented by adjusting the collision energy of the spectrometer Micromass Quattro 11 or 15 (See Table 3.8).

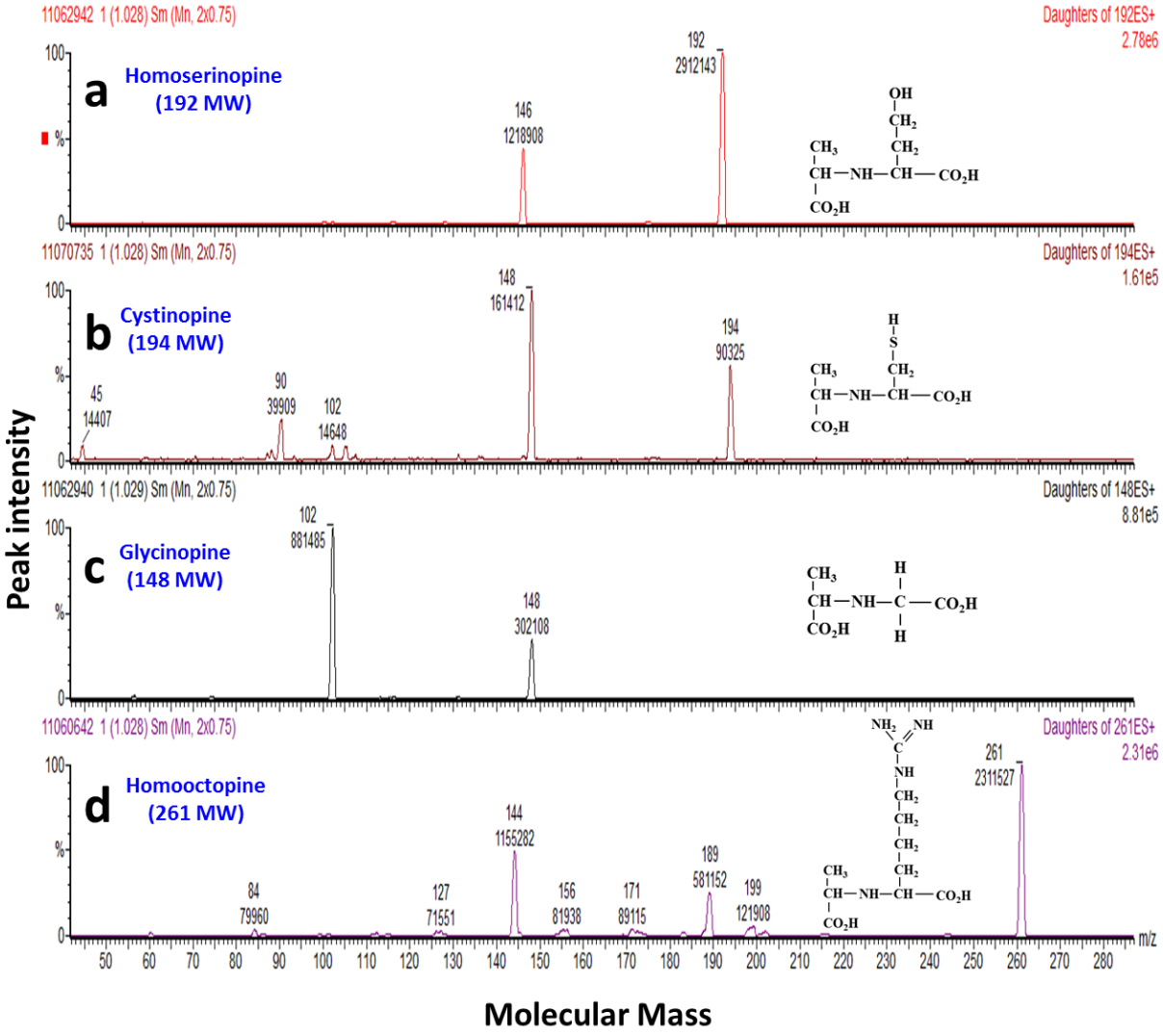


Figure 3.14 ESI-MS/MS analysis of the fragmentation pattern of the enzymatically synthesized opines. a) Homoserinopine; b) cystinopine; c) glycinopine; and d) homooctopine. x-axis molecular weight. y-axis peak intensity. Opines were fragmented by adjusting the collision energy of the spectrometer Micromass Quattro to 11 or 15 (See Table 3.8).

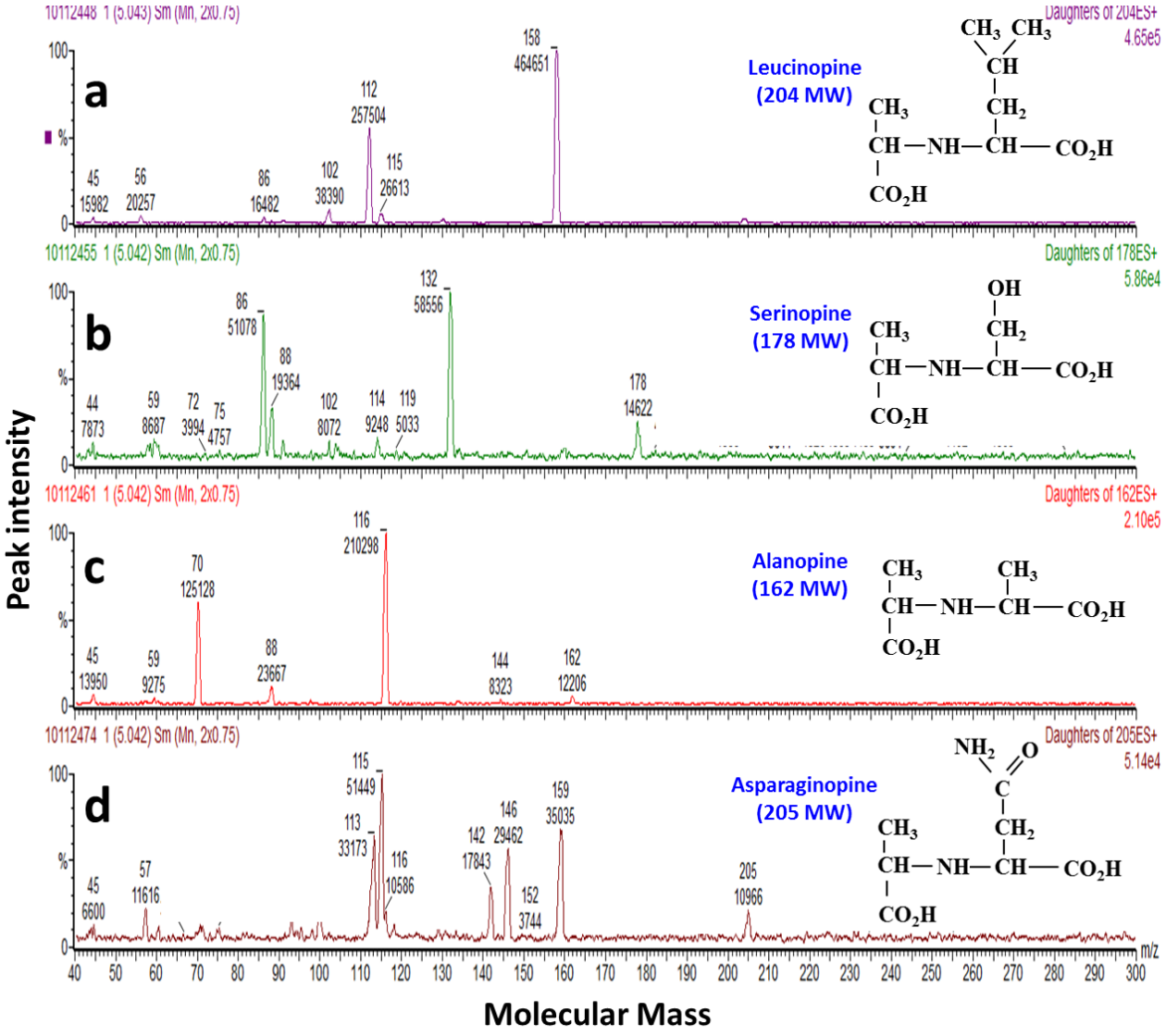


Figure 3.15 ESI-MS/MS analysis of the fragmentation pattern of the enzymatically synthesized opines. a) Leucinopine; b) serinopine; c) alanopine; and d) asparaginopine. x-axis molecular weight. y-axis peak intensity. Opines were fragmented by adjusting the collision energy of the spectrometer Micromass Quattro to 11 or 15 (See Table 3.8).

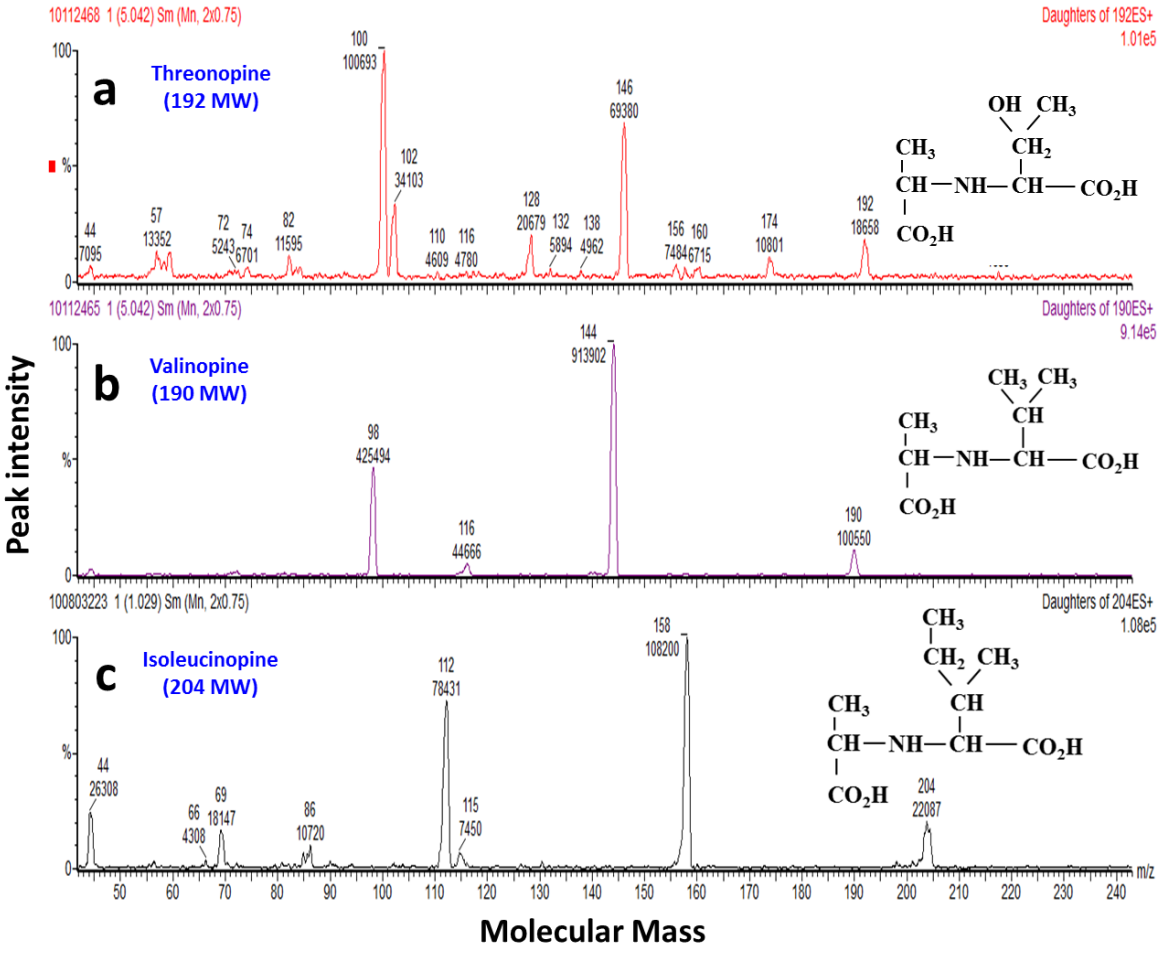


Figure 3.16 ESI-MS/MS analysis of the fragmentation pattern of the enzymatically synthesized opines. a) threonopine; b) valinopine; and c) isoleucine. x-axis molecular weight. y-axis peak intensity. Opines were fragmented by adjusting the collision energy of the spectrometer Micromass Quattro to 11 or 15 (See Table 3.8).

Chapter 4

Methionine represses quorum sensing, conjugation and replication of the Ti plasmid.

4.1 Abstract

Agrobacterium tumefaciens is a plant pathogen that induces the formation of crown gall tumors in wounded plants, directing the infected plant cells to produce nutrients called opines. The colonizing bacteria benefit by using these opines as nutrients and as gene regulatory molecules. Opines activate the transcription of cognate uptake and catabolic genes. Some opines, called conjugative opines, also activate transcription of the master transcriptional regulator *traR*. TraR is a LuxR-type quorum sensing transcriptional factor that regulates replication and conjugation of the Ti plasmid. Previous reports have showed that methionine and cysteine downregulate conjugation of the Ti plasmid. I confirmed this finding and asked whether methionine impacts the expression of *traR*. I found that methionine had no effect on the expression of the first half of the operon, but caused decreased expression of *msh ophABCD,* and *traR.* This finding fully explains how methionine impacts conjugation frequencies. I further mapped a region that comprises 97 bp upstream near the 5' end of *msh* that is critical for repression. *In vitro* assays suggested that the mRNA of this region directly interacts with methionine, suggesting that methionine represses *msh* expression via a riboswitch mechanism. Downregulation of *msh* by methionine should in principle provide for feed-back inhibition, as Msh produces methionine from S-methylmethionine and

homocysteine.

4.2 Introduction

Since the discovery that the Ti plasmid is essential for the oncogenic properties of the *Agrobacterium* (Zaenen et al 1974), many studies have been conducted to understand its mechanism of pathogenesis. During the 1970s it was discovered that the Ti plasmid is also a conjugative plasmid, and that its acquisition can cause an avirulent *Agrobacterium* strain to become virulent (Genetello et al 1977, Kerr 1969, Kerr et al 1977, Van Larebeke et al 1975). It was also found that the conjugation of the Ti plasmid requires certain tumor-released nutrients called opines, and that only some opines can stimulate conjugation (Genetello et al 1977, Klapwijk et al 1978, Petit et al 1978). It was also observed that the transfer of the Ti plasmid in rich media was repressed even in the presence of conjugal opines but not in minimal media. To identify the inhibitory component of rich medium, amino acids, sugars and organic acids were tested, and it was found that methionine and cysteine completely inhibited the conjugation of Ti plasmid. Methionine had no effect on the utilization of octopine as a nitrogen source, or on tumorigenesis (Hooykaas et al 1979). However, the mechanism of this regulation was not described.

In the last three decades, several labs including our lab have intensively studied the transcriptional regulation of *tra* and *trb* genes of the Ti plasmid, whose products direct conjugation (Alt-Morbe et al 1996). These genes are regulated by the TraR and TraI quorum sensing system (Fuqua & Winans 1994, Piper et al 1993). TraR of octopine-

type Ti plasmids is encoded at the 3' end of the octopine catabolism operon (*occ*) and it is expressed only in the presence of octopine-like opines. TraR and TraI belong to the family of LuxR-LuxI quorum sensing proteins. TraR is a LuxR-type transcriptional regulator while TraI is a LuxI-type protein that synthesizes the pheromone 3-oxo-octanoylhomoserine lactone (OOHL) (Fuqua & Winans 1994, Moré et al 1996). TraR requires OOHL for its stability to cellular proteases. Once the complex is formed, TraR-OOHL activates the expression of different genes involved in conjugation and replication of the Ti plasmid (Fuqua & Winans 1996a, Luo & Farrand 2001, Oger & Farrand 2002, Pappas & Winans 2003, Qin et al 2004, Qin et al 2000, Zhu & Winans 1999, Zhu & Winans 2001).

Bacteria have developed diverse mechanisms of gene regulation (Collado-Vides et al 1991, Ishihama 2010, Winkler & Breaker 2005, Zhou & Yang 2006). These mechanisms can be grouped into two main categories, those mediated by proteins and those mediated by RNAs. One category of RNA-mediated gene regulation is referred to as riboswitches, which are located in the 5'-untranslated (5'UTR) region of some mRNAs and form secondary structures that sense different stimuli such as metabolites, cations, temperature (Cromie et al 2006, Groisman et al 2006, Johansson et al 2002, Mandal & Breaker 2004), and other regulatory RNA molecules (Grundy et al 2002, Lease et al 1998). They comprise a ubiquitous system for the regulation of gene expression in bacteria that does not rely on proteins for detection of low molecular mass ligands (Serganov & Patel 2007, Tucker & Breaker 2005).

Riboswitches generally have two domains, an aptamer and an expression platform (Fig. 4.1). The aptamer domain is a highly folded structure and functions to sense the ligand molecule, while the expression platform is located downstream of the aptamer domain, and is directly involved in the regulation of the specific gene (Montange & Batey 2008). These domains may overlap and are linked by a switching sequence, which functions to communicate the binding of the metabolite in the aptamer to the expression platform, resulting in structural changes that ultimately impact some aspect of the expression of downstream genes. Some expression platforms alter transcription elongation, while others alter translation efficiency of the mRNA (Sudarsan et al 2006, Tucker & Breaker 2005).

In this Chapter I show that inhibition of *msh* expression by methionine requires a region near the 5' end of *msh* and downregulates the expression of *msh, ophABCD*, and *traR*. The inhibition of *traR* causes repression of the quorum sensing, conjugation and replication of the Ti plasmid.

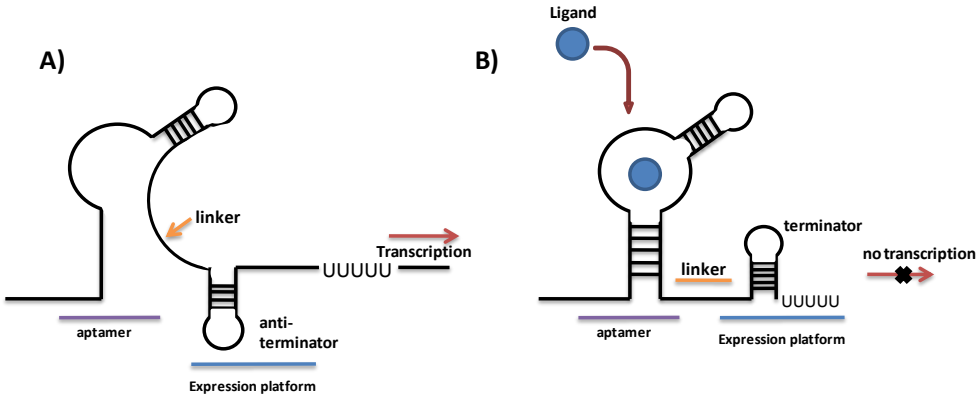


Figure 4.1 Anatomy and transcription regulation of riboswitch: in the (a) absence or (b) presence of a ligand.

4.3 Materials and methods

Bacterial strains, plasmids, and oligonucleotides

The bacterial strains, plasmids and oligonucleotides used in this study are listed in Table 4.1, 4.2 and 4.3, respectively. AT and AB media were used as defined media (Cangelosi et al 1991, Tartoff & Hobbs 1987). *E. coli* strains were cultured at 37°C and *Agrobacterium tumefaciens* strains were cultured at 27°C. Octopine and methionine were purchased from Sigma-Aldrich. 1-methyl-7-nitroisatoic anhydride was chemically synthesized (described below).

DNA manipulation

Recombinant DNA techniques were performed using established procedures (Sambrook et al 1989). Platinum *Taq* DNA Polymerase High Fidelity, T7 RNA polymerase, SuperScript III Reverse Transcriptase, RNaseOUT, and dNTPs were from Invitrogen. Plasmid DNA was isolated using QIAprep spin miniprep kits (Qiagen). DNA fragments generated by PCR or restriction digestion were gel purified using QIAquick Gel Extraction Kit (Qiagen). Restriction enzymes and other DNA modification reagents were purchased from New England Biolabs and used according to the methods described by the manufacturers. Plasmid DNA was introduced into *E. coli* and *A. tumefaciens* by electroporation (Cangelosi et al 1991).

Plasmid construction

The transcriptional *lacZ* vector, pAFM165, which contains P_{lac}- multicloning site

(mcs)-*lacZ* was constructed by digesting pBBRMCS5, a 4.7-Kb Gm^R plasmid, with *BamH*I and *Ale*I, and ligated it with the *lacZYA* genes from pMC1403 digested with *BamH*I and *Stu*I. pAFM165 is a 9.8-kb and compatible with IncP. IncQ, and IncW, as well as with ColE1-and P15a-based replicons (Casadaban et al 1980, Kovach et al 1995, Kovach et al 1994).

Plasmids pAFM166 and pAFM167 were constructed by cloning different fragments from the *ocd-msh* intergenic region and *msh* coding region into pAFM165. These fragments were PCR amplified using ALFM94-ALFM98 for pAFM166, generating a fragment of 428 bp, and ALFM228 and ALFM98 for pAFMN167, generating a fragment of 148 bp. The resulting fragments were inserted between *Xho*I and *Pst*I sites of pAFM165 (Table 4.2).

Mapping of the *occ* operon by *lacZ* fusion by Campbell-type integration

Suicide plasmids containing fragment of *occ* DNA were made by PCR amplification and inserted between the *Kpn*I and *Bam*HI sites of the suicide plasmid pVIK112 (Kalogeraki & Winans 1997), and introduced into *E. coli* strain SY327/λpir by electroporation. The resulting plasmids (pAFM24-25 and pAFM30-pAFM41) were transferred into strain R10 by conjugation, and transconjugants were selected using AB minimal agar plates containing 200 µg/ml of kanamycin (Table 4.2). Since these plasmids cannot replicate in *A. tumefaciens*, they conferred antibiotic resistance by Campbell-type integration into the different regions in the *occ* operon, creating fusions between the *occQ* promoter and *lac*Z in each position. Campbell-type integrations were

confirmed by PCR amplification. The resulting 14 *A. tumefaciens* strains were designated ALFM01 to ALFM14 (Table 4.1)

Table 4.1 Strains used in this study

Strains	Relevant features	Reference
E. coli		
BL21/DE3	*E. coli* P_{lac}-gene 1 of bacteriophage T7	(Studier et al 1990)
S17-1/λpir	RK2, *tra* regulon, *pir*, host for *pir*-dependent plasmid	(Simon et al 1983)
Agrobacterium tumefaciens		
R10	*A. tumefaciens* R10	(Dessaux et al 1989)
KYC55	*A. tumefaciens* R10, Ti plasmid less	(Cho et al 1997)
KYC16	R10 (*ooxA*::Tn5*gus*A7), Km^R	(Cho et al 1996)
KYC17	R10 (*occM*::Tn5*gus*A7), Km^R	(Cho et al 1996)
WCF5a	R10 (*traR*::pCF247) *tra*R-*lac*Z, Km^R	(Fuqua & Winans 1994)
ALFM01	R10::pAFM24, Km^R	This study
ALFM02	R10::pAFM25, Km^R	This study
ALFM03	R10::pAFM30, Km^R	This study
ALFM04	R10::pAFM31, Km^R	This study
ALFM05	R10::pAFM32, Km^R	This study
ALFM06	R10::pAFM33, Km^R	This study
ALFM07	R10::pAFM34, Km^R	This study
ALFM08	R10::pAFM35, Km^R	This study
ALFM09	R10::pAFM36, Km^R	This study
ALFM10	R10::pAFM37, Km^R	This study
ALFM11	R10::pAFM38, Km^R	This study
ALFM12	R10::pAFM39, Km^R	This study
ALFM13	R10::pAFM40, Km^R	This study
ALFM14	R10::pAFM41, Km^R	This study

Table 4.2 Plasmids used in this study

Plasmids	Relevant features	Reference
Cloning vectors		
pVIK112	*lacZY* for transcriptional fusions, Km^R	(Kalogeraki & Winans 1997)
pCC101	P_{lac}-mcs-*lacZ*, for transcriptional *lacZ* fusion, Tet^R	(Chen & Winans 1991)
pAFM165	P_{lac}-mcs- *lacZ*, for transcriptional *lacZ* fusion, $Gent^R$	This study
Plasmids		
pAFM24	pVIK112, last 500 bp from *ocd*, *lacZ* transcriptional fusion	This study
pAFM25	pVIK112, last 500 bp from *msh*, *lacZ* transcriptional fusion	This study
pAFM30	pVIK112, last 500 bp from *ooxA* (+1), *lacZ* transcriptional fusion	This study
pAFM31	pVIK112, 500 bp upstream +100*, *lacZ* transcriptional fusion	This study
pAFM32	pVIK112, 500 bp upstream +200*, *lacZ* transcriptional fusion	This study
pAFM33	pVIK112, 500 bp upstream +300*, *lacZ* transcriptional fusion	This study
pAFM34	pVIK112, 500 bp upstream +400*, *lacZ* transcriptional fusion	This study
pAFM35	pVIK112, 500 bp upstream +500*, *lacZ* transcriptional fusion	This study
pAFM36	pVIK112, 500 bp upstream +600*, *lacZ* transcriptional fusion	This study
pAFM37	pVIK112, 500 bp upstream +700*, *lacZ* transcriptional fusion	This study
pAFM38	pVIK112, 500 bp upstream +100[1], *lacZ* transcriptional fusion	This study
pAFM39	pVIK112, 500 bp upstream +200[1], *lacZ* transcriptional fusion	This study
pAFM40	pVIK112, 500 bp upstream +300[1], *lacZ* transcriptional fusion	This study
pAFM41	pVIK112, 500 bp upstream +400[1], *lacZ* transcriptional fusion	This study
pAFM46	*ocd-msh* intergenic region and 97 bp of *msh* (429 bp) into pCC101, Tet^R	This Study
pAFM166	pAFM165, contains the intergenic region between ocd-msh, and the first 97 bp of the *msh* coding region.	This study
pAFM167	pAFM165, contains 51 bp upstream of msh, and the first 97 bp of the *msh* coding region.	This Study

* +1 represents the first nucleotide downstream of *oo*x*A*.
[1] +1 represents the first nucleotide downstream of *ocd.*

Table 4.3 Oligonucleotides used in this study.

