

Fall 2009

Engineering the Mosquito Symbiont *Pantoea* agglomerans to secrete Anti-Plasmodium Inhibitory Proteins

Dawn C. Bisi

Follow this and additional works at: <https://dsc.duq.edu/etd>

Recommended Citation

Bisi, D. (2009). Engineering the Mosquito Symbiont *Pantoea* agglomerans to secrete Anti-Plasmodium Inhibitory Proteins (Doctoral dissertation, Duquesne University). Retrieved from <https://dsc.duq.edu/etd/323>

This Immediate Access is brought to you for free and open access by Duquesne Scholarship Collection. It has been accepted for inclusion in Electronic Theses and Dissertations by an authorized administrator of Duquesne Scholarship Collection. For more information, please contact phillips@duq.edu.

ENGINEERING THE MOSQUITO SYMBIONT *PANTOEA AGGLOMERANS*
TO SECRETE *PLASMODIUM* INHIBITORY PROTEINS

A Dissertation

Submitted to the Bayer School
of Natural and Environmental Sciences

Duquesne University

In partial fulfillment of the requirements for
the degree of Doctor of Philosophy in Biological Sciences

By

Dawn C. Bisi

August 2009

Copyright by

Dawn C. Bisi

2009

ENGINEERING THE MOSQUITO SYMBIONT *PANTOEA AGGLOMERANS*
TO SECRETE *PLASMODIUM* INHIBITORY PROTEINS

By

Dawn C. Bisi

Approved July 20, 2009

Dr. David J. Lampe
Thesis Advisor

Dr. Mary Alleman
Committee Member

Dr. Joseph R. McCormick
Committee Member

Dr. Marcelo Jacobs-Lorena
Committee Member

Dr. Philip E. Auron
Chair, Department of Biological
Sciences

Dr. David W. Seybert
Dean, Bayer School of Natural and
Environmental Sciences

ABSTRACT

ENGINEERING THE MOSQUITO SYMBIONT *PANTOEA AGGLOMERANS* TO SECRETE *PLASMODIUM* INHIBITORY PROTEINS

By

Dawn C. Bisi

August 2009

Dissertation Supervised by Dr. David J. Lampe

The disease malaria originates from the transmission of the parasite *Plasmodium* to humans by female anopheline mosquitoes. Estimates put the number of deaths at 1-3 million people annually and this number will increase without the establishment of new control strategies. There is currently no vaccine and the effectiveness of insecticides and drugs are thwarted by the gain of resistances for both the insect and parasite. An alternative genetic engineering approach to combating malaria is presented here. The bacterial mosquito symbiont *Pantoea agglomerans*, which resides in the mosquito gut, was chosen to express anti-*Plasmodium* effector gene products that are known to inhibit *Plasmodium* development. A caveat is finding an appropriate protein secretion signal for export of the effectors from the cell. A two-pronged approach to finding a secretion signal involved identifying and testing a native secreted protein signal in

P. agglomerans, and also testing heterologous secretion signals shown to work previously in related species. Proteomic analysis of native secreted proteins in spent growth medium followed by the identification of the corresponding genes revealed the best native candidate for trials was the secreted protein FliC/Flagellin. Various constructs involving the *fliC* 5' UTR and the *fliC* ORF were used for secretion of a test protein (an anti-BSA scFv) in *P. agglomerans* and found to not be a sufficient secretion signal. The heterologous PelB (from *Erwinia carotovora*), and OmpA, TolB, and HlyA (all from *E. coli*) signals were also used in secretion trials with the anti-BSA scFv. The PelB and HlyA signals were shown to secrete the scFv in *P. agglomerans*, however it was only active in the case of PelB-induced secretion. In addition, four anti-*Plasmodium* effector proteins (SM1, Anti-Pbs21, PLA2, and CEL-III) were available for testing in constructs containing the heterologous secretion signals. Varying success was observed with the different combinations of signals and effector genes. The OmpA and TolB signals were not functional in *P. agglomerans*. *P. agglomerans* was able to secrete Anti-Pbs21-HlyA and PLA2 H67N-HlyA fusions and these strains are now available for testing inside malaria-infected anophelines for the inhibition of *Plasmodium* development.

ACKNOWLEDGEMENTS

I'd like to thank my Ph.D. advisor, David Lampe, for his constant support and encouragement over the years. He has become a good friend and I am grateful for the time spent working in his lab. He has taught me everything from cloning to the best way to start a vegetable garden. And I'll never forget how I left him speechless after telling him that I got to drink out of the Stanley Cup once.

My fantastic committee members, Mary Alleman, Joseph McCormick, and Marcelo Jacobs-Lorena, always made themselves available and provided me with limitless advice and help with experiments and ideas. I thank them all for everything.

I'd like to thank the faculty and staff of the Duquesne University Biology Department for the knowledge and guidance that they have bestowed upon me. Throughout all of my courses, the faculty has provided me with so much information and prepared me for my post-graduate career. The staff has always been incredibly kind and helpful and has made my stay at Duquesne an enjoyable one.

I need to acknowledge my fellow graduate students, Rebekah Dedrick and Allen Kotun, and thank them for getting me through graduate school! Their friendships are something that I will always cherish.

I'd also like to thank my amazing husband, Luke Ferdinand. Knowing that he believed in me has made this journey worthwhile. Without him by my side, it is very likely that I would not be writing these words right now.

Finally, I'd like to thank my family and friends. I am incredibly grateful for your love and support. Thank you so much for encouraging me to finish! I'm so lucky to have you all in my life!

TABLE OF CONTENTS

	Page
ABSTRACT.....	iv
ACKNOWLEDGEMENTS.....	vi
LIST OF TABLES.....	x
LIST OF FIGURES.....	xi
CHAPTER 1: INTRODUCTION	1
1.1 Malaria: The Disease and Its Insect Vector.....	1
1.2 The Parasite <i>Plasmodium</i> Causes Malaria.....	3
1.3 The Manifestation of Malaria in Humans.....	6
1.4 Past Treatments and Preventative Measures.....	6
1.5 Combating Insect-Vectored Diseases with Genetic Approaches.....	8
1.6 Paratransgenesis as a Means to Control Diseases.....	14
1.7 Bacterial Candidates for Paratransgenesis.....	18
1.8 Protein Secretion Systems in Gram-negative Bacteria.....	22
1.9 Applying Secretion Systems to Deliver Heterologous Proteins.....	31
CHAPTER 2: IDENTIFYING AND APPLYING NATIVE SECRETED PROTEINS FOR USE IN SECRETION IN <i>PANTOEA AGGLOMERANS</i>	37
Abstract.....	37
Introduction.....	38
Materials and Methods.....	41

TABLE OF CONTENTS - CONTINUED

	Page
Results.....	50
Discussion.....	53
CHAPTER 3: USING HETEROLOGOUS SECRETION SIGNALS FOR SECRETION OF A TEST PROTEIN AND ANTI-<i>PLASMODIUM</i> EFFECTOR PROTEINS IN <i>PANTOEA AGGLOMERANS</i>	74
Abstract.....	74
Introduction.....	75
Materials and Methods.....	79
Results.....	86
Discussion.....	90
CHAPTER 4: SUMMARY AND FUTURE DIRECTIONS	111
REFERENCES	117
APPENDIX 1	128
APPENDIX 2	141
APPENDIX 3	143

LIST OF TABLES

	Page
Table 2.1: Strains Used in this Study.....	61
Table 2.2: Plasmids Used in This Study.....	62
Table 2.3: Oligonucleotides Used in This Study.....	63
Table 2.4: Identification of <i>P. agglomerans</i> secreted proteins by MALDI-TOF.....	65
Table 3.1: N-Terminal Secretion Signal Sequences.....	101
Table 3.2: Strains Used in this Study.....	102
Table 3.3: Plasmids Used in this Study.....	103
Table 3.4: Summary of Expression and Secretion Results Using all Signals and Effector Proteins in <i>E. coli</i> and <i>P. agglomerans</i>	109
Table A1.1: Strains Used in this Study.....	132
Table A1.2: Plasmids Used in this Study.....	132
Table A1.3: Oligonucleotides Used in this Study.....	133

LIST OF FIGURES

	Page
Figure 1.1: The life cycle of <i>Plasmodium</i>	4
Figure 2.1: A representative silver-stained 2D-PAGE showing the secreted protein profile of <i>P. agglomerans</i>	64
Figure 2.2: FliC/Flagellin MALDI-TOF results from Taplin Biological Mass Spec Facility.....	66
Figure 2.3: Peptide fragments matches to <i>E. coli</i> FliC/Flagellin from Taplin Biological Mass Spec Facility.....	67
Figure 2.4: Identification and placement of sequenced <i>P. agglomerans</i> FliC peptide fragments in the <i>E. coli</i> FliC sequence and the <i>P. agglomerans</i> <i>fliC</i> contig assembled after genome walking PCR.....	68
Figure 2.5: Genome walking PCR protocol.....	69
Figure 2.6: <i>fliC</i> secretion constructs used in this study.....	70
Figure 2.7: Attempts to delete <i>fliC</i> from the <i>P. agglomerans</i> chromosome were unsuccessful.....	72
Figure 3.1: Secretion constructs used in this study.....	104
Figure 3.2: Secretion of proteins using PelB in <i>E. coli</i> and <i>P. agglomerans</i>	105
Figure 3.3: Secretion of proteins using OmpA in <i>E. coli</i> and <i>P. agglomerans</i>	106
Figure 3.4: Secretion of proteins using TolB in <i>E. coli</i> and <i>P. agglomerans</i>	107
Figure 3.5: Secretion of proteins using HlyA in <i>E. coli</i> and <i>P. agglomerans</i>	108
Figure 3.6: Proposed polycistronic effector and chaperone construct.....	110
Figure A1.1: FlgL MALDI-TOF results from Taplin Biological Mass Spec Facility.....	134
Figure A1.2: Peptide fragments matches to <i>E. coli</i> FlgL from Taplin Biological Mass Spec Facility.....	135

LIST OF FIGURES - CONTINUED

	Page
Figure A1.3: Ssb MALDI-TOF results from Taplin Biological Mass Spec Facility.....	136
Figure A1.4: Peptide fragments matches to <i>E. coli</i> Ssb from Taplin Biological Mass Spec Facility.....	137
Figure A1.5: Identification and placement of sequenced <i>P. agglomerans</i> FlgL peptide fragments in the <i>E. coli</i> FlgL sequence and the <i>P. agglomerans</i> <i>flgL</i> contig assembled after genome walking PCR.....	138
Figure A1.6: Identification and placement of sequenced <i>P. agglomerans</i> Ssb peptide fragments in the <i>E. coli</i> Ssb sequence and the <i>P. agglomerans</i> <i>ssb</i> contig assembled after genome walking PCR.....	139
Figure A1.7: Secretion constructs using <i>P. agglomerans</i> <i>ssb</i> ORF.....	140
Figure A2.1: Growth curve of wild-type <i>P. agglomerans</i> in LB broth.....	142
Figure A3.1: Parallel gels of late- <i>log</i> protein preparations of <i>P. agglomerans</i> secreted proteins.....	144
Figure A3.2: Additional parallel gels of late- <i>log</i> protein preparations of <i>P. agglomerans</i> secreted proteins.....	145

CHAPTER 1

INTRODUCTION

1.1 Malaria: The Disease and Its Insect Vector

Malaria is a mosquito-vectored disease caused by protozoan parasites of the genus *Plasmodium*. It is a very serious and often fatal disease and is a major worldwide health concern. Around 41% of the human population is at high risk for malaria due to their residence in the subtropical and tropical areas on Earth. The majority of cases occur in sub-Saharan Africa, but the disease is also present in parts of South Asia, the Middle East, and regions of Central and South America. Estimates put the current number of cases of malaria at approximately 500 million and the number of annual deaths at 1-3 million (BREMAN *et al.* 2001; SNOW *et al.* 2005). The Centers for Disease Control and Prevention estimates that as many as 2.7 million lives are lost to malaria each year (CDC 2006). Most of these deaths occur in African children under the age of five. It cannot be underestimated that successful efforts to curb the spread of this disease are desperately needed.

The mosquito vector for malaria, females of the genus *Anopheles*, thrives in the hot and humid climate of subtropical and tropical regions. The females bite humans in order to obtain blood necessary for egg production. If this mosquito is infected with *Plasmodium*, it can transmit the parasite to the human while taking a blood meal. Also, a

naïve mosquito can obtain *Plasmodium* by feeding on an infected human. This exchange between human and mosquito is required for a successful *Plasmodium* life cycle.

Anopheline mosquitoes go through four life cycle stages: egg, larva, pupa, and adult (or imago). Environmental conditions such as ambient temperature and humidity will determine the lifespan of an adult mosquito as well as the successful transmission of malaria. An infected female mosquito must be alive for 9-21 days in order for *Plasmodium* to complete the extrinsic phase of its growth cycle. The cycle will be shorter if ambient temperatures are above 25°C, thus increasing the chance of transmission to humans who live in the warmer regions on Earth.

There are 400 identified species of *Anopheles*, and 30-40 of these species have the capacity to transmit malaria in the wild. *An. gambiae* and *An. funestus* are the most successful malarial vectors in Africa due to their anthropophilic feeding behavior; they strongly prefer to feed on humans rather than on other mammals. Another species, *An. stephensi*, can transmit a malarial parasite called *Plasmodium berghei* that affects rodents and is often used as a model research system.

Mosquitoes carrying malaria incur fitness costs from loss of fecundity to decreased flight distance. As a result, there is significant selective pressure for species like *An. gambiae* to carry refractory alleles to *Plasmodium*. There exists a wide range of refractoriness across different *Anopheles* species and the molecular explanation for this phenomenon is not completely understood. Specific loci that confer resistance to *P. falciparum* infection have been identified in W. African mosquito populations (NIARE *et al.* 2002). Some *An. gambiae* strains do not have vectorial capacity because they can trigger a complement-like immune response against midgut stages of *Plasmodium*, while

other strains undergo melanotic encapsulation of the oocyst stage parasites (BLANDIN *et al.* 2004; GHOSH *et al.* 2000). Melanotic encapsulation is an immune response in certain insects, where foreign objects are encapsulated and targeted for destruction. Even with these immune defenses against *Plasmodium*, most anopheline mosquitoes are still successful malarial vectors, contributing to the rising number of newly infected humans every year.

1.2 The Parasite *Plasmodium* causes Malaria

Malaria is caused by parasites within the genus *Plasmodium*. Depending upon the species of *Plasmodium*, malaria can affect many different animals from birds to reptiles to some mammals. Four species of malarial parasites can infect humans: *P. falciparum*, *P. vivax*, *P. ovale*, and *P. malariae*. *P. falciparum* and *P. vivax* are responsible for the most infections throughout the world with the former causing the most severe and potentially fatal form of the disease. *P. ovale* or *P. malariae* infections can result in dormant liver stage parasites, which can relapse and cause malaria several months to years after the initial infection. In the case of *P. malariae*, long-lasting chronic infections occur, and, if left untreated, can persist asymptotically throughout the lifetime of the host.

The *Plasmodium* parasite has a complicated life cycle involving six distinct developmental forms: female and male gametes, zygote, ookinete, oocyst, and sporozoite, as well as fertilization and invasion of several different types of host tissue (GHOSH *et al.* 2000). A completed life cycle and a successful spread of the parasite require both the insect and human hosts (Figure 1.1).

Plasmodium reproduces sexually when inside the gut of the mosquito and must surpass sequential developmental bottlenecks along the way (Cycle C in Figure 1.1). The *Plasmodium* gametocyte (Step 8 in Figure 1.1), first encounters the mosquito midgut environment and displays proteins on its surface, such as Pbs21 found on *P. berghei*, that are used as targets by researchers for blocking development. A small percentage of gametocytes develop into ookinetes and not much is known about this developmental transition (Step 10).

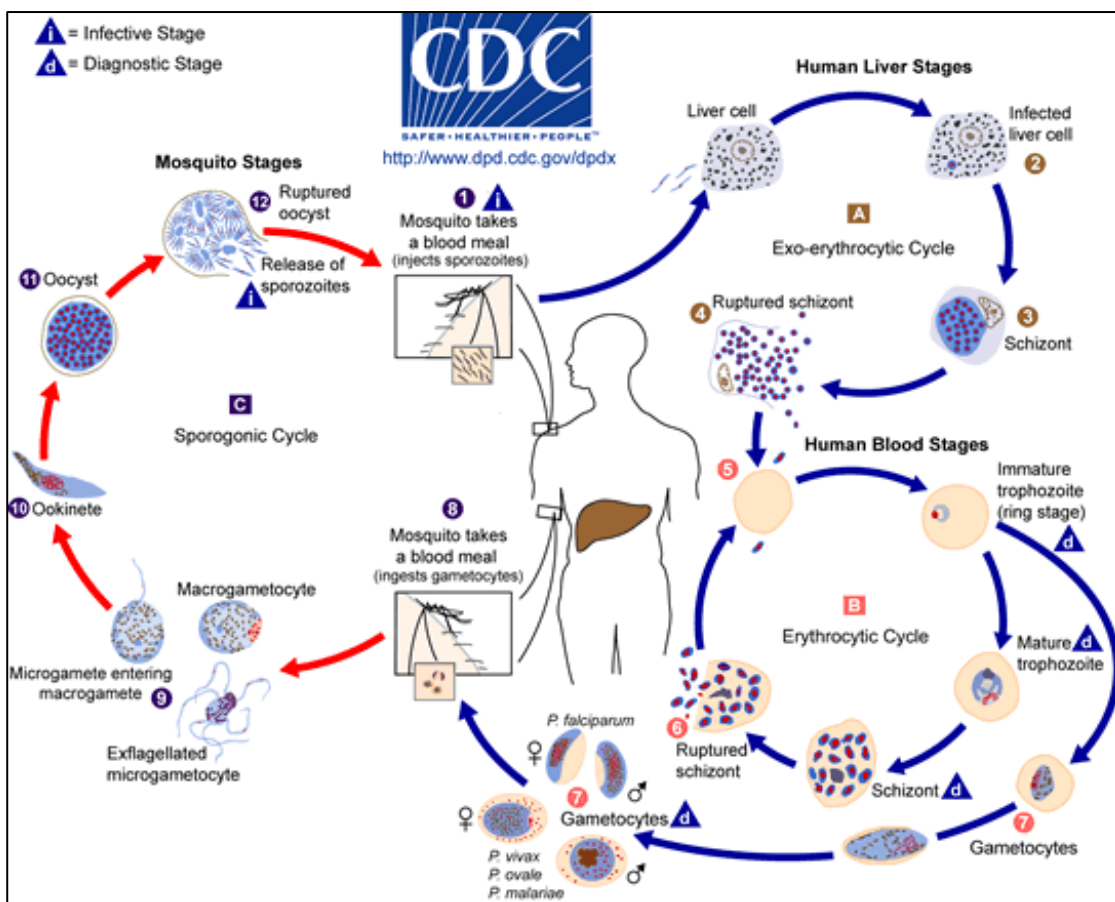


Figure 1.1: The life cycle of *Plasmodium*. *Plasmodium* needs the human and mosquito hosts for a successful life cycle. Stages A and B are the nonsexual stages of *Plasmodium* that occur in the human host. Upon transmission to a naïve human by an infected mosquito, *Plasmodium* will first infect liver cells and eventually invade the bloodstream. When a naïve mosquito takes a blood meal from the infected person, *Plasmodium* can be transmitted and will then undergo Stage C, which entails the sexual stages of development. These occur inside the mosquito gut and eventually the salivary glands. See text for more details. (Figure from Centers for Disease Control and Prevention website: http://www.cdc.gov/malaria/biology/life_cycle.htm)

Within 24 hours, the ookinete makes a difficult journey in traversing both the peritrophic matrix, a chitin-rich inner lining of the gut, and the midgut epithelium in order to implant between the midgut epithelium and basal lamina for further differentiation into oocysts (Step 11). The successful differentiation from ookinete to oocyst can take anywhere from 10-24 days depending on the *Plasmodium* species. Oocysts burst to release thousands of sporozoites that have the ability to invade numerous organs and cell types, but they will only travel through a second epithelial layer, to the salivary glands (Step 12).

From the salivary glands, the sporozoites will soon enter the secretory cavity and will remain there for the life of the mosquito. When the insect takes a blood meal the sporozoites are transferred to the mammalian host in the insect's saliva and will migrate to the liver cells to continue their maturation.

Liver-stage parasites are called schizonts and over the course of a few days they rapidly divide to form merozoites (Cycle A, Step 2). Merozoites are eventually released from the parenchymal cells of the liver into the bloodstream where they invade erythrocytes. Depending on the species, *Plasmodium* can remain dormant in the liver and cause a relapse in an infected individual by invading the bloodstream after weeks or even years. Once present in the blood cells, the parasites reproduce asexually into the trophozoite stage and go on to produce daughter merozoite-stage parasites that will invade more red blood cells (Cycle B). It is these blood stage parasites that cause the clinical manifestations of malaria.

1.3 The Manifestations of Malaria in Humans

Individuals with malaria can display symptoms that range from nonexistent to fatal. For this reason, malaria is classified as being either uncomplicated or complicated/severe. Fortunately, if malaria is diagnosed early and treated properly, it is a curable disease. With uncomplicated malaria, a patient may present any combination of flu-like symptoms ranging from nausea and vomiting to body aches and fever. Proper diagnosis comes only after observing a blood smear under a microscope for the presence of the parasite.

Complicated malaria involves organ failure and drastic changes in the patient's bloodstream and metabolism. This may include hemoglobinuria (hemoglobin in the urine), pulmonary edema (fluid buildup in the lungs), or cardiovascular collapse and shock. These extreme cases are seen in individuals with no immunity to malaria, especially those who reside in areas of low to no malarial transmission. Aggressive treatment and an urgent response is needed for complicated malaria to attempt to save the infected individual's life.

1.4 Past Treatments and Preventive Measures

Prevention and control of malaria thus far has relied primarily on either drug treatments or insect vector control. Plant-derived anti-malarial drugs such as quinine and artemisinin have been used for the past several decades to treat malaria. Quinine is currently only used for treating severe infections with *P. falciparum* due to its horrible side effects like temporary deafness, ringing in the ears, and dizziness. Artemisinin can be administered to patients orally or through intravenous infusion and will target the parasites present in the bloodstream. Artemisinin acts 10 times faster on *Plasmodium*

compared to quinine (WHITTY *et al.* 2006). By destroying the gametocytes in the bloodstream, transmission to a feeding mosquito is somewhat reduced. Over time, however, the effects of artemisinin are reduced because *Plasmodium* evolves resistance. In recent years, the drug chloroquine has become ineffective against resistant strains of *P. falciparum*, but it is still effective against *P. vivax*, *P. ovale*, and *P. malariae*. Chloroquine resistance in *P. falciparum* has been linked to mutations in genes that encode proteins involved in uptake to the parasite vacuole (LE BRAS and DURAND 2003).

Destruction of mosquito larval breeding grounds and the use of physical barriers like insecticide-treated bed nets in high-risk locations can help in minimizing mosquito populations and their interactions with humans. Nets impregnated with pyrethroid permethrin have an increased effectiveness because this insecticide produces a halo that extends beyond the surface of the net. However, the nets can be expensive, are not widely available, and have to be replenished with insecticide twice a year. It is estimated that less than 5% of nets in use are re-treated sufficiently.

Neither method of drug treatment or vector control is completely successful in inhibiting malaria transmission. Most importantly, both *Plasmodium* and *Anopheles* species have evolved resistances to many of the drugs and insecticides in use.

Destruction of natural habitats also comes as a price of insecticide use. As a result, there are a decreasing number of effective and inexpensive means of controlling this disease.

Additional measures must be developed to combat the spread of malaria. There are some estimates that the number of malaria cases will double in the next 20 years if no new preventative measures are put into place (BREMAN *et al.* 2001). To circumvent the inevitable resistance of the mosquito or parasite to any environmental controls and to

minimize the destruction of natural habitats, alternative genetic approaches have been initiated and appear promising.

1.5 Combating Insect-Vectored Diseases with Genetic Approaches

There have been numerous attempts to control malaria by genetically altering the mosquito in a way that would turn it into an ineffective malarial vector (ALPHEY *et al.* 2002; MOREIRA *et al.* 2002a). Using this strategy, the insect is modified to prevent the parasite from traveling from the gut to the salivary glands or from the insect to the human, resulting in a reduction of malarial transmission. The idea of genetic control of vector-borne diseases has origins nearly 40 years old, but it was after the advancements in molecular experiments in the germline of *Drosophila melanogaster* in the 1980s that a reevaluation of the idea emerged (ALPHEY *et al.* 2002; CURTIS 1968). A committee of the World Health Organization's Special Program for Research and Training in Tropical Diseases (WHO/TDR) members in Geneva in 1991 established three parameters to be met before field-testing of any genetically modified mosquitoes (GMMs) could begin. First, the genetic engineering tools for use with malaria vectors had to be established; second, effector genes that would block parasite transmission should be determined and characterized; and lastly, effective methods of dispersing and fixating these effector genes in a wild population must be realized. An effector gene can function by either interfering with parasite development or eradicating them completely inside their host. Added to these parameters is the fact that any modification cannot pose a significant fitness load such that the GMMs would not survive as well as wild mosquitoes in a natural setting.

The last decade has shown considerable progress in the genetic manipulations of anopheline mosquitoes. The *P. falciparum* and *An. gambiae* genomes were both completed in 2002 (GARDNER *et al.* 2002; HOLT *et al.* 2002) and effector genes whose products can hinder parasite development in the mosquito are continually being identified. Because the parasite has to cross two different epithelial layers, it has been suggested that specific interactions exist between *Plasmodium* surface proteins and different molecules on the surfaces of the epithelial layers. Blocking these interactions could result in effective reduction of malarial transmission.

Certain effector proteins have been shown to inhibit parasite development in the mosquito gut (ITO *et al.* 2002; MOREIRA *et al.* 2002b; MOREIRA *et al.* 2004). A dodecapeptide named SM1 (Salivary gland and Midgut peptide 1) was identified through a phage display library designed for phages that bind to both the midgut and salivary gland epithelia (GHOSH *et al.* 2001). *An. stephensi* carrying transgenic *E. coli* that both display SM1 on the surface of their outer membranes and secrete the peptide into the midgut lumen have been shown to reduce *P. berghei* invasion by 41% (RIEHLE *et al.* 2007). Until recently, the mechanism by which SM1 inhibits *Plasmodium* sporozoite invasion of the salivary glands was not known. It has now been shown that SM1 is a conformational analog to the *Plasmodium* protein TRAP (thrombospondin-related anonymous protein) and it competes with TRAP in binding to the salivary gland protein, saglin (GHOSH *et al.* 2009). Sporozoites of all *Plasmodium* species express TRAP and this protein is needed for sporozoite gliding and invasion of mosquito host cells (SULTAN *et al.* 1997). Ghosh *et al.* (2009) have shown that the SM1 peptide and the A-domain of

TRAP both bind to saglin, a protein expressed in the salivary gland distal lobes, which is precisely the site of sporozoite invasion (BRENNAN *et al.* 2000).

Another effector protein identified was a phospholipase isolated from both honeybee and snake venoms called PLA2 (ABRAHAM *et al.* 2005; MOREIRA *et al.* 2002b; ZIELER *et al.* 2001). PLA2 purified from the venom of the eastern diamondback rattlesnake (*Crotalus adamanteus*) was shown to inhibit oocyst formation in both avian and human malarial parasites (*P. gallinaceum* and *P. falciparum*, respectively) when administered to *Aedes aegypti*, *An. stephensi*, and *An. gambiae* mosquitoes following an infectious blood meal (ZIELER *et al.* 2001). Honeybee venom PLA2 was expressed by *An. stephensi* and detected in the midgut epithelia 8-24 h after a blood meal. Importantly, PLA2 was shown to inhibit *P. berghei* oocyst formation by 87% and greatly reduce transmission of the parasite to uninfected mice (MOREIRA *et al.* 2002b). *E. coli* expressing honeybee PLA2 inside the mosquito gut also caused a 23% reduction in *P. berghei* development (RIEHLE *et al.* 2007). It is hypothesized that phospholipases of this class block a specific receptor in the midgut epithelial lining and hamper the parasite's mobility through this layer.

Pbs21 is a 21 kDa protein found on the surface of the ookinete stage of the rodent malarial parasite *P. berghei*. Mouse monoclonal antibodies have been made against this protein and shown to prevent development from gametocytes to oocysts (YOSHIDA *et al.* 1999). Yoshida *et al.* constructed a gene for a single-chain antibody (scFv) by cloning the genes encoding the variable light and heavy chains of the antibody and assembling them into one open reading frame. This Anti-Pbs21 scFv was shown to bind to the surface of *P. berghei* ookinetes and blocked oocyst formation in the mosquito gut by 93%

(YOSHIDA *et al.* 1999). In place of genetically engineered anophelines, *E. coli* expressing the Anti-Pbs21 scFv were introduced into the gut of *An. stephensi* mosquitoes by membrane feeding. The overall number of infected mosquitoes and their oocyst densities were lowered when the mosquitoes were allowed to feed on *P. berghei*-infected mice (YOSHIDA *et al.* 2001). These experiments helped to support the idea of utilizing genetically engineered bacteria in the mosquito gut to fight *Plasmodium* development and transmission.

Another example of an anti-*Plasmodium* effector protein is the CEL-III lectin isolated from the body fluid of the sea cucumber, *Cucumaria echinata*, has hemolytic and cytotoxic activity against human and rat erythrocytes. This protein rapidly creates holes in the erythrocyte cell membrane (HATAKEYAMA *et al.* 1995). It is thought that CEL-III is part of the innate defense system of the sea cucumber. Yoshida *et al.* (2007) hypothesized that with CEL-III present in the mosquito gut during a blood meal, rapid hemolysis would occur and thus the gut environmental conditions would be altered in a way that would inhibit *Plasmodium* development. In addition, it was discovered that CEL-III could bind to ookinetes and thus prevent sporogonic development. Experiments revealed a severe impairment of the rodent malarial parasite *P. berghei* and a moderate blockage of the human parasite *P. falciparum* (YOSHIDA *et al.* 2007).