Name	DNA Sequence	Comments
Mapping *occ* operon		
ALFM54	ccg**gaattc**gaatgcggctccgtcgcc	For generating pAFM24
ALFM55	cgg**ggtacc**ttatccgacggctgtcgg	For generating pAFM24
ALFM56	ccg**gaattc**tcgaaacagatcaggaac	For generating pAFM25
ALFM57	cgg**ggtac**ctcaggctgcggctggctg	For generating pAFM25
ALFM64	ccg**gaattc**gccgaggtcagtgggcgac	For generating pAFM30
ALFM65	cgg**ggtacc**acaggacatatcttatga	For generating pAFM30
ALFM66	ccg**gaattc**cgccggttcctgcgcggaa	For generating pAFM31
ALFM67	cgg**ggtacc**catcgtcagcaacgatac	For generating pAFM31
ALFM68	ccg**gaattc**cgcgaggcaatagcgcttg	For generating pAFM32
ALFM69	cgg**ggtacc**tcatgtcggtgataaaga	For generating pAFM32
ALFM70	ccg**gaattc**gcaaccgtcaccgaaatca	For generating pAFM33
ALFM71	cgg**ggtacc**tcaccaaggtcggcaacg	For generating pAFM33
ALFM72	ccg**gaattc**cgccttgctgaacttgcac	For generating pAFM34
ALFM73	cgg**ggtacc**ggagatctccaataacga	For generating pAFM34
ALFM74	ccg**gaattc**atgacagaacgccttatcc	For generating pAFM35
ALFM75	cgg**ggtacc**gcccagctcttcaagcgc	For generating pAFM35
ALFM76	ccg**gaattc**aagctcgacgcatatctta	For generating pAFM36
ALFM77	cgg**ggtacc**tcttcctccaattttctg	For generating pAFM36
ALFM78	ccg**gaattc**gtgtggaaggaatggttgg	For generating pAFM37
ALFM79	cgg**ggtacc**ttcagtcaggaaggtgtc	For generating pAFM37
ALFM80	ccg**gaattc**gtccatatcaatgcggtcg	For generating pAFM38
ALFM81	cgg**ggtacc**gtgacggtgtcgtcgccg	For generating pAFM38
ALFM82	ccg**gaattc**tatccgccgcagacctgga	For generating pAFM39
ALFM83	cgg**ggtacc**agccttttgaaaacatgg	For generating pAFM39
ALFM84	ccg**gaattc**gtcggatttgcgatcgagg	For generating pAFM40
ALFM85	cgg**ggtacc**gcggcctccagcttacgt	For generating pAFM40
ALFM86	ccg**gaattc**cccgacgagccgcgagac	For generating pAFM41
ALFM87	cgg**ggtacc**cagcgacagcgcagacca	For generating pAFM41
ALFM94	ccg**ctcgag**tttaccgaactcgacctc	For generating pAFM166
ALFM228	ccg**ctcgag**ttatcaaacgccggcaaggat	For generating pAFM167
ALFM98	aa**ctgcag**gcgacagcgcagaccact	For generating pAFM166-167
SHAPE analysis		
ALFM94	ccg**ctcgag**aagtgttgcaccgcctgt	For generating 429 bp fragment
ALFM98	aa**ctgcag**gcgacagcgcagaccact	For generating 429 bp fragment
ALFM105	Taatacgactcactataggggccttcgggcaa gtgttgcaccgcctg	For generating T7 promoter + GGG+ 5'linker + 429 fragment
ALFM108	gaaccggaccgaagcccgatttggatccggcg aaccggatcgagcgacagcgcagacc	For generating 3' linker + 429 fragment

β-galactosidase and β-glucoronidase activity

For the β-galactosidase and β-glucuronidase activity assays, strains were cultured in AT medium supplemented with 200 μg/ml of kanamycin overnight at 37°C. Each culture was diluted 1:100 into AT medium supplemented with and without octopine and with different concentrations of methionine, then incubated with vigorous aeration at 27 °C to an OD_{600} of 0.3-0.4, and assayed for β-galactosidase specific activity either using ortho-Nitrophenyl-β-galactoside (ONPG) for colorimetric measurement or 4-Methylumbelliferyl beta-D-galactopyranoside (MUG) for fluorescence measurement (Miller, 1972). β-glucuronidase specific activity was measured as described previously (Gallagher 1992). Experiments were performed in triplicate with three different isolates for each strain.

Analysis of RNA upstream of *msh in vitro*

A fragment of 429 bp with contains the *ocd-msh* intergenic region together with the 97 bp of *msh* (97-*msh*) gene was PCR amplified using ALFM94 and ALFM98 primers and then cloned into pCC101 between *Xho*I and *Pst*I sites, generating pAFM46. The resulting plasmid was sequenced to ensure the correct nucleotide sequence. The plasmid was purified to use as a DNA template for the following PCR fragment. The DNA sequence was further amplified by using the ALFM105 and ALFM108 primers.

The PCR fragment was gel purified, then 50 pmol of the PCR product was used as a template to produce the mRNA by *in vitro* transcription using T7 RNA polymerase. The reaction was incubated for 2 hours at 37 °C, and treated with DNaseI afterwards. The

produced mRNA was purified by size-fractionation by 8% native PAGE. The mRNA was mixed with 25 µL of 2x gel-loading buffer (18 M urea, 20% (w/v) sucrose, 0.1% (w/v) SDS, 0.05% (w/v) bromophenol blue sodium salt, 0.05% (w/v) xylene cyanol, 90 mM borate, and 10 mM EDTA pH 8.0). The mRNA band was excised and extracted from the gel by incubating it for 1 hour in 400 µl buffer (200 mM NaCl, 10 mM Tris-HCL pH7.5 and 1 mM EDTA pH8.0). The supernatant was removed from the gel and transferred to a new microfuge tube. The mRNA was recovered by ethanol precipitation, by adding 5 µl of 4 M NaCl, 1 µl of 20 mg/ml of glycogen and 2.5 vol of 100% ethanol chilled to -20°C. The solution was mixed and incubated at -80 °C for 30 min. The mRNA was pelleted by centrifugation for 20 min at 14,000 rpm and 4°C. The supernatant was decanted, and the mRNA pellet was washed 3 times with 70% ethanol (chilled at -20 °C) and dried out. Then the pellet was resuspended in 25 µL of DEPC H_20 (Regulski & Breaker 2008).

A solution of 8 µg/ml of mRNA in 24 µl of 0.5X TE (1X TE-10 mM Tris and 1 mM EDTA, pH 8.0) was heated to 95°C for 2 min and then the temperature was decreased to 65 °C for 5 min. Next, the mRNA solution was divided into 2 equal aliquots, 10 mM of methionine was added to one aliquot while DEPC H_20 was added to other one. Then 8 µl of 3.3x folding solution (333 mM HEPES and 333 mM NaCl, pH 8.0) was added to each aliquot and mixed. They were incubated at 37°C for 30 min and were treated with 1/10 vol of 1-methyl-7-nitroisatoric anhydride (1M7) for 3 min and then the mRNA of each treatment was recovered by ethanol precipitation as previously described and resuspended in 10 µl of DEPC H_2O (Mortimer & Weeks 2007).

Ten µl of the mRNA was aliquoted into 200 µl thin-walled PCR tubes, and 1 µl of

100 mM of cDNA reverse primer FAM-labeled, 1 µl of 10 mM dNTP mix (10 mM each dATP,dGTP, dCTP and dTTP) and 1 µl DEPC H_2O were added. The mixture was heated at 65 °C for 5 min and incubated on ice for 1 min. The mixture was collected by a brief centrifugation, and then 4 µl of 5X First-Strand buffer, 1 µl of 0.1 M DTT, 1 µl of RNaseOUT and µl of SuperScript III RT were added. The reactions were incubated at 55 °C for 60 min. After the incubation, the reactions were inactivated by heating at 70 °C for 15 min. The RNA was degraded from the cDNA reactions by adding 1 µl of 4 M NaOH and heating the solution at 95 °C for 5 min. Then, 29 µl of acid stop solution (4:25 (v/v) mixture of 1M unbuffered Tris-HCL and stop solution (85% formimade, 0.5X TE, 50 mM EDTA, pH 8.0,) was added to stop the alkaline reaction. The cDNA from each reaction was purified by ethanol precipitation (as previously described). Fragments were sent to the Cornell University Life Sciences Core Laboratories Center for DNA fragment analysis using the Applied BioSystems 3730xl DNA Analyzer.

Synthesis of 1-methyl-7-nitroisatoric anhydride

The production of 1-methyl-7-nitroisatoric anhydride (1M7, $C_9H_6N_2O_5$ FW 222) was done by following the protocol of (Mortimer & Weeks 2007). Briefly, 0.66 g (3.16 mmoles) of 4-nitroisatoic anhydride (4-NIA, FW 208) and 20 ml of DMF were combined and mixed until dissolved. 0.17 g (4.14 mmoles) of 60% mineral oil suspension of NaH was combined with 20 ml of DMF in a flask covered by a septum. The mixture was flushed with nitrogen and stirred at room temperature. The septum then was removed and the 4-NIA solution was added, giving a clear orange solution after few minutes. Subsequently, 0.2 ml (0.45g; 3.2 mmoles) of CH_3I was added and stirred for 4 hours at

room temperature under nitrogen. The resulting solution was poured and mixed into a solution of chilled 50 ml HCL, giving bright orange particulates of 1M7, which was dissolved in DMSO.

4.4 Results

Effect of methionine on the expression of the *occ* operon

In a previous report it was shown that octopine strongly stimulates Ti plasmid conjugation, and that methionine antagonizes octopine (Hooykaas et al 1979). It is now know that octopine-like opines act through OccR to activate the expression the octopine catabolic operon, which contains *traR* (Cho et al 1996, Cho et al 1997, Fuqua & Winans 1994, Habeeb et al 1991, Valdivia et al 1991, Zhu et al 2000). TraR positively regulates the conjugation and replication of the Ti plasmid (Fuqua & Winans 1994), therefore I hypothesized that methionine could repress expression of TraR.

To address this, I determined whether methionine can inhibit expression of three different three reporter fusions in the *occ* operon. I used the strains KYC17 and KYC16, which have a *gus* fusion in the *occM* and *ooxA*, respectively, and WCF5a, which has a *lacZ* fusion in *traR*. I added octopine in the presence or absence of methionine to assess the effect of methionine in the expression of these fusions. I found that the upstream genes of the operon were not affected by methionine. In contrast, the *traR-lacZ* fusion of WCF5a was significantly downregulated (Fig. 4.2).

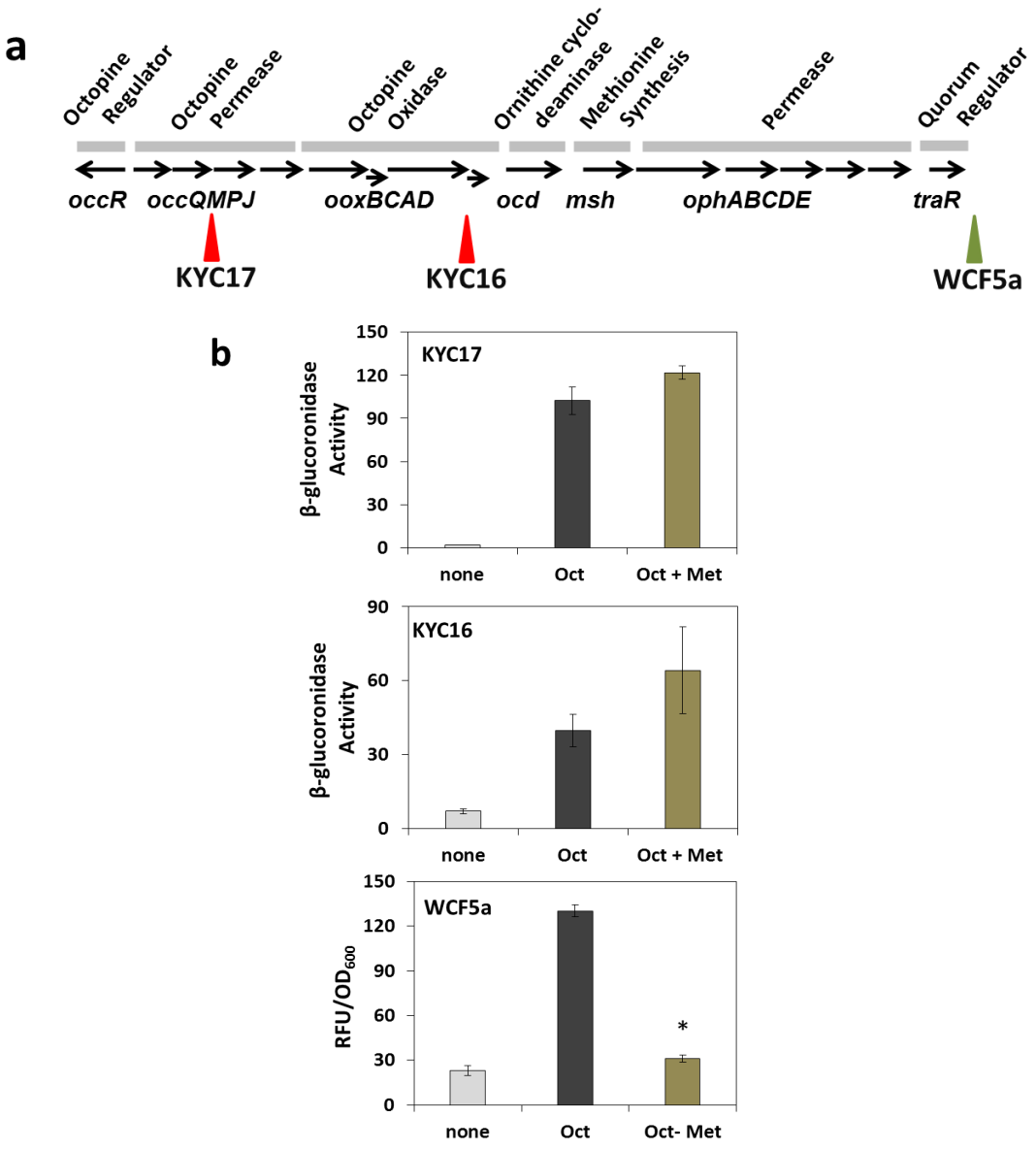


Figure 4.2 Repression of the downstream part of the *occ* operon by methionine. a) *gus* and *lacZ* fusion within *occ* operon. The arrows indicate the position of the fusions: KYC16 and KYC17 have *occ-uidA* fusions made using Tn5*gusA7*, while WCF5a has a *traR- lacZ* fusion made by homologous recombination (Fuqua & Winans 1996b). b) Strains were grown in the absence (none) or presence of octopine containing or lacking methionine. x-axis treatment, y-axis β-glucoronidase activity or β-galactosidase activity expressed in Relative Fluorescence Units (RFU) normalized by the optical density (OD). Mean ± SD of n=3. *P<0.001, *t* test relative to samples with octopine.

This result led me to focus on the downstream half of the *occ* operon, and in particular to investigate two intergenic regions that flank *ocd*. For this experiment, I introduced *lacZ* fusions at 100 nucleotide intervals from *ooxA* to *msh* by Campbell integration. Expression of the resulting 14 fusions (Fig. 4.3a) was measured in the presence of octopine containing or lacking methionine. Fusions upstream of *ocd* were not affected by methionine. In contrast, expression of some of the fusions downstream of *ocd* were inhibited by methionine. The fusion of ALFM2 contains all of *msh* and was as inhibited as strongly as that of WCF5a. The fusion of ALFM14 contains 97 nucleotides of *msh* and was also significantly inhibited. ALFM13 contains the entire intergenic region and none of *msh*, and may have been inhibited somewhat. All other fusions were not affected by methionine. These data indicate that a region near the *msh* start codon is required for inhibition (Fig. 4.3b).

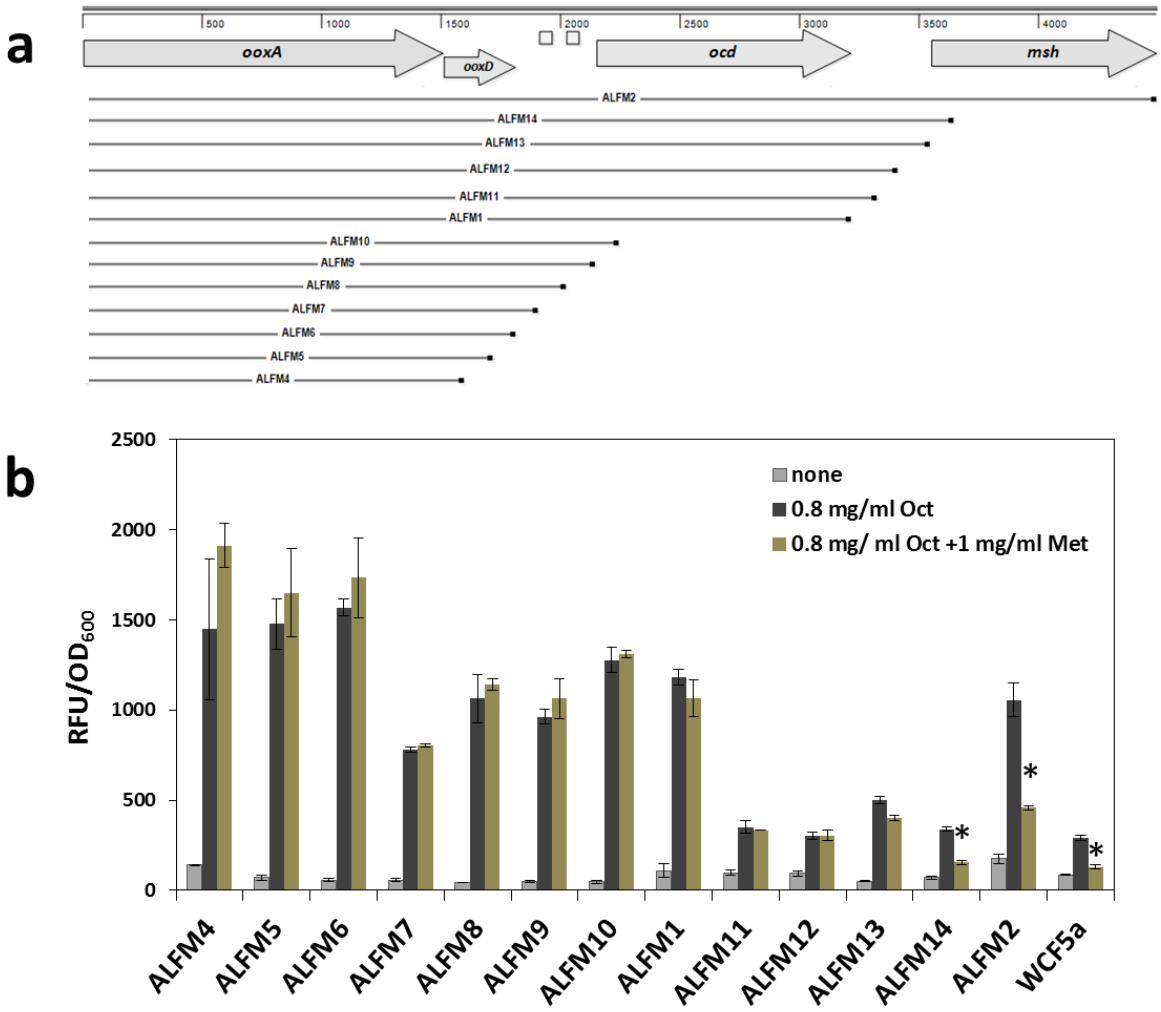


Figure 4.3 Mapping the region involved in the repression of the *occ* operon by methionine. a) Map of the fusions between *ooxA* and *msh*. The small black squares at the end of the lines represent the junction between Ti plasmid sequences and *lacZ*. The gray squares symbolize arginase gene scars. b) Gene expression of the operon in the absence (none) or presence of octopine containing or lacking methionine. x-axis strains in different treatment. y-axis *lacZ* expression presented in Relative Fluorescence Units (RFU) normalized against the optical density (OD) of the cultures. Mean ± SD of n=3. *P<0.001, *t* test relative to samples with octopine.

Once I located the region of repression, I tested 8 fusions (Fig. 4.4a), in the presence of a higher concentration of octopine and different concentrations of methionine to confirm the effect that I described above. For this experiment I used 2 mg/ml of octopine and four different concentrations of methionine. I found that even the lowest concentration of methionine used was sufficient for this downregulation. The results were consistent with those previously found (Fig. 4.3 b), where a sequence from the first 97 bp of *msh* (97-*msh*) coding region was needed for this downregulation. The repression found in the fusion ALFM14, which contains the 97-*msh*, was similar to the repression of the *traR-lacZ* (WCF5a), indicating that a sequence of *msh* is required for this regulation. I tested a smaller set of fusions with even higher methionine concentrations, finding the same results that a region within the *msh* coding sequence plays an important role in repression (Fig. 4.4b, Table 4.4).

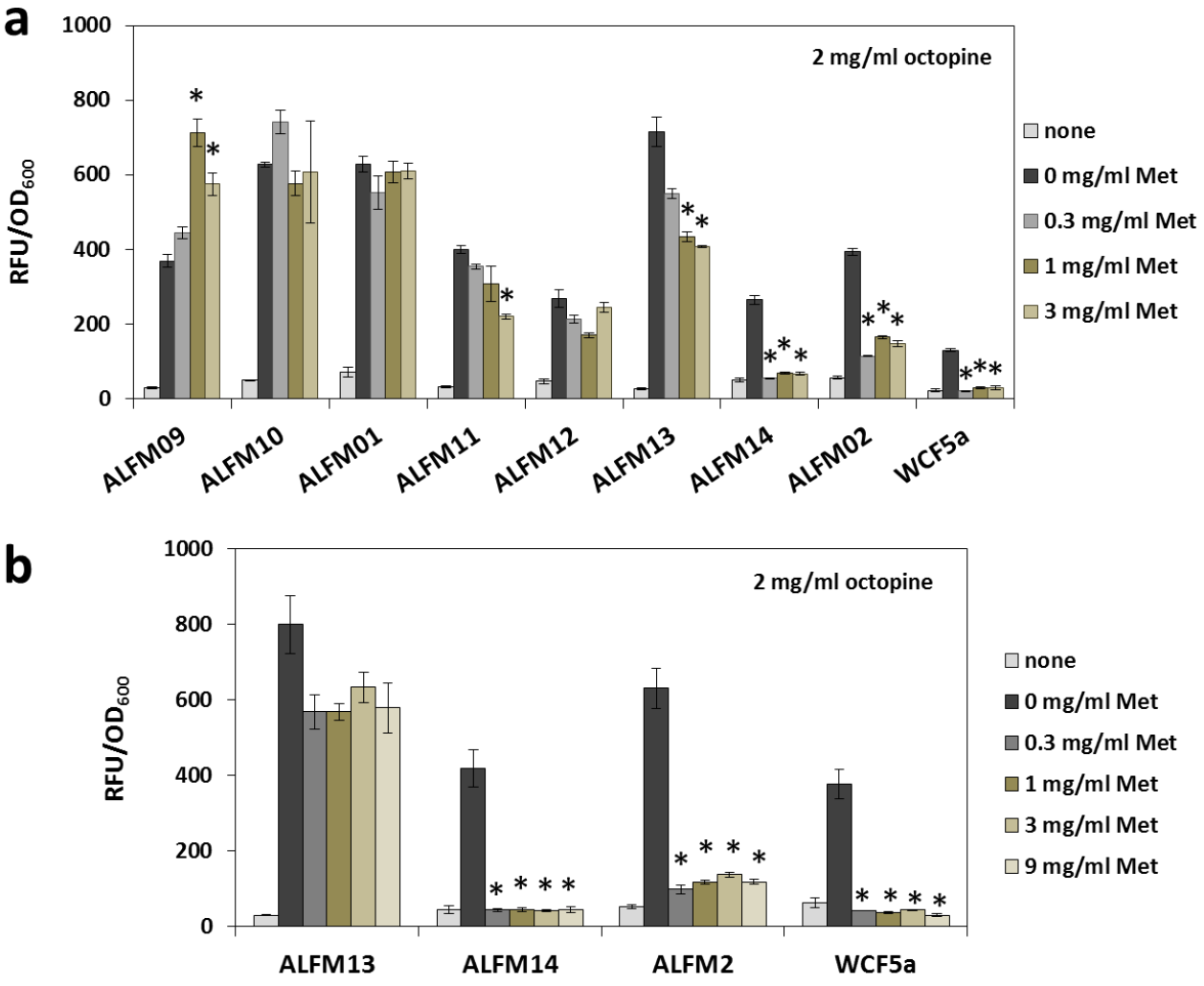


Figure 4.4 Repression of the *msh* and downstream genes by methionine. a) Mapping of the region responsible for the repression. b) Testing a smaller set of fusions with higher concentration of methionine. Gene expression of fusions in the absence (none) or presence of 2 mg/ ml of octopine containing or lacking methionine. x-axis strains in different treatment. y-axis *lacZ* expression presented in Relative Fluorescence Units (RFU) normalized against the optical density (OD). Mean ± SD of n=3. *P<0.001, *t* test relative to samples with octopine.

Table 4.4 Repression ratio of the expression by methionine

	Repression ratio			
	Methionine (mg/ml)			
	0.3	**1**	**3**	**9**
ALFM13	1.4	1.4	1.3	1.2
ALFM14	8.7	8.5	9.0	8.6
ALFM2	6.4	5.3	4.6	5.3
WCF5a	9.0	10.2	8.5	12.4

The intergenic region is enough for the recognition of methionine

The previous experiments with the *lacZ* fusions enabled me to identify the 3' end of the regulatory sequence required for methionine inhibition. However the 5'end sequence was not identified by this approach. To address this two PCR fragments were produced, one of which contains 429 nucleotides that span the entire intergenic region upstream of *msh* and the first 97 nucleotides of *msh*. The second PCR fragment contains 51 nucleotides upstream of *msh* and the first 97 nucleotides of *msh*. These fragments were cloned into pAFM165, generating pAFM166 and pAFM167, respectively (Fig. 4.5a). Plasmid pAFM165 contains a *lac* promoter that is constitutively expressed and a *lacZ* reporter gene, and was designed to construct transcriptional P*lac-lacZ* fusions that contain foreign sequences between the promoter and the reporter. The two plasmids were electroporated into *A. tumefaciens* strain KYC55 (which lacks a Ti plasmid). The resulting strains were cultured in AT medium in the absence or presence of methionine and assayed for β-galactosidase activity. Plasmid pAFM166 was strongly affected by

methionine compared with the vector control. pAFM167 was also affected by methionine but not as strongly, suggesting that 147 bp are not enough for the complete repression of the expression of *msh* (Fig. 4.5b). I conclude that all sequences required for inhibition by methionine are located within the intergenic region, and that pAFM167 may lack sequences required for full inhibition.

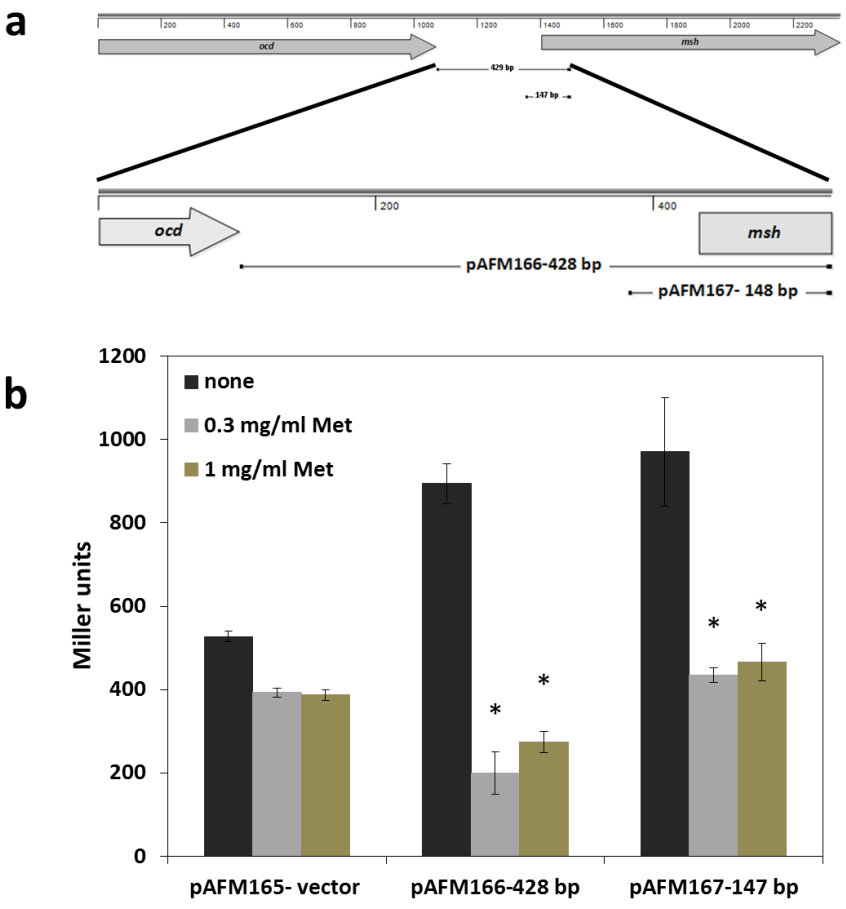


Figure 4.5 Mapping the 5'end of the region regulated by methionine. a) Size of the fragments from the *ocd-msh* sequence region. b) β-galactosidase activity of the fusions in the absence or presence of methionine. x-axis fusions. y-axis miller units. Mean ± SD of n=3. *P<0.001, *t* test relative to none sample.

Methionine interacts directly with the mRNA of the intergenic sequence

I was intrigued by possibility that methionine-inhibition might be RNA-mediated rather than protein-mediated. To address this, RNA of this region was isolated, and analyzed it in the presence or absence of methionine by SHAPE (Selective 2'-Hydroxyl Acylation analyzed by Primer Extension) (Fig. 4.6). This technique involves the use of the reagent 1M7, which covalently modifies the 2' OH residues of RNA bases, leading to strand cleavage. This reagent is selective for unstructured bases whose 2' OH groups are more exposed to solvent than in bases that are structured by interactions with ligands or other portions of the RNA molecule. RNA fragments generated by this reagent are identified by primer extension (Fig. 4.7).

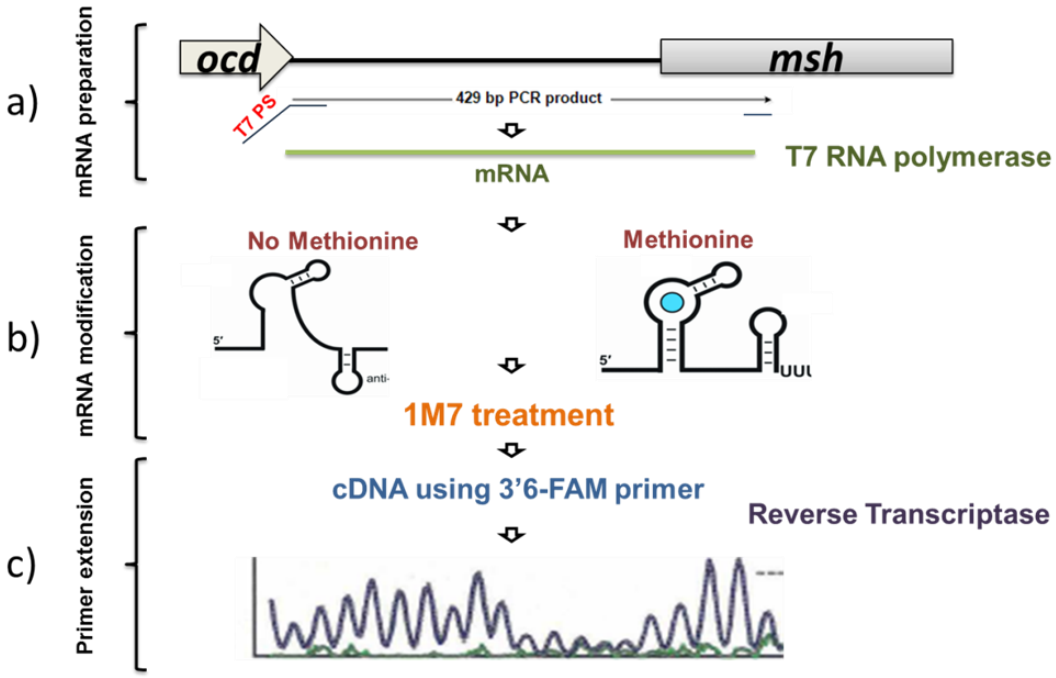
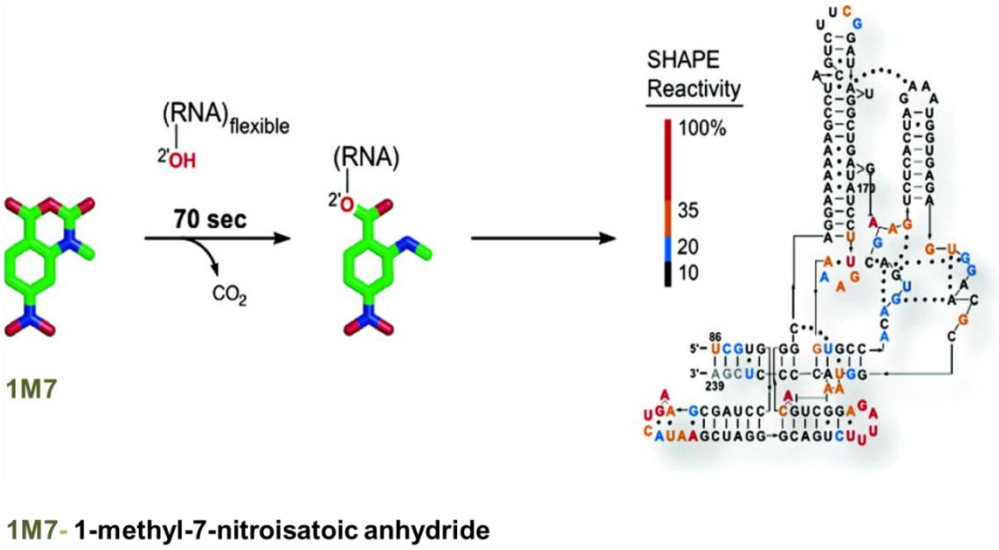


Figure 4.6 Procedure for analyzing the methionine riboswitch by using SHAPE. a) mRNA preparation, b) mRNA modification by adding methionine, and c) primer extension.



1M7- 1-methyl-7-nitroisatoic anhydride

Figure 4.7 Selective 2'-hydroxyl acylation analyzed by primer extension chemistry for resolving mRNA structures. Modified from (Mortimer & Weeks 2007).

I produced an mRNA 429 nucleotides in length that contains the *ocd-msh* intergenic region and the first 97 nucleotides of *msh* (Fig. 4.8). The mRNA was combined with methionine or in a control reaction, with water. Reagent 1M7 was then added, which interacts with the 2' carbon of unconstrained nucleotides, generating a fragmentation pattern. mRNA fragments were converted to cDNA using reverse transcriptase and a primer containing a fluorescent label (FAM) and then analyzed using capillary electrophoresis detecting the FAM (Fig. 4.9a). The SHAPE experiment showed that methionine induced conformational changes in the mRNA, since the fragmentation pattern of the mRNA in the absence of methionine was different from the mRNA that was treated with methionine (Figure 4.9b). I found that the main differences were within the *msh* coding region and in the adjacent 60 bp of the upstream intergenic region (Fig. 4.9c and d).

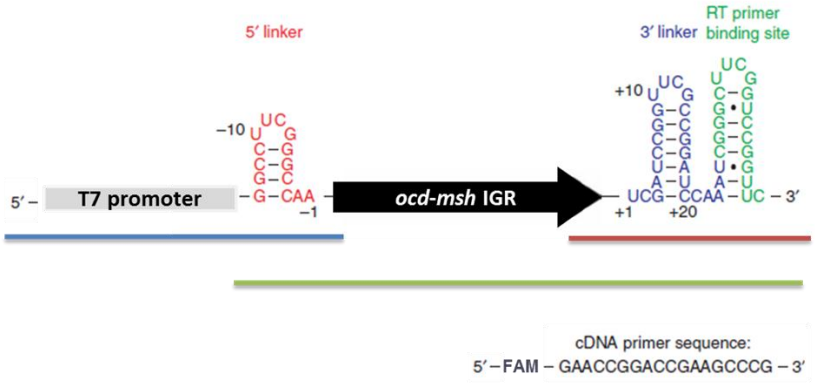
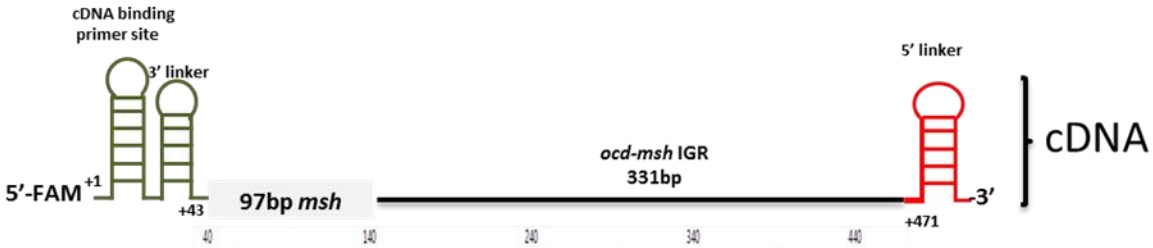
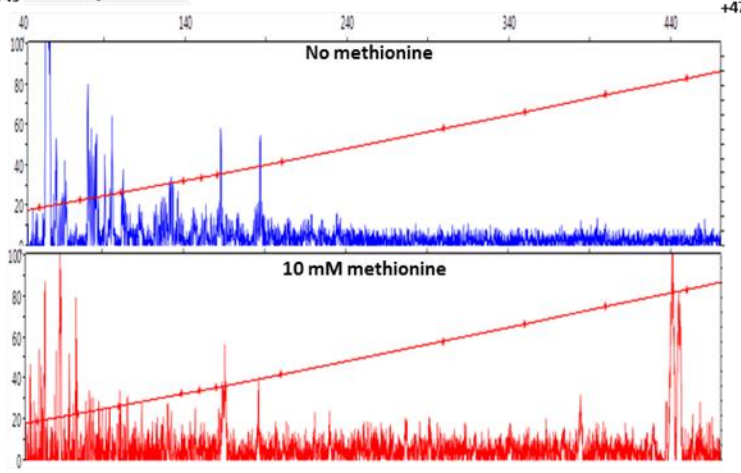


Figure 4.8 Structure of the PCR product used to generate mRNA. The blue line represents the forward primer, which contains a T7 promoter for mRNA synthesis, the 5' linker, and some of the sequence of interest. The red line represents the reverse primer, which contains the RT primer binding site, 3' linker and some of the sequence of interest. The green line represents the generated mRNA sequence. The 5' and 3' linkers are RNA structure cassettes that facilitate the analysis of all nucleotides within the *ocd-msh* intergenic region (IGR). The RT primer binding site is the sequence that cDNA primer recognizes to generate the cDNA. The cDNA primer contains a FAM label. Modified from (Wilkinson et al 2006).

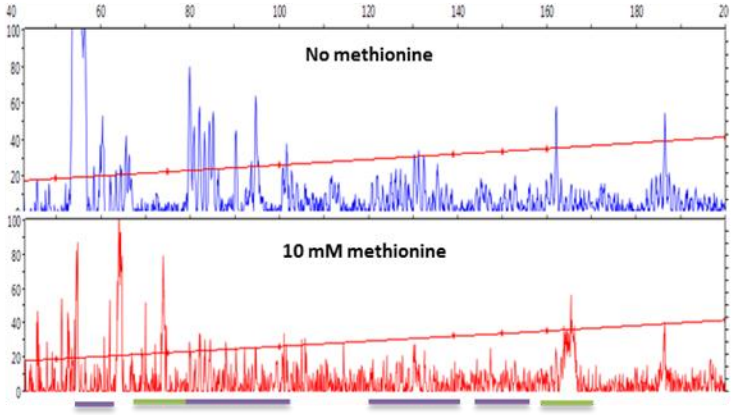
a



b



c



d

```
cDNA 5′-GCGACAGCGCAGACCACTCGGGTTGCCGGAAGGGAGCGCCATTTCTCAG   49
        CAGTTCCCTCCCCATTCCCCCATCCAGGATGGTGACTTTCGATGACATG   98
        CGGCCTCCAGCTTACGTCCGTTGAGAGTTATCCTTGCCGGCGTTTGATA  135
        ATTTAGATCG-3′                                      157
```

Figure 4.9 Comparison of the mRNA conformational folding in the absence or presence of methionine evaluated by SHAPE analysis. a) Shows the map of the cDNA, the 43 bp consists of the cDNA primer binding site and the 3'linker; from 44 bp to 471 bp, which contains the intergenic region (IGR) from *ocd-msh* and the first 97 nucleotides of *msh;* and from 471 to 491 consist of the 5' linker. cDNA was generated by a reverse FAM-labeled primer that recognizes the cDNA binding site of the mRNA. b) Comparison of the fragmentation of the region of interest (428 bp) in the absence or presence of methionine. c) Comparison of the fragmentation of the first 97 bp of *msh* and 60 bp of the *ocd-msh* IGR treated with or without methionine. The purple lines represent sequences that were not protected in the absence of methionine but are protected in the presence of methionine; black lines represent regions that are protected in the absence of methionine but unprotected in the presence of methionine. d) Sequence of cDNA containing 97 nucleotides of *msh* and 60 bp of the *ocd-msh* IGR; in red is the start of *msh* coding region; RBS is underlined; in purple protected bases and in bold green unprotected bases in the presence of methionine. The red line across the spectrograms represents the size standard. x-axis is the base pairs, y-axis peak intensity.

4.5 Discussion

The expression of *tra*, *trb* and *rep* genes is regulated in surprisingly complex ways, given that homologous genes of other plasmids are often expressed constitutively or autoregulated. First, expression of these genes is impacted by octopine, which activates expression of *traR*, the master regulator of other *tra* and *trb* genes (Fuqua & Winans 1994a, Piper et al 1993) . Second, expression of these genes is inhibited by another opine, mannopine, which activates the expression of a defective *traR*-like gene called *trlR* (Chai et al 2001, Oger et al 1998, Zhu & Winans 1998). TrlR is almost identical to the N-terminal domain of TraR, but lacks a DNA binding domain. TrlR forms inactive heterodimers with TraR, blocking TraR activity. Expression of *tra* and *trb* genes also requires a quorum of conjugal donors, which is required to produce sufficient OOHL (Chen et al 2007, Fuqua et al 1995, Fuqua & Winans 1994a). Yet another level of control is mediated by the antiactivator TraM, which forms inactive complexes with TraR and ultimately stimulates TraR proteolysis (Hwang et al 1995, Luo et al 2000, Qin et al 2007). I now show that expression of TraR is also impacted by methionine, and that this inhibition is mediated by a site midway in the *occ* region. It is far from clear what advantage is to be gained from regulating conjugation by two different opines (in opposite ways!), by a quorum of conjugal donors, and a TraR antiactivator, and by methionine.

The downregulation of *msh* by methionine does make some sense from a teleological point of view, in that Msh creates a source of methionine from S-methyl methionine and homocysteine (see Chapter 2). An abundance of methionine could

therefore cause synthesis of Msh to be unnecessary. This reasoning could conceivably also extend to the *ophABCD* genes. These genes appear to encode an ABC-type uptake system. The specificity of this permease is not known, but is most like to include octopine-type opines, given that these genes lie within the *occ* operon. I have shown that octopine-synthase can make at least a dozen or more opines by conjugating pyruvate with the corresponding amino acids (Chapter 3). Perhaps some are internalized by the *occQMPJ* products (including octopine itself) while others are internalized by *ophABCD*. However, only the latter genes are inhibited by methionine. This provides a suggestive hint that perhaps *ophABCD* might direct a methionine-containing opine such as methiopine or sulfonopine.

In an earlier study, other members of the Winans lab screened a library of randomly generated transcriptional fusions made using Tn5*gusA7* for increased expression in response to octopine. Among the set of octopine-induced genes was *cysD,* which encodes a cysteine biosynthetic gene. The induction was rather weak and the mechanism has not been explored. These mutants grew poorly on minimal agar media, due to trace amounts of cysteine in the agar. Curiously, under this form of cysteine limitation, these strains were strongly hyperconjugal (Cho et al 1997). The mechanism has not yet been unraveled, but elevated conjugation by cysteine limitation could have the same root cause as decreased conjugation by exogenous methionine.

In *E. coli*, methionine biosynthesis is regulated by two transcriptional regulators, MetJ and MetR, a repressor and an activator, respectively (Kumar & Gomes 2005). MetJ

interacts with S-adenosylmethionine, and binds to a regulatory sequence, called a *met* box, and represses the transcription of the *met* genes. MetR binds homocysteine, a methionine precursor, and stimulates the expression of *metE* and *metH* genes, which direct methionine synthesis from homocysteine and two different methyl donors. It has also been shown that homocysteine enhances the MetR activation of *metE* expression (Augustus et al 2006, Fritsch et al 2000, LaMonte & Hughes 2006, Liu et al 2001, Weissbach & Brot 1991). I have examined the DNA near *msh* for sequences resembling the MetJ or MetR binding sites, but have not found any such sequences. However, this result is not enough to rule out a protein regulation mechanism yet. Nevertheless, I favor a riboswitch mechanism because by SHAPE analysis I found that methionine directly interacts with this region of the mRNA by changing its structural conformation. I found that the major changes were in the *msh* coding region and 60 bp upstream of *msh*. The regulation of amino acid biosynthesis by riboswitch is not new; the cellular concentrations of lysine and glycine are regulated by a riboswitch-based control (Mandal et al 2004, Serganov & Patel 2009, Sudarsan et al 2003). There are numerous examples where expression of metabolic genes is tightly regulated through multiple mechanisms at several levels (transcriptional, post-transcriptional or post-translational) by their biosynthetic products or intermediates. By analogy I can expect a similar situation for *msh*.

In this Chapter, I showed that methionine indeed represses conjugation as previously reported (Hooykaas et al 1979). Moreover this repression is by an indirect mechanism, since methionine is repressing its biosynthetic gene, *msh*, and in

consequence is also repressing the expression of *traR*. TraR is a master transcriptional regulator that activates the quorum sensing, conjugation, and replication of the Ti plasmid. Therefore nutritional sensing plays a major role in the *traR* regulon.

4.6 References

Alt-Morbe J, Stryker JL, Fuqua C, Li PL, Farrand SK, Winans SC. 1996. The conjugal transfer system of *Agrobacterium tumefaciens* octopine-type Ti plasmids is closely related to the transfer system of an IncP plasmid and distantly related to Ti plasmid vir genes. **Journal of Bacteriology** 178: 4248-57

Augustus AM, Reardon PN, Heller WT, Spicer LD. 2006. Structural basis for the differential regulation of DNA by the methionine repressor MetJ. **Journal of Biological Chemistry** 281: 34269-76

Cangelosi GA, Best EA, Martinetti G, Nester EW. 1991. Genetic analysis of *Agrobacterium*. **Methods in Enzymology** 204: 384-97

Casadaban MJ, Chou J, Cohen SN. 1980. In vitro gene fusions that join an enzymatically active beta-galactosidase segment to amino-terminal fragments of exogenous proteins: *Escherichia coli* plasmid vectors for the detection and cloning of translational initiation signals. **Journal of Bacteriology** 143: 971-80

Chai Y, Zhu J, Winans SC. 2001. TrlR, a defective TraR-like protein of *Agrobacterium tumefaciens*, blocks TraR function in vitro by forming inactive TrlR:TraR dimers. **Mol Microbiol** 40: 414-21

Chen CY, Winans SC. 1991. Controlled expression of the transcriptional activator gene *vir*G in *Agrobacterium tumefaciens* by using the *Escherichia coli lac* promoter. **Journal of Bacteriology** 173: 1139-44

Chen G, Jeffrey PD, Fuqua C, Shi Y, Chen L. 2007. Structural basis for antiactivation in bacterial quorum sensing. **Proc Natl Acad Sci U S A** 104: 16474-9

Cho K, Fuqua C, Martin BS, Winans SC. 1996. Identification of *Agrobacterium tumefaciens* genes that direct the complete catabolism of octopine. **Journal of Bacteriology** 178: 1872-80

Cho K, Fuqua C, Winans SC. 1997. Transcriptional regulation and locations of *Agrobacterium tumefaciens* genes required for complete catabolism of octopine. **Journal of Bacteriology** 179: 1-8

Collado-Vides J, Magasanik B, Gralla JD. 1991. Control site location and transcriptional

regulation in *Escherichia coli*. **Microbiological Reviews** 55: 371-94

Cromie MJ, Shi Y, Latifi T, Groisman EA. 2006. An RNA sensor for intracellular Mg(2+). **Cell** 125: 71-84

Dessaux Y, Petit A, Ellis JG, Legrain C, Demarez M, et al. 1989. Ti plasmid-controlled chromosome transfer in *Agrobacterium tumefaciens*. **Journal of Bacteriology** 171: 6363-66

Fritsch PS, Urbanowski ML, Stauffer GV. 2000. Role of the RNA polymerase alpha subunits in MetR-dependent activation of *metE* and *metH*: Important residues in the C-terminal domain and orientation requirements within RNA polymerase. **Journal of Bacteriology** 182: 5539-50

Fuqua C, Burbea M, Winans SC. 1995. Activity of the *Agrobacterium* Ti plasmid conjugal transfer regulator Trar is inhibited by the product of the *traM g*ene. **Journal of Bacteriology** 177: 1367-73

Fuqua C, Winans SC. 1996a. Conserved cis-acting promoter elements are required for density-dependent transcription of *Agrobacterium tumefaciens* conjugal transfer genes. **Journal of Bacteriology** 178: 435-40

Fuqua C, Winans SC. 1996b. Localization of OccR-activated and TraR-activated promoters that express two ABC-type permeases and the traR gene of Ti plasmid pTiR10. **Mol Microbiol** 20: 1199-210

Fuqua WC, Winans SC. 1994a. A Luxr-Luxi Type Regulatory System Activates Agrobacterium Ti Plasmid Conjugal Transfer in the Presence of a Plant Tumor Metabolite. **Journal of Bacteriology** 176: 2796-806

Fuqua WC, Winans SC. 1994b. A LuxR-LuxI type regulatory system activates *Agrobacterium* Ti plasmid conjugal transfer in the presence of a plant tumor metabolite. **Journal of Bacteriology** 176: 2796-806

Gallagher SR. 1992. *Gus Protocols: Using the Gus gene as a reporter of gene expression*. San Diego: Academic Press.