Experiments involving GMMs appear promising in the laboratory setting, but the fitness and fecundity of these insects will have to equal or exceed that of wild mosquito populations if they are to be fixed successfully in the wild. In five independent laboratory experiments, transgenic mosquitoes expressing PLA2 showed 77-99% inhibition of oocyst formation (MOREIRA *et al.* 2004). As compared to nontransgenic

mosquitoes, the PLA2 transgenics consistently showed lower numbers of sporozoites in salivary glands. Additionally, PLA2 transgenics showed a significantly lower fecundity and ingested 10-50% less blood than nontransgenics (MOREIRA *et al.* 2004). Through a series of cage experiments whereby transgenics and nontransgenics coexisted in the same enclosure and were maintained blindly over five generations, it has been shown that transgenics expressing SM1 showed no significant reduction in fitness as compared to nontransgenics (ABRAHAM *et al.* 2005; MOREIRA *et al.* 2004). Further studies on SM1-expressing mosquitoes fed on *P. berghei*-infected blood revealed that the transgenics actually had a greater fitness over nontransgenics. Over time the transgenics outcompeted nontransgenics and were more fit in terms of fecundity and mortality rates (MARRELLI *et al.* 2007). What these combined experiments suggest is that the influence on fitness of transgenics will be dependent upon the protein product of the particular effector gene analyzed (JACOBS-LORENA 2003).

Before transgenic mosquitoes or bacteria are released into the wild, it is imperative that attempts are made to understand the effects these organisms may have on native mosquito populations. One study used two geographically isolated *An. gambiae* strains (Mbita from western Kenya and Ifakara from Tanzania) to compare hybrid fitness to that of the founder populations (MENGE *et al.* 2005). This study was done in the laboratory with the intent of mimicking the introduction of an exotic transgenic anopheline strain to that of a wild population. These scenarios introduce new alleles into a population in addition to parasite-inhibiting genes. The authors measured traits found to be important in determining an insect vector's transmission capabilities: fecundity, body size, blood meal size, and adult longevity. It was observed that all traits showed

heterosis (increased fitness in the hybrid strain) over each of the founder populations up to the 20th generation. Results such as this suggest that transgenic strains introduced to the wild could become a nuisance by living longer and biting more frequently than indigenous mosquitoes. Also, greater hybrid fitness could lead to stabilizing selection that would make it difficult for refractory genes to become fixed in the population. The authors stress that anopheline strains used to carry transgenes should have very similar genetic makeup to native populations to curb or eliminate the likelihood of increased hybrid fitness in the wild.

A major obstacle in utilizing this technology is determining how to effectively establish the GMMs in the field (RIEHLE and JACOBS-LORENA 2005; RIEHLE *et al.* 2003). One method would be to significantly reduce the numbers of endemic mosquito populations (perhaps through the use of insecticides), and then release of the GMMs on a large scale. The main problem with using genetically modified mosquitoes is that generating enough mosquitoes to replace entire populations on a country or continent-wide scale is nearly impossible. Alternatively, through the use of transposable elements as a genetic drive mechanism, a transgenic population could replace a wild population. However, the effectiveness of this method is not known. Transposable elements could also pose a significant fitness load on the insects because of the frequency and location of transposition events resulting in the disruption of genomic organizations and increase in mutation rates. There is also the possibility that after a certain number of generations, the insects become refractory to transposition, perhaps due to selection of a repressor protein. This would be similar to the *P* (refractory) cytotype in *Drosophila melanogaster* where

the accumulation of a repressor protein inhibits transposition and slows the mobility of the *P* element (CRAIG 2002).

1.6 Paratransgenesis as a Means to Fight Diseases

There is an alternative to creating expensive and time-consuming genetically modified mosquitoes in the fight against *Plasmodium* transmission. The past few years have seen an increased interest in and application of paratransgenesis, which is the use of genetically modified bacteria to treat many different diseases and conditions. For example, paratransgenesis has been effective in a laboratory setting in inhibiting development of the parasite that causes Chagas disease, which is vectored to humans by blood-feeding insects (DURVASULA *et al.* 1997). Paratransgenesis involves genetically modifying a bacterial species that normally resides in the gut of an insect to produce functional molecules that will inhibit the development of the parasite vectored by the insect (ALPHEY *et al.* 2002; BEARD *et al.* 1998; BEXTINE *et al.* 2004). In addition to inhibiting insect-borne diseases, there are also examples of this approach being used to reduce HIV infectivity in mammalian cells (CHANG *et al.* 2003; RAO *et al.* 2005).

Paratransgenic approaches to limiting the spread of the insect-vectored Chagas disease have yielded encouraging results. Chagas disease is a parasitic disease in humans that affects people living in most of Central and South America. It is caused by the parasite *Trypanosoma cruzi*, which is vectored by the blood-feeding insect, *Rhodnius prolixus*, commonly called the kissing bug. These insects colonize thatched homes that are common structures in tropical regions and can transmit *Trypanosoma* to humans who come in contact with insect feces. Contact can occur when the insect is taking a blood meal from a human and deposits a fecal droplet on the skin or even when a person

unknowingly touches their eyes, mouth, or open cuts, with hands contaminated with *Rhodnius* fecal matter. Between 12-15 million cases of Chagas are estimated to occur annually, and approximately 10-30% of cases will progress into a chronic and life-threatening illness of the cardiac and gastrointestinal systems. Chagas disease causes about 50,000 deaths each year in tropical America, making it a leading cause of mortality in these regions (BEARD *et al.* 2002; BEARD *et al.* 2001).

The first example of the disruption of insect vector capabilities via the genetic alteration of endosymbiotic bacteria came in 1997 when an endosymbiont of *Rhodnius prolixus*, the actinomycete *Rhodococcus rhodnii*, was engineered to express cecropin A, a peptide that is lethal to *Trypanosoma* (DURVASULA *et al.* 1997). The recombinant bacterium was fed to *Rhodnius* insects and a significant reduction in *Trypanosoma* survival inside the insect was observed. *R. rhodnii* was a prime choice for genetic modification because it resides in close proximity to *Trypanosoma* when inside the insect's gut. Cecropin A is a small insect immune peptide that creates holes in the membrane of *T. cruzi*. Other optimistic observations in this study included the absence of toxicity to insect tissues due to the presence of the recombinant symbionts and the spread of these symbionts through a colony due to insect coprophagic (the eating of feces) behaviors. These researchers have even developed a paste called CRUZIGARD that contains GM *R. rhodnii* and mimics *Rhodnius* fecal matter in composition, which newly hatched insects will eat and ingest these bacteria early in their development. This gives the GM bacteria ample time to colonize the insect (BEARD *et al.* 2002).

Pierce's disease, a devastating disease that affects grapevines in Southern California and other parts of the world, which is caused by the pathogenic bacterium,

Xylella fastidiosa, is another target for paratransgenic therapies. The glassy-winged sharpshooter (GWSS), *Homalodisca coagulata*, is the vector for this bacterium.

X. fastidiosa is transferred to the xylem of grape plants while the GWSS feeds on the leaves. The plants succumb to this pathogen and as a result, the wineries in these regions can suffer great economic loss. An artificial feeding system can introduce the non-pathogenic bacterium *Alcaligenes xylooxidans denitrificans* to the insect through a liquid medium (BEXTINE *et al.* 2004). This bacterium was chosen from among other culturable bacterial species because it was found to be associated with the xylem of plants. It also resides in close proximity to *X. fastidiosa* in the foregut of the insect, making it an optimum candidate for expression of anti-*Xylella* effector molecules. Successful delivery and colonization of *A. xylooxidans denitrificans* in the foregut of the GWSS suggests that a paratransgenic approach to combating Pierce's disease is within reach.

Paratransgenesis is also being used to combat diseases that do not need an insect vector to infect humans, but the concept of engineering symbiotic bacterial species to express inhibitory molecules remains the same. Human Immunodeficiency Virus, or HIV, is a global health epidemic that demands immediate preventative measures. Women are more at risk for infection due to a greater efficiency of transmission from male to female (ROYCE *et al.* 1997). Alternative methods of virus inhibition include engineering bacteria normally found in the gastrointestinal or cervicovaginal mucosa of mammals to express anti-HIV inhibitor proteins (CHANG *et al.* 2003; RAO *et al.* 2005).

Chang *et al.* (2003) engineered a natural human vaginal isolate, *Lactobacillus jensenii*, to express a secreted form of the HIV-binding protein, CD4. *In vitro* results show that the CD4 produced by these cells was able to inhibit HIV-1 entry into target

cells. The authors believe that these engineered lactobacilli may survive on vaginal mucosa *in vivo* for days to weeks at a time, allowing for intermittent administrations of the bacteria to females at high risk for HIV infection.

Rao et al. (2005) engineered a strain of *E. coli* called Nissle 1917 to secrete an anti-HIV fusion peptide. The inhibitor peptide was a 52 amino acid sequence derived from the C-terminal region of gp41, a transmembrane subunit of the HIV envelope that works in the mechanism of membrane fusion. The authors observed peptide secretion at inhibitory concentrations and also observed the ability of the bacterium to nonpathogenically colonize the gastrointestinal and cervicovaginal tracts of mice for periods of weeks to months. This genetically modified bacterium proves promising as a potential treatment for HIV infection in humans.

A proof-of-concept experiment for paratransgenic malaria control involving the rodent malaria parasite *P. berghei*, showed its development in *An. stephensi* mosquitoes was inhibited 90-95% in the presence of *E. coli* expressing an Anti-Pbs21 scFv (YOSHIDA *et al.* 2001). The Anti-Pbs21 scFv binds to a 21 kDa surface protein on *P. berghei* ookinetes, which occur in the mosquito midgut (YOSHIDA *et al.* 1999). Interestingly, the control for this experiment, administering *E. coli* expressing a non-specific scFv to *P. berghei*-infected mosquitoes, also resulted in a significant reduction of *P. berghei* development (YOSHIDA *et al.* 2001). There may be an indirect negative effect upon *Plasmodium* due to the presence of the *E. coli* or the presence of an scFv (regardless of its specificity). However, because *E. coli* are not naturally found in mosquito guts, and therefore not adapted to this environment, the use of this bacterium may not be efficient in a field setting.

There are many advantages to using paratransgenesis to counter insect vector competence. If the transgene-carrying bacterium colonizes the same areas inside the insect as the targeted parasite, the transgene product affecting the parasite is in close proximity and its effect is increased. Introduction of bacteria into insect populations may be much easier than transgenes alone. Most bacteria can be transformed with DNA quite easily and immense quantities of bacteria can be produced with minimal costs. Also, multiple effector genes can be added to one species simultaneously and if one gene should lose effect over time, it should be relatively easy to introduce an alternate effector gene.

The most important factor in using a paratransgenic approach to fight a disease is the choice of bacterium. Optimally, the chosen bacterial species should be able to be cultured and genetically manipulated in the laboratory setting. This species must also be adapted to or indigenous to the environment that the disease of interest occupies. It is important that the GM bacterium is able to survive, undergo normal cellular operations, and hopefully even replicate in the environment, otherwise, the GM population will have to be continually replenished. Replenishment of the GM bacteria may be a difficult task depending on the disease and its host environment.

1.7 Bacterial Candidates for Paratransgenesis

Certain characteristics of the obligate, intracellular, Gram-negative bacteria called *Wolbachia* suggest it could be a powerful drive mechanism for spreading effector genes through an insect population. This is because these bacteria inhabit the reproductive systems of arthropods and are maternally inherited. *Wolbachia* also cause cytoplasmic incompatibility, which means that females who do not harbor *Wolbachia* will not produce

offspring (SINKINS 2004). It has been observed that *Wolbachia* spread through a *Drosophila simulans* population in California at the rate of 100 km per year (TURELLI and HOFFMANN 1991).

The devastating disease dengue fever, vectored by the mosquito *Aedes aegypti*, may one day be curbed by infecting *Aedes* mosquitoes with strains of *Wolbachia* that shorten the life-span of the mosquito (MCMENIMAN *et al.* 2009). Parasites like *Plasmodium* and the dengue flavivirus require a particular length of time inside their insect host for proper development before they are infective and can be transmitted to a human. This length of time is about 2 weeks for both dengue fever and malaria. For this reason, older mosquitoes are of a higher epidemiological importance. The *Wolbachia* wMelPop strain was previously shown to limit the life span of adult *Drosophila melanogaster* and McMeniman *et al.* tested to see if a similar effect would be observed in *Aedes* mosquitoes (MCMENIMAN *et al.* 2009; MIN and BENZER 1997). These authors report *Aedes* populations were reduced by half with the presence of the life-shortening *Wolbachia* wMelPop strain. With further studies under seminatural conditions, the validity of using *Wolbachia* to reduce the life span, and thus the infectivity, of *Aedes* mosquitoes that carry dengue fever can be assessed.

In the case of malaria, as a cruel twist of fate, *Wolbachia* is not an option for mosquito control. *Wolbachia* has not been observed in anopheline mosquitoes. Additionally, *Wolbachia* colonizes the reproductive structures of insects, which is not in proximity to the gut where *Plasmodium* undergoes development. Therefore, transgenic *Wolbachia* would not be effective in a mosquito control scenario that targets *Plasmodium*, as mentioned above.

Engineering a bacterium that naturally colonizes the gut of the mosquito to express effective anti-malarial genes would greatly increase the likelihood of inhibiting the *Plasmodium* parasite. *Plasmodium* is most vulnerable within the insect midgut. Mosquitoes can ingest 10^3 gametocytes with a blood meal, but less than 10 ookinetes will successfully traverse the midgut epithelium and form oocysts. Parasite numbers increase dramatically only after each oocyst produces thousands of sporozoites (GHOSH *et al.* 2000; RIEHLE and JACOBS-LORENA 2005). Therefore, targeting ookinetes could interfere dramatically with malarial transmission. Additionally, another advantage of this paratransgenic approach is that, unlike with the use of insecticides, the condition of the natural habitat is left intact and the biological niche is filled with mosquitoes that are incapable of transmitting the malarial parasite.

Adult, blood-fed, female *An. stephensi* mosquito colonies in the Jacobs-Lorena laboratory at Johns Hopkins University served as the source for an initial screen for bacterial species that are both well adapted to the gut environment and that are able to be cultured in the lab. Isolates were characterized for Gram staining and cellular morphology, 16S ribosomal DNA sequencing, and resistance to any drug markers (if any). From this screen, Gram-negative *Enterobacter spp.*, *Klebsiella sp.*, *Serratia sp.*, and gram-positive *Bacillus spp.* were identified (RIEHLE *et al.* 2007).

Pantoea (= Enterobacter) agglomerans was chosen as a candidate for downstream applications presented in this dissertation due to its persistence in the mosquito gut environment, its ability to be cultured in the lab, and because it does not pose a pathogenic threat to healthy humans. *P. agglomerans* can also be successively passed through female mosquito guts to obtain strains that will persist in the gut for

weeks at a time and who increase their numbers following a blood meal (RIEHLE *et al.* 2007). Previous studies have shown that *P. agglomerans* is commonly isolated from wild mosquito populations and this suggests the insects can readily obtain these microorganisms from the environment (RIEHLE and JACOBS-LORENA 2005).

P. agglomerans is a rod-shaped Gram-negative γ -proteobacterium that is isolated from diverse environments, including soil, plants, insects, and humans. Classification of this species has primarily relied upon sequence comparison of 16S ribosomal DNA and a handful of protein coding genes (HAUBEN *et al.* 1998; NAUM *et al.* 2008; YOUNG and PARK 2007). Because of the limited phylogenetic analysis, *P. agglomerans* was previously designated as *Erwinia herbicola* or *Enterobacter agglomerans* until identification methods became discriminating enough to discern differences in these species (GAVINI 1989). Classification of this species is quite difficult because the genus is so diverse, and for this reason careful phylogenetic analysis of as many strains as possible at one time is important for proper identification. Fluorescence amplified fragment length polymorphism (FAFLP) has been used to type different strains (BRADY *et al.* 2007). This procedure involves a single PCR reaction that amplified DNA to a concentration that is detectable and sizeable by the laser of an automated sequencer. This technique is somewhat limiting because the source DNA must be in a pure state and that is not always possible when analyzing environmental isolates (MORTIMER and ARNOLD 2001). Multilocus sequence analysis (MLSA) has also been used to classify *P. agglomerans* strains (BRADY *et al.* 2008; DELETOILE *et al.* 2008). This technique allows for genotyping of a more diverse group of prokaryotes and uses several single-

copy protein coding genes in its comparison. As these sophisticated genotyping methods are more routinely used, proper identification of *P. agglomerans* strains will continue.

Another newly identified candidate for malarial paratransgenesis is the α -proteobacterium, *Asaia bogorensis* (YAMADA et al. 2000). This species can be cultured in acidic medium, genetically manipulated, and is associated with the larval and adults stages of anopheline mosquitoes. Additionally, *A. bogorensis* has been reported to be vertically transmitted to offspring, which makes it possible to spread transgenic forms across mosquito populations (FAVIA et al. 2008). A *gfp*-expressing strain of *A. bogorensis* was shown to colonize the female gut and salivary glands of *An. stephensi* (FAVIA et al. 2007). These are the exact locations of *Plasmodium* development and if *A. bogorensis* can be genetically engineered to produce anti-*Plasmodium* effector proteins in these locations, a reduction in transmission could be realized.

1.8 Protein Secretion Systems in Gram-negative Bacteria

One caveat of a paratransgenic approach to fight insect-vectored diseases is the need to deliver the anti-*Plasmodium* effector product to the exterior of the bacterial cell. The effector product must be in proximity to its target (i.e. *Plasmodium* in the mosquito gut environment) in order to have a negative effect upon its development. By fusing a DNA sequence encoding a bacterial secretion signal to the effector gene, the gene product could be exported from the cell upon translation. Currently, there are seven recognized and characterized secretion system types across all bacterial species, but this section will focus primarily on the systems utilized in Gram-negative species like the malarial paratransgenesis candidate, *P. agglomerans*.

Gram-negative bacteria have a large and varied arsenal of protein secretion systems at their disposal. Bacteria export proteins to the extracellular environment in order to participate in processes like pathogenicity, cell movement or cell-cell communication, and each secreted protein is linked to a particular secretion apparatus. Export from the Gram-negative cell involves transport across both the inner membrane (IM) and the outer membrane (OM) and the different pathways involved in traversing these membranes can be divided into two categories: Sec-dependent and Sec-independent (KOSTAKIOTI *et al.* 2005). Sec refers to the secretion (Sec) translocase, a collection of proteins that comprise the IM secretion apparatus common to all Sec-dependent secretion systems (MORI and ITO 2001).

Sec-dependent secretion pathways utilize the Sec translocase for transport across the IM and different machinery for transport across the OM. Examples of Sec-dependent secretion systems include Type II secretion systems (T2SS), Type V or autotransporter secretion systems (T5SS), two-partner secretion (TPS), and the chaperone/usher secretion system (CU) (KOSTAKIOTI *et al.* 2005). The Type IV secretion system (T4SS) is sometimes Sec dependent but is mostly considered Sec-independent (DING *et al.* 2003).

The T2SS exports proteins in a two-step process. Proteins are first sent across the IM via the Sec translocase to the periplasmic space, the site of most extracellular protein folding. Transport across the OM involves several proteins, only one of which is an integral membrane protein. This protein, Protein D is a secretin and is thought to form the translocation channel (KOSTAKIOTI *et al.* 2005). Studies on examples of Type II secretion in *Aeromonas* species and of Pula (pullulanase) secretion in *Klebsiella oxytoca* have led to the suggestion that the IM and OM constituents make contact with one

another transiently in the periplasmic space to form a secretion apparatus that spans the entire cell envelope (LETELLIER *et al.* 1997; POSSOT *et al.* 1997). This is believed to be the case because transport across the OM relies on the proton motive force created in the IM. Proteins to be secreted by T2SS contain N-terminal secretion signal sequences that direct them to Sec-dependent translocation across the IM to the periplasmic space. The signal sequence is removed and the protein is properly folded and further modifications like subunit assembly can be performed. Finally, the protein is secreted across the OM by the T2SS secretion apparatus (SANDKVIST 2001). Examples of T2SS include PulaA (pullulanase) secretion in *Klebsiella oxytoca* and type IV pili formation in *Pseudomonas aeruginosa* (NOUWEN *et al.* 1999; PEABODY *et al.* 2003). The Out system identified in *Erwinia chrysanthemi* is a T2SS that secretes plant cell wall-degrading enzymes, including PelB (LINDEBERG and COLLMER 1992). This same signal sequence isolated from *Erwinia carotovora* has been shown to secrete heterologous proteins in *E. coli* (LINDEBERG and COLLMER 1992; THIE *et al.* 2008; WINTER *et al.* 1994).

Autotransporter or T5SS is one of the most widely distributed secretion systems among Gram-negative bacteria and is mostly used to deliver virulence factors that play a role in pathogenesis. Characteristic of the T5SS is the fact that protein substrates can mediate their own transport across the OM. Proteins targeted for secretion using this pathway contain the Sec-dependent N-terminal signal sequence that aids in transport across the IM as well as a C-terminal β -domain that will insert itself into the OM, directing export of the internal passenger domain (α -domain) to the cell exterior (JACOB-DUBUISSON *et al.* 2004). Representatives of virulence factors secreted by autotransporter

include the IgA1 protease of *Neisseria gonorrhoeae* and the Hap protease of *Haemophilus influenzae* (KOSTAKIOTI *et al.* 2005).

Two-partner secretion (TPS) is very similar to autotransporter secretion and is found across many different Gram-negative bacterial species. The main distinguishing component of the TPS system is its accessory protein involved in translocation of the exoprotein across the OM (JACOB-DUBUISSON *et al.* 2004). Proteins secreted by this system are referred to as TpsA proteins and contain both the Sec-dependent N-terminal signal sequence and a 110-residue N-terminal “TPS domain” that is needed for secretion across the OM through an integral OM channel-forming protein called TpsB. The TpsA protein can remain noncovalently bound to the cell surface or released to the extracellular environment (JACOB-DUBUISSON *et al.* 2004; KOSTAKIOTI *et al.* 2005). TPS secretion is commonly used for the secretion of large virulent proteins including the ShlA cytotoxin of *Serratia marcescens* and the HMW1A adhesin of *Haemophilus influenzae* (JACOB-DUBUISSON *et al.* 2004; ST GEME and GRASS 1998).

CU secretion is reserved for the assembly of a superfamily of virulence-associated surface structures. These include structures involved in adhesion to host cells like P pili and type I pili of uropathogenic *E. coli* as well as fimbriae and capsule structures. CU secretion relies on two proteins that work cooperatively, a periplasmic chaperone and an outer membrane usher protein (KOSTAKIOTI *et al.* 2005). As with all other Sec-dependent substrates, nascent polypeptides are directed to the periplasmic space where they must interact with a chaperone protein. This chaperone aids in proper folding while preventing subunit-subunit interactions from occurring prematurely. Chaperone-subunit complexes interact with the OM usher, and chaperone-subunit complexes are gradually exchanged

for subunit-subunit complexes, which aids in fiber assembly and secretion to the cell surface (SAUER *et al.* 1999).

Sec-independent pathways can export a protein directly from the cytoplasm to the extracellular space without a stop in the periplasmic space because the translocation channel spans both the IM and OM. Type I (T1SS) and Type III (T3SS) secretion systems are categorized as Sec-independent. As mentioned previously, Type IV secretion systems (T4SS) are usually considered to act in a Sec-independent fashion.

T1SS are used by Gram-negative bacteria to secrete proteins such as pore-forming toxins, proteases, and lipases directly from the cytoplasm to the extracellular space without periplasmic intermediates (GENTSCHEV *et al.* 2002; KOSTAKIOTI *et al.* 2005). The mechanism of T1SS is best characterized by the α -hemolysin (HlyA) system of pathogenic *E. coli*. Proteins targeted for secretion by a T1SS contain a noncleavable C-terminal signal sequence that sends them to the secretion apparatus. In the case of HlyA, the C-terminus of the protein acts as the secretion signal. Three proteins form a secretion apparatus that spans the IM and OM. The IM translocase is comprised of a member of the ATP-binding cassette (ABC) superfamily (HlyB) and a member of the membrane fusion protein family (HlyD). An integral OM protein forms a β -barrel with a central hydrophilic core (TolC). In the HlyA system in *E. coli* the last 60 amino acids of HlyA are needed for interaction with the IM translocase, and after this interaction takes place TolC is recruited via HlyD to form a continuous channel through the cell envelope (KORONAKIS *et al.* 1992; WANDERSMAN and DELEPELAIRE 1990).

T3SS utilize an “injectisome” which is comprised of an envelope-spanning channel and a needle-like projection from the bacterial cell that can make contact with

and inject virulence factors directly into host cells (CORNELIS 2006). These systems were first characterized for the Yop proteins of *Yersinia pestis* and are genetically, structurally, and functionally related to bacterial flagella. The main component of the bacterial flagellum, FliC/flagellin, is secreted through the flagellar export apparatus in an identical fashion to a T3SS (YOUNG *et al.* 1999).

The T3SS system consists of three types of proteins, proteins that make up the needle-like projection, the secreted proteins, and regulatory proteins that control the expression of structural and secreted proteins. Approximately 20 proteins make up the needle structure and allow for secretion upon contact with the target host cell. For this reason, T3SS is often referred to as a “contact-dependent pathway” (KOSTAKIOTI *et al.* 2005). It is not exactly certain what signals are involved in targeting a protein for T3SS-mediated secretion. Evidence exists for three different secretion signals. N-terminal amino acid signal sequences target proteins to be secreted via a T3SS. Additionally, the signal can also be located in the 5' of the mRNA molecule for the secreted protein. Finally, it is possible that a chaperone protein can bind the substrate and direct its secretion (ALDRIDGE and HUGHES 2001). An mRNA-mediated secretion would result in a coupling of translation and protein secretion and it is thought that ribosomes in the process of translating secreted proteins are placed in close proximity with the cytoplasmic surface of the secretion apparatus. An N-terminal signal sequence would trigger secretion similar to Sec-dependent secretion pathways, however, the signal is not cleaved from the mature protein, as it does not make a stop in the periplasmic space. Finally, chaperone-mediated secretion requires the chaperone to communicate with the secretion

apparatus. Any substrate that cannot bind to a chaperone will not be secreted (ALDRIDGE and HUGHES 2001; CHENG and SCHNEEWIND 2000).

It is interesting that different species would evolve the use of different signals to secrete the same protein. More than one type of secretion signal for the FliC/Flagellin protein has been characterized in *E. coli* and *Salmonella* species. In *E. coli*, the 5' UTR (untranslated region) of the *fliC* mRNA functions as a protein secretion signal, while the secretion signal in *Salmonella* was identified as 22 residues in the N-terminus of the FliC protein (GAL *et al.* 2006; MAJANDER *et al.* 2005; VEGH *et al.* 2006). Additional and somewhat conflicting evidence for FliC secretion in *E. coli* identifies the first 183 residues in the FliC N-terminus as the secretion signal (KUWAJIMA *et al.* 1989; MAJANDER *et al.* 2005).

T4SS are thought to have evolved from bacterial conjugation machinery and have the capacity to secrete both proteins and single-stranded-DNA-protein complexes directly into host cells (CASCALES and CHRISTIE 2003). Secretion is carried out through the use of *trans*-envelope structures that span the IM and OM and end in a pilus structure at the surface of the bacterial cell. They are generally considered Sec-independent, but exceptions like the Sec-dependent secretion of the *B. pertussis* PT toxin fall under the category of T4SS (CASCALES and CHRISTIE 2003; KOSTAKIOTI *et al.* 2005). Most of what is known about T4SS comes from the T-DNA transfer system of *Agrobacterium tumefaciens*. This Gram-negative soil bacterium causes crown gall tumor disease in plants. It transfers a portion of a plasmid (T-DNA) to the plant cell nucleus, where it is incorporated into the plant chromosome. Protein localization experiments confirm that several of the *Agrobacterium* VirB proteins colocalize with both the inner and outer cell

membranes and contain periplasmic domains, which suggests that these proteins form an envelope-spanning channel (CHRISTIE 1997). In addition to the *trans*-envelope structure, the secretion system contains a conjugative T pilus structure for delivery of the effector to the host cell. There are two models for secretion using this pilus structure. The “channel model” suggests that substrates traverse the OM through the lumen of a pilus-like structure. The “piston model” proposes a retractile pilus that extends from the cell and pushes the substrate through the OM pore (CASCALES and CHRISTIE 2003)

Additional Gram-negative secretion systems are continually being discovered. The Type VI secretion system (T6SS) was identified after researchers reviewed the organization of gene clusters that encode T4SS-like systems. The original name for these clusters was IAHP (IcmF-associated-homologous-proteins) because they contained a gene for an IcmF-like component. IcmF (intracellular multiplication protein F) is an IM protein that is a known participant in T4SSs (FILLOUX *et al.* 2008). However, all of the other genes in the IAHP gene cluster were not homologous to other T4SS genes, leading researchers to think that a novel secretion system gene cluster had been found (CASCALES 2008; FILLOUX *et al.* 2008).

The speculative structural model of a T6SS shows the proteins involved forming a channel that spans both the IM and OM. A protein called VgrG is associated with the lumen side of the OM and is thought to play a dual role in delivery of the secreted protein. VgrG may create a puncture in a host cell for injection of the secreted protein and VgrG may also behave like an autotransporter and carry the secreted protein through the puncturing device into the host cell (FILLOUX *et al.* 2008).