Genetello C, Van Larebeke N, Holsters M, De Picker A, Van Montagu M, Schell J. 1977. Ti plasmids of *Agrobacterium* as conjugative plasmids. **Nature** 265: 561-3

Groisman EA, Cromie MJ, Shi Y, Latifi T. 2006. A Mg2+-responding RNA that controls the expression of a Mg2+ transporter. **Cold Spring Harbor symposia on quantitative biology** 71: 251-8

Grundy FJ, Winkler WC, Henkin TM. 2002. tRNA-mediated transcription antitermination in vitro: codon-anticodon pairing independent of the ribosome. **Proc Natl Acad Sci U S A** 99: 11121-6

Habeeb LF, Wang L, Winans SC. 1991. Transcription of the octopine catabolism operon of the *Agrobacterium* tumor-inducing plasmid pTiA6 is activated by a LysR-type regulatory protein. **Mol Plant Microbe Interact** 4: 379-85

Hooykaas PJJ, Roobol C, Schilperoort RA. 1979. Regulation of the transfer of Ti plasmids of *Agrobacterium tumefaciens*. **J. Gen. Microbiol**. 110: 99-109

Hwang IY, Cook DM, Farrand SK. 1995. A New Regulatory Element Modulates Homoserine Lactone-Mediated Autoinduction of Ti Plasmid Conjugal Transfer. **Journal of Bacteriology** 177: 449-58

Ishihama A. 2010. Prokaryotic genome regulation: multifactor promoters, multitarget regulators and hierarchic networks. **FEMS Microbiology Reviews** 34: 628-45

Johansson J, Mandin P, Renzoni A, Chiaruttini C, Springer M, Cossart P. 2002. An RNA thermosensor controls expression of virulence genes in Listeria monocytogenes. **Cell** 110: 551-61

Kalogeraki VS, Winans SC. 1997. Suicide plasmids containing promoterless reporter genes can simultaneously disrupt and create fusions to target genes of diverse bacteria. **Gene** 188: 69-75

Kerr A. 1969. Transfer of virulence between isolates of *Agrobacterium*. **Nature** 223: 1175-776

Kerr A, Manigault P, Tempe J. 1977. Transfer of virulence *in vivo* and *in vitro* in *Agrobacterium*. **Nature** 265: 560-1

Klapwijk PM, Scheulderman T, Schilperoort RA. 1978. Coordinated regulation of octopine degradation and conjugative transfer of Ti plasmids in *Agrobacterium tumefaciens*: evidence for a common regulatory gene and separate operons. **Journal of Bacteriology** 136: 775-85

Kovach ME, Elzer PH, Hill DS, Robertson GT, Farris MA, et al. 1995. Four new derivatives of the broad-host-range cloning vector pBBR1MCS, carrying different antibiotic-resistance cassettes. **Gene** 166: 175-6

Kovach ME, Phillips RW, Elzer PH, Roop RM, 2nd, Peterson KM. 1994. pBBR1MCS: a broad-host-range cloning vector. **BioTechniques** 16: 800-2

Kumar D, Gomes J. 2005. Methionine production by fermentation. **Biotechnol Adv** 23: 41-61

LaMonte BL, Hughes JA. 2006. In vivo hydrolysis of S-adenosylmethionine induces the met regulon of *Escherichia coli*. **Microbiol-Sgm** 152: 1451-59

Lease RA, Cusick ME, Belfort M. 1998. Riboregulation in *Escherichia coli*: DsrA RNA acts by RNA:RNA interactions at multiple loci. **Proc Natl Acad Sci U S A** 95: 12456-61

Liu RX, Blackwell TW, States DJ. 2001. Conformational model for binding site recognition by the *E. coli* MetJ transcription factor. **Bioinformatics** 17: 622-33

Luo ZQ, Farrand SK. 2001. The *Agrobacterium tumefaciens* rnd homolog is required for TraR-mediated quorum-dependent activation of Ti plasmid tra gene expression. **Journal of Bacteriology** 183: 3919-30

Luo ZQ, Qin YP, Farrand SK. 2000. The antiactivator TraM interferes with the autoinducer-dependent binding of TraR to DNA by interacting with the C-terminal region of the quorum-sensing activator. **Journal of Biological Chemistry** 275: 7713-22

Mandal M, Breaker RR. 2004. Gene regulation by riboswitches. **Nature Reviews: Molecular Cell Biology** 5: 451-63

Mandal M, Lee M, Barrick JE, Weinberg Z, Emilsson GM, et al. 2004. A glycine-dependent riboswitch that uses cooperative binding to control gene expression. **Science** 306: 275-9

Montange RK, Batey RT. 2008. Riboswitches: Emerging themes in RNA structure and function. **Annu Rev Biophys** 37: 117-33

Moré MI, Finger LD, Stryker JL, Fuqua C, Eberhard A, Winans SC. 1996. Enzymatic synthesis of a quorum-sensing autoinducer through use of defined substrates. **Science** 272: 1655-58

Mortimer SA, Weeks KM. 2007. A fast-acting reagent for accurate analysis of RNA secondary and tertiary structure by SHAPE chemistry. **Journal of the American Chemical Society** 129: 4144-5

Oger P, Farrand SK. 2002. Two opines control conjugal transfer of an *Agrobacterium* plasmid by regulating expression of separate copies of the quorum-sensing activator gene *tra*R. *J Bacteriol* 184: 1121-31

Oger P, Kim KS, Sackett RL, Piper KR, Farrand SK. 1998. Octopine-type Ti plasmids code for a mannopine-inducible dominant-negative allele of *tra*R, the quorum-sensing activator that regulates Ti plasmid conjugal transfer. **Molecular Microbiology** 27: 277-88

Pappas KM, Winans SC. 2003. A LuxR-type regulator from *Agrobacterium tumefaciens* elevates Ti plasmid copy number by activating transcription of plasmid replication genes. **Mol Microbiol** 48: 1059-73

Petit A, Tempe J, Kerr A, Holsters M, Van Montagu M, Schell J. 1978. Substrate induction of conjugative activity of *Agrobacterium tumefaciens* Ti plasmids. **Nature** 271: 570-71

Piper KR, Beck von Bodman S, Farrand SK. 1993. Conjugation factor of *Agrobacterium tumefaciens* regulates Ti plasmid transfer by autoinduction. **Nature** 362: 448-50

Qin YP, Luo ZQ, Farrand SK. 2004. Domains formed within the N-terminal region of the quorum-sensing activator TraR are required for transcriptional activation and direct interaction with RpoA from *Agrobacterium*. **Journal of Biological Chemistry** 279: 40844-51

Qin YP, Luo ZQ, Smyth AJ, Gao P, von Bodman SB, Farrand SK. 2000. Quorum-sensing signal binding results in dimerization of TraR and its release from membranes into the cytoplasm. **EMBO Journal** 19: 5212-21

Qin YP, Su SC, Farrand SK. 2007. Molecular basis of transcriptional antiactivation - TraM disrupts the TraR-DNA complex through stepwise interactions. **Journal of Biological Chemistry** 282: 19979-91

Regulski EE, Breaker RR. 2008. In-line probing analysis of riboswitches. **Methods Mol Biol** 419: 53-67

Sambrook J, Fritsch EF, Maniatis T. 1989. *Molecular cloning: a laboratory manual.* . **Cold Spring Harbor**, NY Cold Spring Harbor Laboratory Press.

Serganov A, Patel DJ. 2007. Ribozymes, riboswitches and beyond: regulation of gene expression without proteins. **Nature Reviews: Genetics** 8: 776-90

Serganov A, Patel DJ. 2009. Amino acid recognition and gene regulation by riboswitches. **Biochimica et Biophysica Acta** 1789: 592-611

Simon R, Priefer U, Pühler A. 1983. A broad host range mobilization system for in vivo genetic engineering: Transposon mutagenesis in gram negative bacteria. **Nature Biotechnology** 1: 784-92

Studier FW, Rosenberg AH, Dunn JJ, Dubendorff JW. 1990. Use of T7 RNA-Polymerase to direct expression of cloned genes. **Methods in enzymology** 185: 60-89

Sudarsan N, Hammond MC, Block KF, Welz R, Barrick JE, et al. 2006. Tandem riboswitch architectures exhibit complex gene control functions. **Science** 314: 300-4

Sudarsan N, Wickiser JK, Nakamura S, Ebert MS, Breaker RR. 2003. An mRNA structure in bacteria that controls gene expression by binding lysine. **Genes & development** 17: 2688-97

Tartoff KD, Hobbs CA. 1987. Improved media for growing plasmid and cosmid clones. **Bethesda Res. Lab. Focus** 9: 12

Tucker BJ, Breaker RR. 2005. Riboswitches as versatile gene control elements. **Curr Opin Struct Biol** 15: 342-48

Valdivia RH, Wang L, Winans SC. 1991. Characterization of a putative periplasmic transport system for octopine accumulation encoded by *Agrobacterium tumefaciens* Ti plasmid pTiA6. **Journal of Bacteriology** 173: 6398-405

Van Larebeke N, Genetello C, Schell J, Schilperoort RA, Hermans AK, et al. 1975. Acquisition of tumour-inducing ability by non-oncogenic agrobacteria as a result of plasmid transfer. **Nature** 255: 742-3

Weissbach H, Brot N. 1991. Regulation of methionine synthesis in *Escherichia coli*. **Mol Microbiol** 5: 1593-7

Wilkinson KA, Merino EJ, Weeks KM. 2006. Selective 2 '-hydroxyl acylation analyzed by primer extension (SHAPE): quantitative RNA structure analysis at single nucleotide resolution. **Nature Protocols** 1: 1610-16

Winkler WC, Breaker RR. 2005. Regulation of bacterial gene expression by riboswitches. **Annual Review of Microbiology** 59: 487-517

Zaenen I, Van Larebeke N, Van Montagu M, Schell J. 1974. Supercoiled circular DNA in crown-gall inducing *Agrobacterium* strains. **Journal of Molecular Biology** 86: 109-27

Zhou D, Yang R. 2006. Global analysis of gene transcription regulation in prokaryotes. **Cellular and Molecular Life Sciences: CMLS** 63: 2260-90

Zhu J, Oger PM, Schrammeijer B, Hooykaas PJ, Farrand SK, Winans SC. 2000. The bases of crown gall tumorigenesis. **Journal of Bacteriology** 182: 3885-95

Zhu J, Winans SC. 1998. Activity of the quorum-sensing regulator TraR of *Agrobacterium tumefaciens* is inhibited by a truncated, dominant defective TraR-like protein. **Molecular Microbiology** 27: 289-97

Zhu J, Winans SC. 1999. Autoinducer binding by the quorum-sensing regulator TraR increases affinity for target promoters in vitro and decreases TraR turnover rates in whole cells. **Proc Natl Acad Sci U S A** 96: 4832-7

Zhu J, Winans SC. 2001. The quorum-sensing transcriptional regulator TraR requires its cognate signaling ligand for protein folding, protease resistance, and dimerization. **Proc Natl Acad Sci U S A** 98: 1507-12

CHAPTER 5

Conclusion and Future Work Directions

5.1 Conclusions

In this thesis I show the importance of SMM, a plant produced molecule, in nutrient acquisition by *A. tumefaciens*, as well as its role as a plant-released signal molecule that impacts conjugation and replication of the Ti plasmid. In Chapter 2 I showed that the *msh* gene, located within the octopine catabolism (*occ*) operon, encodes for a methyltransferase, Msh, that synthesizes methionine by transfer of a methyl group from SMM to homocysteine. Msh preferentially uses SMM over other methyl donor molecules, a finding that agrees with other studies done in SMM:homocysteine methyltransferases (Fig. 5.1 step 10).

However, the main question was how SMM accumulates by these cells, and why this gene would be found in an opine catabolic operon. This bacterium is a plant pathogen that directs and integrates a piece of DNA (T-DNA) into the plant chromosome (Fig. 5.1 step 1). This T-DNA contains genes that encode for tumor formation (Fig. 5.1 step 2) and for nutrient production by condensing positively charged amino acid and pyruvate (opines) (Fig. 5.1 step 3). Opines are released from the plant and then they are taken up and catabolized by the colonizing bacteria, inducing the transcription of the *occ* operon, which contains genes for their internalization and utilization. Since *msh* is located in the *occ* operon we hypothesized it plays a role in opine catabolism. If so, then SMM must be

an intermediate in opine breakdown, and that SMM is taken up from the plant in its opine form. In Chapter 3, we demonstrated *in vitro* and *in vivo* that SMM is a substrate for Ocs together with pyruvate to produce a new opine, sulfonopine (Fig. 5.1 step 3). The production of sulfonopine is dependent on both Ocs and the SMM plant concentration in *Arabidopsis* or tobacco plants infected by *A. tumefaciens.* We also showed that in tobacco seedlings once sulfonopine is produced, it is released to the media (Fig. 5.1 step 4). Upon uptake by the bacterium, sulfonopine interacts with OccR to activate the expression of the *occ* operon (Fig. 5.1 steps 5, 6 and 7) and it was used as a sulfur source by *A. tumefaciens*.

Once sulfonopine is taken up into the bacterial cell and catabolized, producing SMM and pyruvate, SMM becomes available for utilization by Msh to produce methionine (Fig. 5.1 steps 9, 10 and 11). A previous study showed that methionine represses the conjugation of the Ti plasmid. In Chapter 4 we showed that methionine represses the transcription of *msh* and downstream genes including *tra*R, and subsequently TraR-regulated genes (Fig. 5.1 8). We mapped a region containing the intergenic sequence between ocd and *msh* together with the first 97 bp of *msh* that is sufficient for this downregulation by methionine (Fig. 5.1 step 12). We also have evidence that methionine interacts directly with the mRNA of interest, suggesting that methionine could be regulating the *msh* and downstream genes by a riboswitch mechanism.

Overall, SMM is taken up by *A. tumefaciens* from infected plant tissue in its opine form, sulfonopine, and uses it as a sulfur source and signal molecule. Sulfonopine is then

catabolized generating SMM which is used by Msh to produce methionine. However, when the concentration of methionine increases, it downregulates its biosynthetic gene, and consequently, the downstream genes including *tra*R. Thus, by inhibiting TraR, quorum sensing, conjugation and replication of the Ti plasmid are down regulated.

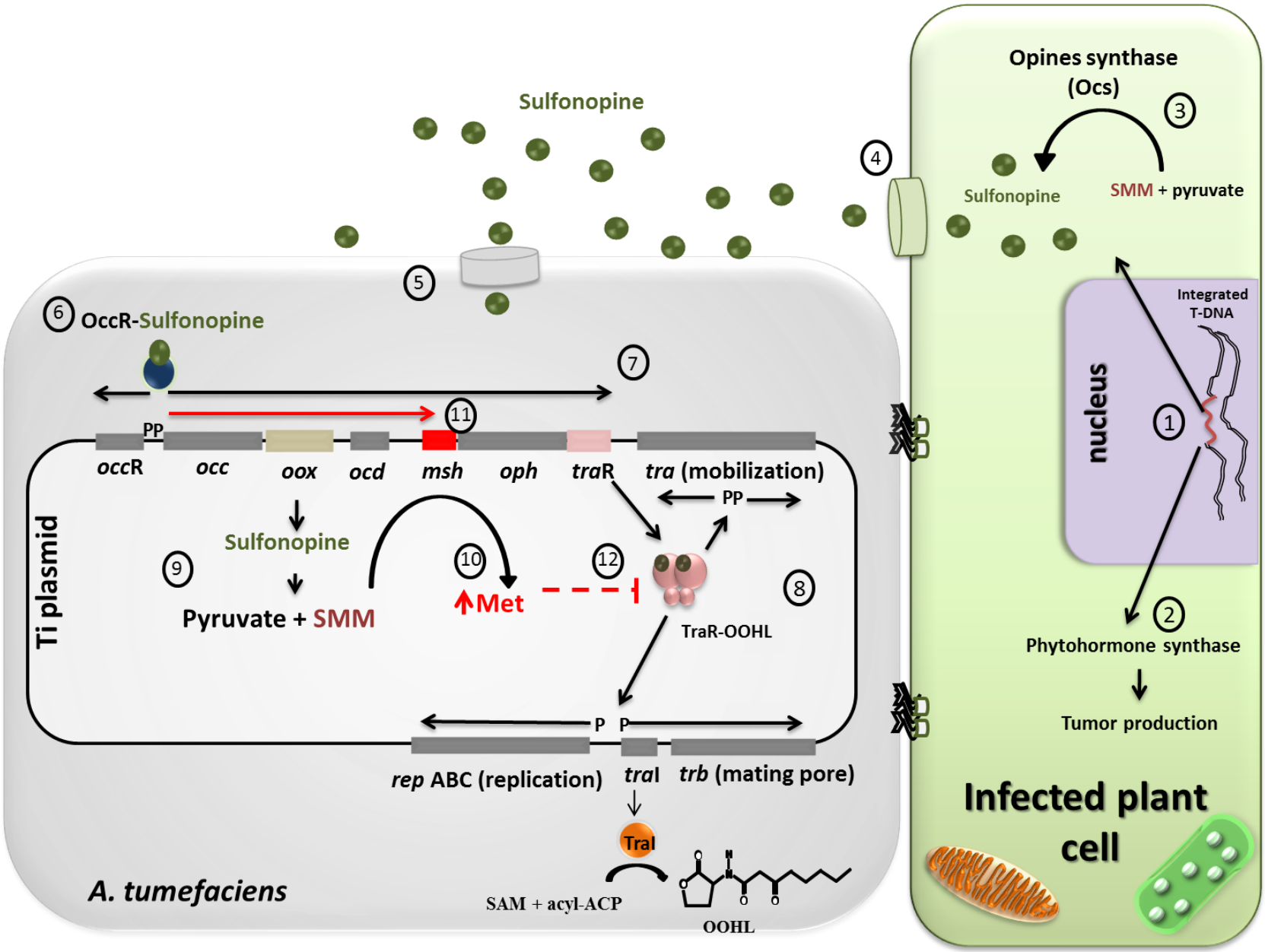

Sulfonopine
Opines synthase
(Ocs)
3
4
Sulfonopine
SMM + pyruvate
5
6
OccR-Sulfonopine
7
11
PP
occR
occ
oox
ocd
msh
oph
traR
tra (mobilization)
PP
Ti plasmid
Sulfonopine
9
10
12
8
Pyruvate + SMM
Met
TraR-OOHL
P P
rep ABC (replication)
traI
trb (mating pore)
TraI
SAM + acyl-ACP
OOHL
A. tumefaciens
nucleus
Integrated
T-DNA
1
2
Phytohormone synthase
Tumor production
Infected plant
cell

Figure 5.1 Synthesis and utilization of sulfonopine. During the infection of the plant cell by *A. tumefaciens*, the T-DNA is directed and integrated inside of the plant chromosome (Step 1). Genes on the T-DNA encode for production of phytohormones for tumor generation (Step 2) and for production of opines for *A. tumefaciens* nutrition (Step 3). The opine production is directed by Octopine synthase (Ocs), this enzyme takes SMM, an abundant plant molecule, together with pyruvate to synthesize a new sulfur-containing nutrient, sulfonopine (Step 3). Sulfonopine is released by the infected plant tissue (Step 4) and taken up by the colonizing bacteria (Step 5). The internalized sulfonopine activates the transcription of the *occ* operon by binding to OccR (Steps 6 and 7). The expressed genes are involved in the internalization and catabolism of the opines, as well as the expression of *tra*R involved in the positively regulation of quorum sensing, conuugation and replication of the Ti plasmid (Step 8). Sulfonopine is catabolized, generating SMM and pyruvate (Step 9), and then SMM is available to be used by Msh to produce methionine (Step 10). However, when the concentration of methionine increases, it downregulates the expression of its biosynthetic gene, msh, and downstream genes including *tra*R (Step 11, red arrow), hence quorum sensing, conjugation and replication of the Ti plasmid are also repressed (Step 12).

5.2 Future work directions

In Chapter 2, we demonstrated that Msh utilizes SMM and homocysteine to produce two molecules of methionine. Msh preferentially uses SMM, however we don't know the mechanism for specificity. A structural analysis of Msh will be useful for characterizing the amino acids that are involved in the catalytic pocket. This can be done by site directed mutagenesis and crystallization of Msh. The results from these experiments will be very informative since the SMM:homocysteine methyl transferase family is not well understood and its strong selection for SMM is unclear.

In Chapter 3, we unveil the role of SMM in the nutrition and gene regulation in *Agrobacterium tumefaciens*. SMM is taken up in its opine form, sulfonopine, by the colonizing bacteria. We know that sulfonopine interacts with OccR to activate the catabolic genes, but in this chapter we did not test the utilization of sulfonopine by the catabolic complex, OoxBCAD. This question is under consideration in the Appendix 3. In addition, we showed that Ocs recognizes a broad range of substrates, and it would be interesting to characterize the residues that formed the catalytic site. The ability of Ocs to utilize different amino acids provides a great advantage to *A. tumefaciens* survival regardless of the nutritional status of the plant.

In Chapter 4, we showed that methionine represses *msh* and the downstream genes. We have evidence that methionine interacts directly with the mRNA, changing the structural conformation, suggesting a regulation by a riboswitch mechanism. However, more experiments should be done to map the specific binding methionine-nucleotides and determine whether the mechanism of regulation, is transcriptional or translational. Even

though methionine interacts directly with the mRNA, regulation by a repressor protein cannot be ruled out, and more experiments will be needed to settle this question. There are numerous examples where expression of metabolic genes is tightly regulated through multiple mechanisms at several levels (transcriptional, post-transcriptional or post-translational) by their biosynthetic products or intermediates.

[1]Appendix 1

Saturation Mutagenesis of a CepR Binding Site as a Means to Identify New Quorum-regulated Promoters in *Burkholderia cenocepacia*

A1.1 Abstract

These experiments were collaborative and carried out primarily by Yuping Wei, a former graduate student in the Winans lab. *Burkholderia cenocepacia* is an opportunistic pathogen of humans that encodes two genes that resemble the acylhomoserine lactone (AHL) synthase gene *luxI* of *Vibrio fischeri* and three genes that resemble the AHL receptor gene *luxR*. Of these, CepI synthesizes octanoylhomoserine lactone (OHL), while CepR is an OHL-dependent transcription factor. Previously, members of the Winans lab identified two direct target promoters of CepR. In the current study we developed a strategy to identify additional genes that are directly regulated by this protein. Yuping Wei systematically altered a CepR binding site (*cep* box) upstream of a target promoter to identify nucleotides that are essential for CepR activity *in vivo* and for CepR binding *in vitro*. She constructed 34 self-complementary oligonucleotides containing altered *cep* boxes, and measured binding affinity for each. These experiments allowed her to identify a consensus CepR binding site. Several hundred similar sequences were identified, some of which were adjacent to probable promoters. I fused 13 such

1 Wei, Y., Ryan, G.T., Flores-Mireles, A.L., Costa, E.D. Scheneider, D.J., and Winans, S.C. (2011) Saturation mutagenesis of a CepR binding site as a mean to identify new quorum-regulated promoters in *Burkholderia cenocepacia*. Mol Microbiol **79**(3):616-632.

promoters to a reporter gene with and without intact *cep* boxes. This allowed me to identify four new regulated promoters that were induced by OHL, and that required CepR and a *cep* box for induction. In collaboration with Ms. Gina Ryan, a current Cornell student, I reconstituted in *E. coli* CepR-dependent, OHL-dependent expression of all four promoters. Ms. Ryan confirmed CepR binding to each of these sites using electrophoretic mobility shift assays. In this thesis, I briefly summarize the work done by others, and describe my contributions in more detail.

A1.2 Introduction

The genus *Burkholderia* encompasses a fascinating collection of diverse β-proteobacteria (Coenye & Vandamme 2003). This genus includes over 50 species, some of which are potentially useful in bioremediation, while other members are capable of forming nitrogen-fixing root nodules with legumes (Bontemps et al 2010, Chen et al 2003). Some members protect host plants against fungal pathogens, while others are themselves pathogenic against plants, animals, and humans (Coenye & Vandamme 2003, Jones & Webb 2003). Seventeen pathogenic species are members of the *Burkholderia cepacia* complex, or BCC (Vandamme et al 1997, Vanlaere et al 2009, Vanlaere et al 2008), two of which are described by the Center for Disease Control as category B select agents (Godoy et al 2003).

B. cenocepacia, previously known as *B. cepacia* genomovar III (Vandamme et al 2003), is recognized as an opportunistic pathogen of humans and is a particular threat to cystic fibrosis (CF) patients (Mahenthiralingam et al 2005, Vandamme et al 1997). Colonization of the CF lung by *B. cenocepacia* (Vandamme et al 2003) tends to occur in

patients already infected with *Pseudomonas aeruginosa*, another opportunistic pathogen of the CF lung (Jones & Webb 2003, Vandamme et al 1997). An infection caused by both organisms can result in serious clinical complications. *B. cenocepacia* strains are resistant to most antibiotics, making them virtually impossible to eradicate (Nzula et al 2002). Infections with *B. cenocepacia* may have variable clinical outcomes ranging from asymptomatic carriage to a sudden fatal deterioration in lung function (Mahenthiralingam et al 2005).

Four strains of *B. cenocepacia* have been sequenced in their entirety, one of which is described in a publication (Holden et al 2009). The Joint Genome Institute is currently sequencing nine additional strains (JGI 2010). All four sequenced isolates have three circular chromosomes that vary in size between 3.9 and 0.88 MB in length. Strains J2315 and HI2424 also have one plasmid, 93 KB and 165 KB in length, respectively.

Many or possibly all *Burkholderia* spp. encode at least one regulatory system that resembles the LuxR and LuxI proteins of *Vibrio fischeri*, where LuxI synthesizes an acylhomoserine lactone (AHL)-type pheromone, also called an autoinducer, and LuxR is an AHL-dependent transcriptional regulator (Choi & Greenberg 1992, Eberhard et al 1981, Engebrecht & Silverman 1984). Regulatory systems of this family are found in countless proteobacteria, where they are thought to allow individual bacteria to coordinate their physiology with their siblings. Collectively, these systems regulate diverse processes, including pathogenesis, biofilm formation, bacterial conjugation, and the production of antibiotics and other secondary metabolites (Whitehead et al 2001). In general, target genes are transcribed preferentially at population densities high enough to

favor AHL accumulation (Eberhard et al 1991), a phenomenon referred to as quorum sensing (Fuqua et al 1994). *B. thailandiensis* has three such systems, one of which is implicated in cell aggregation, while another is required for antibiotic production (Chandler et al 2009, Duerkop et al 2009). A plant growth promoting isolate of *B. ambifaria* uses quorum sensing to regulate the production of the anti-fungal compound pyrrolnitrin (Schmidt et al 2009).

LuxR-type proteins have two domains, an N-terminal pheromone binding domain and a C-terminal DNA binding domain (Pappas et al 2004). Purified LuxR, TraR of *Agrobacterium tumefaciens*, and LasR of *Pseudomonas aeruginosa*, when complexed with their respective AHLs, bind with high specificity to recognition sequences (referred to as *lux, tra,* or *las* boxes, respectively) that are found at target promoters (Schuster et al 2004, Urbanowski et al 2004, Zhu & Winans 1999). LasR is also able to bind to sequences that have no obvious resemblance to canonical *las* boxes. A few members of this family bind DNA only in the *absence* of AHLs (Castang et al 2006, Cui et al 2005, Fineran et al 2005, Minogue et al 2005, Sjoblom et al 2006).

B. cenocepacia J2315 encodes three LuxR homologs and two LuxI homologs (Lewenza et al 1999, Malott et al 2005, Malott et al 2009). Among these, CepR and CepI appear to be well conserved within the BCC (Venturi et al 2004). CepI synthesizes primarily octanoylhomoserine lactone (OHL), and lower levels of hexanoylhomoserine lactone (HHL) (Aguilar et al 2003b, Gotschlich et al 2001, Huber et al 2001, Lewenza et al 1999). Null mutations in *cepI* or *cepR* increased the production of the siderophore ornibactin, and decreased the production of secreted lipases and metalloproteases ZmpA

and ZmpB (Kooi et al 2006, Lewenza et al 1999, Lewenza & Sokol 2001, Sokol et al 2003). CepI and CepR are also required for swarming motility and biofilm formation (Huber et al 2001) and for pathogenicity in several animal models (Kothe et al 2003, Sokol et al 2003). *B. cenocepacia* also expresses the CciI and CciR proteins, which are encoded on a genomic island called *cci* (cenocepacia island), that is associated with epidemic strains (Malott et al 2005). The CepIR and CciIR systems extensively interact, in that CciR negatively regulates *cepI*, while CepR is required for expression of the *cciIR* operon (Malott et al 2005). Transcriptional profiling studies indicate that CepR and CciR regulate many of the same genes, but do so in opposite ways (O'Grady et al 2009). *B. cenocepacia* also encodes an orphan LuxR homolog called CepR2, which represses a cluster of genes that may direct production of an antibiotic or other secondary metabolite (Malott et al 2009).

In addition to transcriptional profiling several other approaches have been used to identify genes whose expression is influenced by CepR and/or OHL. In one study, the proteome of a wild type *B. cenocepacia* was compared to that of a *cepR* mutant. Fifty-five proteins were found to be differentially expressed in the two strains, approximately 10% of all detected proteins (Riedel et al 2003). In a second study, fragments of a *B. cepacia* strain were cloned into a promoter trap plasmid and introduced into an *E. coli* strain that expressed CepR (Aguilar et al 2003a). Twenty-eight promoter fragments were identified as being induced by OHL, and in all cases, induction required CepR. In a third study, a library of *B. cenocepacia* DNA fragments were introduced into a plasmid containing a promoterless *luxCDABE* operon (Subsin et al 2007). That study identified 58 OHL-inducible promoters and 31 OHL-repressible promoters. Regulation of

nine of these genes required CepR, while the others were not tested. Seven OHL-inducible genes were identified by screening a library of *lacZ* fusions (Weingart et al 2005). Induction of all of these genes required CepR. Purified CepR-OHL complexes bound with high affinity and specificity to specific DNA sequences at two target promoters (Weingart et al 2005). These binding sites contained a 16-nucleotide imperfect dyad symmetry and were centered approximately 44 nucleotides upstream of the transcription start sites. These two sites are to date the only experimentally confirmed CepR binding sites.

Most of the studies described above do not distinguish whether a target promoter is controlled by CepR directly or indirectly. CepR could regulate a promoter indirectly, for example, by directly regulating an unknown regulatory gene whose product directly regulates that promoter. Alternatively, a CepR mutation might perturb cellular physiology in such a way that various promoters are affected by secondary effects.

To date, the most comprehensive study attempting to define the optimal CepR binding site was done by Chambers, Sokol, and colleagues (Chambers et al 2006), who approached this question with an impressive combination of genetics and bioinformatics. Mutagenesis of the known CepR binding site within the *cepI* promoter completely abolished induction (Chambers et al 2006). The promoters of six genes known to be induced by OHL were used to formulate a consensus CepR binding motif (Chambers et al 2006). This information was used to test eight additional candidate promoters, six of which were CepR-regulated. Ultimately, ten inducible promoters were used to refine the consensus sequence, and 57 possible CepR binding sites were identified upstream of

various genes.

The consensus motif identified in the Chambers study included the sequence CTG-N_{10}-CAG, which has dyad symmetry. However, several other bases in the consensus did not preserve this symmetry, and some of those non-symmetric bases were said to be highly conserved (Chambers et al 2006). The partial dyad symmetry suggests that CepR binds DNA as dimer and that the two DNA binding domains have rotational symmetry. Although we have no proof of this, structural studies of a related protein support this idea (Vannini et al 2002, Zhang et al 2002). Several other LuxR-type proteins are thought to decode dyad symmetrical sequences (Antunes et al 2008, White & Winans 2007, Whitehead et al 2001). In the present study, we tested the ten putative CepR binding sites described above for the ability to bind purified CepR-OHL complexes. We also systematically resected and mutated a known CepR binding site, and use the resulting information to identify four new promoters that are regulated directly by CepR. All four promoters are regulated by CepR *in vivo*, require their binding sites for regulation, and bind with high affinity to CepR-OHL *in vitro.*

A1.3 Results

Note that only the results corresponding to the experiments that I did for this study are shown here. For the other figures, please refer to (Wei et al 2011).

Identification of new CepR regulated promoters.

In order to identify the nucleotides important for binding, Yuping Wei did a mutation analysis and found that 4 nucleotides near the center of the *cep* box were

important. The dissociation constants were processed by the EnoLOGOS web server that also converted these constants into a log-likelihood matrix. This matrix was used to obtain a similarity score between the *cep* box logo and every 22-nucleotide sequence in the *B. cenocepacia* genome. This was done by Dr. David Scheider (USDA) using the MOODS algorithm (Korhonen et al 2009), which provided a list of 237 possible *cep* boxes and scored their similarity to the consensus sequence. Of these 237 sites, 142 lie within 300 nucleotides upstream of a predicted translation start site. The expression of most of these genes is not affected by a CepR mutation (O'Grady et al 2009). However, 43 genes that have possible nearby *cep* boxes are differentially expressed at least 2-fold by a mutation in *cepR*. Of these 43 genes, 13 were chosen for further analysis (Table A1.1). These 13 were chosen for a variety of reasons. Two were *cepI* and *aidA*, which served as positive controls. We also chose the *cepR* promoter, which was said to be autorepressed in one study (Lewenza & Sokol 2001) but not in another (O'Grady et al 2009). BCAM0188 encodes CepR2, which if induced would show a quorum cascade. Two other genes are adjacent to BCAM0188. BCAM1869 was chosen because it lies adjacent to and divergent from *cepR*. BCAM1413a and BCAM1414 were chosen because they encode three AidA homologs. The promoter of pBCA055 was found to be induced in an earlier study from our lab (Weingart et al 2005).

I fused each of the 13 putative promoters to *lacZ* on a multicopy plasmid. Each fragment extended upstream just far enough to include the putative CepR binding site plus 6-8 extra nucleotides. I also constructed 13 similar fragments that included only the downstream half of the predicted binding sites. All 26 plasmids were introduced into the *cepI* mutant strain K56-2I and tested for induction of β-galactosidase in the absence of

OHL or in the presence of two OHL concentrations, one that causes approximately half-maximal induction of *cepI* and *aidA* (Weingart et al 2005), and one that causes maximal expression. Of the 13 plasmids containing full CepR binding sites, twelve were predicted to be induced by OHL, as one (*cepR*) was predicted to be repressed. Of these twelve, I found that seven were induced at least 4-fold by 1 μM OHL (Table A1.1). Of these seven, induction of six was abolished or severely reduced in isogenic plasmids containing only half of the putative *cep* box (Table A1.1). The exception was the BCAM0186 promoter, which was strongly induced even with only half of its *cep* box. This gene lies close to BCAM0188 (*cepR2*), and may be regulated by the CepR2 protein (Malott et al 2009).

Of the six promoters whose OHL induction required a *cep* box, two were previously characterized (*cepI* and *aidA*), while four were new (BCAL0510, BCAM1869, BCAS0156, and pBCA055-4 (*bqiCD*). Three of these four genes were previously shown to be induced by CepR, and to be unaffected by mutations in CciR or CepR2 (Malott et al 2009, O'Grady et al 2009), although those studies did not show whether CepR acted directly or indirectly. The fourth is BCAS0156, which was not previously reported to be induced. The four newly identified CepR-regulated promoters are divided among the four replicons. BCAL0510 lies on the largest chromosome and its product resembles a group of hypothetical proteins (data not shown). BCAM1869 lies on chromosome 2 and is adjacent to and divergent from *cepR*. The *cep* box lies 114 nucleotide upstream of the BCAM1869 translation start site. Induction was reduced, though not abolished, by removing the promoter-distal half of the *cep* box. BCAM1869 and *cepR* are adjacent in many species of *Burkholderia*, providing suggestive evidence that their proteins are in

some way functionally linked. A protein homologous to BCAM1869 was recently shown to play a role in transcription regulation (Mattiuzzo et al 2010). The *cepR* gene was previously identified as being autorepressed in one study (Lewenza & Sokol 2001), although in another study it was found to be unregulated by CepR (O'Grady et al 2009). Our data support the latter study (Table A1.1).

BCAS0156, which lies on chromosome 3, resembles a family of β-lactamases and other penicillin-binding proteins (pfam00144). pBCA055 (*bqiC*) and pBCA054 (*bqiD*) lie on the 93 KB plasmid and appear to be expressed as an operon. pBCA055 is a multidomain protein whose central domain resembles GGDEF proteins (COG2199) and whose C-terminal domain resembles EAL proteins (pfam00563), and may therefore synthesize and or degrade c-di-GMP (Romling & Amikam 2006). pBCA054 has a C-terminal domain that resembles the DNA binding domains of LuxR proteins (pfam00196), suggesting that it may be a transcription factor.

The fact that these six promoters required OHL and a putative CepR binding site for induction suggested that CepR might directly activate them. In collaboration with Ms. Gina Ryan, I set out to determine whether CepR and OHL could activate these promoters in a heterologous host lacking any other LuxR-type protein. We introduced into *E. coli* strain MC4100 plasmids containing each of the six inducible promoters as well as plasmids containing the same promoters but containing only part of the CepR binding site. These strains also contained a second plasmid expressing CepR or a vector control. Strains containing CepR were cultured in the presence or absence of OHL and assayed for β-galactosidase. All six strains containing CepR and a full *cep* box showed

strong induction of β-galactosidase by OHL (Table A1.2). Strains whose plasmids lacked the full *cep* boxes were either uninduced or weakly induced by OHL, as expected. Strains lacking CepR were not induced by OHL, also as expected. We conclude that for each promoter, induction by OHL requires CepR and a CepR binding site. The reconstitution of CepR-dependent induction of these promoters in *E. coli* provides further evidence that each is directly regulated by CepR.

Table A1.1 Induction of *B. cenocepacia* genes by OHL.

Gene	Sequence	Score	*cep* box	No OHL	OHL (1 nM)	Ratio	OHL (1 μM)	Ratio
BCAL0510	CGCCCGCCAGAATTGACAGGCC	6.338	full	138 ±2	469 ± 117	3.5	747 ± 73	5.4
			half	351± 27	167 ± 40	0.5	303 ± 16	0.9
BCAM0186 (NRPS)	ACCCTGTGATTTGATGCCGGTC	9.398	full	43±7	97 ± 6	2.3	771 ± 67	17.9
			half	63± 6	69 ± 3	1.1	762 ± 34	12.1
BCAM0188 (*cep*R2)	ATCCTGTTCAAAAGGACAGTTT	-1.875	full	543±34	743 ± 141	1.4	1452 ± 105	2.6
			half	669±23	876 ± 27	1.3	1678 ± 223	2.5
BCAM0189 (NRPS)	ATCCTGTTCAAAAGGACAGTTT	-1.875	full	180 ± 17	150 ± 17	0.8	360 ± 16	2
			half	170 ± 5	123 ± 9	0.7	329 ± 40	1.9
BCAM1413a (*aid*C)	TACCTGTCAGGTTTGATGGGGG	6.26	full	134 ± 6	145 ± 6	1.1	408 ± 26	3.1
			half	173 ±16	201± 8	1.1	217 ± 29	1.2
BCAM1414	TACCTGTCAGGTTTGATGGGGG	6.26	full	80 ± 23	83 ± 8	1.03	175 ±52	2.1
			half	101 ± 8	93 ± 9	0.9	114 ± 10	1.1
BCAM1868 (*cep*R)	ACGCTGTCATACTTGTCAGGTT	-8.188	full	205 ± 6	193 ± 4	0.94	224 ± 16	1.1
			half	ND*	ND	ND	ND	ND
BCAM1869	ACGCTGTCATACTTGTCAGGTT	-8.188	full	179 ± 13	1598 ± 193	8.9	3536 ±125	19.7
			half	121 ± 12	153 ± 3	1.3	609 ± 45	5
BCAM1870 (*cep*I)	ACCCTGTAAGAGTTACCAGTTA	-5.885	full	60 ± 3	655 ± 25	10.9	2489 ±162	41.5
			half	40 ±16	39 ± 3	0.9	133 ± 42	3.3
BCAS0155-0153	ATACTGTTAAAACCGGCAGGTT	-9.521	full	67 ± 4	72 ± 6	1.1	176 ± 44	2.6
			half	87 ± 7	61 ± 4	0.7	114 ± 10	1.3
BCAS0156	ATACTGTTAAAACCGGCAGGTT	-9.521	full	110 ± 22	518 ± 64	4.7	1399 ± 144	12.7
			half	101 ± 7	84 ± 6	0.8	116 ± 14	1.1
BCAS0293 (*aid*A)	AAGCTGTAAAAGTAAACAGGTC	-1.315	full	109 ± 9	416 ± 18	3.8	1648 ± 61	15.1
			half	118 ± 8	108 ± 7	0.9	174 ± 32	1.6
pBCA055-054 (*bqi*CD)	CCACTGTCAAATCTACGAGGGC	2.799	full	149 ± 49	775±193	5.2	3126 ± 1052	20.9
			half	167 ± 6	167 ± 1	1	272 ± 106	1.6

* ND- not determined

Table A1.2 Activation of CepR-regulated promoters in the heterologous host *E. coli*.

		+ CepR				
Gene	***cep* box**	**No OHL**	**(1 µM)**	**Ratio**	**- CepR**	**Ratio**
BCAM1870 (*cep*I)	Full	4 ± 0.2	491 ± 6.2	128	4 ± 0.7	120
	Half	4 ± 0.2	12 ± 2.7	3	4± 1.1	3
BCAS0293 (*aid*A)	Full	4 ± 0.2	176 ± 17.7	44	4 ± 0.7	43
	Half	4 ± 0.1	8 ± 1.0	2	3 ± 0.3	3
BCAL0510	Full	4 ± 0.1	49 ± 3.3	13	4 ± 0.0	14
	Half	4 ± 0.1	11 ± 0.4	3	5 ± 0.9	2
BCAM1869	Full	3 ± 0.1	211 ± 6.9	65	4 ± 0.3	52
	Half	4 ± 0.2	29 ± 0.7	7	4 ± 0.1	8
pBCA055-054 (bqiCD)	Full	4 ± 0.1	159 ± 13.4	44	3 ± 0.2	48
	Half	4 ± 0.1	4 ± 0.5	1	3 ± 0.0	1
BCAS0156	Full	4 ± 0.2	41 ± 1.3	10	3 ± 0.4	12
	Half	4 ± 0.1	3 ± 0.1	1	3 ± 0.2	1
(No insert)		4 ± 0.3	3 ± 0.1	1	4 ± 0.2	1

A1.4 Discussion

Members of our laboratory have previously used genetic, biochemical, and structural approaches to study interactions between another LuxR-type protein, TraR, and its DNA binding site. TraR binds DNA as a dimer, and the two DNA binding domains of each dimer have 2-fold rotational symmetry. The DNA binding site also has 2-fold rotational symmetry, and complexes between the TraR-CTD and DNA also have rotational symmetry. This type of symmetry is found in the binding sites of other LuxR-type proteins, and of a great number of other DNA binding proteins. The imperfect dyad symmetry of two CepR binding sites led us to believe that CepR would follow a similar pattern. It was therefore interesting that the consensus sequence previously described had several highly conserved asymmetric bases. Unfortunately, only three of the ten putative binding sites was detectably bound by purified CepR-OHL. Our studies suggest that a core CTG-N10-CAG is critical for binding. Loss of one of these six bases may be tolerated, but loss of any two or more probably is not. Of the ten DNA sequences compiled in the Chambers study, three had all six bases of this consensus and were bound by CepR-OHL, while seven sequences lacked between one and four of these bases, and were not bound. The consensus sequence identified previously (Chambers et al 2006) also has all six of these bases, and was bound. In the Chambers study, there seemed to be an implicit assumption that all CepR-inducible genes would be induced *directly* by CepR. In light of the present data, it seems more likely that induction of some genes could occur indirectly.

In order to study specifically how CepR decodes its binding site, we felt it was

important to move to an *in vitro* system using purified components. The finding that 26 nucleotides are needed for full binding affinity was somewhat surprising, as this sequence would extend 2.5 helical turns, or 1.25 turns per half-site. It is far from clear why such a long sequence would be required. CepR does not strongly discriminate specific sequences at positions -11, -10, +10, or +11, suggesting that the need for bases far removed from the dyad center may not be sequence-specific.

We also tested a set of 22-nucleotide, perfectly symmetric sequences based upon the left half-site of the *cepI cep* box. Mutation of any of the bases from -9 to -5 (and the corresponding bases at +5 to +9) either abolished or strongly impaired binding affinity. From these data, we conclude that the core CepR binding site could be amended from the 16-nucleotide sequence CTG-N_{10}-CAG to the 18 nucleotide sequence CCTGT-N_8-ACAGG. We had previously discounted the bases at positions -9, -5, +5, and +9 as they are not part of the dyad symmetry of the two *cep* boxes previously aligned (Weingart et al 2005). However, these bases are somewhat conserved in a new set of CepR binding sites.

The effects of mutations within the central spacer (-4 to +4) had more variable effects on binding affinity. Mutations from A to C or T at position -4 (and from T to G or A at position +4) enhanced affinity, indicating that the original L-L' sequence was not optimal at this position. Many of the newly identified *cep* boxes contain the bases C or T at position -4 and a G or A at position +4 . We denote the *cep* box residues from -4 to +4 as the central spacer, and predict that there are no sequence-specific protein-DNA contacts in this region, as was shown for TraR (Vannini et al 2002, Zhang et al 2002). It

is well established that non-contacted bases can have large effects on protein affinity, generally via effects on the helical pitch or on the flexibility of the DNA, or by imparting a sequence-directed DNA bend (Sarai & Kono 2005). This phenomenon is sometimes referred to as "indirect readout", while sequence decoding by direct protein-BDNA interactions is called "direct readout".

The identification of the bases essential for CepR binding facilitated a search for new genes that could be regulated directly by CepR. Of the thirteen promoters that were tested, six were significantly induced by OHL and required the putative *cep* box for this induction. These six CepR-regulated genes are distributed across all three chromosomes and the 92 KB plasmid.

BCAS0293 (*aidA*) was reported previously to be OHL-regulated (Aguilar et al 2003b, Chambers et al 2006, Riedel et al 2003, Weingart et al 2005). This gene is in a two-gene operon, and the downstream gene, *aidB,* is homologous to *aidA.* This homology extends to three additional genes that we designate *aidC, aidD,* and *aidE* (BCAM1413a, BCAM1412, and BCAM1414, respectively), none of which was induced more than 2-3 fold by OHL (Table A1.1). BCAM1413a and BCAM1414 are divergent and flank a *cep* box, though the score of this site is weak (Table A1.1). The roles of these proteins is unknown, though AidA was previously reported to play a role in the slow killing of the nematode *C. elegans* (Huber et al 2004). All five proteins are members of the PixA protein family (pfam12306).

In this study, we also examined the regulation of gene BCAM0186. This gene was unusual in that it was strongly induced by OHL and had a possible *cep* box, yet

induction was *cep* box-independent. In hindsight, this should not have been surprising, as the similarity between this putative *cep* box and the consensus is rather weak, and the *cep* box lies almost 600 nucleotides upstream of the BCAM0186 translation start site. Its expression was reported to be inhibited by the product of the nearby BCAM0188 (*cepR2*) (Malott et al 2009). We hypothesize that BCAM0186 may be directly repressed by CepR2, and that repression might be blocked by OHL.

A1.5 Experimental Procedures

Note that only the procedures for the experiments that I did are described in this appendix. For complete experimental procedures for this study please refer to (Wei et al 2011)

Bacterial strains and growth conditions

Strains and plasmids used in this study are described in Table A1.3 and A1.4. As needed, *B. cenocepacia* was cultured in 100 μg/ml of trimethoprim or 300 μg/ml of tetracycline. *E. coli* strains were cultured with 15 μg/ml of tetracycline, 100 μg/ml of streptomycin, or 100 μg/ml of ampicillin.

Identification of six new CepR regulated promoters.

Promoters containing suspected CepR binding sites were PCR amplified using oligonucleotides (Table A1.5). For each promoter, a fragment containing a complete CepR binding site was amplified, as well as a similar fragment containing only the promoter-proximal half of the site. These PCR fragments contained a *Kpn*I site at the

promoter-distal end, and a *Pst*I site at the promoter-proximal end. The PCR fragments were purified by using QIAquick Gel Extraction Kits (Qiagen) and digested with *Kpn*I and *Pst*I (New England Biolabs). The digested fragments were cloned into the promoter probe plasmid pYWN302, generating transcriptional fusions between each promoter and *lacZ*. Plasmid pAFM113 was constructed by cloning the *cep*R gene into the pSRKKm broad-host range vector that had been digested with NdeI and HindIII. The resulting plasmids were introduced into *B. cenocepacia* strain K56-I2 or *E. coli* strain MC4100 by electroporation. For assays of β-galactosidase, *B. cenocepacia* and *E. coli* strains were cultured in LB medium supplemented with either 300 μg/ml of tetracycline or 100ug/ml kanamycin and 12ug/ml tetracycline, respectively, overnight at 37^{O} C. Each culture was diluted 1:100 into LB medium containing the indicated concentrations of OHL, and incubated with aeration at 37^{O} C to an OD_{600} of 0.4, and assayed for β-galactosidase specific activity (Miller 1972). Experiments were performed in triplicate with three different isolates of each strain.

Table A1.3 Strains used in this study

Strains	**Relevant features**	**Reference**
K56-2	*B. cenocepacia* wild type	Lewenza *et al.* 1999
K56-I2	*B. cenocepacia* K56 *cep*I::Tp^{R}	Lewenza *et al.* 1999
DH5α	*E. coli* α-complementation	Stratagene
MC4100	F- *araD*139 Δ*(argF-lac)*U169 *rpsL*150 *relA*1 *deoC1 rbsR fthD*5301 *fruA*25 λ-	Ferenci *et al.*, 2009

Table A1.4 Plasmids used in this study

Plasmids	Relevant features	References
pSRKKm	Broad host range expression vector, Km^R	Khan *et al*, 2008
pYWN302-1	Box less *lac*Z vector, Tet^R	This study
pAFM113	pSRKKm derivative containing *cep*R	This study
pAFM263	BCAL510 full cep box cloned into pYWN302-1	This study
pAFM264	BCAL510 half cep box cloned into pYWN302-1	This study
pAFM290	BCAM0186 (NPR) full cep box cloned into pYWN302-1	This study
pAFM291	BCAM0186 (NPR) half cep box cloned into pYWN302-1	This study
pAFM284	BCAM0188 (*cep*S)full cep box cloned into pYWN302-1	This study
pAFM283	BCAM0188 (*cep*S) half cep box cloned into pYWN302-1	This study
pAFM294	BCAM0189 (NPR) full cep box cloned into pYWN302-1	This study
pAFM295	BCAM0189 (NPR) half cep box cloned into pYWN302-1	This study
pAFM269	BCAM1868 (*cep*R) full cep box cloned into pYWN302-1	This study
pAFM271	BCAM1869 (*cep*J) full cep box cloned into pYWN302-1	This study
pAFM272	BCAM1869 (*cep*J) half cep box cloned into pYWN302-1	This study
pAFM323	BCAM1870 (*cep*I) full cep box cloned into pYWN302-1	This study
pAFM318	BCAM1870 (*cep*I) half cep box cloned into pYWN302-1	This study
pAFM266	BCAM1413a (*aid*C) full cep box cloned into pYWN302-1	This study
pAFM267	BCAM1413a (*aid*C) half cep box cloned into pYWN302-1	This study
pAFM306	BCAM1414 full cep box cloned into pYWN302-1	This study
pAFM307	BCAM1414 half cep box cloned into pYWN302-1	This study
pAFM310	BCAM1415 full cep box cloned into pYWN302-1	This study
pAFM311	BCAM1415 half cep box cloned into pYWN302-1	This study
pAFM298	BCAM0155-0153 full cep box cloned into pYWN302-1	This study
pAFM299	BCAM0155-0153 half cep box cloned into pYWN302-1	This study
pAFM314	BCAM0156 full cep box cloned into pYWN302-1	This study
pAFM315	BCAM0156 half cep box cloned into pYWN302-1	This study
pAFM320	BCAM0293 (*aid*A) full cep box cloned into pYWN302-1	This study
pAFM321	BCAM0293 (*aid*A) half cep box cloned into pYWN302-1	This study
pAFM302	pBCA053-051 full cep box cloned into pYWN302-1	This study
pAFM303	pBCA053-051 half cep box cloned into pYWN302-1	This study
pAFM288	pBCA055-054 (*bqi*CD) full cep box cloned into pYWN302-1	This study
pAFM289	pBCA055-054 (*bqi*CD)half cep box cloned into pYWN302-1	This study

Table A1.5 Oligonucleotides used in this study

Name	DNA sequence	Comments
YWP263	5'-ggGGTACCctgacgcggcctgtcaattctggcgg-3'	BCAL0510 full cep box fusion ,pairs with YWP265
YWP264	5'-ggGGTACCtctggcgggcgcggcgcgca-3'	BCAL0510 half cep box fusion, pairs with YWP265
YWP265	5'-aaCTGCAGagcgagcgggcaggagcgga-3'	BCAL0510 cep box fusion, pairs with YWP263 or 264
YWP266	5'-ggGGTACCaataaatcccccatcaaacctgacag-3'	BCAM1413a (aidC) full cep box fusion, pairs with YWP268
YWP267	5'-ggGGTACCcctgacaggtactaaaacca-3'	BCAM1413a (aidC) half cep box fusion, pairs with YWP268
YWP268	5'-aaCTGCAGggactgagcggctttttcca-3'	BCAM1413a (aidC) cep box fusion, pairs with YWP266 or 267
YWP269	5'-ggGGTACCgagtgcggccagtgcgcctt-3'	BCAM1868 (cepR) full cep box fusion, pairs with YWP270
YWP270	5'-aaCTGCAGcggcgctgaattgttggtag-'3	BCAM1868 (cepR) cep box fusion, pairs with YWP269
YWP271	5'-ggGGTACCcaacgtcacgctgtcatacttgtcag-3'	BCAM1869 (cepJ) full cep box fusion, pairs with YWP273
YWP272	5'-ggGGTACCcttgtcaggtttcagtaccc-3'	BCAM1869 (cepJ) half cep box fusion, pairs with YWP273
YWP273	5'-aaCTGCAGgagtgcggccagtgcgcctt-3'	BCAM1869 (cepJ) box fusion, pairs with YWP271 or 272
ALFM283-Y	5'-ggGGTACCtttgaacaggatacgcgcgggccgca-3'	BCAM0188 (cepS) half cep box fusion, pairs with YWP285
ALFM284-Y	5'-ggGGTACCttcgcctcaaaactgtccttttgaac-3'	BCAM0188 (cepS) full cep box fusion, pairs with YWP284
ALFM286-Y	5'-aaCTGCAGttgatgatcggatcgatggtgacg-3'	BCAM0188 (cepS) box fusion, pairs with ALFM283-Y or ALFM284-Y
ALFM287-Y	5'-ggGGTACCaggccctcgtagatttgacagtgg-3'	pBCA055-054 (bqiCD) full cep box , pairs with ALFM82-Y
ALFM288-Y	5'-ggGGTACCatttgacagtggcggcagattttc-3'	pBCA055-054 (bqiCD) half cep box fusion, pairs with ALFM82-Y
ALFM282-Y	5'-aaCTGCAGcgcgatcgttgcacgttcgatcccct-3'	pBCA055-054 (bqiCD) box fusion, pairs with ALFM87-Y or ALFM88-Y
ALFM290-Y	5'-ggGGTACCtgaccggcatcaaatcacagggt-3'	BCAM0186 (NPRS) full cep box fusion, pairs with ALFM293-Y
ALFM291-Y	5'-ggGGTACCaatcacagggtgaccacggctgtc-3'	BCAM0186 (NPRS) half cep box fusion, pairs with ALFM293-Y
ALFM293-Y	5'-aaCTGCAGcaatttgatgtgctgttgctcggc-3'	BCAM0186 (NPRS) box fusion, pairs with ALFM290-Y or ALFM291-Y
ALFM294-Y	5'-gcTCTAGAatcctgttcaaaaggacagttttg-3'	BCAM0189 (NPRS) full cep box fusion, pairs with ALFM297-Y
ALFM295-Y	5'-gcTCTAGAaggacagttttgaggcgaagccga-3'	BCAM0189 (NPRS) half cep box fusion, pairs with ALFM297-Y
ALFM297-Y	5'-aaCTGCAGctggccttcgacgtcgagtactaa-3'	BCAM0189 (NPRS) box fusion, pairs with ALFM294-Y or ALFM295-Y