T6SS have been identified in species like *E. coli*, *Salmonella enterica* subspecies, *Vibrio cholerae*, and usually play a role in virulence and invasion of eukaryotic cells. *Burkholderia mallei*, which has four T6SS gene clusters, suggests a redundancy in virulence or specific systems for use in particular hosts (FILLOUX *et al.* 2008).

The Gram-positive bacterium, *Mycobacterium* has revealed a seventh secretion system (T7SS). Originally called the ESX-1 system, this T7SS is involved in secretion of virulence factors and conjugation, depending on the *Mycobacterium* strain (ABDALLAH *et al.* 2007). In particular, *Mycobacterium tuberculosis* encodes five T7SSs. There is no sequence homology between the components of a T7SS and other secretion systems, which suggests that this is a novel system. The T7SS is mechanistically unique because it appears that all of the secreted proteins are co-dependent upon one another for secretion. Structural data is lacking, but it is thought that a multiprotein complex spans the *Mycobacterium* cell envelope similar to type I-IV secretion systems.

T7SS appear to only be utilized in Gram-positive bacteria and have been found in pathogenic and non-pathogenic species from *Corynebacterium* and *Staphylococcus* to *Streptomyces* spp. Lots of questions are still being answered regarding the structure and function of this latest secretion system (ABDALLAH *et al.* 2007).

Bacteria also use alternate trafficking pathways for shuttling proteins to the IM as well as the extracellular space that act independent of the Sec machinery. *E. coli* uses a SRP (signal recognition particle) pathway that targets specific proteins to the IM, which is homologous to the eukaryotic SRP pathway that translocates proteins to the endoplasmic reticulum (DE GIER *et al.* 1997; VALENT 2001). The prokaryotic SRP is a ribonucleoprotein complex comprised of a 4.5S RNA and a protein called P48. P48

binds to the hydrophobic signal sequence in a membrane-bound protein in a co-translational fashion and targets the nascent protein to a translocon in the IM through its interaction with a receptor protein FtsY (VALENT 2001). And finally, the Tat pathway (twin-arginine translocation) is a translocon for sending folded proteins across the IM and it recognizes a twin-arginine motif at the end of the N-terminal region of the protein (WEINER *et al.* 1998). Integral membrane proteins TatA, TatB, and TatC use the proton motive force across the IM to acquire the energy needed to translocate the protein. It is TatB or TatC that recognizes the twin-arginine motif in the Tat substrate (PALMER *et al.* 2005). It was initially thought that the Tat pathway was used mainly to export folded proteins that bind to cofactors in the cytoplasm, but it has been shown to play a broader role in delivering virulence factors in several bacterial pathogens (BERKS *et al.* 2003; PALMER *et al.* 2005). Because of its role in virulence, and because there are no homologues in mammalian cells, the Tat pathway could be a target for antimicrobial compounds (DE BUCK *et al.* 2008).

1.9 Applying Secretion Systems to Deliver Heterologous Proteins

Several protein secretion systems have been adapted to deliver heterologous gene products from the bacterial cell. Frequently, these secretion systems have been shown to work in *E. coli*, but other species like *Salmonella* spp. have also proven to be successful at secreting heterologous proteins (GEORGIU and SEGATORI 2005; MERGULHAO *et al.* 2005). Genetically engineering a bacterium to produce heterologous proteins is much easier than modifying a eukaryote and it is also more economically viable to produce large volumes of bacterial cells rather than generate numerous modified eukaryotic organisms. This section will focus on Sec-dependent (T2SS) and Sec-independent

(T1SS, T3SS, and SRP) signals that have been used to secrete heterologous proteins from Gram-negative bacteria.

One of the long-standing technologies for producing large volumes of protein in a bacterium like *E. coli* is antibody phage display (CLACKSON *et al.* 1991). Recombinant antibodies are valuable tools in biomedical research and therapy. These are often secreted post-translationally via a Sec-dependent pathway because the antibodies are only folded correctly when passed through the periplasm. An appropriate leader peptide is added to the N-terminus of the antibody in order for Sec-dependent secretion to occur.

The most frequently used leader peptide is PelB, from the pectate lyase protein of *Erwinia carotovora* (LINDEBERG and COLLMER 1992). It is appealing and convenient to use this leader peptide because the secretion signal is cleaved from the recombinant protein in the periplasm and is not part of the finished product. This leader is used in the human antibody Tomlinson I+J library developed in Greg Winter's laboratory at the MRC Laboratory of Molecular Biology and the MRC Centre for Protein Engineering (Cambridge, UK), as well as in industrial scale production of human proteins with medical importance in *E. coli* (DE WILDT *et al.* 2000; SLETTA *et al.* 2004; SLETTA *et al.* 2007). Commonly, the IPTG-inducible *lac* promoter drives the expression of the PelB-fused passenger protein.

Another frequently used leader peptide is OmpA, from outer membrane protein A of *E. coli* (SMITH *et al.* 2007). This protein is a major component of the outer membrane and is also secreted in a Sec-dependent fashion. OmpA was tested alongside PelB in the industrial scale production of human proteins, and was used as the secretion signal in

phage display constructs (BARBAS et al. 2001; HAIDARIS et al. 2001; SLETTA et al. 2004; SLETTA et al. 2007).

The *E. coli* hemolysin system, the prototype of a T1SS, has long been used as a delivery system for exogenous proteins (FERNANDEZ et al. 2000; HOLLAND et al. 1990b; TZSCHASCHEL et al. 1996). This system provides genes for the IM channel proteins, HlyB and HlyD, and a multiple cloning site to accept the exogenous gene upstream of the coding sequence for the C-terminus (approximately the last 50 amino acids or 23 kDa) of HlyA. HlyA is the secreted protein in the hemolysin system. The OM-associated protein, TolC, is usually provided *in trans* in commonly-tested Gram-negative species (SPRENG et al. 1999). *E. coli* has been engineered to secrete at least 400 proteins, from both prokaryotic and eukaryotic origins, using this pathway and there appears to be no limitations on the size or origin of the secreted protein (GENTSCHEV et al. 2002).

A Shiga-like toxin was successfully secreted from an attenuated *Salmonella typhimurium* vaccine strain using the *E. coli* hemolysin system (TZSCHASCHEL et al. 1996). In this example, one plasmid provides the *hlyB* and *hlyD* genes as well as the cloning site for fusion to the 3'-end of *hlyA* (*'hlyA*). The IPTG-inducible *lac* promoter drives expression of these genes. The authors note that this plasmid is replicated using a low-copy number *ori* to eliminate the degradation of the secreted protein as well as to maintain the stability of the recombinant *S. typhimurium* clones, but they make no attempt to prove that the recombinant toxin is functional.

A two-plasmid hemolysin system was shown to work efficiently in *E. coli* in the secretion of functional single-chain antibodies (scFvs) (FERNANDEZ et al. 2000; TZSCHASCHEL et al. 1996). One plasmid carries *hlyB* and *hlyD* and the second carries a

multiple cloning site and the *hlyA* gene sequence. The IPTG-inducible *lac* promoter is also used in these constructs. The reasons the authors give for using the hemolysin system for production of scFvs over the more widely used N-terminal signal peptide and Sec-dependent secretion methods are the possible toxic accumulation of scFvs in the periplasmic space leading to the lysis of cells, as well as clogging of the Sec pathway or greater outer membrane permeability. The hemolysin system seems like an ideal alternative to these problems because the product is exported directly to the cell exterior without a periplasmic intermediate. Fernandez *et al.* (2000) compared the activity of scFvs with affinity for transmissible gastroenteritis virus that were secreted by either the hemolysin system or exported in a Sec-dependent fashion in an ELISA assay. The authors found that the scFvs had the same binding activity. It is important to note that active proteins were secreted using the hemolysin system even though there was no stop in the periplasm, where it is thought that most proteins are correctly folded.

The *E. coli* flagellar export pathway, which closely resembles a prototypical T3SS, has been shown to secrete heterologous proteins in *E. coli* (MAJANDER *et al.* 2005). The *E. coli* flagellar apparatus is a multisubunit protein complex that spans both the OM and IM of the Gram-negative bacterium with a whip-like flagellum that extends from the surface of the cell that can propel the cell through its aqueous environment. The exterior flagellum is made of 14 different proteins and can contain up to 20,000 subunits of the protein FliC (flagellin) (CHEVANCE and HUGHES 2008). In order to evade host defense mechanisms, FliC proteins can display antigenic variation by altering the sequence of its central domain, which is exposed on the filament surface. The N-terminal and C-

terminal regions are highly conserved so that the FliC monomers can recognize one another and polymerize to form the flagellum (YONEKURA *et al.* 2003).

In the case of *E. coli* FliC secretion, two different signals have been characterized. The N-terminus (the first 183 amino acids of the protein) of FliC and the *fliC* 5' UTR were characterized as secretion signals and were used for the secretion of several different heterologous proteins in *E. coli* (MAJANDER *et al.* 2005). The latter is an important finding because the resultant secreted product is not fused to an endogenous protein.

Majander *et al.* (2005) observed the secretion of three different heterologous proteins (ranging in size from 115-434 residues) under the control of the *fliC* 5' UTR as well as an N-terminal FliC fusion with the eukaryotic green fluorescent protein (238 residues). These proteins were secreted in a $\Delta fliC$ and $\Delta fliD$ (encodes FliD, the flagellum capping protein) *E. coli* strain. The yield of secreted product was reported at the mg/L level, which greatly exceeds yields seen with the hemolysin system ($\mu\text{g/L}$) (CHOI and LEE 2004).

The SRP (signal recognition pathway) mode of secretion in Gram-negative bacteria has not been tested extensively for secretion of heterologous proteins. In this pathway, proteins are secreted co-translationally to the IM and then transported across the IM to the periplasmic space (VALENT 2001). The N-terminal leader peptides from three different proteins secreted by the SRP pathway (DsbA (periplasmic protein disulphide isomerase I), TorT (regulatory protein of TorCAD), and TolB (a periplasmic protein involved in the TonB-independent uptake of group A colicins) were fused to the coding

sequence for a soluble antibody fragment and shown to successfully secrete an active antibody (THIE *et al.* 2008).

Finally, it is the common perception that laboratory bacterial strains, like *E. coli* K12, do not secrete proteins when grown under routine conditions. However, it was reported that *E. coli* K12 secretes a small (10.8 kDa) protein called YebF when grown in the laboratory (ZHANG *et al.* 2006). The function of YebF and the pathway that secretes it were unknown upon its discovery in the growth medium, nonetheless the authors saw this protein as a potential tool for secretion of exogenous proteins. N-terminal amino acid sequence analysis revealed a 21 residue Sec leader sequence and cellular fractionation revealed localization of YebF in the periplasm, which altogether suggest a Sec-dependent secretion pathway (ZHANG *et al.* 2006).

Three different passenger proteins were fused to the C-terminal end of YebF and were successfully secreted in an active state. The passengers: human interleukin-2 (15 kDa), α -amylase (48 kDa), and alkaline phosphatase (94 kDa), differ in their sizes and hydrophobicities, which suggests that YebF-coupled secretion is adaptable to different proteins (ZHANG *et al.* 2006). Recently, Athena Environmental Science, Inc. (Baltimore, MD) patented a YebF secretion kit for use in *E. coli* indicating the potentially broad application of this protein in the production of heterologous proteins (<http://www.athenaes.com>).

CHAPTER 2

IDENTIFYING AND APPLYING NATIVE SECRETED PROTEINS FOR USE IN SECRETION IN *PANTOEA AGGLOMERANS*

ABSTRACT

Pantoea agglomerans is a candidate for the paratransgenic control of malaria because it is routinely isolated from the gut of the anopheline mosquito that transmits this disease. Malaria is a serious and often fatal disease that affects millions of people around the world. The genetic manipulation of bacteria to fight diseases is a reasonable alternative to drug therapies and the genetic modification of eukaryotic disease carriers. One caveat of this approach is the need to secrete the anti-*Plasmodium* effector gene product from the bacterial cell to the mosquito gut environment. It was hypothesized that the coding sequence of a native secreted protein could serve as the secretion signal. A search for native secreted proteins was performed using MALDI-TOF to identify 2D-PAGE spots isolated from spent growth medium. The corresponding genes were isolated using a genome walking PCR technique and identified based on homology to sequenced bacterial species. The best candidate from this search was FliC/Flagellin, which is the major component of the bacterial flagellum and is known to be secreted from the cell in a Type III secretion system-like manner. Constructs were made using the *P. agglomerans* *fliC* 5'UTR and the *fliC* 5'UTR and complete *fliC* ORF each as the upstream signal for

secretion of a test protein (an anti-BSA single-chain antibody). These constructs were based on previously published results showing the *fliC* 5' UTR and the N-terminus of FliC function as sufficient signals for heterologous secretion in *E. coli*. Attempts to delete the endogenous *fliC* gene from the *P. agglomerans* chromosome using Lambda Red recombination failed multiple times, which could suggest that this deletion is fatal to the cell or that the recombination system is non-functional in this species. Overall, these constructs were not successful in the expression or secretion of the test protein. Reasons may include the presence of the endogenous *fliC* gene, or it may simply be that the flagellar export system in *P. agglomerans* is not suitable for the secretion of exogenous proteins.

INTRODUCTION

The number of annual deaths as a result of the anopheline mosquito-transmitted disease malaria is estimated at 1-3 million people. This is more than the number of deaths as a result of HIV and tuberculosis combined (BREMEN *et al.* 2001). The majority of these cases are young African children. Preventive measures have changed little over the years: the main strategies are still vector eradication through the use of insecticides and drugs to combat the parasite in infected people. Both the mosquitoes and the *Plasmodium* parasite have evolved resistances to the chemicals, clearly showing that current strategies are not completely successful (LE BRAS and DURAND 2003). New approaches to limiting the spread of this disease are desperately needed.

An alternate approach is to render the mosquito an ineffective vector by genetically engineering it to express anti-*Plasmodium* proteins. Several effector proteins that inhibit *Plasmodium* development in the mosquito have already been characterized

(ITO *et al.* 2002; MOREIRA *et al.* 2002b; MOREIRA *et al.* 2004; YOSHIDA *et al.* 2007).

Translating these effectors into practical control reagents, however, has been challenging. Anti-*Plasmodium* transgenic mosquitoes have been engineered and shown to hinder *Plasmodium* development, but it was a costly endeavor and the tools to drive such genes into entire wild populations are lacking (GHOSH *et al.* 2001; MOREIRA *et al.* 2004; YOSHIDA *et al.* 2001; YOSHIDA *et al.* 2007).

To circumvent the hurdles involved with engineering an insect to express anti-*Plasmodium* effector proteins, a bacterial symbiont of mosquitoes could be engineered to manufacture the proteins. A prime candidate would be a bacterium well adapted to the mosquito gut environment, where the gametocyte and ookinete developmental stages of the *Plasmodium* life cycle take place. These forms of *Plasmodium* are key targets for inhibition of development. The transgenic bacterium could be introduced into a mosquito population and express the anti-*Plasmodium* genes while circulating in the gut environment.

This paratransgenic approach, engineering a bacterial symbiont to produce antagonistic proteins, could prove to be an efficient tool against the spread of malaria to humans (RIEHLE and JACOBS-LORENA 2005). This approach has already been shown to be successful in a laboratory setting in reducing the spread of Chagas disease. This disease is caused by *Trypanosoma cruzi*, which is vectored by *Rhodnius prolixus* (“kissing bugs”) in parts of Central and South America. Because of the coprophagic (eating of feces) behavior of these insects, researchers have been able to introduce genetically modified bacteria to the insects in a paste that mimics adult *R. prolixus* fecal matter (BEARD *et al.* 2002; BEARD *et al.* 2001).

In this study, the Gram-negative γ -proteobacterium *P. agglomerans*, a species routinely found in anopheline mosquitoes, is a paratransgenic candidate (RIEHLE and JACOBS-LORENA 2005; RIEHLE *et al.* 2007; STRAIF *et al.* 1998). *P. agglomerans* is established in the gut microbiota of tephritid flies and locusts and observations suggest a “moderately mutualistic” relationship (DILLON *et al.* 2000; DILLON *et al.* 2002; PELOQUIN *et al.* 2002). *P. agglomerans* has been isolated from diverse environments from plant material to soil to food preparations, and is not a pathogen in healthy humans (LOIRET *et al.* 2004; SCHEIRLINCK *et al.* 2008). Currently there is no completed genome sequence for this bacterium, but it is related to *E. coli* and *Salmonella* spp. and can be cultured and manipulated in the laboratory using common molecular biology cloning techniques.

In order to properly target *Plasmodium* development in the mosquito gut, the effector protein must be secreted from the bacterial cell. I hypothesized that a native *P. agglomerans* secreted protein could act as the secretion signal in an anti-*Plasmodium* construct. The proteins secreted by *P. agglomerans* were precipitated from spent growth medium by methods similar to those previously published and identified using 2D-PAGE and MALDI-TOF analysis (CHITLARU *et al.* 2006; KAZEMI-POUR *et al.* 2004). Because of the lack of genome sequence data for *P. agglomerans*, analysis of the protein profile in spent growth medium and the subsequent isolation of the corresponding genes was dependent upon sequence homology to related species whose genomic information is present in the GenBank database. A genome walking PCR using degenerate primers was utilized to isolate the secreted protein genes from the *P. agglomerans* chromosome (GUO and XIONG 2006).

In this study, the best *P. agglomerans* secretion candidate (*fliC* which encodes FliC/Flagellin, the main component of the bacterial flagellum) was cloned into vectors along with the gene for a test protein (a single chain antibody (scFv) with anti-BSA activity). Because it was previously reported that the 5' UTR (untranslated region) of *E. coli fliC* was a sufficient signal for secretion, the *P. agglomerans fliC* 5' UTR was also included in some of the test constructs (MAJANDER *et al.* 2005).

P. agglomerans was unable to express or secrete the anti-BSA scFv test protein under the control of the *fliC* 5' UTR or as a *fliC-scFv* fusion. It should be noted that these experiments were done in the presence of the endogenous *fliC* gene because attempts to delete endogenous *fliC* from the *P. agglomerans* chromosome were unsuccessful, which may indicate that this gene is necessary for viability, or that the Lambda Red recombination machinery used to delete this gene was nonfunctional in *Pantoea* species (KATASHKINA *et al.* 2009) In conclusion, *fliC* was not able to be used as a secretion signal and now further attempts are being made to find an appropriate secretion signal for use in this paratransgenic candidate.

MATERIALS AND METHODS

Media and Antibiotics *E. coli* and *P. agglomerans* were grown in Luria-Bertani broth or agar (LB). *S. cerevisiae* cells were grown on YPD agar plates (20 g tryptone, 10 g yeast extract, 20 g dextrose, 20 g agar per liter) or Minimal Drop-Out Media excluding uracil (2% glucose) when selecting for yeast recombinants (Sigma Y1501). Final concentrations of antibiotics (Sigma or Fisher Scientific) were as follows: ampicillin (Ap), 150 µg/ml; apramycin (Apr) 80 µg/ml; chloramphenicol (Cam), 30 µg/ml; gentamycin (Gent), 30 µg/ml; streptomycin sulfate (Str), 100 µg/ml; tetracycline (Tc), 15

µg/ml. Antibiotic stocks were filter-sterilized through a 0.2 µM filter syringe and stored at -20°C.

Storage of bacterial cells, DNA stocks. All bacterial cell stocks were stored in LB broth +7% DMSO in 1 ml screw-cap tubes at -80°C. All DNA stocks (plasmid or genomic) were kept on ice during usage and stored at -20°C when not in use.

Growth of cells in preparation for bulk protein precipitation. A freshly streaked *P. agglomerans* colony was used to inoculate 50 ml of LB broth and was grown overnight (16h) at 250 rpm, 30°C. Thirty milliliters of this culture was used to inoculate 3 L of LB and was grown to late-*log* phase (approximately 8 h, see Appendix 2 for growth curve of wild-type *P. agglomerans*). The cells were removed by two sequential centrifugation steps at 10,000 x g, 4°C for 15 min. The spent medium was carefully transferred to a clean flask and trichloroacetic acid was added to a final concentration of 10%. The flask was set on a slow rocking platform and incubated overnight at 4°C. After centrifugation (12,000 rpm, 4°C, 30 min), the protein precipitate was washed with copious amounts of very cold 95% ethanol and allowed to dry for several minutes on the benchtop. The precipitate was collected directly in approximately 200 µl of isoelectric focusing buffer (8M urea, 2% CHAPS, 50 mM dithiothreitol, 0.2% (w/v) Bio-Lyte® 3/10 ampholytes, trace bromophenol blue (BioRad)) using vigorous pipetting and stored at -20°C (CHITLARU *et al.* 2006).

Isoelectric focusing and 2D-PAGE. Isoelectric focusing was performed using the BioRad Protean® IEF cell. Briefly, 11 cm pH 3-10 strips were passively rehydrated with 185 µl of precipitated protein (volume = full capacity of the strip) overnight at room temperature. The focusing program contained three steps: 250 V for 20 min with a

linear ramp; 8,000 V for 2.5 h with a linear ramp; and finally 8,000 V for 20,000 V-hr with a rapid ramp. Total running time was approximately 5.3 hr or 30,000 V-hr. The strip was equilibrated for 2D-PAGE with a 10 min incubation in 6 M urea, 2% SDS, 0.375 M Tris-HCl (pH 8.8), 20% glycerol, 2% dithiothreitol and a 10 min incubation in 6 M urea, 2% SDS, 0.375 M Tris-HCl (pH 8.8), 20% glycerol and then placed across the top of a 10% acrylamide Tris-Glycine gel and subjected to 200 V for approximately 2 h. If IEF strips were not used immediately for 2D-PAGE they were stored at -80°C.

MALDI-TOF analysis. 2D-PAGE gels were visualized by silver stain (Pierce 24612, SilverSNAP Stain Kit II). Spots were cored from the gel using sterile wide-bore pipette tips and stored in deionized H₂O at 4°C until they could be shipped overnight to the MALDI-TOF facility. In order to increase the amount of protein for each sample to be analyzed, parallel gels were run and identical spots were combined into one tube (Appendix 3). MALDI-TOF analysis was done at the Taplin MS Facility at Harvard Medical School (<http://gygi.med.harvard.edu/taplin/>). This facility performs in-gel trypsin digestion of the gel spots, microcapillary LC/MS/MS analysis, and searching of the NCBI Enterobacteriaceae protein database. When results were obtained from a sample they were reported in a web-based format that includes the identity of the spot along with the sequence of the peptide fragments and where they align to a known protein in the NCBI Enterobacteriaceae protein database (Figure 2.3).

Genome walking PCR. Degenerate PCR primers based on the potential codons for the peptide fragments identified with MALDI-TOF were used to amplify a small portion of the corresponding gene from the *P. agglomerans* chromosome. The PCR products were cloned using TOPO technology (Invitrogen), sequenced (Big Dye, ABI 3101), and

additional primers were designed for a nested “genome walking” PCR as described in Guo and Xiong (2006). A gene-specific primer was paired with each of four arbitrary primers in a PCR reaction to acquire flanking sequences (Table 2.3). The arbitrary primers, designed by Guo and Xiang (2006) were designed to have enough degeneracy in them as to allow the primer to anneal somewhere in the flanking DNA of the sequence of interest. The initial PCR reaction was carried out in a 100 μ l volume containing 20 ng chromosomal DNA as template, 1 μ M each of the specific primer and arbitrary primer, 250 μ M dNTPs, *Taq* polymerase, and 1X buffer (10 mM Tris-HCl (pH 8.8), 50 mM KCl, 1.5 mM MgCl₂). The cycling conditions were comprised of two stages: In stage 1, an initial 95°C for 3 min was followed by denaturation at 95°C for 45 sec, annealing at a range from 60°C to 47.5°C for 45 sec with a stepwise decreasing gradient of 0.5°C per cycle, and elongation at 72°C for 2 min for 25 cycles. Stage 2 consisted of 95°C for 45 sec, 50°C for 45 sec, and 72°C for 2 min for 35 cycles. One microliter from the first PCR was used in a second 50 μ l volume reaction with 1 μ M of a specific nested primer and of the same arbitrary primer. Cycling conditions were 95°C for 3 min followed by 35 cycles of 95°C for 45 min, 50°C for 45 sec, and 72°C for 2 min and a final 7 min elongation step at 72°C. PCR products were resolved on a 1% agarose gel, gel-purified (Zymoclean), and cloned into pCR®2.1-TOPO using the TOPO system (Invitrogen). Approximately 300 ng of the recombinant TOPO plasmid and 0.1 μ g of the M13 Forward (-20) or M13 Reverse primer were used in a fluorescent dye-labeled dideoxynucleotide sequencing reaction (BigDye, ABI 3101). This process was repeated, generating new specific primers as more gene sequence was identified until the characterized sequence

encompassed the entire ORF and intergenic sequence up to and including the neighboring genes.

“Recombineering” to delete the *P. agglomerans fliC* gene from the chromosome. In order to remove the *fliC* ORF, 60-mer oligos (Ent deltaFliC Left and Ent deltaFliC Right, Table 2.3) with 40 bp of homology to the DNA sequence directly flanking the ORF were designed to amplify the Kan^R cassette from pKD4 (DATSENKO and WANNER 2000). After amplification of the cassette using *Pfx* polymerase (Invitrogen), residual plasmid template was digested by adding *DpnI* directly to the PCR reaction and incubating at 37°C for 1 h. The cassette was agarose gel-purified (Zymoclean) and eluted in 20 µl of deionized H₂O.

P. agglomerans was electro-transformed with either 10 ng of pIJ790, pSIM5, pSIM7, or pSIM9 (Table 2.2) at 1.8 kV in a 0.1 mm cuvette. Electro-transformed cells were grown in a shaking incubator (250 rpm) for 1 h at 30°C in 1 ml of LB broth and then plated on selective LB agar and incubated overnight at 30°C. Before using these cells in a recombineering experiment, PCR was performed to ensure presence of the plasmid (a portion of the *cat* gene was amplified from the plasmid). Fifty milliliter cultures of *P. agglomerans* cells expressing a Lambda Red plasmid were grown at 30°C in LB broth containing selective antibiotics and 10 mM L-arabinose until an OD₆₀₀ of 0.6 was reached. The cells were removed from the media and washed twice in 50 ml of very cold 10% glycerol. The remaining pellet of cells was resuspended in the drop of remaining 10% glycerol and 50 µl was transferred to a 0.1 mm cuvette and transformed at 1.8 kV with 100-300 ng of Kan^R cassette. After incubating the cells in a shaking

incubator (250 rpm) for 1 h at 30°C, different amounts were plated on LB agar plates containing kanamycin and incubated 24-48 h at 30°C.

P. agglomerans that grew on LB Kan plates were picked and assessed for successful *fliC* deletion. This was done using the different combinations of *fliC* and Kan^R primers that are listed in Table 2.3 and shown in Figure 2.7.

Plasmid construction. The plasmids in this study were constructed using a yeast gap repair method (SHANKS *et al.* 2006). Briefly, the vector to be modified was digested with restriction enzymes and treated with calf intestinal phosphatase according to manufacturer's directions (New England Biolabs). DNA inserts were amplified with oligos that contained 40 bp of homology to the digested vector (up to 2000 bases from the digestion site). Agarose gel-purified vector (~20-200 ng) and inserts (50-500 ng), along with herring sperm DNA (100 µg) were co-transformed into 8-10 large colonies of *S. cerevisiae* INVSc-1 (Invitrogen) cells, harvested directly from a YPD plate and resuspended in 500 µl of PLATE solution (50% PEG, 100 mM LiAc, 10 mM Tris, pH 7.4, 1 mM EDTA). The cells were incubated on the benchtop for ~2 d. After heat shock at 42°C for 30 min, the yeast cells were resuspended in 150 µl of deionized H₂O and plated on uracil drop-out medium (6.7 g/L yeast nitrogen base minus amino acids, 1.92 g/L yeast synthetic drop-out media supplement without uracil, 2% glucose (w/v) and 20 g/L bacteriological agar) and incubated at 30°C for up to 2 d. Total yeast DNA was purified from the colonies using the "Yeast Smash and Grab DNA Miniprep" protocol (ROSE *et al.* 1990). Fifty nanograms of total yeast DNA was transformed into an appropriate strain of *E. coli* and bacterial clones were verified for the resultant plasmid by restriction enzyme digestion and DNA sequencing.

The yeast-replicating vector, pDB14, was built as follows. pACYC184 was digested with *XmnI* for 1.5 h followed by treatment with calf intestinal phosphatase for 1 h at 37°C. The 2 μ m *ori* and *URA3* gene from the yeast cloning vector pMQ64 were amplified as a single PCR product with 60-mer oligos that contained 40 bp of homology to the site of insertion in pACYC184 (CHANG and COHEN 1978; SHANKS *et al.* 2006).

To make pDB19, pDB14 was digested with *HindIII* and *EagI* and the inserts (*fliC* 5' UTR from *P. agglomerans* and the anti-BSA *scFv* ORF from pIT2-*scFv*) were amplified with 60-mer oligos that contained 40 bp of homology that targeted the amplicons to recombine with pDB14 near the restriction sites. pDB19 allows for expression of the anti-BSA *scFv* under the control of the *P. agglomerans fliC* promoter.

Similarly, pDB20 was made using pDB14/*HindIII-EagI* as the vector and the *fliC* 5' UTR and *fliC* ORF (as one amplicon from *P. agglomerans* genomic DNA), and the anti-BSA *scFv* gene from pIT2-*scFv* as the inserts. pDB20 allows for expression of the fusion protein FliC-anti-BSA *scFv* under the control of the *P. agglomerans fliC* promoter.

pDB22 was made by digesting pDB19 within the *fliC* 5' UTR using *EcoNI* and replacing the UTR with the P_{tac} promoter from pMALTM-cRI (New England Biolabs) using yeast gap repair. With this plasmid, the expression of anti-BSA *scFv* was under the control of the P_{tac} promoter and acted as a negative control for protein secretion.

pDB24 was made by digesting pDB14 with *HindIII-EagI* and recombining in the *P. agglomerans fliC* 5' UTR with an *AscI* site at the 3' end, and the 6His and myc epitopes and stop codon from pIT2-*scFv*. With pDB24, the unique *AscI* site allows for the cloning of genes downstream of the *P. agglomerans fliC* promoter. Then, pDB27 was constructed by digesting pDB24 with *AscI* and recombining in the *malE* and *scFv*

ORFs. pDB27 allows for the expression of the fusion protein MalE-anti-BSA scFv under the control of the *P. agglomerans fliC* promoter.