Table A1.5 (Continuation)

Name	DNA sequence	Comments
ALFM298-Y	5'-ggGGTACCaaaacctgccggttttaacagtat-3'	BCAS0155 full cep box fusion, pairs with ALFM301-Y.
ALFM299-Y	5'-ggGGTACCttttaacagtatcgaatccggcgg-3'	BCAS0155 half cep box fusion, pairs with ALFM301-Y
ALFM301-Y	5'-aaCTGCAGccgacgcggccacggtttcggtca-3'	BCAS0155 box fusion, pairs with ALFM298-Y or ALFM299-Y
ALFM302-Y	5'-ggGGTACCtctcgcggtaatatcagacaagat-3'	pBCA053 full cep box fusion, pairs with ALFM305-Y
ALFM303-Y	5'-ggGGTACCatcagacaagattttcattcgcca-3'	pBCA053 half cep box fusion, pairs with ALFM305-Y
ALFM305-Y	5'-aaCTGCAGaactcgaagcgcgcggaaacgtgc-3'	pBCA053 box fusion, pairs with ALFM302-Y or 303-Y
ALFM306-Y	5'-ggGGTACCtacctgtcaggtttgatgggggat-3'	BCAM1414 box fusion, pairs with ALFM309-Y
ALFM307-Y	5'-ggGGTACCttgatgggggatttattttaatga-3'	BCAM1414 half box fusion, pairs with ALFM309-Y
ALFM309-Y	5'-aaCTGCAGgctcggattcggatacgttttctt-3'	BCAM1414 box fusion, pairs with ALFM306-Y or ALFM307-Y
ALFM310-Y	5'-ggGGTACCcgcctgtaggcgggtacagggcgc-3'	BCAM1415 box fusion, pairs with ALFM313-Y
ALFM311-Y	5'-ggGGTACCggtacagggcgcgatgtcccggcc-3'	BCAM1415 half box fusion, pairs with ALFM313-Y
ALFM313-Y	5'-aaCTGCAGcgcgttcgtcgcaccgtttacgta-3'	BCAM1415 box fusion, pairs with ALFM310-Y or ALFM311-Y
ALFM314-Y	5'-ggGGTACCatactgttaaaaccggcaggtttt-3'	BCAS0156 box fusion, pairs with ALFM317-Y
ALFM315-Y	5'-ggGGTACCccggcaggttttcccggaaaacgg-3'	BCAS0156 half box fusion, pairs with ALFM317-Y
ALFM317-Y	5'-aaCTGCAGgacgatgccgatcgccatgccggg-3'	BCAS0156 box fusion, pairs with ALFM314-Y or ALFM315-Y
ALFM323-Y	5'-ggGGTACCcaccctgtaagagttaccagttacag-3'	BCAM1870 (*cep*I) full cep box fusion, pairs with ALFM319-Y
ALFM318-Y	5'-ggGGTACCttaccagttacaggctcctcgtgc-3'	BCAM1870 (*cep*I) half cep box fusion, pairs with ALFM319-Y
ALFM319-Y	5'-aaCTGCAGtttcgcgcgaacacgtagacggta-3'	BCAM1870 (*cep*I) cep box fusion pairs with ALFM318-Y and ALFM323-Y
ALFM320-Y	5'-ggGGTACCccgacctgtttacttttacagctt-3'	BCAS0293 (*aid*A) full cep box fusion pairs with ALFM321-Y
ALFM321-Y	5'-ggGGTACCcttttacagcttccatgattaacc-3'	BCAS0293 (*aid*A) half cep box fusion pairs with ALFM321-Y
ALFM322-Y	5'-aaCTGCAGgtcgccgacctgcgccttcaggtc-3'	BCAS0293 (*aid*A) cep box fusion pairs with ALFM320-Y or ALFM321Y
ALFM328-Y	5'-ggaattcCATATGgaactgcgctggcagg-3'	Forward primer for *cep*R gene to be cloned in to pSRKG (NdeI)
ALFM329-Y	5'-cccAAGCTTaaaagcatttccagcgtgagc-3'	Reverse primer for *cep*R gene to be cloned in to pSRKG (HindIII)

A1.5 References

Aguilar C, Bertani I, Venturi V. 2003a. Quorum-sensing system and stationary-phase sigma factor (*rpoS*) of the onion pathogen *Burkholderia cepacia* genomovar I type strain, ATCC 25416. **Appl Environ Microbiol** 69: 1739-47

Aguilar C, Friscina A, Devescovi G, Kojic M, Venturi V. 2003b. Identification of quorum-sensing-regulated genes of *Burkholderia cepacia*. **Journal of Bacteriology** 185: 6456-62

Antunes LC, Ferreira RB, Lostroh CP, Greenberg EP. 2008. A mutational analysis defines *Vibrio fischeri* LuxR binding sites. **Journal of Bacteriology** 190: 4392-7

Bontemps C, Elliott GN, Simon MF, Dos Reis Junior FB, Gross E, et al. 2010. *Burkholderia* species are ancient symbionts of legumes. **Mol Ecol** 19: 44-52

Castang S, Reverchon S, Gouet P, Nasser W. 2006. Direct evidence for the modulation of the activity of the *Erwinia chrysanthemi* quorum-sensing regulator ExpR by acylhomoserine lactone pheromone. **J Biol Chem** 281: 29972-87

Chambers CE, Lutter EI, Visser MB, Law PP, Sokol PA. 2006. Identification of potential CepR regulated genes using a cep box motif-based search of the *Burkholderia cenocepacia* genome. **BMC Microbiol** 6: 104

Chandler JR, Duerkop BA, Hinz A, West TE, Herman JP, et al. 2009. Mutational analysis of *Burkholderia thailandensis* quorum sensing and self-aggregation. **Journal of Bacteriology**

Chen WM, Moulin L, Bontemps C, Vandamme P, Bena G, Boivin-Masson C. 2003. Legume symbiotic nitrogen fixation by beta-proteobacteria is widespread in nature. **Journal of Bacteriology** 185: 7266-72

Choi SH, Greenberg EP. 1992. Genetic dissection of DNA-binding and luminescence gene activation by the *Vibrio fischeri* LuxR protein. **Journal of Bacteriology** 174: 4064-69

Coenye T, Vandamme P. 2003. Diversity and significance of *Burkholderia* species occupying diverse ecological niches. **Environ Microbiol** 5: 719-29

Cui Y, Chatterjee A, Hasegawa H, Dixit V, Leigh N, Chatterjee AK. 2005. ExpR, a LuxR homolog of *Erwinia carotovora* subsp. *carotovora*, activates transcription of *rsmA*, which specifies a global regulatory RNA-binding protein. **Journal of Bacteriology** 187: 4792-803

Duerkop BA, Varga J, Chandler JR, Peterson SB, Herman JP, et al. 2009. Quorum-sensing control of antibiotic synthesis in *Burkholderia thailandensis*. **Journal of Bacteriology** 191: 3909-18

Eberhard A, Burlingame AL, Eberhard C, Kenyon GL, Nealson KH, Oppenheimer NJ. 1981. Structural identification of autoinducer of *Photobacterium fischeri* luciferase. **Biochemistry** 20: 2444-49

Eberhard A, Longin T, Widrig CA, Stranick SJ. 1991. Synthesis of the *lux* gene autoinducer in *Vibrio fischeri* Is positively autoregulated. **Arch Microbiol** 155: 294-97

Engebrecht J, Silverman M. 1984. Identification of genes and gene products necessary for bacterial bioluminescence. **Proc Natl Acad Sci U S A** 81: 4154-58

Fineran PC, Slater H, Everson L, Hughes K, Salmond GP. 2005. Biosynthesis of tripyrrole and beta-lactam secondary metabolites in *Serratia*: integration of quorum sensing with multiple new regulatory components in the control of prodigiosin and carbapenem antibiotic production. **Molecular Microbiology** 56: 1495-517

Fuqua WC, Winans SC, Greenberg EP. 1994. Quorum sensing in bacteria - the LuxR-LuxI family of cell density-responsive transcriptional regulators. **Journal of Bacteriology** 176: 269-75

Godoy D, Randle G, Simpson AJ, Aanensen DM, Pitt TL, et al. 2003. Multilocus sequence typing and evolutionary relationships among the causative agents of melioidosis and glanders, *Burkholderia pseudomallei* and *Burkholderia mallei*. **J Clin Microbiol** 41: 2068-79

Gotschlich A, Huber B, Geisenberger O, Togl A, Steidle A, et al. 2001. Synthesis of multiple N-acylhomoserine lactones is wide-spread among the members of the *Burkholderia cepacia* complex. **Syst Appl Microbiol** 24: 1-14

Holden MT, Seth-Smith HM, Crossman LC, Sebaihia M, Bentley SD, et al. 2009. The genome of *Burkholderia cenocepacia* J2315, an epidemic pathogen of cystic fibrosis patients. **Journal of Bacteriology** 191: 261-77

Huber B, Feldmann F, Kothe M, Vandamme P, Wopperer J, et al. 2004. Identification of a novel virulence factor in *Burkholderia cenocepacia* H111 required for efficient slow killing of *Caenorhabditis elegans*. **Infect Immun** 72: 7220-30

Huber B, Riedel K, Hentzer M, Heydorn A, Gotschlich A, et al. 2001. The *cep* quorum-sensing system of *Burkholderia cepacia* H111 controls biofilm formation and swarming motility. **Microbiology** 147: 2517-28

JGI. 2010. http://www.jgi.doe.gov/genome-projects/.

Jones AM, Webb AK. 2003. Recent advances in cross-infection in cystic fibrosis: *Burkholderia cepacia* complex, *Pseudomonas aeruginosa*, MRSA and *Pandoraea* spp. **J R Soc Med** 96: 66-72

Kooi C, Subsin B, Chen R, Pohorelic B, Sokol PA. 2006. *Burkholderia cenocepacia* ZmpB is a broad-specificity zinc metalloprotease involved in virulence. **Infect Immun**

74: 4083-93

Korhonen J, Martinmaki P, Pizzi C, Rastas P, Ukkonen E. 2009. MOODS: fast search for position weight matrix matches in DNA sequences. **Bioinformatics** 25: 3181-2

Kothe M, Antl M, Huber B, Stoecker K, Ebrecht D, et al. 2003. Killing of *Caenorhabditis elegans* by *Burkholderia cepacia* is controlled by the *cep* quorum-sensing system. **Cell Microbiol** 5: 343-51

Lewenza S, Conway B, Greenberg EP, Sokol PA. 1999. Quorum sensing in *Burkholderia cepacia*: Identification of the LuxRI homologs CepRI. **Journal of Bacteriology** 181: 748-56

Lewenza S, Sokol PA. 2001. Regulation of ornibactin biosynthesis and N-acyl-L-homoserine lactone production by CepR in *Burkholderia cepacia*. **Journal of Bacteriology** 183: 2212-18

Mahenthiralingam E, Urban TA, Goldberg JB. 2005. The multifarious, multireplicon *Burkholderia cepacia* complex. **Nat Rev Microbiol** 3: 144-56

Malott RJ, Baldwin A, Mahenthiralingam E, Sokol PA. 2005. Characterization of the *cciIR* quorum-sensing system in *Burkholderia cenocepacia*. **Infect Immun** 73: 4982-92

Malott RJ, O'Grady EP, Toller J, Inhulsen S, Eberl L, Sokol PA. 2009. A *Burkholderia cenocepacia* orphan LuxR homolog is involved in quorum-sensing regulation. **Journal of Bacteriology** 191: 2447-60

Mattiuzzo M, Bertani I, Ferluga S, Cabrio L, Bigirimana J, et al. 2010. The plant pathogen Pseudomonas fuscovaginae contains two conserved quorum sensing systems involved in virulence and negatively regulated by RsaL and the novel regulator RsaM. **Environ Microbiol**

Miller J. 1972. *Cold Spring Harbor Laboratory Press*. **Cold Spring Harbor**, NY.

Minogue TD, Carlier AL, Koutsoudis MD, von Bodman SB. 2005. The cell density-dependent expression of stewartan exopolysaccharide in *Pantoea stewartii* ssp. *stewartii* is a function of EsaR-mediated repression of the rcsA gene. **Molecular Microbiology** 56: 189-203

Nzula S, Vandamme P, Govan JR. 2002. Influence of taxonomic status on the in vitro antimicrobial susceptibility of the *Burkholderia cepacia* complex. **J Antimicrob Chemother** 50: 265-9

O'Grady EP, Viteri DF, Malott RJ, Sokol PA. 2009. Reciprocal regulation by the CepIR and CciIR quorum sensing systems in *Burkholderia cenocepacia*. **BMC Genomics** 10: 441

Pappas KM, Weingart CL, Winans SC. 2004. Chemical communication in

proteobacteria: biochemical and structural studies of signal synthases and receptors required for intercellular signalling. **Molecular Microbiology** 53: 755-69

Riedel K, Arevalo-Ferro C, Reil G, Gorg A, Lottspeich F, Eberl L. 2003. Analysis of the quorum-sensing regulon of the opportunistic pathogen *Burkholderia cepacia* H111 by proteomics. **Electrophoresis** 24: 740-50

Romling U, Amikam D. 2006. Cyclic di-GMP as a second messenger. **Curr Opin Microbiol** 9: 218-28

Sarai A, Kono H. 2005. Protein-DNA recognition patterns and predictions. **Annu Rev Biophys Biomol Struct** 34: 379-98

Schmidt S, Blom JF, Pernthaler J, Berg G, Baldwin A, et al. 2009. Production of the antifungal compound pyrrolnitrin is quorum sensing-regulated in members of the *Burkholderia cepacia* complex. **Environ Microbiol** 11: 1422-37

Schuster M, Urbanowski ML, Greenberg EP. 2004. Promoter specificity in *Pseudomonas aeruginosa* quorum sensing revealed by DNA binding of purified LasR. **Proc Natl Acad Sci U S A** 101: 15833-9

Sjoblom S, Brader G, Koch G, Palva ET. 2006. Cooperation of two distinct ExpR regulators controls quorum sensing specificity and virulence in the plant pathogen *Erwinia carotovora*. **Molecular Microbiology** 60: 1474-89

Sokol PA, Sajjan U, Visser MB, Gingues S, Forstner J, Kooi C. 2003. The CepIR quorum-sensing system contributes to the virulence of *Burkholderia cenocepacia* respiratory infections. **Microbiology** 149: 3649-58

Subsin B, Chambers CE, Visser MB, Sokol PA. 2007. Identification of genes regulated by the cepIR quorum-sensing system in *Burkholderia cenocepacia* by high-throughput screening of a random promoter library. **Journal of Bacteriology** 189: 968-79

Urbanowski ML, Lostroh CP, Greenberg EP. 2004. Reversible acyl-homoserine lactone binding to purified *Vibrio fischeri* LuxR protein. **Journal of Bacteriology** 186: 631-7

Vandamme P, Holmes B, Coenye T, Goris J, Mahenthiralingam E, et al. 2003. *Burkholderia cenocepacia* sp. nov.--a new twist to an old story. **Res Microbiol** 154: 91-6

Vandamme P, Holmes B, Vancanneyt M, Coenye T, Hoste B, et al. 1997. Occurrence of multiple genomovars of *Burkholderia cepacia* in cystic fibrosis patients and proposal of *Burkholderia multivorans* sp. nov. **Int J Syst Bacteriol** 47: 1188-200

Vanlaere E, Baldwin A, Gevers D, Henry D, De Brandt E, et al. 2009. Taxon K, a complex within the *Burkholderia cepacia* complex, comprises at least two novel species, *Burkholderia contaminans* sp. nov. and *Burkholderia lata* sp. nov. **Int J Syst Evol Microbiol** 59: 102-11

Vanlaere E, Lipuma JJ, Baldwin A, Henry D, De Brandt E, et al. 2008. *Burkholderia latens* sp. nov., *Burkholderia diffusa* sp. nov., *Burkholderia arboris* sp. nov., *Burkholderia seminalis* sp. nov. and *Burkholderia metallica* sp. nov., novel species within the *Burkholderia cepacia* complex. **Int J Syst Evol Microbiol** 58: 1580-90

Vannini A, Volpari C, Gargioli C, Muraglia E, Cortese R, et al. 2002. The crystal structure of the quorum sensing protein TraR bound to its autoinducer and target DNA. **EMBO J** 21: 4393-401

Venturi V, Friscina A, Bertani I, Devescovi G, Aguilar C. 2004. Quorum sensing in the *Burkholderia cepacia* complex. **Res Microbiol** 155: 238-44

Wei Y, Ryan GT, Flores-Mireles AL, Costa ED, Schneider DJ, Winans SC. 2011. Saturation mutagenesis of a CepR binding site as a means to identify new quorum-regulated promoters in Burkholderia cenocepacia. **Molecular Microbiology** 79: 616-32

Weingart CL, White CE, Liu S, Chai Y, Cho H, et al. 2005. Direct binding of the quorum sensing regulator CepR of *Burkholderia cenocepacia* to two target promoters in vitro. **Molecular Microbiology** 57: 452-67

White CE, Winans SC. 2007. The quorum-sensing transcription factor TraR decodes its DNA binding site by direct contacts with DNA bases and by detection of DNA flexibility. **Molecular Microbiology** 64: 245-56

Whitehead NA, Barnard AML, Slater H, Simpson NJL, Salmond GPC. 2001. Quorum-sensing in gram-negative bacteria. **FEMS Microbiology Reviews** 25: 365-404

Zhang R, Pappas T, Brace JL, Miller P, Oulmassov T, et al. 2002. Structure of a bacterial quorum-sensing transcription factor complexed with pheromone and DNA. **Nature** 417: 971–74

Zhu J, Winans SC. 1999. Autoinducer binding by the quorum-sensing regulator TraR increases affinity for target promoters in vitro and decreases TraR turnover rates in whole cells. **Proc Natl Acad Sci U S A** 96: 4832-7

[2]Appendix 2

RepC Protein of the Octopine-type Ti Plasmid Binds to the Probable Origin of Replication Within repC and Functions only in cis

A2.1 Abstract

This collaborative study of the RepC protein was carried out primarily by Dr. Uelinton Pinto, with help from Dr. Esther Costa and myself. Relative contributions are indicated in the text. As this work has been published, the contributions of others are briefly summarized. Vegetative replication and partitioning of many plasmids and some chromosomes of alpha-proteobacteria are directed by their repABC operons. RepA and RepB proteins direct the partitioning of replicons to daughter cells, while RepC proteins are replication initiators, though they do not resemble any characterized replication initiation protein. Here we show that the replication origin of an Agrobacterium tumefaciens Ti plasmid resides fully within its repC gene. Purified RepC bound to a site within repC with moderate affinity, high specificity, and with two-fold cooperativity. The binding site was localized to an AT-rich region that contains a large number of GANTC sites, which have been implicated in replication regulation in related organisms. A fragment of RepC containing residues 26-158 was sufficient to bind DNA, though with limited sequence specificity. This portion of RepC is predicted to have structural homology to members of the MarR family of transcription factors. Overexpression of

2 Pinto, U. M., Flores-Mireles, A. L., Costa E.D., and Winans, S.C. (2011) RepC protein of the octopine-type Ti plasmid binds to the probable origin of replication within *rep*C and functions only in cis. Mol Microbiol **81**(6): 1593-1606.

RepC in A. tumefaciens caused large increases in copy number in *cis* but did not change the copy number of plasmids containing the same *oriV* sequence in *trans*, confirming other observations that RepC functions only in *cis*.

A2.2 Introduction

Plasmids are semiautonomous genetic elements that reside within many genera of bacteria, and can be found in all three domains of life (del Solar et al 1998). Though by definition they are non-essential for host cell viability, many provide their hosts with remarkable new survival strategies (Crossman 2005, Johnson & Nolan 2009, Nikaido 2009, Phale et al 2007). Plasmids have evolved a variety of mechanisms to replicate frequently enough to maintain their copy number. Low-copy plasmids also have mechanisms to faithfully distribute newly replicated plasmids to both daughter cells prior to cell division (Bingle & Thomas 2001, Funnell 2005, Thomas 2000).

Most plasmids have an essential replication initiator protein that plays both positive and negative roles in replication initiation (Chattoraj 2000, Paulsson & Chattoraj 2006). These proteins bind to the origin of replication (*oriV*), which in many cases contains directly repeated DNA sequences called iterons (del Solar et al 1998). One of the ways that replication is limited is by restricting the concentration of the initiator protein, either via negative autoregulation, antisense RNA, protein instability, covalent modification, or dimerization (Paulsson & Chattoraj 2006). In several plasmids, initiator proteins are thought to mediate formation of so-called “handcuffs” in which they bridge iterons of two sister plasmids, preventing further rounds of replication (Das & Chattoraj

2004, Zzaman & Bastia 2005). In some cases, initiators at low concentrations form active monomers, while at high concentration they form inactive dimers (Chattoraj, 2000).

The alpha-proteobacteria are a diverse group that includes photoautotrophs, saprophytes, mutualists, and pathogens. Plasmids play central roles in many of these survival strategies. A prime example is found in the plant pathogen *Agrobacterium tumefaciens,* which causes crown gall tumors on dicotyledonous plants (Smith & Townsend 1907). Pathogenesis requires a 200 kb plasmid called the Ti (tumor-inducing) plasmid, a portion of which is transferred into plant nuclei via a mechanism that clearly evolved from conjugative transfer (Alvarez-Martinez & Christie 2009). Ti plasmids can replicate at low copy number and are partitioned to daughter cells with extremely high fidelity.

All Ti plasmids as well as many other *Agrobacterium* plasmids and secondary chromosomes contain *repABC* operons (Slater et al 2009, Wood et al 2001). RepA and RepB proteins resemble components of plasmid partitioning systems that are widespread in bacterial plasmids, some plasmid-like prophages, and some chromosomes (Funnell 2005). Both proteins are dispensable for plasmid propagation, though null mutations in either gene reduce the efficiency of plasmid inheritance, resulting in the accumulation of cured cells within a culture (Chai & Winans 2005a). RepC is the only protein that is essential for plasmid propagation. A plasmid containing *repC* and a small amount of upstream DNA is capable of autonomous replication (Chai &

Winans 2005b, Izquierdo et al 2005). The origin of replication must therefore lie within this region. To our knowledge no RepC protein from any bacterium has been studied biochemically.

Expression of the *repABC* operon of the octopine-type Ti plasmid is regulated in complex ways. First, there are four promoters upstream of *repA* (Fig. A2.1A). The P4 promoter provides basal levels of expression sufficient to ensure plasmid replication (at approximately one copy per circular or linear chromosome) and partitioning (Pappas & Winans 2003b). This promoter is also inducible by phosphorylated VirG, and is therefore stimulated by phenolic compounds that induce the *vir* regulon. Consequently, phenolic compounds increase Ti plasmid copy number approximately 4-fold per chromosome (Cho & Winans 2005). All four promoters are activated by the quorum-sensing protein TraR in the presence of the pheromone 3-oxo-octanoyl-homoserine lactone, resulting in a 7-8 fold increase in plasmid copy number per chromosome (Li & Farrand 2000, Pappas & Winans 2003a, Pappas & Winans 2003b).

The *repABC* operon is also negatively autoregulated. RepA binds directly downstream of P4, and represses its transcription (Pappas & Winans 2003b). RepB potentiates the ability of RepA to bind to this site by forming a complex. Conversely, RepB can bind to a site between *repA* and *repB* inhibiting transcription of *repB* and *repC* (Chai & Winans 2005a). RepA potentiates the ability of RepB to do so. Finally, an antisense RNA, *repE,* is found directly upstream of *repC*. This small RNA hybridizes with complementary mRNA, blocking transcription and translation of *repC* (Cervantes-Rivera

et al 2010, Chai & Winans 2005b). In previous studies, we have shown that a plasmid with a fragment containing *repE* and *repC* can replicate in *A. tumefaciens* when expressed from the P*lac* promoter of the vector (Chai & Winans 2005b). The same plasmid containing a fragment with just a P*lac-repC* failed to do so, indicating either (i) that such a plasmid replicated at lethally high levels preventing colony formation, or (ii) that *repE* plays some essential role in replication, or (iii) that *oriV* lies within *repE.* RepE RNA can also act in *trans* to downregulate other copies of *repC* in the same cell, and is therefore the major mediator of plasmid incompatibility (Chai & Winans 2005b).

Another member of the alphaproteobacteria whose cell cycle has received considerable attention is *Caulobacter crescentus*. This organism has a DNA methylase, CcrM (cell cycle regulated methylase) that methylates the adenine residue of GANTC sites (Collier et al 2007). The abundance of this methylase is cell-cycle regulated, with a burst of synthesis and activity directly before replication initiation. This causes hemimethylated sites to become fully methylated, which appears to alter the activity of several promoters involved in the cell cycle (McAdams & Shapiro 2003). *A. tumefaciens* has a gene predicted to encode an orthologous methylase (Slater et al 2009, Wood et al 2001). It may be significant that there are five tightly clustered GANTC sites near the middle of the *repC* gene, and three more in the promoter region of *repE* (Fig. A2.1B). This is reminiscent of the chromosome origin of *E. coli,* which is rich in GATC sites, which are methylated by Dam. Methylation of GATC sites is delayed due to binding of the SeqA protein, which is thought to sequester these sites from Dam (Waldminghaus & Skarstad 2009).

As described above, virtually nothing is known about the biochemical properties of any RepC protein. The origin of replication has not been identified, although it must lie within *repC* or directly upstream. There are no directly repeated DNA sequences anywhere in this region, so replication evidently does not require iterons. However, there is an AT-rich region within *repC* that could be part of a replication origin. In an effort to understand this protein, we purified RepC from pTiR10 from an overexpressing strain of *E. coli*, and tested it for binding to sites within *repE* and *repC*. This protein is identical or virtually identical to orthologous proteins of other octopine-type Ti plasmids. We have also tested fragments of RepC protein for DNA binding. Finally, we tested the ability of overexpressed RepC to function in *cis* and in *trans*. To our knowledge this is the first study showing interaction of a purified RepC protein with the putative origin of replication among the large group of *repABC* replicons.

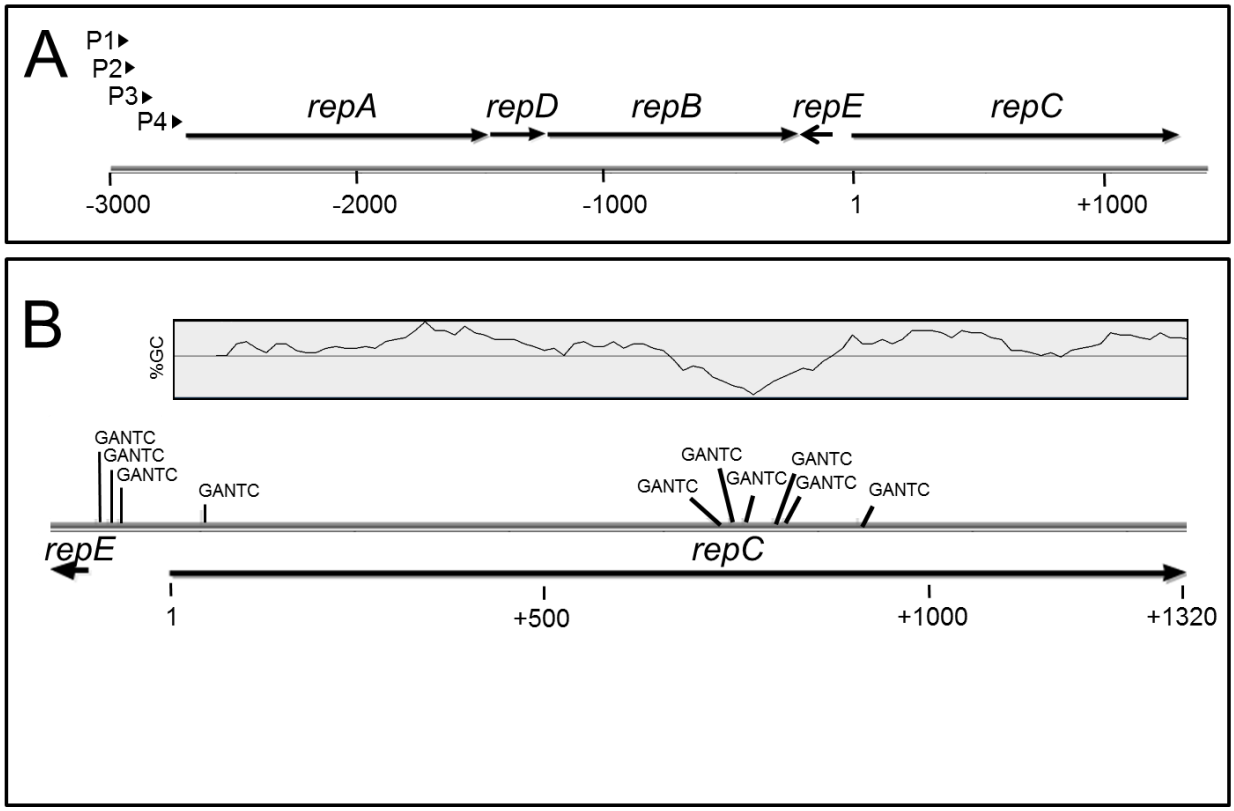


Figure A2.1 Genetic map of the *repABC* region. A) RepA and RepB resemble proteins of plasmid partitioning systems. RepC is the only essential replication initiator. RepE is an antisense RNA that downregulates RepC expression. RepD has no known function. Promoters P1-P4 are activated by TraR-OOHL complexes, while P4 provides basal levels of transcription, is activated by phospho-VirG. All four promoters are autorepressed by RepA and RepB. **B)** Genetic map of *repC* showing the clusters of GANTC sites at the *repE* promoter and at the middle of *repC*. The GC content is shown in the box, and shows an AT-rich region near the middle of the gene.

A2.3 Results

Here I present only the results of my own contributions. For the other results, please refer to (Pinto et al 2011).

RepC overexpression causes strong increases in plasmid copy number.

In previous studies, it was shown that mild overexpression of the native *repABC* operon by activated forms of TraR or VirG causes a mild increase in copy number (Cho & Winans 2005, Li & Farrand 2000, Pappas & Winans 2003a), and that downregulation of the antisense RNA *repE* also increased copy number, probably by increasing the accumulation of RepC (Chai & Winans 2005b). Several plasmids containing P*lac-repC* fusions and lacking *repE* also failed to transform *A. tumefaciens,* possibly because of lethal overreplication (Chai & Winans 2005b). To test this idea Uelinton constructed two new P*lac-repC* plasmids (both lacking *repE*) in the broad host range plasmid pPZP201. The first of these, pUP450, retains the weak native *repC* ribosome binding site (RBS). In this plasmid, translation of *repC* mRNA might possibly be further weakened by translational occlusion, as the *lacZ* translation start site would cause ribosomes to cross the native *repC* ribosome binding site, translating in a different reading frame. The second construct, pUP455, contains a translational fusion between amino acids 1-14 of *lacZ* and codons 2-440 of *repC*. RepC translation in this plasmid begins at the strong *lacZ* ribosome binding site. Note that pUP450 and pUP455 have broad host range replication systems of pPZP201.

Electroporation of strain UIA143 with pUP450 yielded approximately as many colonies per microgram of DNA as with the vector control (pPZP201), though these colonies grow more slowly. Plasmid DNA was recovered in extremely high amounts

(approximately 40 fold higher than the vector pPZP201, Fig. A2.2A). In contrast, electroporation of pUP455 yielded no colonies. In a control experiment, all three plasmids gave rise to colonies when transformed into *E. coli*, and plasmid DNA was recoverable from *E. coli* in similar yields. The most likely explanation is that pUP455 expresses excessively high amounts of RepC, leading to lethal runaway replication, while pUP450, with its weaker ribosome binding site, made sub-lethal amounts of RepC. To test this, we transformed all three plasmids into UIA143(pSRKKm), as pSRKKm encodes *lacI* (Khan et al 2008). Electroporation of pUP455 into this host yielded colonies that grew at normal rates (Table A2.1). In the presence of high levels of IPTG (which are required to inactivate Lac repressor in *A. tumefaciens* (Chen & Winans 1991)), cells containing pUP455 formed slow-growing colonies, while cells containing pUP450 formed normal growing colonies (Table A2.1). Addition of IPTG to broth cultures of UIA143(pSRKKm) (pUP455) also caused a strong increase in copy number of pUP455 (Table A2.1). Taken together, these findings confirm that overexpression of *repC* can cause such high levels of replication that host viability is impaired.

The plasmids described above were designed to have P*lac*-*repC* fusions that would make elevated levels of RepC in response to IPTG. Nevertheless, we wanted to make sure that this was the case. I therefore made derivatives of these plasmids that contained epitope tags at the C-terminus of RepC. This was done by PCR amplifying *repC* using an oligonucleotide that encodes a Flag tag. The Flag tags derivatives of pUP450 and pUP455 are pAFM162 and pAFM163, respectively. I also made a similar Flag-tag derivative of pYC183, which is similar to pUP450, but contains *repE*, which

downregulates *repC* at a post transcriptional level. The Flag-tag derivative of pYC183 is designated pAFM164. All three plasmids were introduced into UIA143 (pSRKKm), and the resulting strains were cultured in the presence or absence of IPTG, and RepC was detected by Western immunoblots. RepC-Flag was easily detectable from strains carrying either of the first two plasmids when cultured in the presence of IPTG. It was not detectable from a strain carrying pAFM164 (Fig. A2.2B). These results demonstrate that pUP450 and pUP455 synthesize elevated levels of RepC in the presence of IPTG, as predicted. These results also underscore the importance of RepE (present in pYC183 and pAFM164) in downregulation of *repC*, as the presence of *repE* caused RepC-Flag to be undetectable. I also conclude that a FLAG epitope tag is tolerated at the C-terminus of RepC, as pAFM164 requires RepC function for replication.

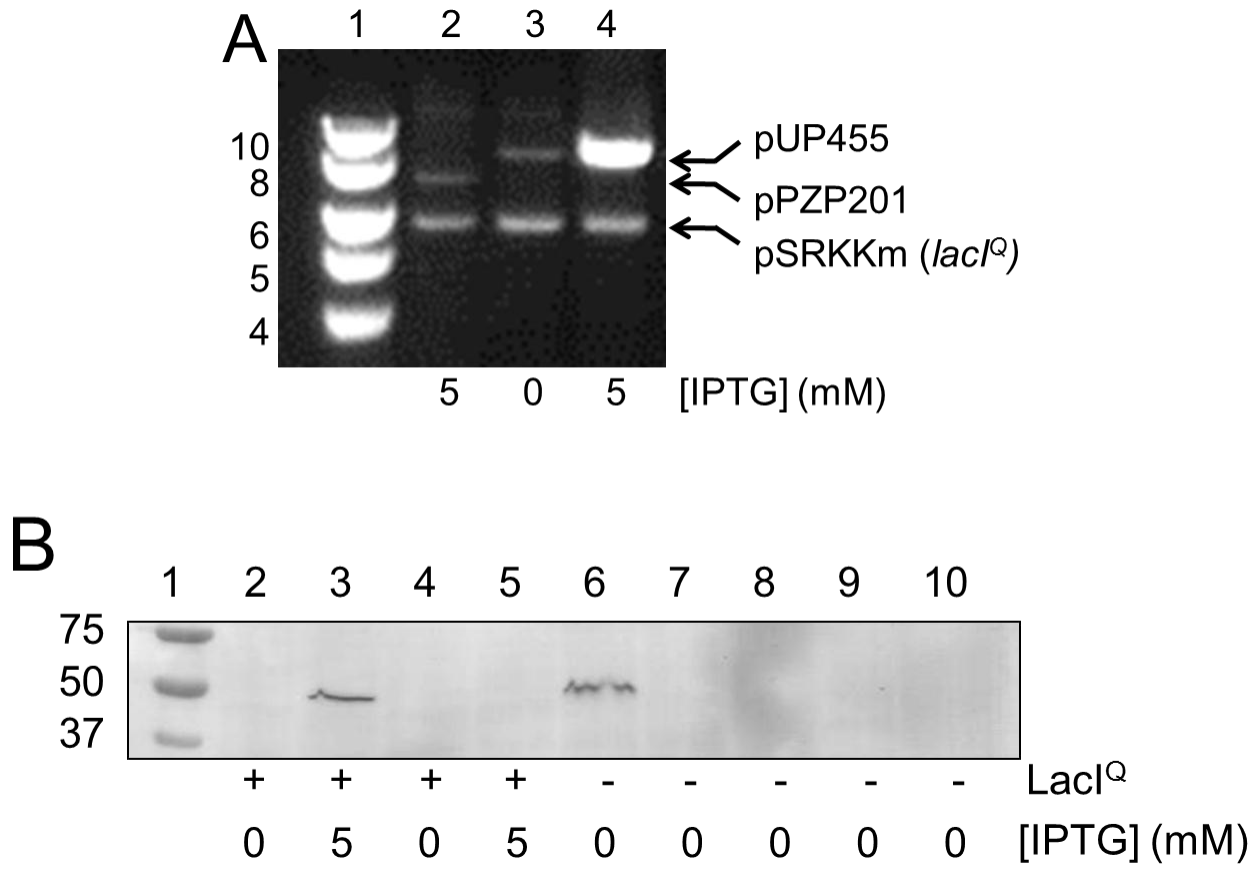


Figure A2.2 Increased expression of RepC causes elevated copy number.

A) Regulated expression of RepC using the P*lac* promoter and Lac repressor. Lane 1: molecular weight standards; lane 2: pPZP201 (broad host range cloning vector); lane 3: Plasmid pUP455 (a derivative of pPZP201 containing a P*lac-repC* fusion) cultured without IPTG; lane 4: pUP455 cultured in the presence of 5 mM IPTG. Copy number of plasmid pUP455 is sharply increased by addition of IPTG. All strains also contain pSRKKm, which expresses the LacI repressor. These plasmids were digested using *Kpn*I, which generates a band of 5.8 kb for pSRKKm, 7.1 Kb for pPZP201, and 8.5 kb for pUP455.

B) Immunodetection of RepC proteins that were tagged with FLAG epitopes.

1: Molecular weight standards; **2**: UIA143(pSRKKm)(pAFM163);

3: UIA143(pSRKKm)(pAFM163); **4**: UIA143(pSRKKm)(pPZP201);

5: UIA143(pSRKKm)(pPZP201); **6**: UIA143(pAFM162); **7**: UIA143 (pPZP201); **8**: UMP01;

9: UIA143(pAFM164); **10**: UIA143. IPTG and LacI^{Q} were present or absent as indicated.

Table A2.1. Uncontrolled expression of RepC causes lethal runaway replication.

pPZP201 derivative	Lac Repressor[1]	Colony formation without IPTG	Colony formation with IPTG (5 mM)	Change in plasmid concentration with IPTG
pPZP201 (no *repC*)	-	+	+	No change
pUP450 (*repC* expressed weakly)	-	+	+	No change
pUP455 (*repC* expressed strongly)	-	-	-	n. a.[2]
pPZP201 (no *repC*)	+	+	+	No change
pUP450 (*repC* expressed weakly)	+	+	+	No change
pUP455 (*repC* expressed strongly)	+	+	+[3]	Strong increase

1: Lac Repressor provided using pSRKKm.
2: not applicable.
3: small colonies.

A2.4 Discussion

See published paper (Pinto et al 2011).

A2.4 Experimental Procedures

Bacterial strains and plasmids

Bacterial strains and plasmids used in this study are listed in Table A2.2. *E. coli* strains were cultured in Luria-Bertani broth (LB) or solid medium at 37°C (Miller, 1972). *A. tumefaciens* strains were cultured in liquid or solid AT minimal medium at 28°C (Cangelosi et al., 1991) or in LB. Spectinomycin and kanamycin were added at 100 µg ml^{-1} to select for resistance in *E. coli* and *A. tumefaciens*. Ampicillin was used at 100 µg ml^{-1} in *E. coli*, while carbenicilin was used at 50 µg ml^{-1} for *A. tumefaciens.*

DNA manipulations

Recombinant DNA techniques were performed using established procedures (Sambrook & Russel 2001). Plasmid DNA was isolated from *E. coli* cells using QIAprep spin miniprep kits (Qiagen). DNA fragments generated by PCR or restriction digestion were gel purified using GeneJET purification columns (Fermentas). Restriction endonucleases were obtained from New England Biolabs and used according to methods described by the manufacturers. Plasmids were introduced into *E. coli* by transformation and into *A. tumefaciens* strains by electroporation (Cangelosi et al 1991). Oligonucleotides were purchased from IDT (Integrated DNA technologies). Site directed

mutagenesis was performed as previously described (Cho et al 2009). Quantification of DNA was done using a NanoDrop (Thermo Scientific).

Construction of flag-tagged RepC

Plasmids pAFM162, pAFM163 and pAFM164 were constructed by using pUP450, pUP455, and pYC183 as a template to generate flag-tagged RepC, respectively. PCR fragments were amplified using oligonucleotides (Table A2.3) and FLAG epitope tag (TTTCGGCCTCGACTTG) was inserted in-frame at the 3'- end of the coding. PCR product were purified, digested with *BamH*I and *Hind*III and cloned into pPZP201 for pAFM162 and pAFM163. Plasmid pAFM164 was constructed in a similar fashion but it was cloned into pBluescript SK+. pAFM162, pAFM163 and pAFM164 were electroporated onto UIA143 strain and UIA143 strain containing pSRKKm.

Immunodetection of RepC

Immunodetection of Flag-RepC protein was tested by growing cultures in 10 ml of AT medium supplemented with appropriate antibiotics and in the absence or presence of 5 mM IPTG. When cultures reached an OD600 of 0.4, they were centrifuged, resuspended in 5% of their original volume in a buffer containing 125 mM Tris pH 6.8, 4% SDS, 20% glycerol, 200 mM DTT, 0.02% bromphenol blue, and disrupted by boiling for 5 min, cooling to -80^{O} C, and boiling for another 5 min. A fraction of each sample was size-fractionated using 12% SDS polyacrylamide gels, and electroblotted onto nitrocellulose membranes (BIORAD). The membranes were blocked using TBS (20

mM Tris pH 7.9, 500 mM NaCl, 0.05% Tween 20) with 5% skim milk, and immunodetected in TBS with monoclonal anti-Flag M2 antibody (Sigma-Aldrich). Goat anti-mouse IgG conjugated with alkaline phosphatase (Jackson immunoResearch Laboratories) was used as the secondary antibody, and the membranes were stained with BCIP (5-bromo-4-chloro-3-indoyl phosphate p-toluidine salt) and NBT (p-nitro blue tetrazolium chloride) (BIORAD).

Table A2.2 Strains and plasmids used in this study

Strains	**Relevant features**	**References**
DH5α	*E. coli*, α-complementation	Stratagene
UIA143	A. tumefaciens C58 strain, cured of Ti plasmid, *rec*A::*ery*, Ery^R	Farrand *et al*, 1989
UMP01	*A. tumefaciens* R10 strain, *repC-flag* tag in the Ti plasmid	This study
Plasmids		
pBluescript SK+	Cloning vector, ColE1 ori, Ap^R	Stratagene
pPZP201	Broad-host-range cloning vector, replication and partitioning genes of pVS1, Sp^R	Hajdukiewicz *et al*. 1994
pYC183	*rep*E-*rep*C of pTiR10 cloned into pBluescript SK+	Chai and Winans, 2005b
pUP450	*rep*C cloned into pPZP201 with native *rep*C ribosomal binding site	This study
pUP455	*rep*C cloned into pPZP201, translational fusion between lacZα and *rep*C	This study
pAFM162	*repC-flag* tag cloned into pPZP201 with native repC ribosomal binding site	This study
pAFM163	*repC-flag* tag cloned into pPZP201, translational fusion between *lacZ*α and *repC*	This study
pAFM164	*repE-repC-flag* tag of pTiR10 cloned into pBluescript SK+	This study

Table A2.3 Oligonucleotides used in this study

Name	Sequence	Comments
ALFM282	5'-cgcggatccagtatcgtttcgacgagctgatttctt-'3	Forward primer for amplification of N- terminus RepC from pUP450, pairs with ALFM183.
ALFM283	5'-ccaagctttcacttatcatcatcgtctttatagtc tttcggcctcgacttg-'3	Reverse primer for addition of FLAG-tag in the C-terminus of RepC into pUP450, pairs with ALF282, ALFM284, and ALFM285.
ALFM284	5'-cgcggatccacagacgcatttatcaacgacgccct-'3	Forward primer for amplification of N- terminus RepC for translational fusion in pPZP201 from pUP455, pairs with ALFM183.
ALFM285	5'-cgcggatccgatgaaatgaaccgcaaaagaaa aaggcccc-'3	Forward primer for amplification of RepE- RepC from pYC183, pairs with ALFM183.

A2.5 References

Alvarez-Martinez CE, Christie PJ. 2009. Biological diversity of prokaryotic type IV secretion systems. **Microbiol Mol Biol Rev** 73: 775-808

Bingle LE, Thomas CM. 2001. Regulatory circuits for plasmid survival. **Curr Opin Microbiol** 4: 194-200

Cangelosi GA, Best EA, Martinetti G, Nester EW. 1991. Genetic analysis of *Agrobacterium*. **Methods Enzymol** 204: 384-97

Cervantes-Rivera R, Romero-Lopez C, Berzal-Herranz A, Cevallos MA. 2010. Analysis of the mechanism of action of the antisense RNA that controls the replication of the *repABC* plasmid p42d. **Journal of Bacteriology** 192: 3268-78

Chai Y, Winans SC. 2005a. RepB protein of an *Agrobacterium tumefaciens* Ti plasmid binds to two adjacent sites between *repA* and *repB* for plasmid partitioning and autorepression. **Molecular Microbiology** 58: 1114-29

Chai Y, Winans SC. 2005b. A small antisense RNA downregulates expression of an essential replicase protein of an *Agrobacterium tumefaciens* Ti plasmid. **Molecular Microbiology** 56: 1574-85

Chattoraj DK. 2000. Control of plasmid DNA replication by iterons: no longer paradoxical. **Molecular Microbiology** 37: 467-76

Chen CY, Winans SC. 1991. Controlled expression of the transcriptional activator gene *virG* in *Agrobacterium tumefaciens* by using the *Escherichia coli lac* promoter. **Journal of Bacteriology** 173: 1139-44

Cho H, Pinto UM, Winans SC. 2009. Transsexuality in the rhizosphere: quorum sensing reversibly converts *Agrobacterium tumefaciens* from phenotypically female to male. **Journal of Bacteriology** 191: 3375-83

Cho H, Winans SC. 2005. VirA and VirG activate the Ti plasmid *repABC* operon, elevating plasmid copy number in response to wound-released chemical signals. **Proc Natl Acad Sci U S A** 102: 14843-8

Collier J, McAdams HH, Shapiro L. 2007. A DNA methylation ratchet governs progression through a bacterial cell cycle. **Proc Natl Acad Sci U S A** 104: 17111-6

Crossman LC. 2005. Plasmid replicons of *Rhizobium*. **Biochem Soc Trans** 33: 157-8

Das N, Chattoraj DK. 2004. Origin pairing ('handcuffing') and unpairing in the control of

P1 plasmid replication. **Molecular Microbiology** 54: 836-49

del Solar G, Giraldo R, Ruiz-Echevarria MJ, Espinosa M, Diaz-Orejas R. 1998. Replication and control of circular bacterial plasmids. **Microbiol Mol Biol Rev** 62: 434-64

Funnell BE. 2005. Partition-mediated plasmid pairing. **Plasmid** 53: 119-25

Izquierdo J, Venkova-Canova T, Ramirez-Romero MA, Tellez-Sosa J, Hernandez-Lucas I, et al. 2005. An antisense RNA plays a central role in the replication control of a *repC* plasmid. **Plasmid** 54: 259-77

Johnson TJ, Nolan LK. 2009. Pathogenomics of the virulence plasmids of *Escherichia coli*. **Microbiol Mol Biol Rev** 73: 750-74

Khan SR, Gaines J, Roop RM, 2nd, Farrand SK. 2008. Broad-host-range expression vectors with tightly regulated promoters and their use to examine the influence of TraR and TraM expression on Ti plasmid quorum sensing. **Appl Environ Microbiol** 74: 5053-62

Li PL, Farrand SK. 2000. The replicator of the nopaline-type Ti plasmid pTiC58 is a member of the *repABC* family and is influenced by the TraR-dependent quorum-sensing regulatory system. **Journal of Bacteriology** 182: 179-88

McAdams HH, Shapiro L. 2003. A bacterial cell-cycle regulatory network operating in time and space. **Science** 301: 1874-7

Nikaido H. 2009. Multidrug resistance in bacteria. **Annu Rev Biochem** 78: 119-46

Pappas KM, Winans SC. 2003a. A LuxR-type regulator from *Agrobacterium tumefaciens* elevates Ti plasmid copy number by activating transcription of plasmid replication genes. **Molecular Microbiology** 48: 1059-73

Pappas KM, Winans SC. 2003b. The RepA and RepB autorepressors and TraR play opposing roles in the regulation of a Ti plasmid *repABC* operon. **Molecular Microbiology** 49: 441-55

Paulsson J, Chattoraj DK. 2006. Origin inactivation in bacterial DNA replication control. **Molecular Microbiology** 61: 9-15

Phale PS, Basu A, Majhi PD, Deveryshetty J, Vamsee-Krishna C, Shrivastava R. 2007. Metabolic diversity in bacterial degradation of aromatic compounds. **Omics** 11: 252-79

Pinto UM, Flores-Mireles AL, Costa ED, Winans SC. 2011. RepC protein of the octopine-type Ti plasmid binds to the probable origin of replication within *rep*C and functions only in *cis*. **Molecular Microbiology** 81: 1593-606

Sambrook J, Russel DW. 2001. *Molecular cloning - a Laboratory manual.* **Cold Spring**

Harbor, NY: Cold Spring Harbor Laboratory Press.

Slater SC, Goldman BS, Goodner B, Setubal JC, Farrand SK, et al. 2009. Genome sequences of three *Agrobacterium* biovars help elucidate the evolution of multichromosome genomes in bacteria. **Journal of Bacteriology** 191: 2501-11

Smith EF, Townsend CO. 1907. A plant-tumor of bacterial origin. **Science** 25: 671-73

Thomas CM. 2000. Paradigms of plasmid organization. **Molecular Microbiology** 37: 485-91

Waldminghaus T, Skarstad K. 2009. The *Escherichia coli* SeqA protein. **Plasmid** 61: 141-50

Wood DW, Setubal JC, Kaul R, Monks DE, Kitajima JP, et al. 2001. The genome of the natural genetic engineer *Agrobacterium tumefaciens* C58. **Science** 294: 2317-23

Zzaman S, Bastia D. 2005. Oligomeric initiator protein-mediated DNA looping negatively regulates plasmid replication in vitro by preventing origin melting. **Mol Cell** 20: 833-43

Appendix 3

Initial Characterization of the Octopine Oxidase Complex

A3.1 Rationale and purpose

As I described in Chapter 3, plants infected by *A. tumefaciens* synthesize a group of nutrients, opines, which are metabolized by the colonizing bacteria. The utilization of octopine, a type of opine, by *A. tumefaciens* is conferred by the activity of the octopine oxidase genes, located within the octopine catabolism (*occ*) operon (Knauf & Nester 1982, Stachel et al 1985). Utilization of opines is an important part of the *A. tumefaciens*-plant interaction and presumably was the driving force in the evolution of the system (Guyon et al 1980, Zanker et al 1992). One of the questions that arose from Chapter 3 was if sulfonopine is catabolized by octopine oxidase to generate SMM and pyruvate. I demonstrated that sulfonopine activates the expression of the genes in the *occ* operon, but the biochemical evidence that the octopine oxidase uses sulfonopine is lacking.

A3.2 Results and Discussion

The catabolic genes were first described by Schroder and colleagues (Zanker et al 1994). Two open reading frames were identified as responsible for the catabolism of octopine, designated as *ooxA* and *ooxB*. These genes were located between the octopine permease (*occQMPT*) and ornithine cyclodeaminase (*ocd*) genes in the *occ* operon (Zanker et al 1992). Considerable intergenic regions were identified flanking ooxA (300 nucleotides upstream and 225 downstream, see Fig. A3.1a. This contrasted with the tight

translational coupling the four permease genes *occQMPT* upstream of *ooxB* (Zanker et al 1992). Amino acid sequence analysis of the subunits suggested that the proteins might utilize FAD, FMN or $NAD(P)^+$ as cofactors (Zanker et al 1994).