Finally, the *fliC* 5'UTR in pDB27 was replaced by digesting the vector with *EcoNI* and recombining the P_{tac} promoter from pMALTM-cRI to make pDB28. With pDB28, expression of the MalE-anti-BSA scFv fusion protein under the control of the P_{tac} promoter serves as a negative control for secretion.

Plasmid maps (Figure 2.6) were made using the free program XPlasMap available from <http://www.iayork.com/XPlasMap/>.

Protein preparations and Western Blot analysis on *P. agglomerans* pellet and spent growth medium samples. *P. agglomerans* was electro-transformed with 10 ng of a *fliC* secretion construct (1.8 kV, 0.1 mm cuvette) and incubated in a shaking incubator for 1 h at 30°C before plating on selective LB agar and incubating overnight at 30°C. A fresh colony was used to inoculate a 5 ml overnight culture in selective LB broth. Pellet and spent growth medium samples were harvested from the overnight cultures and separated using sodium dodecyl sulfate-polyacrylamide electrophoresis (SDS-PAGE).

Specifically, 100 µl of overnight culture were centrifuged to pellet the cells (10,000 rpm, 30 sec). Seventy-five microliters of the supernatant was transferred to a new tube containing 25 µl of 3X Laemmli sample buffer (BioRad 161-0737). The remaining pellet was resuspended in 100 µl of 3X Laemmli sample buffer and all of the samples were boiled for 10 min before resolving on a 4% stacking and 10% separating acrylamide gel. Proteins were transferred to polyvinylidene difluoride membrane (PVDF) in a standard Western Blot transfer apparatus (BioRad) with 10% methanol in the transfer buffer (95V, 350 mA, 75 min). The membranes were blocked in Tris-

buffered saline (1X TBS, 10 mM Tris, 150 mM NaCl, pH 7.4) containing 1% (w/v) bovine serum albumin (BSA) for 3 h at room temperature. For immunodetection of myc-tagged proteins, membranes were incubated overnight at 4°C with α -myc Ab (1 μ g/10 ml or 1:10,000; Invitrogen 46-0603) diluted in the blocking buffer. After four 15 min washings in 1X TBS/0.05% Tween-20, the membranes were incubated in stabilized goat anti-mouse HRP-conjugated secondary antibody (0.01 μ g/100 ml or 1:100,000; Pierce 1858413) for 1 h at room temperature. The washing steps were repeated and the bound antibody-HRP conjugate was detected using a chemiluminescent reaction (SuperSignal West Femto Maximum Sensitivity Substrate; Pierce 34095) and autoradiographic film.

Nickel column purification of 6XHis-tagged proteins from *P. agglomerans* cell

lysates and clarified overnight supernatants. To concentrate any 6XHis-tagged anti-BSA antibodies present in the *P. agglomerans* cells or overnight supernatants, a small-scale nickel column purification was performed using a kit from Pierce (B-PER[®] 6XHis Fusion Protein Spin Purification Kit 78300) and the eluates were assayed by either Western blot or ELISA as described.

Enzyme-linked immunosorbent assays (ELISAs) using spent growth medium from

induced *E. coli* or *P. agglomerans*. The activity of the secreted anti-BSA scFv was tested using an ELISA assay. For each sample analyzed, duplicate MaxiSorp wells (Nunc) were coated with 100 μ l of bovine serum albumin (2 mg/ml) along with two negative control wells, one coated with 100 μ l of 1X PBS and one empty well. The plate was stored at 4°C overnight (approx. 12 h). Each well was washed three times with 200 μ l of 1X PBS before blocking for 2 h at room temperature with 2% dry milk in 1X PBS. The wells were washed three times with 200 μ l of 1X PBS and 100 μ l of clarified spent

growth medium from an overnight culture was added to each well and left to incubate for 1 h at room temperature. The wells were then washed eight times with 200 μ l of 1X PBS/0.1% Tween-20. A α -myc-HRP antibody diluted in blocking reagent (1:2000, Roche 11-814-15-0001) was added to the wells and left to incubate for 1 h at room temperature. A second set of eight washes was performed before the α -myc-HRP antibody was detected using 50 μ l/well of the chromogenic substrate 1-Step Ultra TMB (3,3',5,5'-tetramethyl benzidine, Pierce 34028). When a sufficient blue-colored signal was reached (anywhere from 10-45 min), the reaction was stopped with the addition of 50 μ l of 2 M H₂SO₄ and the absorbance was read at 450 nm with the background subtracted at 650 nm in a BioRad 3550 plate reader.

RESULTS

2D-PAGE and MALDI-TOF analysis of P. agglomerans spent growth medium.

Large-scale cultures (3 L) of *P. agglomerans* were grown to late-*log* phase (OD₆₀₀ ~ 0.8) and the cells were removed by centrifugation. Any secreted proteins present in the medium at that time were precipitated using trichloroacetic acid and samples were resolved by 2D-PAGE and visualized using silver stain. Figure 2.1 shows a representative silver-stained acrylamide gel with the secreted proteins present in a late-*log* culture. In order to have enough material per spot for MALDI-TOF, parallel gels were run and identical spots were combined to increase the amount of protein (Appendix 3). In total, sixteen spots were abundant enough for subsequent analysis.

The identity of the highlighted spots (Figure 2.1) was determined by MALDI-TOF analysis and peptide sequence comparison to *E. coli* and other related species whose peptide sequence data was present in the GenBank Enterobacteriaceae protein database.

Table 2.4 shows the match for the sixteen spots tested. Six of the spots did not have a significant match to any sequences present in the database. The remaining spots that were identified matched to *E. coli* proteins, such as OmpA (outer membrane protein), periplasmic proteins, the intracellular protein Ssb (single-stranded DNA-binding protein), and various intracellular enzymes. FliC/Flagellin was chosen as the prime secreted protein gene candidate. Figures 2.2 and 2.3 show the FliC/Flagellin MALDI-TOF results from the Taplin Biological Mass Spec Facility. FlgL and Ssb were also chosen as secondary candidates and the information regarding them is located in Appendix 1.

Isolating secreted protein genes based on peptide sequence.

The *fliC* gene was isolated from the *P. agglomerans* chromosome using a genome walking PCR technique (GUO and XIONG 2006). A schematic explaining this technique is given in Figure 2.5. Figure 2.4A shows the *P. agglomerans* FliC peptide fragments that were identified and their relative placement in the *E. coli* FliC sequence (shown in green), as well as the site of degenerate PCR primer design (red arrows). The primers were based on Gram-negative codon usage tables and allowed for as much degeneracy as possible to ensure a successful amplification of the corresponding DNA sequence (SAMBROOK and RUSSELL 2001). The sequence of these primers and the series of universal “walking primers” are given in Table 2.3. PCR products ranging in size from 200-600 bp were cloned using TOPO technology (Invitrogen). The DNA sequences from these PCR products were used to generate a *P. agglomerans fliC* contig from which additional specific primers could be designed and sequential rounds of genome walking PCR could be performed. This procedure was repeated until the entire *P. agglomerans fliC* gene and the flanking DNA up to and including some of the neighboring genes was

isolated and sequenced (Figure 2.4B). The *P. agglomerans fliC* gene is 1,287 bp long and is neighbored by the *fliD* and *fliB* genes. There are 336 bases of intergenic DNA between the 5' of *fliC* and the 5' of the divergently transcribed *fliD*. The intergenic region between the 3' of *fliC* and the 5' of *fliB* is 160 bp long. The sequence was deposited as GenBank accession number GQ150763.

Attempts to delete the P. agglomerans fliC gene.

The *fliC* gene was unable to be deleted from the *P. agglomerans* chromosome. Trials were performed using four different Lambda Red recombination plasmids in the hope that one of the host range origins of replication available would be compatible with this species. When pIJ790 was the source of the Lambda Red genes (*exo*, *bet*, and *gam*), no *P. agglomerans*/pIJ790 colonies grew post-recombination with the Kan^R cassette. This was attempted twice, with the second trial using a higher concentration of cassette. Again, no colonies grew after this second attempt.

When the alternate Lambda Red plasmid pSIM5 or pSIM7 was used, there were also no colonies recovered. When pSIM9 was used, which has 5-8 copies per cell and a broad host range *ori*, 3 colonies were recovered. Each of these colonies were tested for the deletion of *fliC* using primers that amplify a portion of *fliC*, a portion of the Kan^R gene, and a fusion between the 5' UTR of *fliC* and the Kan^R gene (Figure 2.7). These results unambiguously show that the cassette recombined into another region of the chromosome because a portion of the *fliC* gene was amplified along with a portion of the Kan^R gene. No 5' UTR-Kan^R product was seen in any of the colonies tested. This inability to recover *P. agglomerans* after recombineering is similar to the results seen in

Pantoea ananatis, where the Lambda Red system was shown to be toxic due to the simultaneous expression of the *exo*, *bet*, and *gam* genes (KATASHKINA *et al.* 2009)

Testing for fliC-mediated secretion of a test protein in P. agglomerans.

P. agglomerans transformed with the various *fliC* secretion constructs were grown for approximately 16 h in LB broth containing antibiotics and tested for the presence of the scFv or MalE-scFv or FliC-scFv fusions, depending on the construct, in the bacterial pellet and/or clarified overnight supernatant by Western blot analysis. Additionally, clarified overnight supernatant was tested for the presence of an active scFv antibody against BSA in an ELISA assay. Presence of the proteins in the supernatant would indicate *fliC*-mediated secretion.

The Western blot results can be summarized as follows. Cells transformed with pDB19 (5' UTR-*scFv*), pDB20 (5' UTR-*fliC-scFv*), or pDB27 (5' UTR-*malE-scFv*) (Figure 2.4) failed to show any protein in the pellet or supernatant samples by Western blot analysis followed by detection with α -myc antibodies (data not shown, blank film). A nickel column 6His purification of overnight cell lysates and clarified overnight supernatants followed by Western blot analysis and detection with α -myc antibodies also failed to show any proteins present in either sample (data not shown, blank film). The constructs for negative control of secretion, pDB22 (P_{tac} -*scFv*) and pDB28 (P_{tac} -*malE-scFv*), also failed to show expression of the proteins in the overnight pellet samples (data not shown, blank film).

DISCUSSION

Malaria is a very serious disease that claims the lives of over 1 million people every year. New strategies for curbing the spread of this disease from mosquitoes to

humans are desperately needed. Proposed here is a paratransgenic approach whereby a bacterial symbiont of the mosquito vector is engineered to express anti-*Plasmodium* gene products. To be effective, the product must be secreted from the bacterial cell so steps to find a native secreted protein for use as the secretion signal were taken.

Secreted protein profile of *P. agglomerans*

P. agglomerans secreted sixteen proteins that were detectable with silver stain and reproducible across multiple protein preparations and gels while growing to late-*log* phase in standard laboratory culture conditions. It was recommended by the Taplin Biological Mass Spec Facility that identical silver-stained spots be combined across multiple gels in order to increase the amount of protein available for MALDI-TOF. And, if a spot (despite its abundance) was not reproduced across more than one gel, it was considered an artifact and not chosen for analysis. Overall, there was not much variety in the types of proteins being secreted and in fact, multiple spots matched to the same protein in more than one instance. MALDI-TOF analysis was successful in identifying 10 of the spots based on their homology to *E. coli*. The remaining spots may not have been abundant enough for peptide sequencing or perhaps they simply had no significant match to any protein sequence in the NCBI Enterobacteriaceae protein database. These proteins may be unique to *P. agglomerans* and could be identified with additional protein identification techniques.

In the collection of identified proteins, several matched to intracellular or periplasmic proteins. This could be due to misidentification of the peptide fragments or simply due to cell lysis during cell growth, or the collection of cells for protein precipitation from the spent growth medium. Spots A and I (Figure 2.1) matched to a D-

ribose binding periplasmic precursor and a D-ribose binding periplasmic protein, respectively (RbsB). Noting the similar migration of these spots in the gel, it is fair to say that these spots may be isoforms of the same protein. As its name implies, this protein localizes to the periplasm and is involved in sugar transport through the membrane and acts as the primary chemoreceptor for chemotaxis (GROARKE *et al.* 1983).

Spots B and K both matched to OmpA (outer membrane protein A). These two spots are very close to one another and migrated to the same molecular weight on the gel, which would suggest covalent modifications of OmpA resulting in different IEF values. OmpA is a major component of the outer membrane of *E. coli* and has several roles including stabilization of mating aggregates in conjugation and acting as a receptor for bacteriophages (KOEJNIK 1999).

Spot H matched to the intracellular protein Ssb (single-stranded DNA-binding protein), which is a DNA helix destabilizing protein that is involved in several DNA processes in the cell, including DNA replication, recombination, and repair (MEYER and LAINE 1990). Appendix 1 provides information on the isolation of *P. agglomerans* *ssb*, but it was decided that this protein was not a viable candidate for secretion of the test protein.

Spots M and N were given the vague identities of “putative hydrogenase” and “short chain dehydrogenase,” respectively. Although no specific gene names were assigned to them, proteins that fall under these categories are normally intracellular and are therefore not optimal secreted protein candidates. Hydrogenases are thought to play a role in hydrogen cycling during fermentative growth while certain dehydrogenases participate in the metabolism of lipids and steroids (BLATTNER *et al.* 1997).

Spots C and O matched to FlgL (flagellar hook-associated protein 3) and could be an example of degradation products of one protein because while both spots migrated to roughly the same isoelectric point, spot O is at a higher molecular weight than spot C. FlgL is a protein that is secreted via the flagellar export pathway and is a component of the flagellum hook structure (MINAMINO and NAMBA 2004). As explained elsewhere (Appendix 1), *P. agglomerans flgL* was not able to be amplified and thus could not be used as secreted protein candidate.

Finally, spot P matched to FliC (Flagellin), which is the main subunit of the bacterial flagellum. FliC monomers polymerize end-to-end to form the flagellum (WONG *et al.* 2007). FliC is a bona fide secreted protein that is exported by the flagellar export pathway and was previously shown in *E. coli* to mediate secretion of passenger proteins (MAJANDER *et al.* 2005; MINAMINO and NAMBA 2004). This makes FliC the most optimal candidate identified from the *P. agglomerans* secreted protein profile to mediate secretion of the test protein.

It has also been observed that the secreted proteins visible on a stained acrylamide gel of the Johns Hopkins mosquito isolate of *P. agglomerans* used in this study are highly similar to that of another *P. agglomerans* strain (“E325” isolated from plant matter) grown under the same laboratory conditions (PUSEY 2002). The similar 2D-PAGE profiles infer that secreted protein gene candidates found in the mosquito isolate may work in the plant isolate to export proteins from the cell. The Lampe laboratory is currently engineering E325 to produce proteins that will inhibit the spread of Pierce’s disease. Pierce’s disease may also be controlled using paratransgenic methods. Similar to malaria, Pierce’s disease is also an insect-vectored disease. A pathogenic bacterium

called *Xylella fastidiosa* is transmitted to plant xylem by an insect called the glassy-winged sharpshooter (*Homalodisca vitripennis*, formerly *H. coagulata*) (REDAK *et al.* 2004). This results in rapid damage to the grapevines in wineries in California, so there is an effort in place to control the spread of *X. fastidiosa* in order to save the wine industry in this region of the United States (www.piercesdisease.org).

Isolating the *P. agglomerans fliC* gene sequence

The sequence for *P. agglomerans fliC* contig (Figure 2.4) was completed after several rounds of genome walking PCR. The upstream gene is *fliD* and the downstream gene is *fliB*. This is the same genomic organization as seen in *E. coli* (BLATTNER *et al.* 1997). In *E. coli* and *Salmonella* spp., *fliD* encodes the flagellum capping protein, FliD. This cap serves to prevent the FliC subunits from spilling out of the central channel of the flagellum instead of polymerizing at the flagellum distal end (CHEVANCE and HUGHES 2008; MINAMINO and NAMBA 2004). The gene *fliB* encodes a lysine methylase, which is involved in post-translational modification of the flagellum in related species (WONG *et al.* 2007). With the complete *P. agglomerans fliC* sequence determined, secretion constructs were made to test for the secretion of the anti-BSA antibody.

Failure to delete the endogenous *fliC* gene from the *P. agglomerans* chromosome.

A *P. agglomerans* $\Delta fliC$ strain was unable to be made despite the use of several different Lambda Red recombination plasmids (DATSENKO and WANNER 2000; DATTA *et al.* 2006; GUST *et al.* 2004). Amplification of a portion of the *cat* gene in the Lambda Red plasmids (data not shown) confirmed the presence of each of the Lambda Red plasmids in *P. agglomerans* transformants. With repeated attempts to electro-transform a Lambda Red-expressing *P. agglomerans* with increasing amounts of the Kan cassette and

still not obtaining transformants, it was hypothesized that a *fliC* deletion is fatal to the cell and cannot be recovered.

A recent paper reports that the simultaneous expression of the *exo*, *bet*, and *gam* genes in a related species, *Pantoea ananatis*, is highly toxic to the cells and they are not able to grow into normal-sized colonies on a plate (KATASHKINA *et al.* 2009). In order to circumvent this problem and to attain the ability to perform rapid genomic modifications in this species, the authors screened 10^6 transformants for mutants that were resistant to the effects of the Lambda Red proteins. These were colonies that grew to a normal size on a plate and that then had a similar growth rate in LB broth (before and after induction of the Lambda Red genes) as compared to Lambda Red control strains (KATASHKINA *et al.* 2009). In the experiments reported here, it is uncertain if after induction of the Lambda Red genes with arabinose toxic levels of the proteins caused the *P. agglomerans* cells to die because the OD₆₀₀ reading of the induced cultures continued to increase over the 6 h incubation in a fashion similar to a wild-type *P. agglomerans* culture. Regardless, recombinant colonies were not recovered on LB agar plates post-recombination.

If this technique were to be tried again, it may be necessary to screen several Lambda Red *P. agglomerans* colonies for the ability to grow at a normal rate in broth and as colonies on a plate after induction of the recombineering genes. If such a strain is recovered and recombineering is still unsuccessful, longer stretches of homology on the oligos used to amplify the Kan cassette may increase the chances of recombination occurring in the exact desired location. Alternatively, instead of deleting the entire gene, recombination could be performed to insert a point mutation in the *fliC* coding sequence.

FliC would be made by the cell and may not result in the inability of the cell to grow, if indeed, that is the result of a complete *fliC* deletion.

Testing for the *fliC*-mediated secretion of a test protein with anti-BSA activity.

The Western blot results obtained in this study do not show expression and/or secretion of the test protein with any of the constructs tested. Therefore it cannot be concluded that the *P. agglomerans fliC* sequence is not a sufficient signal for secretion of heterologous proteins. Several constructs were tested using the 5' UTR region of *fliC* and the test protein, whether by itself, or as a fusion to FliC or MalE (Figure 2.6). In all cases, protein expression was not observed let alone protein secretion to the growth medium. *P. agglomerans* expressing negative control constructs utilizing the P_{tac} promoter to drive expression of the test protein also failed to produce protein.

Future Directions

In order for an anti-*Plasmodium* strain of *P. agglomerans* to be effective against the development of *Plasmodium* in the mosquito gut, the effector proteins that it produces need to be exported from the cell. If a native secreted protein cannot be used as the signal for secretion, then perhaps a heterologous secretion signal that has been well characterized in related species like *E. coli* can be used in its stead. Examples include the Sec-dependent PelB leader peptide from the *Erwinia carotovora* pectate lyase gene, and the HlyA (hemolysin A) system from pathogenic strains of *E. coli*, which is secreted from the cell by a Sec-independent Type 1 secretion system (HOLLAND *et al.* 1990a; KOSTAKIOTI *et al.* 2005). Both of these signals have been used extensively in *E. coli* to produce active secreted protein products (GENTSCHEV *et al.* 2002; HOLLAND *et al.* 1990b; THIE *et al.* 2008). Because *P. agglomerans* is a close relative to *E. coli*, these alternate

secretion signals should be tested for their efficacy in delivering anti-*Plasmodium* gene products to the cell exterior.

TABLE 2.1 Strains Used in This Study

Strain	Relevant characteristics ^a	Source or Reference
<i>E. coli</i> Top10	F ⁻ <i>mcrA</i> Δ(<i>mrr-hsdRMS-mcrBC</i>) ø80 <i>lacZ</i> ΔM15 Δ <i>lacX74 recA1 araD139</i> Δ (<i>ara-leu</i>)7697 <i>galU</i> <i>galK rpsL</i> (Str ^R) <i>endA1 nupG</i>	Invitrogen Corporation
<i>E. coli</i> Top10 F'	Top 10 with F' [<i>lacI</i> ^q Tn10 (Tc ^R)]	Invitrogen Corporation
<i>S. cerevisiae</i> INVSc-1	Sc1: <i>MATa his3D1 leu2 trp1-289 ura3-52</i> <i>MATα his3D1 leu2 trp1-289 ura3-52</i>	Invitrogen Corporation
<i>P. agglomerans</i>	Wild-type strain isolated from Johns Hopkins U. mosquitoes	(RIEHLE <i>et al.</i> 2007)
<i>P. agglomerans</i> /pIJ790	Johns Hopkins isolate carrying λ Red plasmid	This study
<i>P. agglomerans</i> /pSIM5	Johns Hopkins isolate carrying narrow host range λ Red plasmid	This study
<i>P. agglomerans</i> /pSIM7	Johns Hopkins isolate carrying broad host range λ Red plasmid	This study
<i>P. agglomerans</i> /pSIM9	Johns Hopkins isolate carrying broad host range λ Red plasmid	This study
<i>E. coli</i> BW25113	<i>lacI</i> ^q <i>rrnB</i> Δ <i>lacZ hsdR514</i> Δ <i>araBAD</i> Δ <i>rhaBAD</i>	(DATSENKO and WANNER 2000)

^a Str^R, streptomycin resistance; Tc^R, tetracycline resistance

TABLE 2.2 Plasmids Used in This Study

Plasmid	Relevant Characteristics ^a	Source or Reference
pACYC184	Cam ^R , Tc ^R ; cloning vector with p15A low copy <i>ori</i>	(CHANG and COHEN 1978)
pMQ64	Gent ^R ; yeast recombination vector and source of yeast <i>ori</i> (2μm) and yeast <i>URA3</i> gene	(SHANKS <i>et al.</i> 2006)
pMAL TM -c2X	Ap ^R ; source of <i>malE</i> (maltose binding protein) gene	New England Biolabs
pMAL TM -cRI	Ap ^R ; source of P _{<i>tac</i>}	New England Biolabs
pIT2- <i>scFv</i>	Ap ^R ; source of anti BSA <i>scFv</i> with 6His and myc epitopes	(WINTER <i>et al.</i> 1994)
pDB14	Cam ^R ; pACYC184/2 μm <i>ori</i> and <i>URA3</i> gene	This study
pDB19	Cam ^R ; pDB14/ <i>P. agg fliC</i> 5'UTR-anti BSA <i>scFv</i>	This study
pDB20	Cam ^R ; pDB14/ <i>P. agg fliC</i> 5'UTR- <i>fliC</i> ORF-anti BSA <i>scFv</i>	This study
pDB22	Cam ^R ; pDB14/P _{<i>tac</i>} -anti BSA <i>scFv</i>	This study
pDB24	Cam ^R ; pDB14/ <i>P. agg fliC</i> 5'UTR- <i>AscI</i> -6His-myc-STOP	This study
pDB27	Cam ^R ; pDB24/ <i>P. agg fliC</i> 5'UTR- <i>malE</i> -anti BSA <i>scFv</i>	This study
pDB28	Cam ^R ; pDB14/P _{<i>tac</i>} - <i>malE</i> -anti-BSA <i>scFv</i>	This study
pIJ790	Cam ^R ; source of λ Red recombination genes	(GUST <i>et al.</i> 2004)
pIJ799	Apr ^R ; source of <i>oriT-aac(3)IV</i> recombination cassette	(GUST <i>et al.</i> 2004)
pKD4	Kan ^R ; source of FRT- <i>aphII</i> -FRT recombination cassette	(DATSENKO and WANNER 2000)
pSIM5	Cam ^R ; low copy pSC101 <i>ori</i> , narrow host range λ Red plasmid	(DATTA <i>et al.</i> 2006)
pSIM7	Cam ^R ; broad host range pBBR1 <i>ori</i> , λ Red plasmid	(DATTA <i>et al.</i> 2006)
pSIM9	Cam ^R ; broad host range RK2 <i>ts ori</i> , λ Red plasmid	(DATTA <i>et al.</i> 2006)

^a Ap^R, ampicillin resistance; Cam^R, chloramphenicol resistance; Gent^R, gentamycin resistance; Kan^R, kanamycin resistance; Tc^R, tetracycline resistance; *P. agg*, *P. agglomerans*

Table 2.3: Oligonucleotides Used in this Study

Oligonucleotide	Sequence 5'-3' ^a	Purpose
16S Forward	AGAGTTTGATCCTGGCTCAG	To verify <i>P. agg</i> genomic DNA
16S Reverse	ACGGCTACCTTGTTACGACTT	To verify <i>P. agg</i> genomic DNA
M13 Forward (-20)	GTAAAACGACGGCCAG	Sequencing TOPO inserts
M13 Reverse	CAGGAAACAGCTATGAC	Sequencing TOPO inserts
Flagel38-53 For	GCNAARGAYGAYGCNGCNGGNC	Amplifying <i>P. agg fliC</i> fragment
Flagel65-89 For	AGRAAYGCNAAAYGAYGGNAT	Amplifying <i>P. agg fliC</i> fragment
Flagel65-89 For 2	YTNGCNCARACNACNGARGG	Amplifying <i>P. agg fliC</i> fragment
Flagel126-140 Rev	NCCRTTRAAYTGNGTYTGNC	Amplifying <i>P. agg fliC</i> fragment
Flagel582-595 Rev	NACYTGyTGNGGNACYTGRTTNGC	Amplifying <i>P. agg fliC</i> fragment
fliC walk to 3'1	TTCAGGCAGCGACAACATATG	Gathering <i>fliC</i> flanking DNA
fliC walk to 3'2	GATGATGCGACCGGTAAAGTTG	Gathering <i>fliC</i> flanking DNA
Semi-1	GCCAATTCGGATNGAYKSNNGNTC	Arbitrary walking primer
Semi-2	GCCTTAAGGCCTANGARMSNCCNAG	Arbitrary walking primer
Semi-3	CGGTAAAGGCCTANYTCSKNGANGC	Arbitrary walking primer
Semi-4	GCCAATTCGGATNSAGYMNCTNCG	Arbitrary walking primer
fliC L1	CAATTGCCAACCGTTTCAC	Used with arb primers, $\Delta fliC$ check
fliC L2	GTAACGCCAACGACGGTATC	Used with arb primers
fliC R1	AATCAGAGTTGGTGCCGTTT	Used with arb primers, $\Delta fliC$ check
fliC R2	ACAGTCAGCTCACGAACACG	Used with arb primers
fliC contig to 3' gene1	ATTACTGTTGTGGAACCGCTG	Nested Genome Walking PCR
fliC contig to 3' gene2	AGCGCTATCTGAAATCCTCAC	Nested Genome Walking PCR
fliC contig to 5' gene1	TTTTCTGCTTCAGCCTGTCACG	Nested Genome Walking PCR
fliC contig to 5' gene2	TGTAGCCGCTGAATTACCCTTC	Nested Genome Walking PCR
Ent deltaFliC Left	ACACCTTAACCGAAAGACTTGATTAACAGGAAA...	To delete <i>P. agg fliC</i> (70-mer)
Ent deltaFliC Right	ACCCCGCCGAGGCGGGTTTTGAATTTGCGTT...	To delete <i>P. agg fliC</i> (70-mer)
<i>fliC</i> 5' nest 1	GTAAATTCCTTTATAAGGTTCG	Check for deletion of <i>fliC</i>
k1	CAGTCATAGCCGAATAGCCT	Check for deletion of <i>fliC</i>
k2	CGGTGCCCTGAATGAACTGC	Check for deletion of <i>fliC</i>
kt	CGGCCACAGTCGATGAATCC	Check for deletion of <i>fliC</i>

^a R = A or G; Y = C or T; N = any nucleotide; K = G or T; M = A or C; S = G or C; *P. agg*, *P. agglomerans*

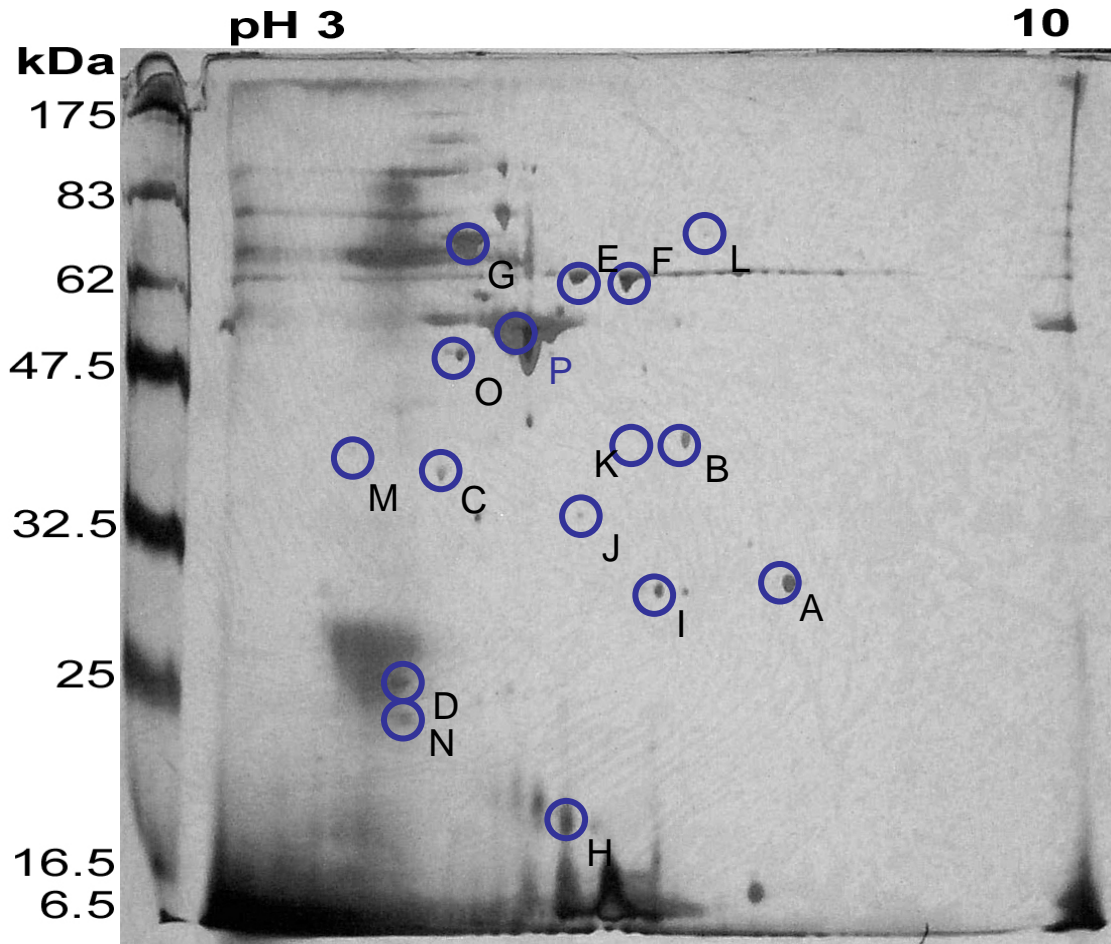


Figure 2.1: A representative silver-stained 2D-PAGE showing the secreted protein profile of *P. agglomerans*. Precipitated secreted proteins from a 3 L late-log culture of *P. agglomerans* were separated by IEF (isoelectric focusing) in a pH range of 3-10 in the first dimension. The proteins were resolved in the second dimension on a 10% acrylamide gel and visualized using silver stain. The spots circled in blue were analyzed using MALDI-TOF. Table 2.4 lists the identities of the spots that could be identified. Appendix 3 shows the additional parallel gels that were generated in order to combine identical spots for MALDI-TOF.