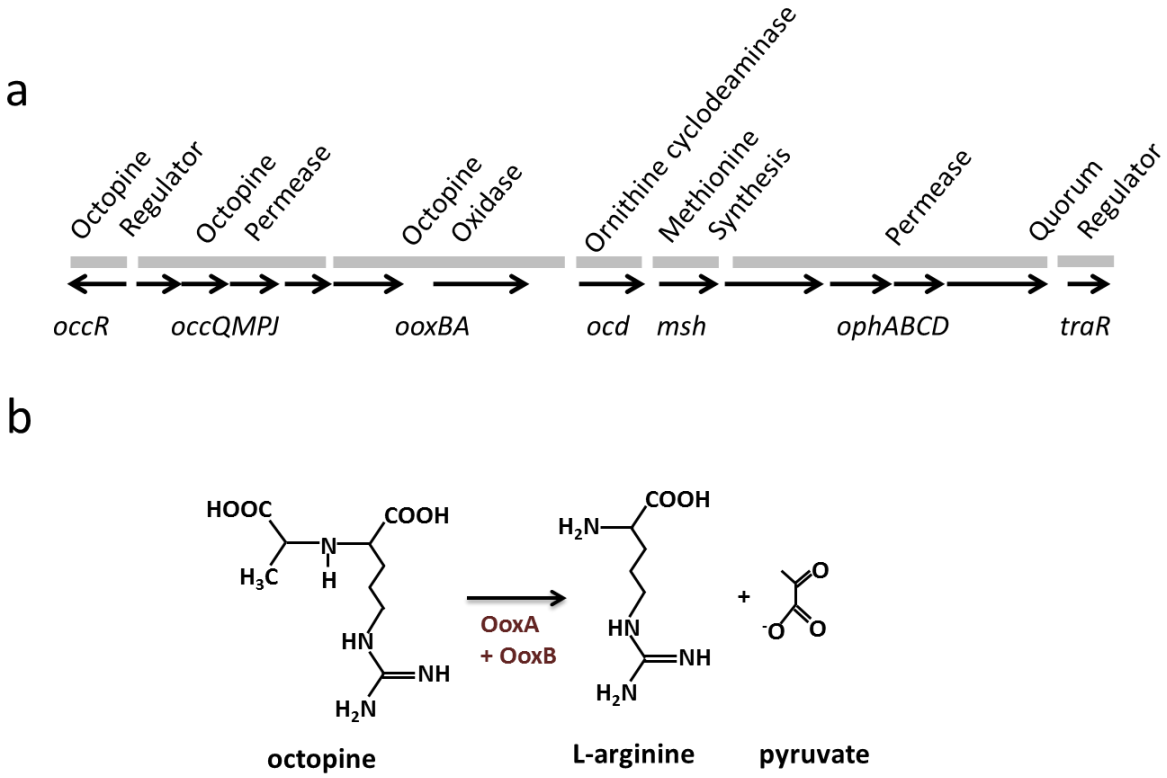


Figure A3.1 Organization and functions of the catabolic regions. a) Location of the catabolic gene, *oox*B and *oox*A. b) Catabolism of octopine by OoxA-OoxB, producing L-arginine and pyruvate.

Zanker and collaborators performed a preliminary characterization of octopine oxidase, showing that crude extracts containing high amounts of these proteins can produce arginine and pyruvate at the expense of octopine (Fig. A3.1b). They found no significant effect on the enzymatic reaction from FAD, FMN, NAD^+ or $NADP^+$, and some combinations even led to reduced activities. Furthermore, they were unable to purify octopine oxidase in an active form, making it difficult to determine both the mechanism of the oxidase reaction and the cofactors involved (Zanker et al 1994).

I have begun an investigation of the ability of octopine oxidase to utilize sulfonopine, with the ultimate goal of reconstituting the reaction with purified components. I first re-examined the DNA sequence and predicted protein sequences of this region. OoxA and OoxB show sequence similarity to a large family of putative oxidases, most of which are uncharacterized. In most cases, *ooxB* homologs are genetically linked to *ooxA* homologs. The close homologs all contain two additional predicted genes that flank the *ooxA* homologs, and one directly downstream (Fig. A3.2). This led me to reexamine the putative intergenic regions that flank *ooxA*. I found one rather short additional open reading frame in each of these regions (Fig. A3.3a-c), and the amino acid sequences of these reading frames is conserved (Fig. A3.4). The stop codons of each gene is tightly linked to the start codon of the next gene downstream (Fig. A3.3c), providing further evidence that these reading frames are functional genes. These genes are provisionally named *ooxC* and *ooxD*.

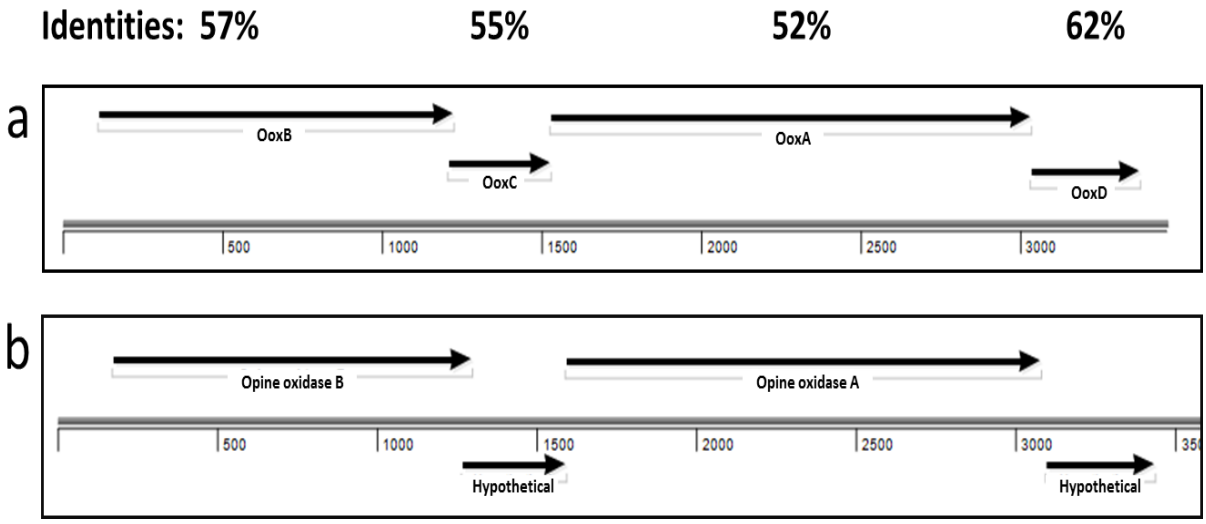


Figure A3.2 Amino acid similarity and synteny between (a) *ooxBCAD* and (b) genes designated EEQ92902.1, EEQ92901.1, EEQ92900.1, and EEQ92899.1 of *Ochrobactrum intermedium* LMG 3301. Amino acid identities are indicated at the top. Homologous genes have similar sizes and are syntenic.

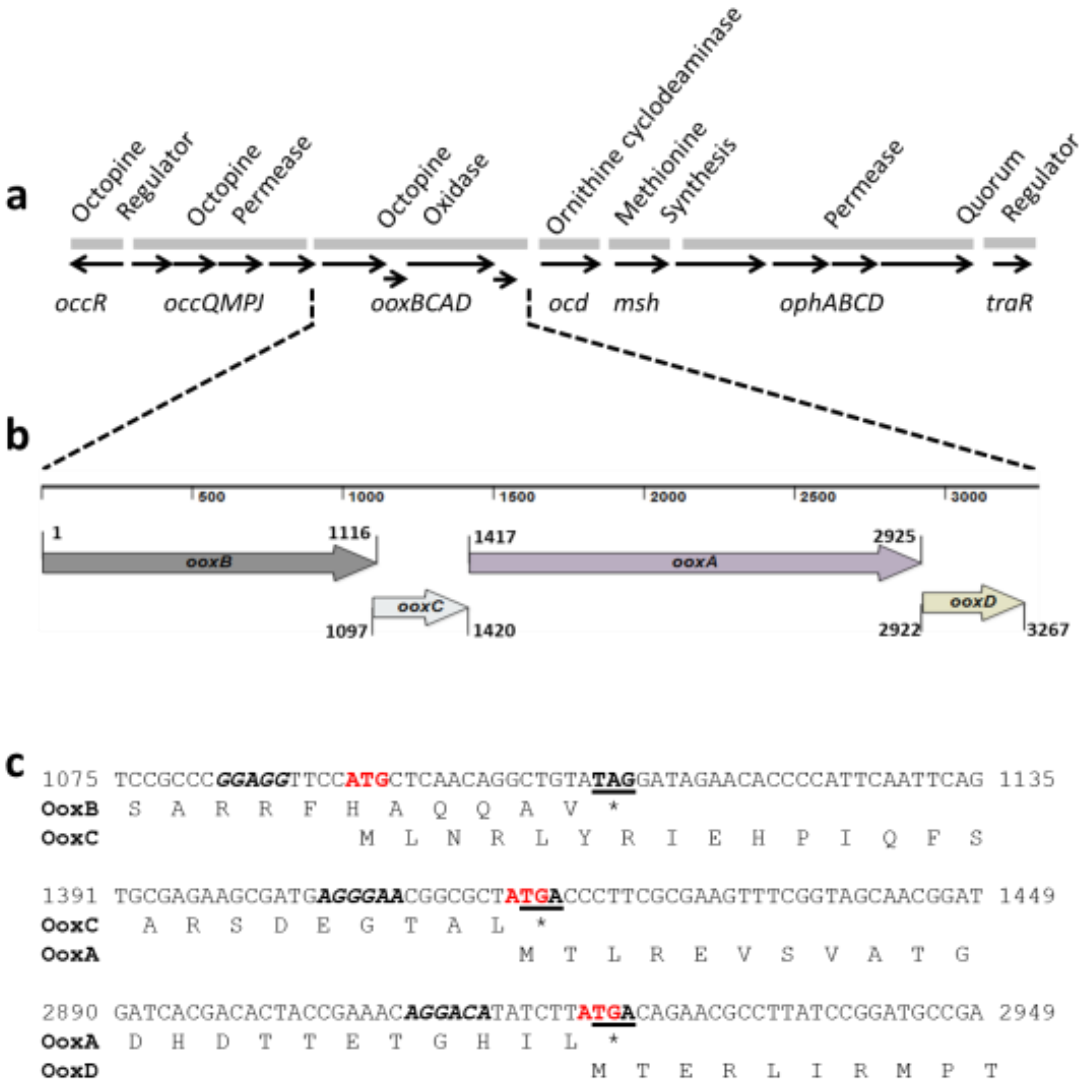


Figure A3.3 Organization of opine catabolic genes and nearby downstream sequence. a) Localization of the opine catabolic genes within the *occ* operon. b) Arrangement of the octopine oxidase subunit genes *ooxB, ooxC, ooxA, and ooxD*. c) Nucleotide sequence at 5' end of the octopine catabolic genes *C, A* and *D*; and the intergenic regions corresponding to the deduced amino acid sequence. N-terminal sequence of each subunit is in red, the ribosomal binding site is in *italics* and stop codon is underlined. Numbers correspond to the nucleotide position in the opine catabolic genes sequence showed in (b).

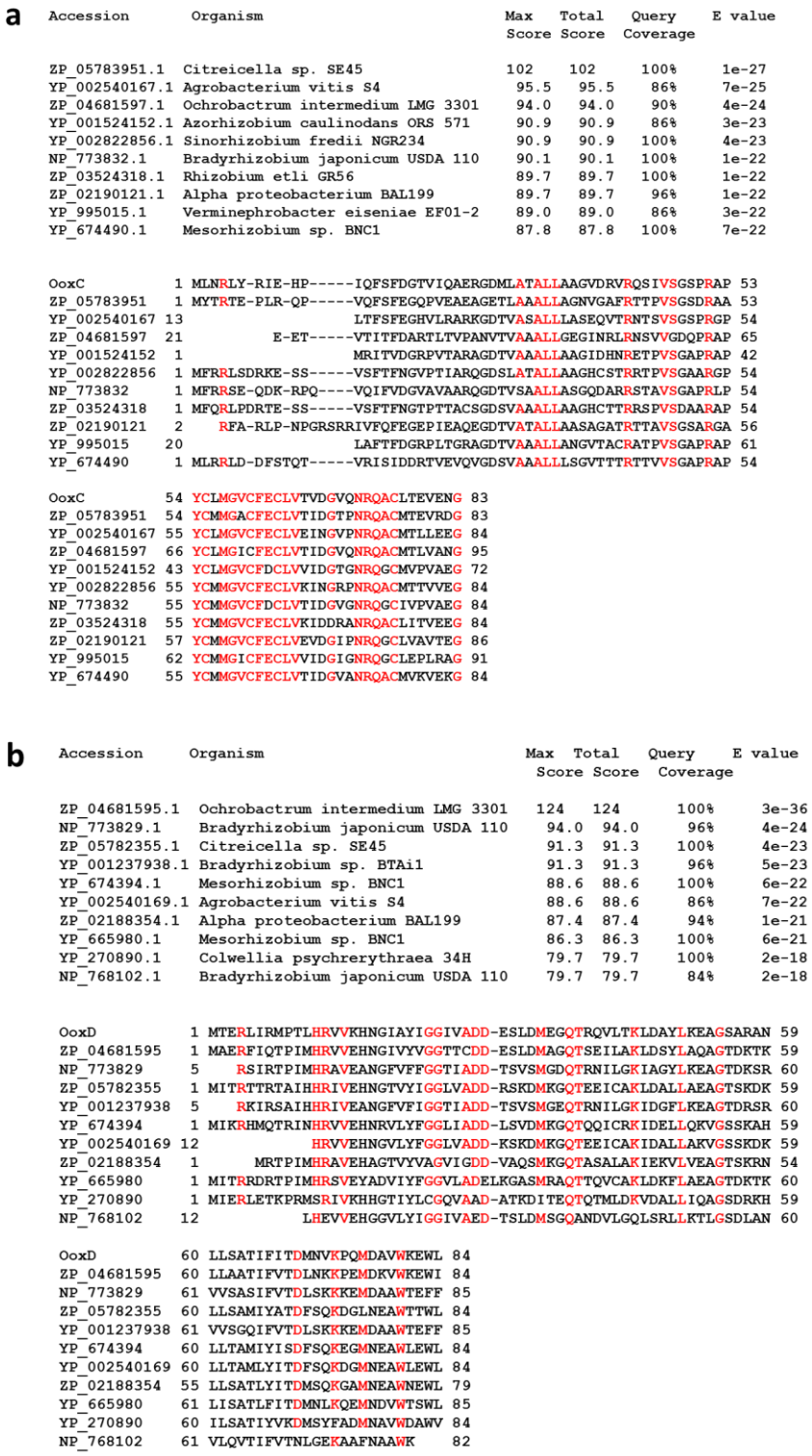


Figure A3.4 OoxC and OoxD amino acids resemble many proteins in the Genbank protein database. Alignments of a)OoxC and b) OoxD. Conserved amino acids are in

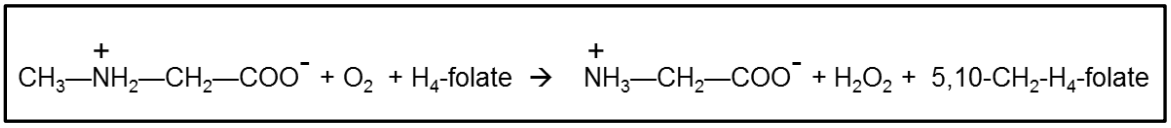
red.

Directly downstream of *ooxD* are small stretches of sequence that resemble genes that encode arginases. Arginases convert arginine to ornithine and urea (Schindler et al 1989, Schrell et al 1989). Arginase is required for the complete catabolism of octopine, but is encoded on another replicon. It is noteworthy that the nopaline-type Ti plasmid pTiC58 contains a functional arginase gene located within the nopaline catabolism operon (Sans et al 1987, Schindler et al 1989, Schrell et al 1989). The finding of fragments of a homologous arginase downstream of *ooxD* suggests that an ancestor of the octopine-type Ti plasmid may have contained a functional arginase, most of which was subsequently deleted.

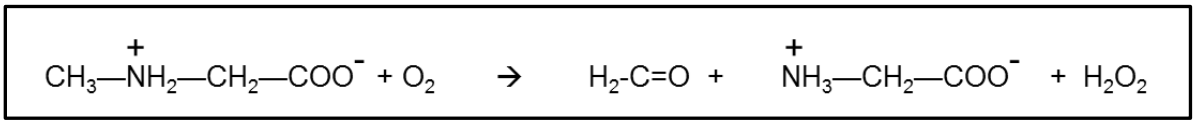
Among the family of octopine oxidase homologs, the best characterized is the heterotetrameric sarcosine oxidase of *Corynebacterium spp*., which has been studied at the biochemical and structural levels. Sarcosine oxidase is encoded by four genes *soxB, soxD, soxA, and soxG* (Chlumsky et al 1995, Hassan-Abdallah et al 2005). Among these SoxB strongly resembles OoxB, while OoxA resembles the N-terminal half of SoxA. OoxC and OoxD do not significantly resemble SoxD or SoxG. Sarcosine oxidase contains three coenzymes (NAD, FMN and FAD), and a binding site for a fourth coenzyme (tetrahydrofolate, H_4folate) that acts as a substrate. NAD is noncovalently bound to the A subunit, while the B subunit has covalently bound FMN and noncovalently bound FAD (Chlumsky et al 1995, Hassan-Abdallah et al 2005).

Sarcosine oxidase demethylates sarcosine (N-methyl-glycine). In the presence of

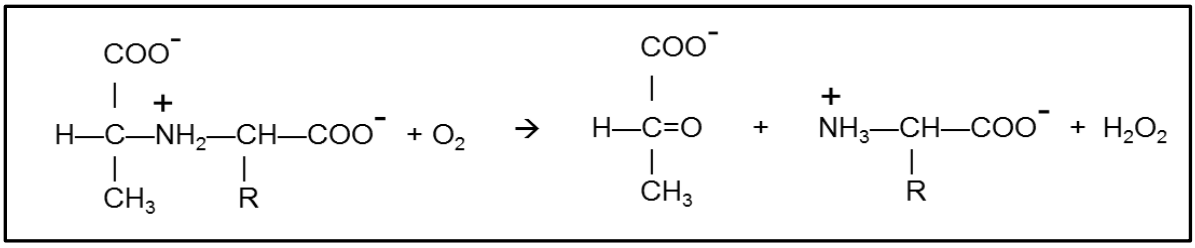
tetrahydrofolate, the overall reaction is:



In the absence of tetrahydrofolate, the reaction proceeds at the same rate and releases formaldehyde:



The C-terminal half of SoxB contains the binding site for tetrahydrofolate. This region of SoxB is absent from OoxB, suggesting that octopine catabolism does not involve tetrahydrofolate. Accordingly, I propose that octopine catabolism proceeds by a reaction similar to the second reaction above:



If this is correct, then octopine catabolism generates arginine and pyruvate, as long predicted. However, it would not convert NAD(P)+ to NAD(P)H. Rather, it would generate hydrogen peroxide at the expense of oxygen. This reaction would presumably occur spontaneously as written.

Sarcosine oxidase complex has been purified by His-tagging the SoxG subunit and then using IMAC to purify the complex (Hassan-Abdallah et al 2005, Suzuki et al 2005). In order to test if the octopine oxidase complex also contains four subunits, the four predicted genes were introduced into plasmid pET11a in such a way that they would be expressed from the strong T7 promoter, and the OoxD protein would have a His_6 tag. The resulting plasmid, pAFM103, was introduced into BL21/DE3 by transformation and then the *ooxBCDA* operon was expressed by addition of IPTG at 24˚C overnight. The cells then were concentrated by centrifugation, disrupted using a French pressure cell, and the lysate was clarified by ultracentrifugation. This lysate was compared by SDS PAGE to one identically prepared from a strain containing pET11a. The predicted size of OoxB, OoxC, OoxA, and OoxD-His6 are 39.9 kDa, 11.5 kDa, 53 kDa, and 13.4 kDa, respectively.

Three major proteins were detected in the lysate made from the strain containing pAFM103 but not in the control lysate. The largest of these was 55 KDa, in reasonable agreement with the predicted mass of OoxA. The second largest band was 33 KDa, somewhat smaller than the predicted mass of OoxB. The third band had an apparent mass of approximately 13 KDa, reasonably close to the predicted masses of both OoxC and OoxD (Fig. A3.5a).

The cleared lysate from BL21/DE3(pAFM103) was purified by IMAC. This was done for two reasons. First, I wanted to unambiguously detect OoxD. Second, I wanted to determine whether the four predicted proteins formed a complex that could be purified by IMAC. Bound complexes were eluted using imidazole. The eluted proteins from

BL21/DE3(pAFM103) contained extremely large amounts of a protein approximately 13 KDa. This protein was absent from the negative control, and I therefore conclude that this protein is OoxD-His_6, confirming the existence of OoxD (Fig. A3.5b). The eluted proteins did not include high levels of OoxA or OoxB (OoxC if present, would have been obscured by OoxD). Therefore, under the conditions used, these four proteins did not form a complex that was stable during IMAC.

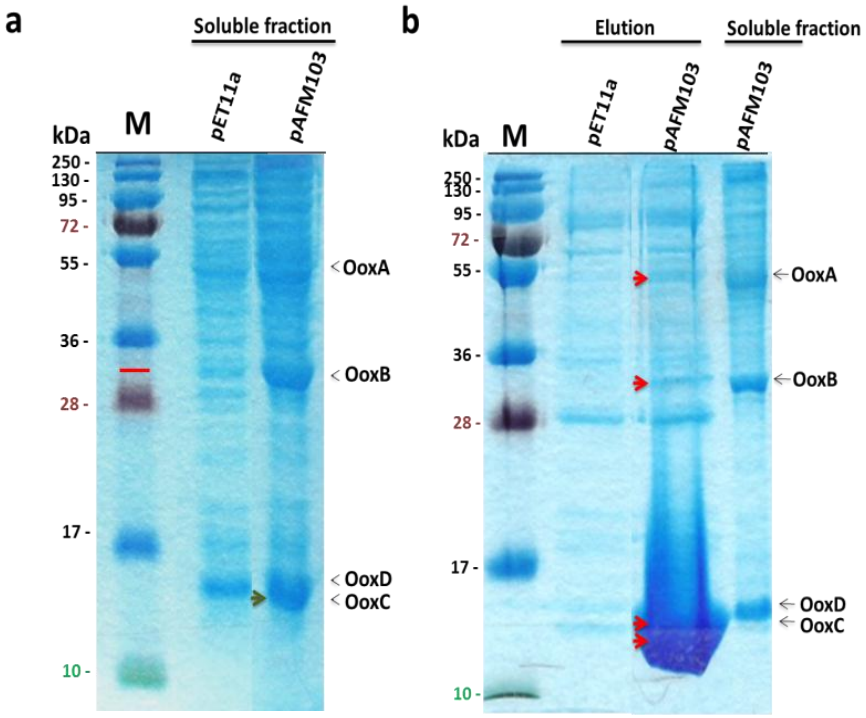


Figure A3.5 Coomasie blue stained-PAGE gel of the soluble fraction and purification of the octopine oxidase subunits. a) Shows the soluble fraction comparison between the pET11a (vector control) and the pAFM103 (which contains *ooxA, ooxB, ooxC*, and *ooxD-His6*). b) IMAC purification of lysates of BL21/DE3(pET11a) and BL21/DE3(pAFM103); the fourth lane is the pAFM103 soluble fraction, which is used as reference for the subunits. M refers to the size standard marker.

In future work, new approaches will be tested to optimize the purification of the OoxABCD proteins. Fortunately, at least three of these proteins, and probably all four, are strongly expressed in soluble forms. A complex containing all four proteins is predicted to have a significant negative charge, and so their ability to bind to Q-Sepharose matrix will be tested. If all subunits bind and all co-elute in a gradient of increasing salt, this will provide evidence for complex formation. Whether they form a complex or not, they should be purifiable by ion chromatography, especially given their high abundance. Cofactors of each protein can be detected by protein mass spectroscopy at the Cornell Life Sciences Core Facility.

The purified OoxABCD subunits will be tested for the ability to consume produce arginine and pyruvate at the expense of octopine. All three can be quantified by ESI-MS/MS. The production of arginine and pyruvate is almost a foregone conclusion. More interesting is the identity of the required oxidant. The data from sarcosine oxidase predicts that the reaction will consume oxygen and produce H_2O_2. The production of H_2O_2 can be quantified spectrophotometrically using commercially available reagents (for example the Red Hydrogen Peroxide Assay Kit, from Enzo Life Sciences, which generates a fluogenic substrate). The enzyme and octopine will be combined in a buffer lacking oxygen. I predict that little if any H_2O_2 will be produced, and that production will increase greatly when oxygen is introduced as a stream above the reaction. Similar reactions will also be done using sulfonopine or other opines in place of octopine.

A3.3 References

Chlumsky LJ, Zhang L, Jorns MS. 1995. Sequence analysis of sarcosine oxidase and

nearby genes reveals homologies with key enzymes of folate one-carbon metabolism. **The Journal of biological chemistry** 270: 18252-9

Guyon P, Chilton MD, Petit A, Tempe J. 1980. Agropine in "null-type" crown gall tumors: Evidence for generality of the opine concept. **Proceedings of the National Academy of Sciences of the United States of America** 77: 2693-7

Hassan-Abdallah A, Zhao G, Eschenbrenner M, Chen ZW, Mathews FS, Jorns MS. 2005. Cloning, expression and crystallization of heterotetrameric sarcosine oxidase from Pseudomonas maltophilia. **Protein expression and purification** 43: 33-43

Knauf VC, Nester EW. 1982. Wide host range cloning vectors: a cosmid clone bank of an *Agrobacterium* Ti plasmid. **Plasmid** 8: 45-54

Sans N, Schroder G, Schroder J. 1987. The Noc region of Ti plasmid C58 codes for arginase and ornithine cyclodeaminase. **European Journal of Biochemistry / FEBS** 167: 81-7

Schindler U, Sans N, Schroder J. 1989. Ornithine cyclodeaminase from octopine Ti plasmid Ach5: identification, DNA sequence, enzyme properties, and comparison with gene and enzyme from nopaline Ti plasmid C58. **Journal of Bacteriology** 171: 847-54

Schrell A, Alt-Moerbe J, Lanz T, Schroeder J. 1989. Arginase of Agrobacterium Ti plasmid C58. DNA sequence, properties, and comparison with eucaryotic enzymes. **European journal of biochemistry / FEBS** 184: 635-41

Stachel SE, An G, Flores C, Nester EW. 1985. A Tn3 lacZ transposon for the random generation of beta-galactosidase gene fusions: application to the analysis of gene expression in *Agrobacterium*. **EMBO journal** 4: 891-8

Suzuki H, Tamamura R, Yajima S, Kanno M, Suguro M. 2005. Corynebacterium sp. U-96 contains a cluster of genes of enzymes for the catabolism of sarcosine to pyruvate. **Bioscience, Biotechnology, and Biochemistry** 69: 952-6

Zanker H, Lurz G, Langridge U, Langridge P, Kreusch D, Schroder J. 1994. Octopine and nopaline oxidases from Ti plasmids of *Agrobacterium tumefaciens*: molecular analysis, relationship, and functional characterization. **Journal of Bacteriology** 176: 4511-7

Zanker H, von Lintig J, Schroder J. 1992. Opine transport genes in the octopine (occ) and nopaline (noc) catabolic regions in Ti plasmids of *Agrobacterium tumefaciens*. **Journal of Bacteriology** 174: 841-9