TABLE 2.4: Identification of *P. agglomerans* secreted proteins by MALDI-TOF

Spot	Identity ^a
A	RbsB (D-ribose periplasmic protein)
B	OmpA
C	FlgL
D	No significant matches
E	No significant matches
F	No significant matches
G	No significant matches
H	Ssb
I	RbsB (D-ribose periplasmic protein)
J	No significant matches
K	OmpA
L	No significant matches
M	Putative hydrogenase
N	Short chain dehydrogenase
O	FlgL
P	FliC/Flagellin

^a Identifications of Spots C, H, and O (FlgL, Ssb, and FliC) are based on protein match scores as determined by Taplin Biological Mass Spec Facility. See Figures 2.2 (FliC), A1.1 (FlgL), and A1.3 (Ssb) for scores.

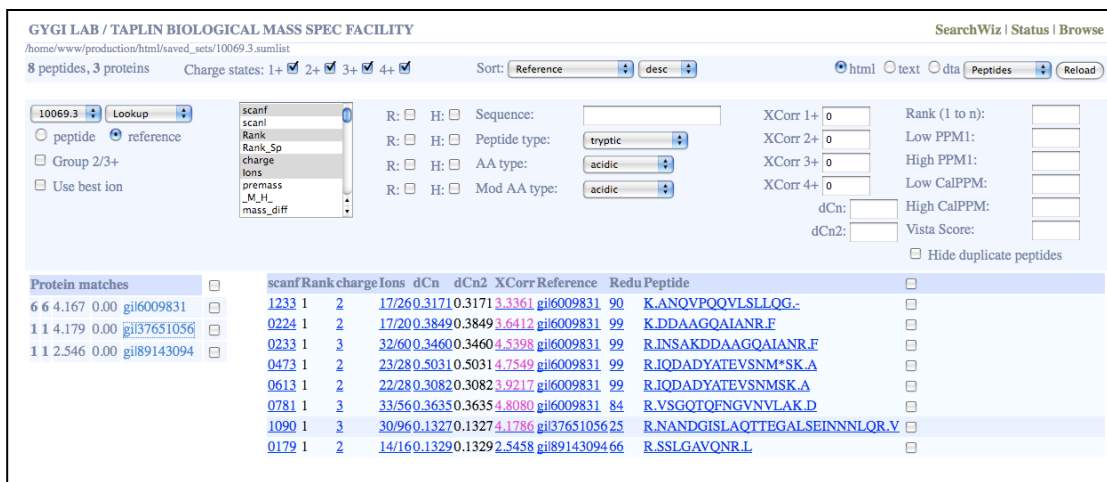


Figure 2.2 FliC/Flagellin MALDI-TOF Results from Taplin Biological Mass Spec Facility. Results from MALDI-TOF analysis of each individual spot are viewed by signing in to the Taplin Biological Mass Spec Facility website with a provided username and password. In the lower left-hand corner is a menu of protein matches that shows the number of peptide fragments isolated for this particular spot (8 total). The ID numbers are a link to an image that shows where each of the peptide fragments align to a known protein in the NCBI Enterobacteriaceae database. In this case, the peptide fragments matched to *E. coli* FliC/Flagellin.

GYGI LAB / TAPLIN BIOLOGICAL MASS SPEC FACILITY

Protein and Peptide Details for reference gi|6009831

Analyzing cached database: /tmp/httpd-tmp/Ecoli.fasta-gi|6009831

Create protein database : /lab/tmp/gi_6009831_dbj_BAA85080_.fasta

gi|6009831|dbj|BAA85080.1| flagellin [Escherichia coli]

MAQVINTSL SLITQNNINK NQALSSSIE RLSSGLRINS AKDDAAGQAI ANRPSTNIKG LTQAARNAND GISVAQTTEG ALSEINNNLQ RIRELTVQAS TGTNSDSDLD SIQDEIKSRL DEIDRVSGQZ
 QFNGVNVLLAK DGSMKIQVGA NDGQTIIDL KRIDSDTLGL NGFNVNNGSGT IANKAATISD LTAAKMDAAT NTITTTNNAL TASKALDQLK DGDVTIKAD AAQATATVYTY NASAGNFSFS NVSNNTSAKA
 GDVAASLEPP AGQTASGVYK AASGEVNFV DANGKITIGG QEAYLTSNGN LTTNDAGGAT AATLDGLEFK AGDQSGIGFN KTASVTMGOT TYNFKTGADA GAATANAGVS FTDTASKETV LNKVATAKQG
 TAVAAAGDPS ATITFYKGVQ TQAVFAAGD GTASARYADN TDVSNATATY TDADGEMTII GSYFTYISID ANNGKVTVDG GTGTGKIAPK VGAEVTVSAN GLTITDASE GTVTKDPLKA LDEAISSIDK
 FRSSLGAIQN RLDSAVTNLN NTTTTLSEMQ SRIQADYATF EVSNMKSQI IQQAGNSVLA KANQVPOQVL SLLQG

Position	MH+	Sequence (click to highlight)
43-53	1102.1489	DDAAGQAIANR
582-595	1495.7188	ANQVPOQVLSLLQG
126-140	1562.7660	VSGQTFNGVNVLLAK
38-53	1615.7432	INSAKDDAAGQAIANR
553-567	1672.8042	IQADYATEVSNMKS

Search Links

- [BLAST entire protein sequence](#)
- Master Database Sources
- unknown
- Search on [Google](#)

Protein coverage by amino acid count: 60/595 = 10.1%
 Protein coverage by mass: 6289/60938 = 10.3%

Highlight special amino acid

C K R Color it

Figure 2.3: Peptide fragment matches to *E. coli* FliC/Flagellin from Taplin

Biological Mass Spec Facility. In the case of the spot that matched to *E. coli*

FliC/Flagellin, eight peptide fragments were generated upon tryptic digestion. This

figure shows where 5 of the 8 fragments matched to the *E. coli* flagellin protein sequence

present in the NCBI Enterobacteriaceae protein database. The amino acid positions are

given as well as the amino acid sequence of the fragment.

A.

```
1  maqvintnsl slitqnnink nqsalsssie rlssglrins akddaagqai anrftsnikg
61  ltqaarnand qisvaqtteg alseinlnlq rirelvtqas tgnstdsldd siqdeiksrl
121 deidrvsgqt qfngvnlak dgsmkiqvga ndgqtitidl kkidsdtlgl ngfnnvngsgt
181 iankaatisd ltaakmdaat ntitttnnal taskaldqlk dgdtvtikad aaqtatvyty
241 nasagnfsfs nvsntsaka gdvaasllpp agqtasgvyk aasgevnfdv dangkitigg
301 qeayltsdgn lttndaggat aatldglfkk agdgqsigfn ktasvtmngt tynfktgada
361 gaatanagvs ftdtasketv lnkvatakqg tavaangdts atityksgvq tyqavfaagd
421 gtasakyadn tdvsnataty tdadgemtti gsyttkysid annkvvtvds gtgtgkyapk
481 vgaevyvsan gtlttatse gvtkdpkka ldeaissidk frsslgaiqn rldsavtnln
541 ntttnlseaq sriqdayat evsnmskaqi iqgagnsvla kanqvpqqvl sllqg
```

B.



Figure 2.4: Identity and placement of sequenced *P. agglomerans* FliC peptide fragments in the *E. coli* FliC sequence and the *P. agglomerans* *fliC* contig assembled after genome walking PCR. Eight *P. agglomerans* FliC peptide fragments were sequenced using MALDI-TOF and identified based on homology to the *E. coli* FliC sequence. Figure 2.4A shows where the fragments (2 fragments are overlapping) match to the *E. coli* FliC sequence (residues in green). The red arrows indicate where the degenerate primers were designed. 2.4B shows the resulting *P. agglomerans* *fliC* contig assembled after genome walking PCR. The blue bracket roughly indicates the length of sequence obtained.

Genome Walking PCR Protocol

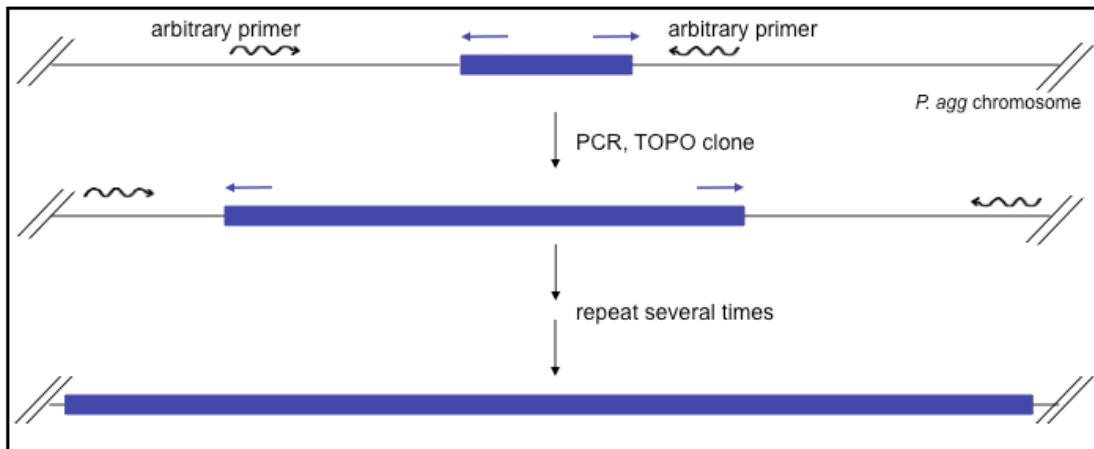


Figure 2.5: Genome walking PCR protocol. After a small fragment of DNA was identified using degenerate primers (top blue box), specific primers were designed on either end (blue arrows). Flanking sequence could then be obtained by pairing the specific primers with arbitrary primers in a genome walking PCR reaction. The TOPO-cloned and sequenced PCR products can be compiled into a contig that can be made longer by repeating the process of creating specific primers on the ends of the known sequence and performing another round of PCR. (Figure adapted from Guo and Xiong 2006).

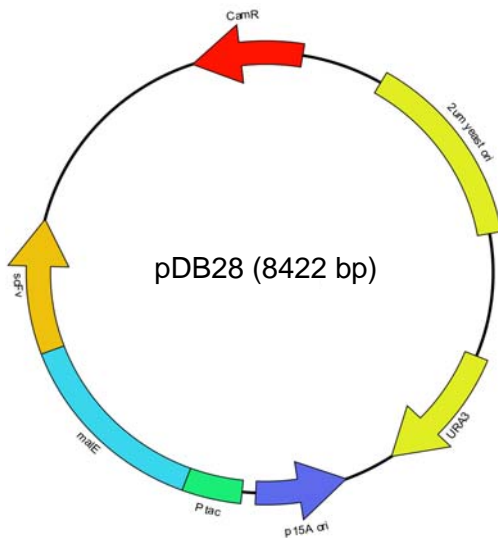
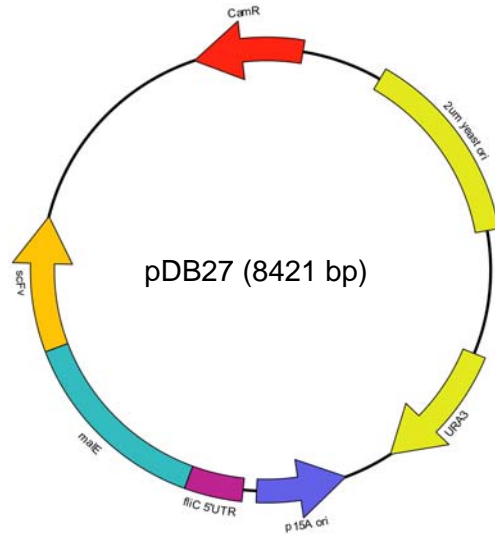
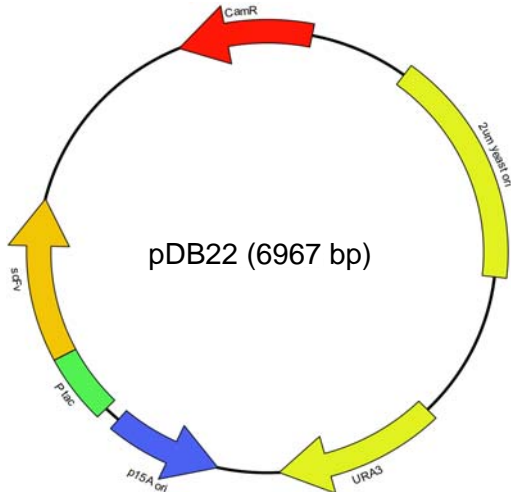
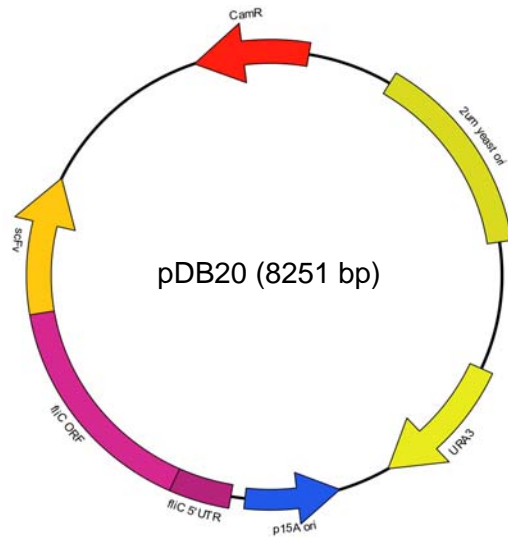
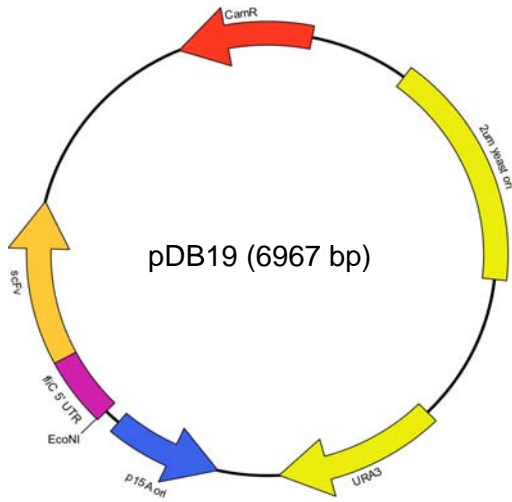
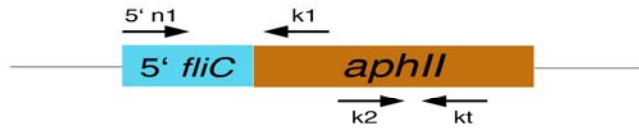


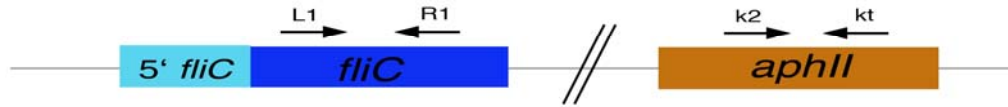
Figure 2.6: *fliC* secretion constructs used in this study. This figure shows the suite of constructs tested in *P. agglomerans* for the *fliC*-mediated secretion of a test protein (the *scFv* with anti-BSA activity). These plasmids are yeast recombination vectors due to the presence of the 2 μ m *ori* and *URA3* gene and also contain a bacterial low-copy *ori* (p15A *ori*). The antibody was expressed by itself (pDB19, pDB22), as a fusion to FliC (pDB20), or as a fusion to MalE (maltose-binding protein) (pDB27, pDB28). Plasmids pDB22 and pDB28 served as negative controls for secretion because an antibody expressed under the control of the P_{tac} promoter would not be secreted from the bacterial cell. (Plasmid maps were made using the free program XPlasMap available from <http://www.iayork.com/XPlasMap/>)

A.

if *fliC* was deleted:



if *fliC* was not deleted:



B.

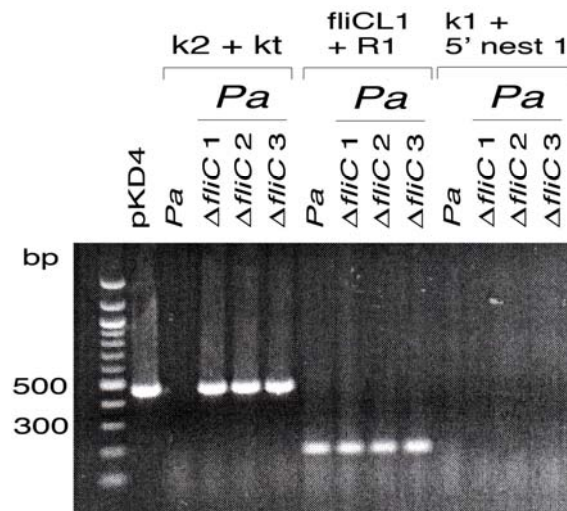


Figure 2.7: Attempts to delete *fliC* from the *P. agglomerans* chromosome were unsuccessful. A. If the *fliC* gene was replaced by the *aphII* gene (confers kanamycin resistance) using recombineering techniques, then a PCR product (from the primers 5' nest1 and k1) comprising part of the *fliC* promoter and part of *aphII* would be generated. Concurrently, a portion of *aphII* would also be amplified (k2 and kt). However, if *fliC* was not deleted, and *aphII* was recombined elsewhere in the chromosome, the only PCR products generated would be from an interior portion of the *fliC* gene (fliCL1 and R1) and a portion of *aphII* (k2 and kt). B. An ethidium bromide stained agarose gel shows the results from a *fliC* recombineering attempt where the *aphII* gene did not recombine into the *fliC* gene as expected, but rather somewhere else in the chromosome. Genomic DNA was isolated from three separate Kan^R *P. agglomerans* colonies post-recombineering (Δ *fliC* 1-3) and used as the template in three different PCR reactions. The product from pKD4 serves as a positive control for the k2 + kt PCR reaction. “Pa” indicates genomic DNA from cells not used in the recombineering experiments as a negative control template in the PCR reaction.

CHAPTER 3

USING HETEROLOGOUS SECRETION SIGNALS FOR SECRETION OF A TEST PROTEIN AND ANTI-*PLASMODIUM* EFFECTOR PROTEINS IN *PANTOEA AGGLOMERANS*

ABSTRACT

Malaria originates from the transmission of the parasite *Plasmodium* to humans by female anopheline mosquitoes. Estimates put the number of deaths at 1-2 million people annually and this number will increase without the establishment of new control strategies. The effectiveness of insecticides and drugs are thwarted by the eventual gain of resistances for both the insect and parasite, therefore as an alternative, a genetic engineering approach is presented here. The bacterial mosquito symbiont *Pantoea agglomerans* was engineered to express effector gene products that are known to inhibit *Plasmodium* development. A caveat of this approach is finding an appropriate protein secretion signal for export of the effectors from the cell. N-terminal secretion signals (PelB from *Erwinia carotovora*, and OmpA and TolB from *E. coli*) and the C-terminal signal from *E. coli* hemolysin A (HlyA) were tested for secretion of a test protein (an anti-BSA scFv) in *E. coli* and *P. agglomerans*. The scFv was secreted from both species using the PelB and HlyA signals, however it was only active in the case of PelB-induced secretion. OmpA-induced secretion of the scFv was only seen in *E. coli*, and the

antibody was not active. Subsequently, four anti-*Plasmodium* effector proteins (SM1, Anti-Pbs21 scFv, PLA2-H67N, and CEL-III) were chosen for secretion and there was varied expression and/or secretion of the different combinations of secretion signals and effector proteins in both species. In particular, SM1 appeared to be somewhat toxic to *P. agglomerans* as those cultures grew poorly. Additionally, CEL-III-containing plasmids recovered after cloning contained mutations in the *pelB* sequence that abolished transcription, which would suggest that a functional and secretable CEL-III is toxic to *E. coli* cells. As a result, this effector was abandoned. The TolB signal was not successful in secretion of any of the proteins tested in either species. The anti-*Plasmodium* Anti-Pbs21-HlyA and PLA2-H67N-HlyA fusions were successfully secreted from *P. agglomerans* and these strains are now available for testing inside malaria-infected anophelines for the inhibition of *Plasmodium* development.

INTRODUCTION

Malaria is arguably the most prevalent insect-vector-borne disease on the planet. It is estimated that 500-700 million people are diagnosed every year and approximately 1-3 million of those will lose their lives to the disease (BREMEN *et al.* 2001). To curb the spread and symptoms of this disease, insecticide-treated bed nets and plant-derived medications can be provided to those in need, but the effects are not a permanent solution. Eventually, mosquitoes and *Plasmodium* species will evolve resistances to these methods. Efforts have been made to genetically modify mosquitoes that can disrupt *Plasmodium* development in the gut, but this approach is time-consuming and fitness costs to the transgenic insect population are a concern (ITO *et al.* 2002; MENGE *et al.* 2005). A paratransgenic approach, engineering a bacterial resident of the mosquito gut to

produce proteins antagonistic to *Plasmodium*, could prove to be an efficient tool against the spread of malaria to humans during a blood meal (RIEHLE and JACOBS-LORENA 2005).

Established examples of paratransgenesis include the use of modified bacteria to fight Chagas disease, a parasitic disease vectored by kissing bugs (*Rhodnius prolixus*) that affects people mainly in Central and South America. An actinomycete endosymbiont of *R. prolixus*, *Rhodococcus rhodnii*, was engineered to express cecropin A, a peptide that is lethal to the Chagas parasite *Trypanosoma cruzi*. (BEARD *et al.* 2001). Methods are being developed to spread the transgenic bacteria through an insect population via a food source for newly hatched insects (DURVASULA *et al.* 1997).

In addition to insect-borne diseases, there are examples of paratransgenic strategies to reduce HIV infectivity in mammalian cells (CHANG *et al.* 2003; RAO *et al.* 2005). *Lactobacillus jensenii*, a bacterium normally found in the intestinal and reproductive mucosa of mammals was engineered to express a secreted form of CD4, an anti-HIV inhibitor protein. *In vitro* results showed reduced HIV entry in cultured cells. Transgenic lactobacilli may survive on vaginal mucosa *in vivo* for days to weeks, so routine inoculation of patients at high risk for HIV infection is a possible form of preventive treatment (CHANG *et al.* 2003). Additionally, an anti-HIV fusion peptide expressed by a strain of *E. coli* was able to colonize the gastrointestinal and cervicovaginal tracts of mice for prolonged periods and produce the peptide at inhibitory levels (RAO *et al.* 2005).

The bacterial candidate for malaria paratransgenesis is *Pantoea agglomerans* (formerly *Enterobacter agglomerans*), a Gram-negative, γ -proteobacterium that is closely

related to *E. coli* and *Salmonella* spp. (GAVINI 1989). The isolate used in this study originated from adult bloodfed female *Anopheles stephensi* mosquitoes at Johns Hopkins University and has been passaged several times through the mosquito to select for bacteria that will survive for longer periods of time in the gut environment (RIEHLE *et al.* 2007).

In order to deliver anti-*Plasmodium* effector proteins to the cell exterior, the effector gene must be fused to DNA coding for a protein secretion signal. This could ensure export of the effector protein into the gut environment where it would interfere with *Plasmodium* development. Several secretion signals have been used to secrete heterologous proteins in *E. coli*. The coding sequences for the first ca. 20 amino acids of PelB (from the pectate lysase gene of *Erwinia carotovora*) or OmpA (from the outer membrane protein A of *E. coli*), which are secreted from the cell in a Type II-dependent manner, have been used for heterologous protein secretion in *E. coli* (LINDEBERG and COLLMER 1992; SLETTA *et al.* 2007; THIE *et al.* 2008; WINTER *et al.* 1994). The first 20 amino acids of a periplasmic protein involved in colicin uptake (TolB), secreted via the bacterial SRP pathway, were also used for *E. coli* protein production (DE GIER *et al.* 1997; THIE *et al.* 2008; VALENT 2001). With these secretion signals, the N-terminal sequence was cleaved from the final product when it was first sent to the periplasmic space. Finally, the *E. coli* hemolysin A signal, mapped to the last 60 amino acids of the protein, has been used to secrete active scFv antibodies in a Type I-dependent manner (FERNANDEZ *et al.* 2000; GENTSCHIEV *et al.* 2002; TZSCHASCHEL *et al.* 1996). The C-terminus of HlyA remains fused to the passenger protein upon secretion.

The four secretion signals mentioned above were tested for the secretion of a test passenger protein (scFv with anti-BSA activity) followed by testing for the secretion of four different anti-*Plasmodium* effector proteins in *E. coli* and *P. agglomerans*. These included the dodecapeptide SM1 (salivary and midgut peptide 1; PCQRAIFQSICN), which was discovered in a phage display library and is analogous to the *Plasmodium* TRAP protein that is used to invade the mosquito salivary glands and midgut epithelium (GHOSH *et al.* 2009; GHOSH *et al.* 2001). Anti-Pbs21 is a 21 kDa single-chain antibody that binds to a surface protein of *P. berghei* and inhibits the developmental transition from gametocyte to ookinete (YOSHIDA *et al.* 1999). Phospholipase A2 (PLA2 H67N) was isolated from honeybee venom, and while its exact anti-*Plasmodium* mechanism is unknown, it is believed that intercalation of PLA2 (20 kDa) into the mosquito midgut lining prevents *Plasmodium* from traversing this membrane on its migration to the salivary glands (MOREIRA *et al.* 2002b). Finally, CEL-III, a lectin isolated from sea cucumber, was discovered to block the transition from ookinete to sporozoite in both *P. berghei* and *P. falciparum* (YOSHIDA *et al.* 2007). Previously, *E. coli* expressing and displaying SM1, Anti-Pbs21 scFv, or PLA2 H67N on their outer membranes while inside the mosquito gut have resulted in the inhibition of *Plasmodium* development, but due to its inability to thrive for long periods of time in the gut environment, the alternate species *P. agglomerans* was chosen for experimentation (RIEHLE *et al.* 2007; YOSHIDA *et al.* 2001).

Varying success was found with these secretion signals and effector proteins. PelB and HlyA directed secretion of the anti-BSA scFv in both species, however it was only active after PelB-induced secretion. The fusion protein OmpA-Anti-BSA scFv was

only expressed and secreted in *E. coli* and it was not an active protein. Of the anti-*Plasmodium* effectors tested, the PelB-SM1 fusion was expressed but not secreted in both species. Fusion proteins PelB-Anti-Pbs21 and OmpA-Anti-Pbs21 were expressed and secreted only by *E. coli*. PelB-PLA2-H67N was expressed but not secreted by both species. Finally, Anti-Pbs21-HlyA and PLA2-H67N-HlyA were secreted from both *E. coli* and *P. agglomerans*. The CEL-III lectin gene could not be cloned without the introduction of mutations in the promoter sequence so it was abandoned as a candidate.

It would appear that the pairing of a secretion signal to a particular effector protein has to be determined empirically. There are several factors to consider when attempting to secrete a heterologous protein, including proper protein folding, clogging of the bacterial membrane translocation machinery, and proteolytic degradation (GEORGIU and SEGATORI 2005). Some effector proteins may be produced at levels toxic to some bacterial species. Other proteins may not be able to be exported via a certain secretion system due to its size or conformation. Even with all of these factors to consider, the pursuit of an anti-*Plasmodium* paratransgenic bacterium should continue. The need for new measures to fight the spread of malaria remains and a transgenic bacteria approach could be a cost-effective and efficient means of attack (RIEHLE and JACOBS-LORENA 2005).

MATERIALS AND METHODS

Media Bacteria were grown in Luria-Bertani broth or agar (LB). *S. cerevisiae* cells were grown on YPD agar plates (20 g tryptone, 10 g yeast extract, 20 g dextrose, 20 g agar per liter) or Minimal Drop-Out Media excluding uracil (2% glucose) when selecting for yeast recombinants (Sigma Y1501). Final concentrations of antibiotics (Sigma or Fisher

Scientific) were as follows: ampicillin (Amp), 150 µg/ml; apramycin (Apr), 80 µg/ml; chloramphenicol (Chl), 30 µg/ml; gentamycin (Gent), 10 µg/ml; nalidixic acid (Nal), 30 µg/ml; rifampicin (Rif), 30 µg/ml; streptomycin sulfate (Str), 100 µg/ml; tetracycline (Tc) 15 µg/ml (unless otherwise indicated). Antibiotic stocks were filter-sterilized through a 0.2 µm filter syringe and stored at -20°C.

Storage of bacterial cells, DNA stocks. All bacterial cell stocks were stored in LB broth +7% DMSO in 1 ml screw-cap tubes at -80°C. All DNA stocks (plasmid or genomic) were kept on ice during usage and stored at -20°C when not in use.

Plasmid construction. The secretion signal plasmids in this study (with the exception of pDB36 and the HlyA group, Table 3.2) were constructed using a yeast gap repair method (SHANKS *et al.* 2006). Briefly, the vector to be modified was digested with restriction enzymes and treated with calf intestinal phosphatase according to manufacturer's directions (New England Biolabs). DNA inserts were amplified with 60-mer oligos that contained 40 bp of homology to the digested vector (up to 2000 bases from the digestion site). Agarose gel-purified vector (~20-200 ng) and inserts (50-500 ng), along with herring sperm DNA (100 µg) were co-transformed into 8-10 large colonies of *S. cerevisiae* INVSc-1 (Invitrogen) cells, harvested directly from a YPD plate and resuspended in 500 µl of PLATE solution (50% PEG, 100 mM LiAc, 10 mM Tris, pH 7.4, 1 mM EDTA). The cells were incubated on the benchtop for ~2 d. After heat shock at 42°C for 30 min, the yeast cells were resuspended in 150 µl of deionized H₂O and plated on uracil drop-out medium (6.7 g/L yeast nitrogen base minus amino acids, 1.92 g/L yeast synthetic drop-out media supplement without uracil, 2% glucose (w/v) and 20 g/L bacteriological agar) and incubated at 30°C for up to 2 d. Total yeast DNA was

purified from the colonies using the “Yeast Smash and Grab DNA Miniprep” protocol (ROSE *et al.* 1990). Fifty nanograms of total yeast DNA was transformed into an appropriate strain of *E. coli* and bacterial clones were verified for the resultant plasmid by restriction enzyme digestion and DNA sequencing.

PelB plasmid construction. Replacing the *bla* gene in pIT2-*scFv* with the *aac(3)IV* gene using “recombineering” techniques resulted in pDB36 (DATSENKO and WANNER 2000; GUST *et al.* 2004). This plasmid allows for the cloning of genes in frame with the *pelB* secretion signal and 6His/myc epitope tag sequence in an apramycin resistance background. Briefly, *E. coli* BW25113 replicating the Lambda Red plasmid pIJ790 and pIT2-*scFv* was grown at 30°C in LB broth containing selective antibiotics and 10 mM L-arabinose until an OD₆₀₀ of 0.6 was reached. The cells were pelleted and washed twice in 50 ml of ice cold 10% glycerol. The cells were resuspended in the drop of remaining 10% glycerol and 50 µl was transformed (1.8 kV, 0.1 mm cuvette) with 500 ng of the *aac(3)IV-oriT* cassette from pIJ799. The *aac(3)IV* gene is flanked by sequence homologous to the *bla* gene. After incubating in 1 ml of LB broth in a shaking incubator (250 rpm) for 1 h at 30°C, different amounts were plated on LB agar plates containing apramycin and incubated overnight at 30°C.

BW25113/pDB36 transformants were also Amp^R (due to pIT2-*scFv* background) so a conjugation step was performed to transfer pDB36 into the *E. coli* recipient LL308. Plasmid DNA was purified from BW25113/pDB36 and used to transform the donor strain *E. coli* ET12567/pUZ8002. An overnight culture of ET12567/pUZ8002 and of recipient strain *E. coli* LL308 were grown at 37°C in selective LB broth. On the next day, 3 ml LB cultures were inoculated with 1/100 dilutions of the overnights and

incubated until an OD₆₀₀ reading of 0.4-0.5 was reached (1-2 h). Five hundred microliter aliquots of the donor and recipient strains were resuspended in fresh LB broth without antibiotics, combined, and incubated for 20 min at 37°C at a low shaking speed. Then, the shaking speed was increased and incubation was continued for an additional 1 h. Varying amounts of the conjugation mixture were plated on LB agar containing nalidixic acid and apramycin and incubated overnight at 37°C. The resultant colonies were grown on LB plates containing apramycin or ampicillin to ensure loss of the pIT2-*scFv* plasmid.

pDB48 contains the *pelB* secretion signal and the 6His and myc affinity tags along with a unique *AscI* site for cloning effector genes in frame between the signal and tags. The *lac* promoter drives expression of the resultant protein. To make pDB48, pMQ64 was digested with *HindIII* and the entire MCS was replaced with *pelB-AscI-6His-myc-STOP* by yeast recombination. This was accomplished with two inserts, one containing *pelB* and the other containing the epitope tags. Both inserts were amplified from pIT2-*scFv* using 60-mer oligos that contained 40 bp of homology to the pMQ64 site of insertion. The left-hand primer for the epitope tags insert contained the *AscI* recognition sequence. pDB51 through pDB54 were made by digesting pDB48 with *AscI* and cloning in each effector gene as one insert (encoding Anti-BSA *scFv*, SM1, Anti-Pbs21, or PLA2 H67N, respectively) using yeast recombination with 60-mer oligos containing 40 bp of homology to the site of insertion.

OmpA plasmid construction. pDB67 contains the *ompA* secretion signal and the 6His and myc affinity tags along with a unique *AscI* site for cloning effector genes in frame between the signal and tags. The *lac* promoter drives expression of the resultant protein. To make pDB67, pMQ64 was digested with *HindIII* and the entire MCS was replaced

with *ompA-AscI-6His-myc-STOP* by yeast recombination. This was accomplished with two inserts, one containing *ompA* (amplified from *E. coli* genomic DNA) and the other containing the epitope tags (amplified from pIT2-*scfv*). Both inserts were amplified using 60-mer oligos that contained 40 bp of homology to the pMQ64 site of insertion. The left-hand primer for the epitope tags insert provided the *AscI* recognition sequence. pDB69 and pDB71 through pDB73 were made by digesting pDB67 with *AscI* and cloning in each effector gene as one insert (encoding Anti-BSA scFv, SM1, Anti-Pbs21, or PLA2 H67N, respectively) using yeast recombination with 60-mer oligos containing 40 bp of homology to the site of insertion.

TolB plasmid construction. pDB68 contains the *tolB* secretion signal and the 6His and myc affinity tags along with a unique *AscI* site for cloning effector genes in frame between the signal and tags. The *lac* promoter drives expression of the resultant protein. To make pDB68, pMQ64 was digested with *HindIII* and the entire MCS was replaced with *tolB-AscI-6His-myc-STOP* by yeast recombination. This was accomplished with two inserts, one containing *tolB* (amplified from *E. coli* genomic DNA) and the other containing the epitope tags coding sequence. Both inserts were amplified with 60-mer oligos that contained 40 bp of homology to the pMQ64 site of insertion. pDB70 and pDB75 through pDB77 were made by digesting pDB68 with *AscI* and cloning in each effector gene as one insert (encoding Anti-BSA scFv, SM1, Anti-Pbs21, or PLA2 H67N, respectively) using yeast recombination with 60-mer oligos containing 40 bp of homology to the site of insertion.

HlyA plasmid construction. Replacing the *bla* gene in pEHLA2-SD with the *aac(3)IV* gene resulted in pDB47. This plasmid allows for the cloning of genes upstream

of the coding sequence for the C-terminus of HlyA and E-tag epitope in an apramycin resistance background. To construct pDB47, “recombineering” techniques followed by a conjugation to eliminate the original plasmid were performed identical to the construction of pDB36.

To make pDB49, pDB50, and pDB58-60, the vector pDB47 was digested with *NheI/XmaI* and treated with a calf intestinal phosphatase according to the manufacturer’s instructions (New England Biolabs). Each insert (encoding Anti-BSA scFv, Male-Anti BSA scFv, SM1, Anti-Pbs21, or PLA2 H67N, respectively) was amplified with 20-mer oligos that incorporated *NheI* and *XmaI* recognition sites on the 5’ and 3’ ends of the amplicons, respectively. The vector and inserts were agarose gel-purified (Zymoclean) and eluted in deionized H₂O. Ligation reactions containing ~150 ng of vector, varying amounts of insert (~300-800 ng), and T4 DNA ligase (New England Biolabs) were incubated overnight (~16h) at 16°C. The reactions were stopped with the addition of 20 µl of deionized H₂O and heat inactivation at 65°C for 20 min. One microliter of the ligation reaction was used to transform the *E. coli* Top10 recipient strain (1.8 kV, 0.1 mm cuvette). The cells were incubated at 37°C for 1 h in LB broth and then aliquots were plated on selective LB agar and incubated overnight at 37°C.

Induction conditions for secretion constructs in *E. coli* and *P. agglomerans*.

Individual colonies were used to inoculate 5 ml of LB broth containing antibiotics and 1% glucose (glucose is only needed for *E. coli* as *P. agglomerans* is Lac⁻) and grown at 30°C overnight (12-16 h). On the next day, a fresh 5 ml LB culture containing antibiotics and glucose was inoculated with 50 µl of the overnight culture and grown to an OD₆₀₀ of

0.5. The bacteria were harvested by centrifugation and resuspended in the same amount of LB broth containing 1 mM IPTG and incubated further at 30°C overnight (12-16 h).

Protein preparations and Western Blot analysis on bacterial pellet and spent growth medium samples. A 100 µl aliquot from an induced overnight culture of either *E. coli* or *P. agglomerans* was centrifuged to pellet the cells (10,000 rpm, 30 sec). Seventy-five microliters of the supernatant was transferred to a new tube containing 25 µl of 3X Laemmli sample buffer (BioRad 161-0737). The remaining pellet was resuspended in 100 µl of 3X Laemmli sample buffer and all of the samples were boiled for 10 min before resolving on a 4% stacking and 10% separating acrylamide gel. Proteins were transferred to polyvinylidene difluoride membrane (PVDF) in a standard Western Blot transfer apparatus (BioRad) with 10% methanol in the transfer buffer (25 mM Tris, 150 mM glycine, 20% methanol, pH 8.3; 95V, 350 mA, 75 min). The membranes were blocked in 1X Tris-buffered saline + 0.05% Tween-20 (1X TBS: 10 mM Tris, 150 mM NaCl, pH 7.4) containing 1% (w/v) bovine serum albumin (BSA) for 3 h at room temperature. For immunodetection of myc-tagged proteins, membranes were incubated overnight at 4°C with an α -myc Ab (1 µg/ml or 1:10,000; Invitrogen 46-0603) diluted in the blocking buffer. After four 15 min washings in 1X TBS/0.05% Tween-20, the membranes were incubated in stabilized goat anti-mouse HRP-conjugated secondary antibody (0.01 µg/100 ml or 1:100,000; Pierce 1858413) for 1 h at room temperature. The washing steps were repeated and the bound antibody-HRP conjugate was detected using a chemiluminescent reaction (SuperSignal West Femto Maximum Sensitivity Substrate; Pierce 34095) and autoradiographic film.

Enzyme-linked immunosorbent assays (ELISAs) using spent growth medium from induced *E. coli* or *P. agglomerans*. The activity of the secreted anti-BSA scFv was tested using an ELISA assay. For each sample analyzed, duplicate MaxiSorp wells (Nunc) were coated with 100 μ l of bovine serum albumin (2 mg/ml) along with two negative control wells, one coated with 100 μ l of 1X PBS and one empty well. The plate was stored at 4°C overnight (ca. 12 h). Each well was washed three times with 200 μ l of 1X PBS before blocking for 2 h at room temperature with 2% dry milk in 1X PBS. The wells were washed three times with 200 μ l of 1X PBS and 100 μ l of clarified spent growth medium from an overnight culture was added to each well and left to incubate for 1 h at room temperature. The wells were then washed eight times with 200 μ l of 1X PBS/0.1% Tween-20. An α -myc-HRP antibody (1:2000, Roche 11-814-15-0001) diluted in blocking reagent was added to the wells and left to incubate for 1 h at room temperature. A second set of eight washes was performed before the α -myc-HRP antibody was detected using 50 μ l/well of the chromogenic substrate 1-Step Ultra TMB (3,3', 5,5'-tetramethyl benzidine, Pierce 34028). When a sufficient blue-colored signal was reached (anywhere from 10-45 min), the reaction was stopped with the addition of 50 μ l of 2 M H₂SO₄ and the absorbance was read at 450 nm with the background subtracted at 650 nm in a BioRad 3550 plate reader.

RESULTS

The secretion constructs

Figure 3.1A shows the organization of the Type II secretion-based constructs (PelB, OmpA, and TolB) created for this study. Effector genes were cloned in to a unique *Asc*I site between the secretion signal and epitope tags. The N-terminal secretion

signal is directly upstream of the effector gene. The effector protein is tagged with 6His and myc epitope tags at its C-terminus. The *lac* promoter drives expression of the ORF. The plasmid replicates using the broad-host range *colE1* origin and confers resistance to apramycin (*aac(3)IV*).

Figure 3.1B shows the hemolysin secretion-based construct used in this study. The effector gene, tagged with 6His and myc at its C-terminus, was cloned by restriction digestion (*NheI/XmaI*) between the *lac* promoter and the 3' end of *hlyA* (*'hlyA*). There is also the E-tag epitope available for immunodetection purposes. This plasmid also replicates using the *colE1* origin and confers resistance to apramycin. It must be co-expressed with pVDL9.3, which provides the membrane channel proteins HlyB and HlyD.

Expression and secretion using the PelB leader

The PelB leader allowed for the expression and secretion of an active anti-BSA scFv in both *E. coli* HB2151 and *P. agglomerans* as shown by Western blot analysis and ELISA assay (Figure 3.2A and C). The anti-BSA scFv secreted by *E. coli* is three times more active than the scFv secreted by *P. agglomerans* as shown by ELISA assay (Figure 3.2B). The *P. agglomerans* E325 strain was also able to secrete an active anti-BSA scFv, although at an even lower level compared to the other strains tested (Figure 3.2A).

Of the three anti-*Plasmodium* effectors tested for secretion under the PelB leader, PelB-Anti-Pbs21 was secreted, but only by *E. coli* HB2151 (Figure 3.2C). PelB-SM1 was expressed by both species, but was not visualized in the spent growth medium (Figure 3.2C). Finally, PelB-PLA2-H67N was neither expressed nor secreted in either

species. All of these strains grew at a normal rate before sample collection as compared to wild-type *E. coli* HB2151 and *P. agglomerans* strains.

Expression and secretion using the OmpA leader

The OmpA leader allowed for the secretion of an inactive anti-BSA scFv, but only in *E. coli* HB2151 (Figure 3.3). OmpA-Anti-Pbs21 was expressed and secreted in *E. coli*, but not in *P. agglomerans* (Figure 3.3C). OmpA-PLA2-H67N was expressed in both species, but did not secrete (Figure 3.3C). Finally, OmpA-SM1 was neither expressed nor secreted in either species. The *P. agglomerans* OmpA-SM1 strains grew at a very slow rate compared to the other OmpA strains, so it may be that the OmpA-induced secretion of the SM1 protein was preventing normal growth of the cultures.

Expression and secretion using the TolB leader

The TolB constructs were not successful in the expression or secretion of the majority of the proteins tested. The only strain to express a protein (anti-BSA scFv) was *E. coli* HB2151 carrying the pDB70 plasmid. The anti-BSA scFv protein was made, but was not visible in the spent growth medium (Figure 3.4C). Despite the inability to detect the anti-BSA scFv protein via Western blot, a slight amount of activity was detected for this protein in an ELISA assay, however, the level of activity was not statistically significant when compared to background levels (Figure 3.4A and B). The *P. agglomerans* TolB strains grew extremely poorly and at a very slow rate, however the *E. coli* TolB strains grew at a normal rate.

Expression and secretion using the C-terminus of HlyA

The anti-BSA scFv-HlyA construct pDB49 was expressed in *E. coli* HB2151 and *P. agglomerans*. Both species were able to secrete the protein at high levels (Figure

3.5C), however, neither antibody was active in an ELISA assay (Figure 3.5A). A MalE-anti-BSA scFv fusion was made (pDB50) in the hopes that antibody activity would be restored, but this fusion protein was only expressed in *E. coli* and was not secreted (Figure 3.5C).

The anti-*Plasmodium* effectors Anti-Pbs21 and PLA2-H67N were both expressed and secreted in high levels as HlyA fusions in *E. coli* and *P. agglomerans* (Figure 3.5C). SM1-HlyA was not expressed or secreted in either species (Figure 3.5C). This does not seem to be the result of poor growth of the cells, as the cultures grew at a normal rate as compared to wild-type *E. coli* HB2151 and *P. agglomerans* strains.

Difficulties in cloning the anti-Plasmodium effector gene CEL-III

The CEL-III lectin gene from sea cucumber was chosen for expression and secretion in *E. coli* and *P. agglomerans* (YOSHIDA et al. 2007). The *CEL-III* ORF (1.2 kb) was amplified from pAgCP-CEL-III (Table 3.3) using 60-mer oligos that contained 40 bp of homology for yeast recombination to create pDB55. The yeast strain in which the recombination took place and the subsequent culture of transformed *E. coli* Top10 grew similar to other yeast cloning experiments and it was expected that the cloning was successful. Five *E. coli* Top10 colonies were checked by restriction digest for the presence of the recombinant plasmid. All 5 clones contained the recircularized vector. A second trial of pDB55 yeast cloning was performed and much higher amounts of *CEL-III* insert were added to the reaction and the length of yeast incubation time in PLATE solution was increased to three days. Eight recombinant *E. coli* colonies were chosen for restriction digestion and PCR reactions to ensure presence of the *CEL-III* ORF. Only one recombinant contained the *CEL-III* ORF in the resultant plasmid, however, after DNA

sequencing, it was revealed that a deletion mutation in the secretion signal abolished transcription of a *pelB-CEL-III* fusion. Therefore, a functional CEL-III protein would not be translated nor secreted. Because of this difficulty in cloning and recovering the *pelB-CEL-III* construct, it was abandoned as a candidate for secretion using the other heterologous signals.

DISCUSSION

The disease malaria claims the lives of millions of people each year. Without the advent of new and effective preventative strategies this number will certainly rise. The introduction of a bacterial symbiont expressing anti-*Plasmodium* gene products into the mosquito vector could be a targeted method of inhibiting *Plasmodium* development prior to transfer to a human host. In order to be effective, the gene product must be delivered to the cell exterior. Steps to find a suitable secreted protein signal from *E. coli* for secretion of anti-*Plasmodium* effector proteins in the paratransgenesis candidate *P. agglomerans* were undertaken in this study.

The suite of anti-*Plasmodium* effector genes

Four anti-*Plasmodium* effector genes were available for secretion tests in *E. coli* and *P. agglomerans*. Previously, three of these effectors (SM1, Anti-Pbs21, and PLA2 H67N) were expressed and/or surface-displayed by *E. coli* while inside the mosquito gut (RIEHLE *et al.* 2007; YOSHIDA *et al.* 2001). Impaired *Plasmodium* development has been seen with transgenic mosquitoes expressing CEL-III (YOSHIDA *et al.* 2007). *E. coli* has been used to express fragments of CEL-III for characterization of its functional domains (KOUZUMA *et al.* 2003). Expression of the full-length *CEL-III* gene in *E. coli* has proven difficult; the fusion to proteins like thioredoxin are needed to increase yield and

solubility, and even then the amount of functional and soluble protein made is very small (Y. Kouzuma, personal correspondence).

With this prior knowledge, it was hypothesized that *E. coli* and *P. agglomerans* could manufacture the SM1, Anti-Pbs21, and PLA2 H67N effector proteins (8 kDa, 21 kDa, and 20 kDa, respectively, as estimated on a protein gel) without difficulty. It was hypothesized that the CEL-III protein (47 kDa) might pose a problem for prokaryotic protein expression and secretion. As will be explained in each of the following sections, not all of the effector proteins were expressed and/or secreted by *E. coli* or by *P. agglomerans*.

The PelB leader

The PelB leader, isolated from the pectate lyase gene of *Erwinia carotovora*, is frequently used as a signal for heterologous protein secretion in *E. coli* (LINDEBERG and COLLMER 1992; THIE *et al.* 2008; WINTER *et al.* 1994). It was encouraging when *E. coli* and *P. agglomerans* secreted a functional anti-BSA scFv (Figure 3.2), and it was hypothesized that the anti-*Plasmodium* effector proteins would behave in a similar fashion when using this secretion signal. However, the only effector to be secreted, and only by *E. coli*, was Anti-Pbs21. SM1 was expressed by both species, but not detected in the spent growth medium. Encouragingly, there may be explanations for these mixed results and additional methods that may improve the expression and yield of these recombinant proteins.

Proteins that are tagged with the PelB leader are secreted from the cell via a Type II secretion system, which involves passage through the periplasmic space before export via the Sec translocation machinery present in the outer membrane (OM) (SANDKVIST

2001). Proteins are folded into their final conformation and the PelB leader is removed in the periplasmic space (CHOI and LEE 2004). SM1 was detected in the cell pellet but not in the spent growth medium, so it may have encountered a problem in the periplasm. This could include periplasmic inclusion bodies, errors in folding or disulfide bond formation, or degradation (KOLAJ *et al.* 2009; SCHLAPSCHY *et al.* 2006). In the case of Anti-Pbs21, this protein may be toxic or misfolded when made by *P. agglomerans*, although this does not seem to be likely because Anti-Pbs21 is an scFv similar to the anti-BSA scFv and this protein was secreted by *P. agglomerans* without any difficulty.

When dealing with difficult proteins destined for the periplasm and beyond, a group of chaperones that prevent misfolding in the *E. coli* periplasm, including seventeen kiloDalton protein (Skp), have been used to improve production, yield, and activity of scFvs (HAYHURST and HARRIS 1999; MAVRANGELOS *et al.* 2001). Co-expression of Skp and an scFv against the herbicide atrazine, whether under the same promoter or on separate plasmids, increased the amount of soluble and active scFv made by *E. coli* (HAYHURST and HARRIS 1999). The yield and activity of two anti-mouse granulocyte scFvs was improved when the scFv was expressed on a plasmid that also contained *skp* (MAVRANGELOS *et al.* 2001). Mavrangelos *et al.* (2001) also employed an alternate Shine-Delgarno sequence upstream of the scFv that resulted in tighter ribosome binding and enhanced expression.

The oxidizing environment of the periplasm is also the site of disulfide bond formation, which is carried out by the Dsb (*disulfide bond formation*) family of proteins (NAKAMOTO and BARDWELL 2004). DsbA transfers a disulfide bond to the target protein and is reoxidized by the membrane-bound DsbB. The isomerization of incorrect

disulfide pairs is carried out by DsbC, which is maintained in an active state by the membrane-bound DsbD (KOLAJ *et al.* 2009). Considerable success has been achieved with the co-expression of DsbABCD and recombinant proteins (JOLY *et al.* 1998; LEE *et al.* 2004; WULFING and RAPPUOLI 1997). This also includes the co-production of Dsb proteins and glutamate racemase, which has no disulfide bridges in its secondary structure (KOHDA *et al.* 2002). As with all of these modifiers, the optimal combination for each recombinant protein had to be determined empirically.

A helper plasmid containing genes for DsbAC and FkpA and SurA was utilized for the improvement of secretion of two recombinant secreted proteins in *E. coli* (SCHLAPSCHY *et al.* 2006). These proteins are PPIases, peptidyl-prolyl *cis/trans* isomerases that aids in the *cis-trans* isomerization of peptidyl-prolyl bonds in newly translated polypeptides. This combination of disulfide bond formation proteins and isomerases could be tested for the improved expression and secretion of all of the effector proteins used in this study.

It is curious why PLA2 H67N was not expressed by either species, as previous studies showed expression of this protein in *E. coli* without incident (RIEHLE *et al.* 2007). Because PLA2 H67N was not detected in either species in this study, it may be too complex of a protein to be expressed in this particular type of construct and then secreted in a Type II-dependent fashion. Perhaps changing the plasmid origin of replication to a lower copy *ori* would help in the production of PelB-PLA2 H67N. It is also worth co-expressing PLA2 H67N with cytoplasmic folding chaperones like DnaK, periplasmic chaperones like Skp, FkpA, and/or a DsbAC protein to determine if a bottleneck is reached pre- or post-translocation into the periplasmic space.

Finally, the addition of thioredoxin, a small 12 kDa protein that participates in disulfide bond formation, has been shown to improve production of some difficult proteins that were normally found in inclusion bodies (GARCIA-ORTEGA *et al.* 2000; YUAN *et al.* 2004). Whether as a separate entity or as a fusion to the recombinant protein (this option is available in a Novagen vector, EMD North America), thioredoxin could improve production of an effector like SM1, PLA2 H67N, or as mentioned in the previous section, CEL-III.

The OmpA leader

The OmpA leader has been used to secrete and/or surface display heterologous proteins from *E. coli* (EARHART 2000; LANG 2000; SLETTA *et al.* 2007; THIE *et al.* 2008). It was expected that this leader would be useful in secreting the anti-BSA scFv as well as some of the effector proteins. The opposite was observed. Only *E. coli* secreted an inactive OmpA-anti-BSA scFv (Figure 3.3). The Anti-Pbs21 protein was expressed and secreted by *E. coli*, as was seen in the case of PelB- and HlyA-induced secretion of this protein (Figures 3.2 and 3.5), however *P. agglomerans* was only able to secrete an Anti-Pbs21-HlyA protein. PLA2 H67N was expressed in both species, but was not detected in the spent growth medium (Figure 3.3C). As was seen with the previous signals tested, SM1 was not detected in the cell pellet or spent growth medium.

In the case of the *P. agglomerans* OmpA-SM1 cultures, the growth rate was significantly affected by the presence of this construct. Explanations for the loss of detection of SM1 such as improper folding may not apply in this instance. The fact that the cultures grew so poorly suggests that the accumulation of SM1 in the cells may have been too toxic for *P. agglomerans* to maintain growing. Perhaps the OmpA leader

resulted in a jamming of the translocation machinery and the cultures ceased to grow at a normal rate.

Because recombinant proteins tagged with either the OmpA or PelB leader peptide are secreted in a similar fashion, the result suggested modifiers in the PelB section above would apply here. The lack of expression of the anti-BSA scFv, SM1, and Anti-Pbs21 in *P. agglomerans* (Figure 3.3C) could be reversed with the introduction of any of the cytoplasmic chaperones or periplasmic modifiers mentioned above. The lack of secretion of PLA2 H67N from *E. coli* and *P. agglomerans* may be reversed with the co-expression of the helper plasmid developed by Schlapschy *et al.* (2006). Again, each of these options would have to be tested for their efficacy.

The TolB leader

An overall pattern of failure was seen with TolB; it was the least successful in expression and secretion and also negatively affected the growth of *P. agglomerans* (Figure 3.4). TolB, an *E. coli* periplasmic protein involved in colicin uptake, is secreted via the bacterial SRP (signal recognition particle) pathway, which translocates proteins in a co-translational fashion (DE GIER *et al.* 1997; VALENT 2001). This transport involves interaction with a ribonucleoprotein complex that shuttles the passenger protein to the IM Sec translocon and into the periplasmic space before secretion via an OM SRP translocon (VALENT 2001).

With the exception of detecting the anti-BSA scFv in the *E. coli* cell pellet, none of the other proteins were visualized by Western blot in either species. There was a very faint reaction in the *E. coli* TolB-anti-BSA scFv ELISA assay, however, the signal

detected was not statistically significant compared to background readings and was likely due to technical error during the procedure.

To date, there are few examples of successful SRP-mediated secretion or surface display of heterologous proteins in the literature (STEINER *et al.* 2006; THIE *et al.* 2008). The SRP pathway may be too esoteric for optimizing secretion of each unique effector protein, or it may be that a different SRP signal (ex. DsbA or TorT) would be more successful (STEINER *et al.* 2006; THIE *et al.* 2008). Each signal along with each effector protein would have to be tested empirically.

The *E. coli* TolB leader may not be recognized by *P. agglomerans* as a secretion signal. Conversely, failure to detect expression or secretion may be due to clogging of the SRP secretion machinery. If so, precursors or altered forms of the proteins would accumulate and could be detected. This could explain why the cultures grew poorly; accumulation of aggregated heterologous proteins or the formation of inclusion bodies would cause the cells to become sick (HOFFMANN and RINAS 2004). Because SRP-mediated secretion occurs co-translationally, it is hard to know if the co-expression of cytoplasmic chaperones and folding catalysts like trigger factor (TF, a PPIase) and heat shock proteins in the Hsp60 and Hsp70 families (such as the Hsp70 trio DnaJ-DnaK-GrpE or the Hsp60 proteins GroEL and GroES, which all act upon newly translated polypeptides) would improve the yield and delivery of passenger proteins (KOLAJ *et al.* 2009).

The HlyA secretion signal

Although the anti-BSA scFv test protein fused to the C-terminus of HlyA was secreted in high levels in both species, neither was active in an ELISA assay (Figure 3.5).

This could be because fusion to the 30 kDa C-terminus of HlyA abolished function of the scFv. However, this was not the case for other researchers who secreted functional scFvs fused to the HlyA C-terminus (FERNANDEZ *et al.* 2000).

Regardless, HlyA was the most successful of the four signals tested in terms of delivery of anti-*Plasmodium* effector proteins to the extracellular space. Both Anti-Pbs21 and PLA2 H67N were expressed and secreted as HlyA fusions by *P. agglomerans* (Figure 3.5C). The efficacy of these strains in inhibition of *Plasmodium* development has yet to be tested, but in the case of PLA2 H67N (20 kDa), a functional phospholipase is not needed for inhibition to occur (MOREIRA *et al.* 2002b). Anti-Pbs21 is a 21 kDa antibody that binds to a *P. berghei* surface protein; so proper folding of the protein is necessary for function (YOSHIDA *et al.* 1999). It will be interesting to see whether this scFv can function after being secreted in one step to the cell exterior via this Type I secretion system and remaining as a fusion to the HlyA C-terminus. As with the N-terminal signals, SM1 (8 kDa) was not detected in the spent growth medium. The fact that it was not detected in the cell pellet indicates a problem with expression of this effector protein under these conditions.

The co-expression of a cocktail of cytoplasmic chaperones with PPIase activity and/or folding properties have shown improvement in the yield, activity, and secretion of some recombinant proteins in *E. coli* (DE MARCO 2007; DE MARCO *et al.* 2007; NISHIHARA *et al.* 2000). Trigger factor (TF), a member of the superfamily of cytoplasmic PPIases and the first chaperone to bind a polypeptide upon its exit from the ribosome, aids in the *cis-trans* isomerization of peptidyl-prolyl bonds (KOLAJ *et al.* 2009). This isomerization is often a rate-limiting step in the folding process and overexpression of TF

can sometimes improve production of a protein (NISHIHARA *et al.* 2000). Hsp60s and Hsp70s are ubiquitous proteins with the ability to aid in folding and trafficking of polypeptides (KOLAJ *et al.* 2009). The overexpression of different combinations of DnaK-DnaJ-GrpE (Hsp70s), GroESL (Hsp60s) and other accessory proteins gave a considerable boost in the production of soluble heterologous proteins in *E. coli* (DE MARCO 2007; DE MARCO *et al.* 2007).

Activity and solubility of various scFvs were improved after coexpression of DnaK-DnaJ-GrpE (CHOI *et al.* 2004; HU *et al.* 2007). Some of these researchers also reported a reduced solubility of the scFv after co-overproduction of GroESL, which provides a warning that certain additives may end up hindering production (HU *et al.* 2007). It is important to note that in each of these studies different combinations of modulators worked best for each of the recombinant proteins tested; it is unlikely that a universal “repair kit” for recombinant production exists (KOLAJ *et al.* 2009).

Finally, if the HlyA C-terminus fusion reveals itself as a problem, a linker containing the outer membrane protease OmpT cleavage sequence could be cloned in between the effector gene and the '*hlyA* sequence (HANKE *et al.* 1992). Upon secretion, the effector protein, unencumbered by HlyA, may behave more as expected and go on to successfully inhibit *Plasmodium* in the mosquito gut.

Summary

Table 3.4 summarizes the results obtained using the four secretion signals and four passenger proteins. Overall, the results were mixed and no apparent pattern of success for one signal or one passenger was seen. Across the board, SM1 was not secreted using any of the signals and the *ompA*-SM1 constructs appeared to make

P. agglomerans cells sick. Anti-Pbs21 was secreted with most of the signals, but only in *E. coli*. A protein similar in structure to Anti-Pbs21, the anti-BSA scFv, was secreted as an active protein in both species, but only under the PelB signal. PLA2 H67N was only secreted using the HlyA signal, which may indicate something inherent in this protein that makes it easiest to secrete in one step (Type I secretion). TolB was the least successful signal tested, with the majority of proteins neither expressed nor secreted.

Future Directions

Realistically, as new effector genes are discovered or created, a case-by-case empirical testing to obtain a functional and secretable effector product will be needed. This study shows that there is no predicted response for how a secretion signal or effector protein will behave in *P. agglomerans*, let alone the most-studied bacterium in the world, *E. coli*. It is just as probable that an effector protein would require no modifiers as it would need a unique series of modifiers in order to see it expressed and/or secreted from the cell.

Luckily, there are methods that are continually being developed for improving the yield and activity of a secreted protein. With the co-expression of chaperone proteins that act in the cytoplasm or periplasm to aid in folding and processing of nascent polypeptides, the yield of soluble and secreted protein could increase dramatically (KOLAJ *et al.* 2009). These chaperone proteins could first be provided on helper plasmids to see if their presence improves production and secretion.

Because there are concerns regarding paratransgenic bacterium carrying multiple plasmids in the wild, there are ways to limit the amount of extraneous DNA carried by the bacterium. As a first test, the chaperone protein genes and effector protein genes

could be combined on one plasmid, possibly under control of the same promoter (HAYHURST and HARRIS 1999). Organizing these genes in an operon would ensure simultaneous production of the effector protein and the chaperones that it needs (Figure 3.6). Secondly, this operon could be cloned into a transposable element that would incorporate into the chromosome. As long as a single-copy version of this operon showed promise as an effective means of *Plasmodium* inhibition, this would reduce the already rare chance of horizontal gene transfer. This organization may also improve the expression of some of the effector proteins that may be too toxic to the cell if they are translated from a multicopy plasmid.

TABLE 3.1: N-terminal Secretion Signal Sequences

Signal	Sequence ^a	Reference
PelB	MKYLLPTAAAGLLLLAAQPAMA↓EV...	(WINTER <i>et al.</i> 1994)
OmpA	MKKTATAIAVALAGFATVAQA↓ <u>MA</u> EV...	(THIE <i>et al.</i> 2008)
TolB	MKQALRVAFGFLILWASVLHA↓ <u>AQPAMA</u> EV...	(THIE <i>et al.</i> 2008)

^aThe ↓ indicates the predicted cleavage site. The residues in italics represent the amino terminus of a secreted protein. Additional amino acids not belonging to the signal sequence are underlined.

TABLE 3.2 Strains Used in This Study

Strain	Relevant characteristics ^a	Source or Reference
<i>E. coli</i> Top10	F <i>mcrA</i> Δ (<i>mrr-hsdRMS-mcrBC</i>) ϕ 80 <i>lacZ</i> Δ M15 <i>lacX74 recA1 araD139</i> Δ (<i>ara-leu</i>)7697 <i>galU galK</i> <i>rpsL</i> (Str ^R) <i>endA1 nupG</i>	Invitrogen Corporation
<i>E. coli</i> Top10 F'	Top 10 with F' [<i>lacI^q</i> Tn10 (Tc ^R)]	Invitrogen Corporation
<i>S. cerevisiae</i> INVSc-1	Sc1: <i>MATa his3D1 leu2 trp1-289 ura3-52</i> <i>MATα his3D1 leu2 trp1-289 ura3-52</i>	Invitrogen Corporation
<i>E. coli</i> BW25113	<i>lacI^q rrmB</i> Δ <i>lacZ hsdR514</i> Δ <i>araBAD</i> Δ <i>rhaBAD</i>	(DATSENKO and WANNER 2000)
<i>E. coli</i> ET12567	<i>dam-13::Tn9 dcm-6 hsdM hsdR recF143</i> <i>zjj201::Tn10 galK2 galT22 ara-14 lacY1 xyl-5 leuB6</i> <i>thi-1 tonA31 rpsL136 hisG4 tsx-78 mtl-1 glnV44</i> F'	(MACNEIL <i>et al.</i> 1992)
<i>E. coli</i> LL308	Δ (<i>pro-lac</i>) <i>recA nalA supE thi/F' pro⁺ lacI^q</i> <i>lacZ</i> Δ M15	(ZENDEL <i>et al.</i> 1980)
<i>E. coli</i> HB2151	Δ <i>lac-pro ara nalA thi F' (proAB lacI^q lacZ</i> Δ M15)	(WINTER <i>et al.</i> 1994)
<i>P. agglomerans</i>	Wild-type strain isolated from Johns Hopkins U. mosquitoes	(RIEHLE <i>et al.</i> 2007)
<i>P. agglomerans</i> E325	Commercial strain isolated from plant matter; Rif ^R	(PUSEY 2002)

^a Str^R, streptomycin resistance; Tc^R, tetracycline resistance; Nal^R, nalidixic acid resistance; Rif^R, rifampicin resistance

TABLE 3.3 Plasmids Used in This Study

Plasmid	Relevant Characteristics ^a	Source or Reference
pIJ790	Cam ^R ; λ Red (<i>gam</i> , <i>bet</i> , <i>exo</i>) <i>araC rep101</i> ^{ts}	(GUST <i>et al.</i> 2004)
pIJ799	Apr ^R ; source of <i>aac(3)IV-oriT</i> cassette with <i>bla</i> homology	(GUST <i>et al.</i> 2004)
pUZ8002	Kan ^R ; RK2 derivative, nontransmissible plasmid in <i>E. coli</i>	(PAGET <i>et al.</i> 1999)
pMQ64	Gent ^R ; yeast recombination vector containing <i>colE1 ori</i>	(SHANKS <i>et al.</i> 2006)
pDB27	Cam ^R ; source of <i>malE</i> -anti BSA <i>scFv</i>	Bisi, D.C., unpublished
pHA-2-5-1	source of SM1 effector gene (encodes 2 tandem copies)	(GHOSH <i>et al.</i> 2001)
pDL200.3	source of Anti-Pbs21 effector gene	Lampe, D.J., unpublished
pTOPO-PLA2-H67N	source of PLA2-H67N effector gene	(MOREIRA <i>et al.</i> 2002b)
pAgCP-CEL-III	source of CEL-III effector gene	(YOSHIDA <i>et al.</i> 2007)
PelB		
pIT2- <i>scFv</i>	Ap ^R ; pIT2 with α -BSA <i>scFv</i> between <i>pelB</i> and epitope tags	(WINTER <i>et al.</i> 1994)
pDB36	Apr ^R ; pIT2- <i>scFv</i> with <i>bla</i> gene replaced with <i>aac(3)IV</i>	This study
pDB48	Gent ^R ; <i>pelB-AscI</i> -6His-myc-STOP cloned in pMQ64 MCS	This study
pDB51	Gent ^R ; pDB48/Anti-BSA <i>scFv</i>	This study
pDB52	Gent ^R ; pDB48/2-SM1	This study
pDB53	Gent ^R ; pDB48/Pbs21	This study
pDB54	Gent ^R ; pDB48/PLA2-H67N	This study
OmpA		
pDB67	Gent ^R ; <i>ompA-AscI</i> -6His-myc-STOP cloned in pMQ64 MCS	This study
pDB69	Gent ^R ; pDB67/ α -BSA <i>scFv</i>	This study
pDB71	Gent ^R ; pDB67/2-SM1	This study
pDB72	Gent ^R ; pDB67/Pbs21	This study
pDB73	Gent ^R ; pDB67/PLA2 H67N	This study
TolB		
pDB68	Gent ^R ; <i>tolB-AscI</i> -6His-myc-STOP cloned in pMQ64 MCS	This study
pDB70	Gent ^R ; pDB68/ α -BSA <i>scFv</i>	This study
pDB75	Gent ^R ; pDB68/2-SM1	This study
pDB76	Gent ^R ; pDB68/Pbs21	This study
pDB77	Gent ^R ; pDB68/PLA2 H67N	This study
HlyA		
pVDL9.3	Cam ^R ; production of HlyB and HlyD transporters	(TZSCHASCHEL <i>et al.</i> 1996)
pEHLA2-SD	Ap ^R ; polylinker for cloning ORFs in frame with E-tagged 'hlyA (23-kDa C-terminal domain of HlyA)	(FERNANDEZ <i>et al.</i> 2000)
pDB47	Apr ^R ; pEHLA2-SD with <i>bla</i> gene replaced with <i>aac(3)IV</i>	This study
pDB49	Apr ^R ; pDB47/ α -BSA <i>scFv</i>	This study
pDB50	Apr ^R ; pDB47 <i>malE</i> - α -BSA <i>scFv</i>	This study
pDB58	Apr ^R ; pDB47/SM1	This study
pDB59	Apr ^R ; pDB47/Pbs21	This study
pDB60	Apr ^R ; pDB47/PLA2-H67N	This study

^a Ap^R, ampicillin resistance; Apr^R, apramycin resistance; Cam^R, chloramphenicol resistance; Gent^R, gentamycin resistance; MCS, multiple cloning site

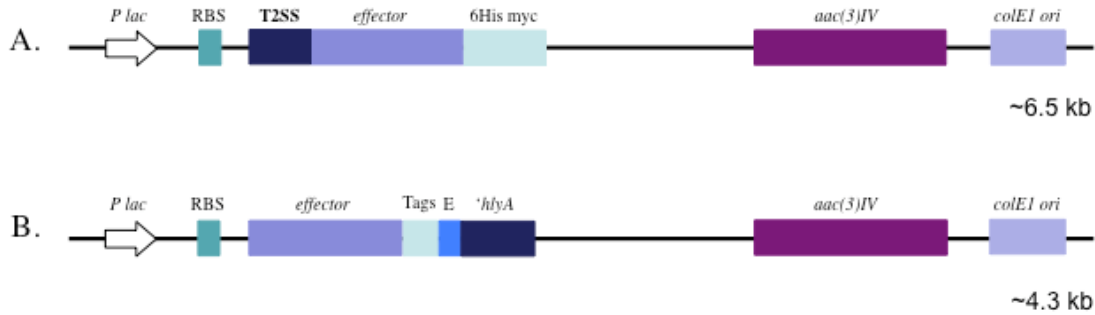


Figure 3.1: Secretion constructs used in this study. The Type II secretion-based construct (PelB, OmpA, and TolB) is shown in Figure 3.1A. The *lac* promoter drives expression of the effector gene, which was tagged with an N-terminal secretion signal (T2SS) and C-terminal epitope tags (6His myc). The plasmid carries the *colE1* origin of replication and the *aac(3)IV* gene (apramycin resistance). Figure 3.1B shows the hemolysin (*hlyA*) secretion-based construct. Again, the *lac* promoter drives expression of the effector gene, which is tagged at its C-terminal end with the 6His and myc epitopes (Tags), as well as the E-tag (E). The effector is fused to the 3' end of hemolysin A (*hlyA*). The hemolysin construct must be co-expressed with pVDL9.3, which provides the membrane channel proteins HlyB and HlyD. (RBS = ribosome binding site)

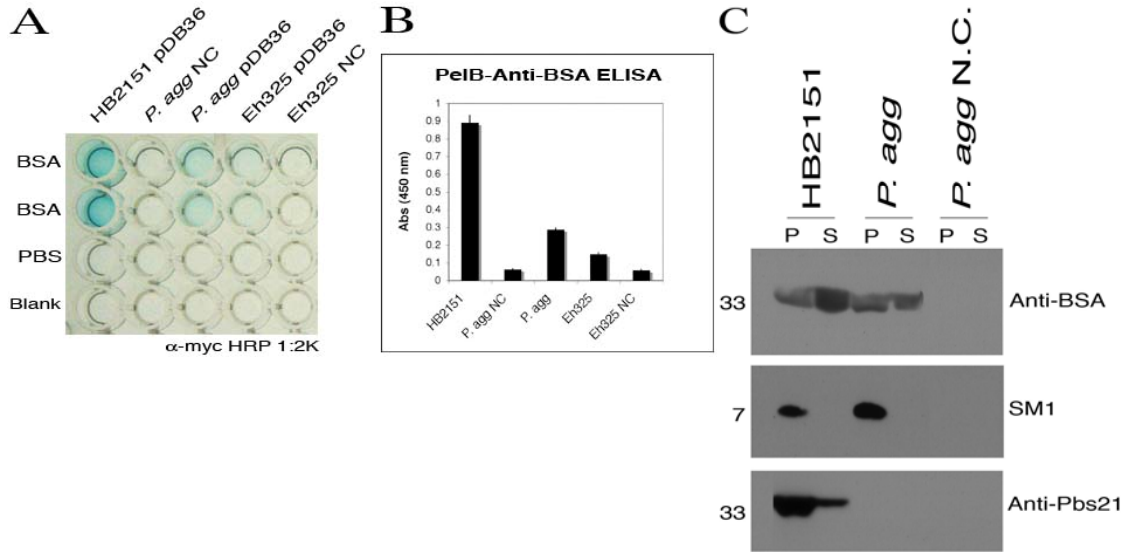


Figure 3.2: Secretion of proteins using PelB in *E. coli* and *P. agglomerans*. Figure 3.2A shows an ELISA analysis of overnight supernatants from *E. coli*, *P. agglomerans*, and the commercial *P. agglomerans* strain E325 expressing pDB36 (*pelB*-anti-BSA *scFv*). The negative controls are *P. agglomerans* and Eh325 with no plasmid (N.C.). Figure 3.2B compares the averaged signal from the BSA-coated wells for each strain. Figure 3.2C shows the presence or absence of various proteins in the cell pellet (P) or spent growth medium supernatant (S) collected from overnight cultures of *E. coli* and *P. agglomerans*. The anti-BSA *scFv* antibody is secreted in both species. The anti-*Plasmodium* effector SM1 is expressed in both species both not secreted, and Anti-Pbs21 is expressed and secreted in *E. coli* only. *P. agglomerans* without a plasmid served as a negative control for the Western blot analysis (N.C.). The proteins were detected using a α -myc antibody.

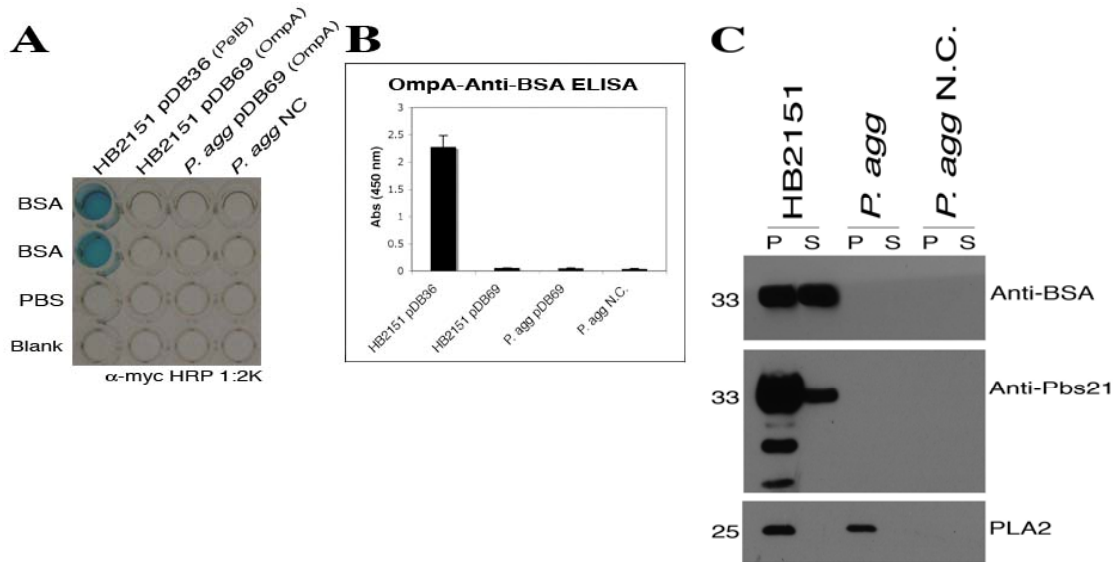


Figure 3.3: Secretion of proteins using OmpA in *E. coli* and *P. agglomerans*. Figure 3.3A shows an ELISA analysis of overnight supernatants from *E. coli* and *P. agglomerans* carrying pDB69 (*ompA*-anti-BSA *scFv*). The proteins were not active in either species. The positive control is HB2151 transformed with pDB36 and a negative control is *P. agglomerans* with no plasmid (N.C.). The averaged signal from the BSA-coated wells for each strain is shown in Figure 3.3B. Figure 3.3C shows the presence or absence of various proteins in the cell pellet (P) or spent growth medium supernatant (S) collected from overnight cultures of *E. coli* and *P. agglomerans*. The anti-BSA scFv antibody is expressed and secreted only in *E. coli*. The anti-*Plasmodium* effector Anti-Pbs21 was expressed and secreted by *E. coli*, and PLA2 H67N were expressed, but not secreted, by both species. *P. agglomerans* with no plasmid is the negative control for Western analysis (N.C.). The proteins were detected using a α -myc antibody.

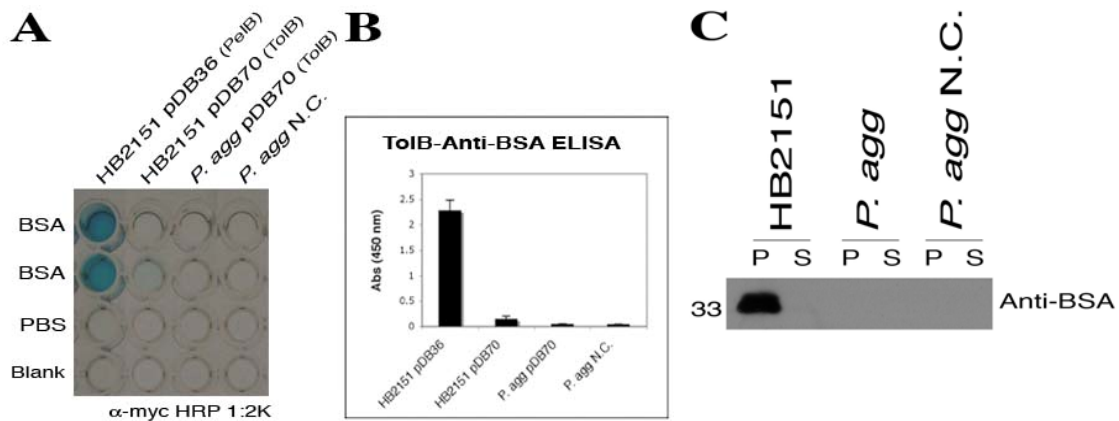


Figure 3.4: Secretion of proteins using TolB in *E. coli* and *P. agglomerans*. Figure 3.4A shows an ELISA analysis of overnight supernatants from *E. coli* and *P. agglomerans* carrying pDB70 (*tolB*-anti-BSA *scFv*). The proteins were not active in either species. The positive control is HB2151 transformed with pDB36 and a negative control is *P. agglomerans* with no plasmid (N.C.). Figure 3.4B shows the averaged signal from the BSA-coated wells for each strain. Figure 3.4C shows the presence or absence of various proteins in the cell pellet (P) or spent growth medium supernatant (S) collected from overnight cultures of *E. coli* and *P. agglomerans*. The anti-BSA *scFv* antibody is expressed, but not secreted in *E. coli*. None of the anti-*Plasmodium* effectors tested (SM1, Anti-Pbs21, or PLA2 H67N) were expressed or secreted in either species. *P. agglomerans* with no plasmid is the negative control for Western analysis (N.C.). The proteins were detected using a α -myc antibody.

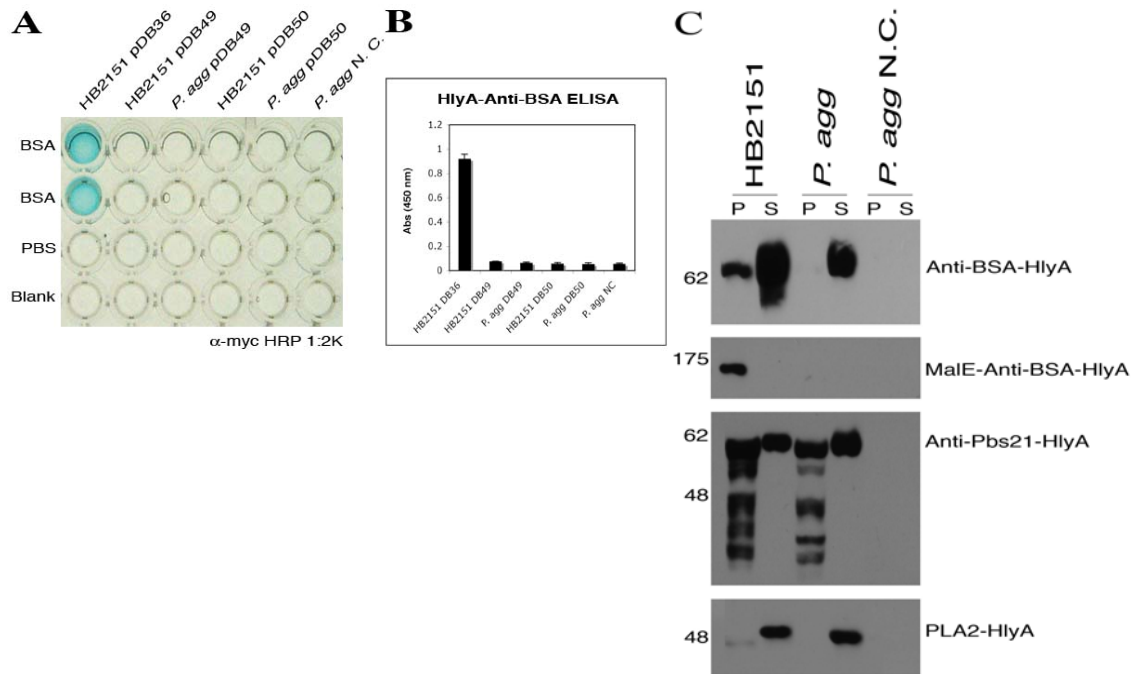


Figure 3.5: Secretion of proteins using HlyA in *E. coli* and *P. agglomerans*. Figure 3.5A shows an ELISA analysis of overnight supernatants from *E. coli* and *P. agglomerans* carrying pDB49 (Anti-BSA scFv-*hlyA*) or pDB50 (*malE*-Anti-BSA scFv-*hlyA*). The proteins were not active in either species. The positive control is HB2151 transformed with pDB36 and a negative control is *P. agglomerans* with no plasmid (N.C.). The averaged signal from the BSA-coated wells for each strain is shown in Figure 3.5B. Figure 3.5C shows the presence or absence of various proteins in the cell pellet (P) or spent growth medium supernatant (S) collected from overnight cultures of *E. coli* and *P. agglomerans*. The Anti-BSA scFv antibody is secreted in both species, but the MalE-Anti-BSA scFv fusion is only expressed in *E. coli*. The anti-*Plasmodium* effectors Anti-Pbs21 and PLA2 H67N were secreted as HlyA fusions by both species. *P. agglomerans* with no plasmid is the negative control for Western analysis (N.C.). The proteins were detected using a α -myc antibody.

TABLE 3.4: Summary of Expression and Secretion Results Using all Signals and Effector Proteins in *E. coli* and *P. agglomerans*

	Protein	PelB		HlyA		OmpA		TolB	
		Exp. ^a	Secr. ^b	Exp.	Secr.	Exp.	Secr.	Exp.	Secr.
<i>E. coli</i>	Anti-BSA	Y	Y	Y	Y	Y	Y	Y	N
<i>P. agg</i>		Y	Y	Y	Y	N	N	N	N
<i>E. coli</i>	SM1	Y	N	N	N	N	N	N	N
<i>P. agg</i>		Y	N	N	N	N	N	N	N
<i>E. coli</i>	Pbs21	Y	Y	Y	Y	Y	Y	N	N
<i>P. agg</i>		N	N	Y	Y	N	N	N	N
<i>E. coli</i>	PLA2-H67N	N	N	Y	Y	Y	N	N	N
<i>P. agg</i>		N	N	Y	Y	Y	N	N	N

^a Expression of the protein as visualized by Western Blot; ^b Secretion of the protein as visualized by Western Blot

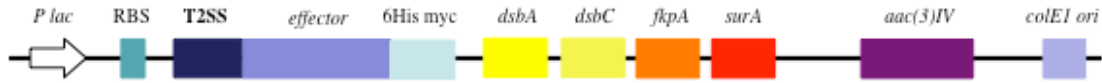


Figure 3.6: Proposed polycistronic effector and chaperone construct. This figure presents a hypothetical operon for expression of a secreted effector protein and its necessary chaperones for optimal secretion and activity. The *lac* promoter would drive expression of the effector gene as well as the chaperones *dsbA*, *dsbC*, *fkpA*, and *surA* (SCHLAPSCHY et al. 2006). The proteins would all be translated at the same time, thus increasing the likelihood of the chaperone proteins interacting with the effector protein and optimizing its potential for secretion and inhibition of *Plasmodium*.

CHAPTER 4

SUMMARY AND FUTURE DIRECTIONS

The research presented in this dissertation is part of a paratransgenic approach to combating the spread of malaria. Malaria is caused by the protist *Plasmodium* and is vectored to humans by female anopheline mosquitoes. A bacterial symbiont of the mosquito gut environment, *Pantoea agglomerans*, is a candidate for expression and delivery of anti-*Plasmodium* effector products. An important caveat of this approach is that the effector product must be secreted from the bacterial cell in order to be effective against *Plasmodium*. This research focused on characterizing and utilizing secretion signals for use in this bacterium.

P. agglomerans was tested, along with the more-characterized species *E. coli*, for the expression and secretion of various anti-*Plasmodium* effector products using several protein secretion signals. These included a native *P. agglomerans* signal within the FliC/Flagellin protein that was characterized after a search for native secreted proteins was performed using spent growth medium. Also, a collection of secretion signals (PelB from *Erwinia carotovora*; OmpA, TolB, and HlyA from *E. coli*) was assessed for their ability to mediate secretion in *P. agglomerans*. The FliC/Flagellin sequences did not work as a secretion signal in *P. agglomerans*. And ultimately, there was no predictable pattern of success when pairing a secretion signal with an effector gene. The results were mixed within each signal used, each

effector gene tested, and even within the behavior of *E. coli* or *P. agglomerans* expressing a particular construct. However, *P. agglomerans* was able to secrete the effectors Anti-Pbs21 and PLA2 H67N as fusions to the HlyA C-terminal secretion signal, and these strains have been sent to Johns Hopkins University. Trials are currently underway there to determine the efficacy of these strains in inhibiting *Plasmodium* development in *Plasmodium*-infected mosquitoes.

Native Secretion Signals

Chapter two summarizes the search for native *P. agglomerans* secreted proteins and their corresponding genes. It was hypothesized that the secretion signal within these genes could be used for secretion of a passenger anti-*Plasmodium* protein. Most signals used for secretion of heterologous proteins come from well-characterized proteins in *E. coli*, and it was expected that a secreted protein from a bacterium like *P. agglomerans* (an environmental isolate with no completed genome sequence) may not behave in the same fashion or perform at all.

The native secreted protein search did not provide a great variety of secreted proteins. Despite the detection of 16 secreted proteins by 2D-PAGE, only 10 could be identified based on sequence homology to protein sequences in the NCBI Enterobacteriaceae database. Within these 10, three pairs of spots matched to the same protein, indicating that isoforms of one protein were being treated as unique proteins. Moreover, most identities matched to putative intracellular proteins with vague descriptors like “hydrogenase” or “periplasmic protein” (Table 2.4). Finally, one protein, FliC/Flagellin, was identified as a bona fide secreted protein (ALDRIDGE and HUGHES 2001; CHEVANCE and HUGHES 2008; MINAMINO and NAMBA 2004). Previously published data shows FliC/Flagellin functions as a signal for

recombinant protein secretion in *E. coli* and *Salmonella typhimurium* (MAJANDER *et al.* 2005; VEGH *et al.* 2006; YOUNG *et al.* 1999).

The *P. agglomerans fliC* 5'UTR and *fliC* ORF were tested for their ability to mediate secretion of an anti-BSA scFv in *P. agglomerans* cells that still had the intact *fliC* gene. There was a complete inability to detect expression of the test protein in the pellet, as well as secretion into the growth medium. The *fliC* gene was abandoned as a secretion signal for mediating heterologous secretion.

As the sequencing of bacterial genomes and proteomes becomes more routine, and databases are updated, the results generated in this study can be re-evaluated. The identity of the remaining unknown proteins (Figure 2.1 and Table 2.4) may be revealed after a search through a more current database. Also, the treatment performed at the Taplin facility, namely trypsin digestion followed by LC-MS/MS (liquid chromatography-tandem-mass-spectrometry), which generates mass fingerprinting and provides users with tryptic peptide masses in a mass spectrum that can be matched to calculated tryptic peptide masses in a database, is sometimes not enough to identify an unknown protein (STEEN and MANN 2004). Using alternate peptide sequencing techniques like *de novo* peptide sequencing may also identify more secreted proteins for testing in secretion trials.

De novo peptide sequencing takes the mass spectrometry output data and uses that to directly determine the primary amino-acid sequence, and the success of identification depends greatly on the quality of data. However, experts warn that the mass accuracy and resolution of the instrument can skew interpretations (STEEN and MANN 2004).

In summary, a native secretion protein from a species like *P. agglomerans* may never work as a signal for recombinant protein production. Additional secreted proteins could be

identified and tested as with *fliC*, however, it may be more time-saving and successful to optimize a heterologous secretion signal for use in this species (Chapter 3).

Heterologous Secretion Signals

Four secretion signals, known to work in *E. coli*, were tested in *E. coli* and *P. agglomerans* for secretion of four different proteins (Anti-BSA scFv, and the anti-*Plasmodium* effectors SM1, Anti-Pbs21 scFv, and PLA2 H67N). PelB and OmpA are N-terminal signal sequences that result in Type II-mediated secretion. TolB is a signal used by the bacterial SRP secretion pathway. When the C-terminus of HlyA is fused to a passenger protein, it is exported from the cell in a Type I secretion pathway. The results from Chapter 3 reveal no clear pattern of successful secretion with one signal or effector protein. It is unpredictable whether the pairing of a signal and effector protein will result in a functional, secreted protein. Fortunately, there are additional tools to test in conjunction with the secretion constructs made in this study that could aid in secretion and restore function and/or folding of the effector protein.

It may be that asking an environmental isolate like *P. agglomerans* to express and secrete heterologous proteins is too tall an order. The data presented in Chapter 3 shows that it often is too much to ask even of a long-domesticated laboratory strain like *E. coli*! It may be possible to take advantage of the considerable efforts by the biotechnology industry to devise ways to improve secretion. For example, the co-expression of a number of folding and/or protein modifying chaperones can greatly improve the yield, solubility, or activity of the desired protein product (KOLAJ *et al.* 2009). One major difficulty is that improvement cannot be predicted and it takes time to determine the correct “cocktail” of chaperones needed for a particular recombinant protein.

As a starting point, the Lampe laboratory has received a helper plasmid called pTUM4 that encodes four periplasmic molecular chaperones, DsbA, DsbC, FkpA, and SurA (SCHLAPSCHY *et al.* 2006). The chloramphenicol resistance gene is compatible with *P. agglomerans*. The co-expression of pTUM4 with each of the anti-*Plasmodium* effector constructs (Table 3.2) should be tested. Perhaps these chaperones will aid in the periplasmic folding (disulfide bond formation and *cis/trans* isomerization) of the effector proteins and they will be more efficiently processed and secreted. In addition to the proteins provided on pTUM4, there are several other chaperones that could improve expression or secretion for each particular signal/effector protein pair, and these are presented in the discussion section of Chapter 3.

If a “magic bullet” is found that improves secretion and processing of an effector protein, it would be most efficient to organize the corresponding gene(s) along with the secretion signal and effector gene in a polycistronic operon (Figure 3.6). If one promoter drove expression of the secretable effector gene and corresponding chaperones, this would greatly ensure the co-production of these proteins and enhance the chaperone’s availability to the effector protein directly upon translation.

What is most important to note is that *P. agglomerans* is capable of secreting an active single-chain antibody (the anti-BSA scFv) using the PelB signal (Figure 3.2). This is an important finding because, in principle, the anti-BSA scFv in the pDB36 construct can be replaced with anti-*Plasmodium* scFvs and tested for secretion and activity against *Plasmodium* parasites present in infected mosquito guts.

It is encouraging that *P. agglomerans* is able to secrete an active scFv via the pDB36 construct, because there are several anti-*Plasmodium* mouse monoclonal antibodies known to

block *Plasmodium* transmission that can be converted to scFv form and tested in pDB36 and the other secretion constructs made in this study (BARR *et al.* 1991; LI *et al.* 2005; QUAKYI *et al.* 1987; RENER *et al.* 1983). Pfs25 is expressed on the surface of zygote and ookinete forms of *P. falciparum* and monoclonal antibodies against this antigen have resulted in the complete failure of the parasite to transition to the oocyst stage inside the mosquito gut (BARR *et al.* 1991). The Pfs48/45 antigens are found on the surface of *P. falciparum* gametocytes and monoclonal antibodies have reduced the infectivity of this parasitic stage by preventing fertilization of gametes taken up during a blood meal (RENER *et al.* 1983). The Pfs230 monoclonal antibody is also active against gametocyte and zygote stages of *P. falciparum* and interferes with the infectivity of the parasite (QUAKYI *et al.* 1987). Finally, an anti-chitinase monoclonal antibody (PfCHT1) has already been converted to an scFv form and been shown to reduce the transmission of *P. falciparum* and *P. gallinaceum* (avian malarial parasite) to mosquitoes (LI *et al.* 2005). The three Pfs antibodies can be converted to scFv using standard procedures and primers designed for mouse V_L, V_k, and V_H genes (TOLEIKIS *et al.* 2004).

REFERENCES

- ABDALLAH, A. M., N. C. GEY VAN PITTIUS, P. A. CHAMPION, J. COX, J. LUIRINK *et al.*, 2007 Type VII secretion--mycobacteria show the way. *Nat Rev Microbiol* **5**: 883-891.
- ABRAHAM, E. G., M. DONNELLY-DOMAN, H. FUJIOKA, A. GHOSH, L. MOREIRA *et al.*, 2005 Driving midgut-specific expression and secretion of a foreign protein in transgenic mosquitoes with *AgAper1* regulatory elements. *Insect Mol Biol* **14**: 271-279.
- ALDRIDGE, P., and K. T. HUGHES, 2001 How and when are substrates selected for type III secretion? *Trends Microbiol* **9**: 209-214.
- ALPHEY, L., C. B. BEARD, P. BILLINGSLEY, M. COETZEE, A. CRISANTI *et al.*, 2002 Malaria control with genetically manipulated insect vectors. *Science* **298**: 119-121.
- BARBAS, C. F., III, D. R. BURTON, J. K. SCOTT and G. J. SILVERMAN, 2001 *Phage Display: A Laboratory Manual*. Cold Spring Harbor Laboratory Press.
- BARR, P. J., K. M. GREEN, H. L. GIBSON, I. C. BATHURST, I. A. QUAKYI *et al.*, 1991 Recombinant Pfs25 protein of *Plasmodium falciparum* elicits malaria transmission-blocking immunity in experimental animals. *J Exp Med* **174**: 1203-1208.
- BEARD, C. B., C. CORDON-ROSALES and R. V. DURVASULA, 2002 Bacterial symbionts of the triatominae and their potential use in control of Chagas disease transmission. *Annu Rev Entomol* **47**: 123-141.
- BEARD, C. B., E. M. DOTSON, P. M. PENNINGTON, S. EICHLER, C. CORDON-ROSALES *et al.*, 2001 Bacterial symbiosis and paratransgenic control of vector-borne Chagas disease. *Int J Parasitol* **31**: 621-627.
- BEARD, C. B., R. V. DURVASULA and F. F. RICHARDS, 1998 Bacterial symbiosis in arthropods and the control of disease transmission. *Emerg Infect Dis* **4**: 581-591.
- BERKS, B. C., T. PALMER and F. SARGENT, 2003 The Tat protein translocation pathway and its role in microbial physiology. *Adv Microb Physiol* **47**: 187-254.
- BEXTINE, B., C. LAUZON, S. POTTER, D. LAMPE and T. A. MILLER, 2004 Delivery of a genetically marked *Alcaligenes* sp. to the glassy-winged sharpshooter for use in a paratransgenic control strategy. *Curr Microbiol* **48**: 327-331.
- BLANDIN, S., S. H. SHIAO, L. F. MOITA, C. J. JANSE, A. P. WATERS *et al.*, 2004 Complement-like protein TEP1 is a determinant of vectorial capacity in the malaria vector *Anopheles gambiae*. *Cell* **116**: 661-670.
- BLATTNER, F. R., G. PLUNKETT, 3RD, C. A. BLOCH, N. T. PERNA, V. BURLAND *et al.*, 1997 The complete genome sequence of *Escherichia coli* K-12. *Science* **277**: 1453-1474.
- BRADY, C., I. CLEENWERCK, S. VENTER, M. VANCANNEYT, J. SWINGS *et al.*, 2008 Phylogeny and identification of *Pantoea* species associated with plants, humans and the natural environment based on multilocus sequence analysis (MLSA). *Syst Appl Microbiol* **31**: 447-460.
- BRADY, C., S. VENTER, I. CLEENWERCK, M. VANCANNEYT, J. SWINGS *et al.*, 2007 A FAFLP system for the improved identification of plant-pathogenic and plant-associated species of the genus *Pantoea*. *Syst Appl Microbiol* **30**: 413-417.

- BREMAN, J. G., A. EGAN and G. T. KEUSCH, 2001 The intolerable burden of malaria: a new look at the numbers. *Am J Trop Med Hyg* **64**: iv-vii.
- BRENNAN, J. D., M. KENT, R. DHAR, H. FUJIOKA and N. KUMAR, 2000 *Anopheles gambiae* salivary gland proteins as putative targets for blocking transmission of malaria parasites. *Proc Natl Acad Sci U S A* **97**: 13859-13864.
- CASCALES, E., 2008 The type VI secretion toolkit. *EMBO Rep* **9**: 735-741.
- CASCALES, E., and P. J. CHRISTIE, 2003 The versatile bacterial type IV secretion systems. *Nat Rev Microbiol* **1**: 137-149.
- CDC, 2006 Malaria: Special Topics, pp.
- CHANG, A. C., and S. N. COHEN, 1978 Construction and characterization of amplifiable multicopy DNA cloning vehicles derived from the P15A cryptic miniplasmid. *J Bacteriol* **134**: 1141-1156.
- CHANG, T. L., C. H. CHANG, D. A. SIMPSON, Q. XU, P. K. MARTIN *et al.*, 2003 Inhibition of HIV infectivity by a natural human isolate of *Lactobacillus jensenii* engineered to express functional two-domain CD4. *Proc Natl Acad Sci U S A* **100**: 11672-11677.
- CHENG, L. W., and O. SCHNEEWIND, 2000 Type III machines of Gram-negative bacteria: delivering the goods. *Trends Microbiol* **8**: 214-220.
- CHEVANCE, F. F., and K. T. HUGHES, 2008 Coordinating assembly of a bacterial macromolecular machine. *Nat Rev Microbiol* **6**: 455-465.
- CHITLARU, T., O. GAT, Y. GOZLAN, N. ARIEL and A. SHAFFERMAN, 2006 Differential proteomic analysis of the *Bacillus anthracis* secretome: distinct plasmid and chromosome CO₂-dependent cross talk mechanisms modulate extracellular proteolytic activities. *J Bacteriol* **188**: 3551-3571.
- CHOI, G. H., D. H. LEE, W. K. MIN, Y. J. CHO, D. H. KWEON *et al.*, 2004 Cloning, expression, and characterization of single-chain variable fragment antibody against mycotoxin deoxynivalenol in recombinant *Escherichia coli*. *Protein Expr Purif* **35**: 84-92.
- CHOI, J. H., and S. Y. LEE, 2004 Secretory and extracellular production of recombinant proteins using *Escherichia coli*. *Appl Microbiol Biotechnol* **64**: 625-635.
- CHRISTIE, P. J., 1997 *Agrobacterium tumefaciens* T-complex transport apparatus: a paradigm for a new family of multifunctional transporters in eubacteria. *J Bacteriol* **179**: 3085-3094.
- CLACKSON, T., H. R. HOOGENBOOM, A. D. GRIFFITHS and G. WINTER, 1991 Making antibody fragments using phage display libraries. *Nature* **352**: 624-628.
- CORNELIS, G. R., 2006 The type III secretion injectisome. *Nat Rev Microbiol* **4**: 811-825.
- CRAIG, N. L., CRAIGIE, R., GELLERT M., LAMBOWITZ, A. M., 2002 *Mobile DNA II*. ASM Press.
- CURTIS, C. F., 1968 Possible use of translocations to fix desirable genes in insect pest populations. *Nature* **218**: 368-369.
- DATSENKO, K. A., and B. L. WANNER, 2000 One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products. *Proc Natl Acad Sci U S A* **97**: 6640-6645.
- DATTA, S., N. COSTANTINO and D. L. COURT, 2006 A set of recombineering plasmids for gram-negative bacteria. *Gene* **379**: 109-115.

- DE BUCK, E., E. LAMMERTYN and J. ANNE, 2008 The importance of the twin-arginine translocation pathway for bacterial virulence. *Trends Microbiol* **16**: 442-453.
- DE GIER, J. W., Q. A. VALENT, G. VON HEIJNE and J. LUIRINK, 1997 The *E. coli* SRP: preferences of a targeting factor. *FEBS Lett* **408**: 1-4.
- DE MARCO, A., 2007 Protocol for preparing proteins with improved solubility by co-expressing with molecular chaperones in *Escherichia coli*. *Nat Protoc* **2**: 2632-2639.
- DE MARCO, A., E. DEUERLING, A. MOGK, T. TOMOYASU and B. BUKAU, 2007 Chaperone-based procedure to increase yields of soluble recombinant proteins produced in *E. coli*. *BMC Biotechnol* **7**: 32.
- DE WILDT, R. M., C. R. MUNDY, B. D. GORICK and I. M. TOMLINSON, 2000 Antibody arrays for high-throughput screening of antibody-antigen interactions. *Nat Biotechnol* **18**: 989-994.
- DELETOILE, A., D. DECRE, S. COURANT, V. PASSET, J. AUDO *et al.*, 2008 Phylogeny and Identification of *Pantoea* Species and Typing of *Pantoea agglomerans* Strains by Multilocus Gene Sequencing. *J Clin Microbiol*.
- DILLON, R. J., C. T. VENNARD and A. K. CHARNLEY, 2000 Exploitation of gut bacteria in the locust. *Nature* **403**: 851.
- DILLON, R. J., C. T. VENNARD and A. K. CHARNLEY, 2002 A note: gut bacteria produce components of a locust cohesion pheromone. *J Appl Microbiol* **92**: 759-763.
- DING, Z., K. ATMAKURI and P. J. CHRISTIE, 2003 The outs and ins of bacterial type IV secretion substrates. *Trends Microbiol* **11**: 527-535.
- DURVASULA, R. V., A. GUMBS, A. PANACKAL, O. KRUGLOV, S. AKSOY *et al.*, 1997 Prevention of insect-borne disease: an approach using transgenic symbiotic bacteria. *Proc Natl Acad Sci U S A* **94**: 3274-3278.
- EARHART, C. F., 2000 Use of an Lpp-OmpA fusion vehicle for bacterial surface display. *Methods Enzymol* **326**: 506-516.
- FAVIA, G., I. RICCI, C. DAMIANI, N. RADDADI, E. CROTTI *et al.*, 2007 Bacteria of the genus *Asaia* stably associate with *Anopheles stephensi*, an Asian malarial mosquito vector. *Proc Natl Acad Sci U S A* **104**: 9047-9051.
- FAVIA, G., I. RICCI, M. MARZORATI, I. NEGRI, A. ALMA *et al.*, 2008 Bacteria of the genus *Asaia*: a potential paratransgenic weapon against malaria. *Adv Exp Med Biol* **627**: 49-59.
- FERNANDEZ, L. A., I. SOLA, L. ENJUANES and V. DE LORENZO, 2000 Specific secretion of active single-chain Fv antibodies into the supernatants of *Escherichia coli* cultures by use of the hemolysin system. *Appl Environ Microbiol* **66**: 5024-5029.
- FILLOUX, A., A. HACHANI and S. BLEVES, 2008 The bacterial type VI secretion machine: yet another player for protein transport across membranes. *Microbiology* **154**: 1570-1583.
- GAL, P., B. VEGH, P. ZAVODSZKY and F. VONDERVISZT, 2006 Export signals. *Nat Biotechnol* **24**: 900-901; author reply 901-902.
- GARCIA-ORTEGA, L., J. LACADENA, V. LACADENA, M. MASIP, C. DE ANTONIO *et al.*, 2000 The solubility of the ribotoxin alpha-sarcin, produced as a recombinant protein in *Escherichia coli*, is increased in the presence of thioredoxin. *Lett Appl Microbiol* **30**: 298-302.

- GARDNER, M. J., N. HALL, E. FUNG, O. WHITE, M. BERRIMAN *et al.*, 2002 Genome sequence of the human malaria parasite *Plasmodium falciparum*. *Nature* **419**: 498-511.
- GAVINI, F., J. MERGAERT, A. BEIJ, C. MIELCAREK, D. IZARD, K. KERSTERS, J. DE LEY, 1989 Transfer of *Enterobacter agglomerans* (Beijerinck 1888) Ewing and Fife 1972 to *Pantoea* gen. nov. as *Pantoea agglomerans* comb. nov. and Description of *Pantoea dispersa* sp. nov. *Int. J. Syst. Bacteriol.* **39**: 337-345.
- GENTSCHCHEV, I., G. DIETRICH and W. GOEBEL, 2002 The *E. coli* alpha-hemolysin secretion system and its use in vaccine development. *Trends Microbiol* **10**: 39-45.
- GEORGIU, G., and L. SEGATORI, 2005 Preparative expression of secreted proteins in bacteria: status report and future prospects. *Curr Opin Biotechnol* **16**: 538-545.
- GHOSH, A., M. J. EDWARDS and M. JACOBS-LORENA, 2000 The journey of the malaria parasite in the mosquito: hopes for the new century. *Parasitol Today* **16**: 196-201.
- GHOSH, A. K., M. DEVENPORT, D. JETHWANNEY, D. E. KALUME, A. PANDEY *et al.*, 2009 Malaria parasite invasion of the mosquito salivary gland requires interaction between the *Plasmodium* TRAP and the anopheles saglin proteins. *PLoS Pathog* **5**: e1000265.
- GHOSH, A. K., P. E. RIBOLLA and M. JACOBS-LORENA, 2001 Targeting *Plasmodium* ligands on mosquito salivary glands and midgut with a phage display peptide library. *Proc Natl Acad Sci U S A* **98**: 13278-13281.
- GROARKE, J. M., W. C. MAHONEY, J. N. HOPE, C. E. FURLONG, F. T. ROBB *et al.*, 1983 The amino acid sequence of D-ribose-binding protein from *Escherichia coli* K12. *J Biol Chem* **258**: 12952-12956.
- GUO, H., and J. XIONG, 2006 A specific and versatile genome walking technique. *Gene* **381**: 18-23.
- GUST, B., G. CHANDRA, D. JAKIMOWICZ, T. YUQING, C. J. BRUTON *et al.*, 2004 Lambda red-mediated genetic manipulation of antibiotic-producing *Streptomyces*. *Adv Appl Microbiol* **54**: 107-128.
- H Aidaris, C. G., J. MALONE, L. A. SHERRILL, J. M. BLISS, A. A. GASPARI *et al.*, 2001 Recombinant human antibody single chain variable fragments reactive with *Candida albicans* surface antigens. *J Immunol Methods* **257**: 185-202.
- HANKE, C., J. HESS, G. SCHUMACHER and W. GOEBEL, 1992 Processing by OmpT of fusion proteins carrying the HlyA transport signal during secretion by the *Escherichia coli* hemolysin transport system. *Mol Gen Genet* **233**: 42-48.
- HATAKEYAMA, T., H. NAGATOMO and N. YAMASAKI, 1995 Interaction of the hemolytic lectin CEL-III from the marine invertebrate *Cucumaria echinata* with the erythrocyte membrane. *J Biol Chem* **270**: 3560-3564.
- HAUBEN, L., E. R. MOORE, L. VAUTERIN, M. STEENACKERS, J. MERGAERT *et al.*, 1998 Phylogenetic position of phytopathogens within the Enterobacteriaceae. *Syst Appl Microbiol* **21**: 384-397.
- HAYHURST, A., and W. J. HARRIS, 1999 *Escherichia coli* *skp* chaperone coexpression improves solubility and phage display of single-chain antibody fragments. *Protein Expr Purif* **15**: 336-343.
- HOFFMANN, F., and U. RINAS, 2004 Stress induced by recombinant protein production in *Escherichia coli*. *Adv Biochem Eng Biotechnol* **89**: 73-92.

- HOLLAND, I. B., B. KENNY and M. BLIGHT, 1990a Haemolysin secretion from *E coli*. *Biochimie* **72**: 131-141.
- HOLLAND, I. B., B. KENNY, B. STEIPE and A. PLUCKTHUN, 1990b Secretion of heterologous proteins in *Escherichia coli*. *Methods Enzymol* **182**: 132-143.
- HOLT, R. A., G. M. SUBRAMANIAN, A. HALPERN, G. G. SUTTON, R. CHARLAB *et al.*, 2002 The genome sequence of the malaria mosquito *Anopheles gambiae*. *Science* **298**: 129-149.
- HU, X., L. O'HARA, S. WHITE, E. MAGNER, M. KANE *et al.*, 2007 Optimisation of production of a domoic acid-binding scFv antibody fragment in *Escherichia coli* using molecular chaperones and functional immobilisation on a mesoporous silicate support. *Protein Expr Purif* **52**: 194-201.
- ITO, J., A. GHOSH, L. A. MOREIRA, E. A. WIMMER and M. JACOBS-LORENA, 2002 Transgenic anopheline mosquitoes impaired in transmission of a malaria parasite. *Nature* **417**: 452-455.
- JACOB-DUBUISSON, F., R. FERNANDEZ and L. COUTTE, 2004 Protein secretion through autotransporter and two-partner pathways. *Biochim Biophys Acta* **1694**: 235-257.
- JACOBS-LORENA, M., 2003 Interrupting malaria transmission by genetic manipulation of anopheline mosquitoes. *J Vector Borne Dis* **40**: 73-77.
- JOLY, J. C., W. S. LEUNG and J. R. SWARTZ, 1998 Overexpression of *Escherichia coli* oxidoreductases increases recombinant insulin-like growth factor-I accumulation. *Proc Natl Acad Sci U S A* **95**: 2773-2777.
- KATASHKINA, J. I., Y. HARA, L. I. GOLUBEVA, I. G. ANDREEVA, T. M. KUVAEVA *et al.*, 2009 Use of the lambda Red-recombineering method for genetic engineering of *Pantoea ananatis*. *BMC Mol Biol* **10**: 34.
- KAZEMI-POUR, N., G. CONDEMINE and N. HUGOUVIEUX-COTTE-PATTAT, 2004 The secretome of the plant pathogenic bacterium *Erwinia chrysanthemi*. *Proteomics* **4**: 3177-3186.
- KOEBNIK, R., 1999 Structural and functional roles of the surface-exposed loops of the beta-barrel membrane protein OmpA from *Escherichia coli*. *J Bacteriol* **181**: 3688-3694.
- KOHDA, J., Y. ENDO, N. OKUMURA, Y. KUROKAWA, K. NISHIHARA *et al.*, 2002 Improvement of productivity of active form of glutamate racemase in *Escherichia coli* by coexpression of folding accessory proteins. *Biochemical Engineering Journal* **10**: 39-45.
- KOLAJ, O., S. SPADA, S. ROBIN and J. G. WALL, 2009 Use of folding modulators to improve heterologous protein production in *Escherichia coli*. *Microb Cell Fact* **8**: 9.
- KORONAKIS, V., P. STANLEY, E. KORONAKIS and C. HUGHES, 1992 The HlyB/HlyD-dependent secretion of toxins by gram-negative bacteria. *FEMS Microbiol Immunol* **5**: 45-53.
- KOSTAKIOTI, M., C. L. NEWMAN, D. G. THANASSI and C. STATHOPOULOS, 2005 Mechanisms of protein export across the bacterial outer membrane. *J Bacteriol* **187**: 4306-4314.
- KOUZUMA, Y., Y. SUZUKI, M. NAKANO, K. MATSUYAMA, S. TOJO *et al.*, 2003 Characterization of functional domains of the hemolytic lectin CEL-III from the marine invertebrate *Cucumaria echinata*. *J Biochem* **134**: 395-402.

- KUWAJIMA, G., I. KAWAGISHI, M. HOMMA, J. ASAKA, E. KONDO *et al.*, 1989 Export of an N-terminal fragment of *Escherichia coli* flagellin by a flagellum-specific pathway. *Proc Natl Acad Sci U S A* **86**: 4953-4957.
- LANG, H., 2000 Outer membrane proteins as surface display systems. *Int J Med Microbiol* **290**: 579-585.
- LE BRAS, J., and R. DURAND, 2003 The mechanisms of resistance to antimalarial drugs in *Plasmodium falciparum*. *Fundam Clin Pharmacol* **17**: 147-153.
- LEE, D. H., M. D. KIM, W. H. LEE, D. H. KWEON and J. H. SEO, 2004 Consortium of fold-catalyzing proteins increases soluble expression of cyclohexanone monooxygenase in recombinant *Escherichia coli*. *Appl Microbiol Biotechnol* **63**: 549-552.
- LETELLIER, L., S. P. HOWARD and J. T. BUCKLEY, 1997 Studies on the energetics of proaerolysin secretion across the outer membrane of *Aeromonas* species. Evidence for a requirement for both the protonmotive force and ATP. *J Biol Chem* **272**: 11109-11113.
- LI, F., K. P. PATRA and J. M. VINETZ, 2005 An anti-Chitinase malaria transmission-blocking single-chain antibody as an effector molecule for creating a *Plasmodium falciparum*-refractory mosquito. *J Infect Dis* **192**: 878-887.
- LINDEBERG, M., and A. COLLMER, 1992 Analysis of eight out genes in a cluster required for pectic enzyme secretion by *Erwinia chrysanthemi*: sequence comparison with secretion genes from other gram-negative bacteria. *J Bacteriol* **174**: 7385-7397.
- LOIRET, F. G., E. ORTEGA, D. KLEINER, P. ORTEGA-RODES, R. RODES *et al.*, 2004 A putative new endophytic nitrogen-fixing bacterium *Pantoea* sp. from sugarcane. *J Appl Microbiol* **97**: 504-511.
- MACNEIL, D. J., K. M. GEWAIN, C. L. RUBY, G. DEZENY, P. H. GIBBONS *et al.*, 1992 Analysis of *Streptomyces avermitilis* genes required for avermectin biosynthesis utilizing a novel integration vector. *Gene* **111**: 61-68.
- MAJANDER, K., L. ANTON, J. ANTIKAINEN, H. LÅNG, M. BRUMMER *et al.*, 2005 Extracellular secretion of polypeptides using a modified *Escherichia coli* flagellar secretion apparatus. *Nat Biotechnol* **23**: 475-481.
- MARRELLI, M. T., C. LI, J. L. RASGON and M. JACOBS-LORENA, 2007 Transgenic malaria-resistant mosquitoes have a fitness advantage when feeding on Plasmodium-infected blood. *Proc Natl Acad Sci U S A* **104**: 5580-5583.
- MAVRANGELOS, C., M. THIEL, P. J. ADAMSON, D. J. MILLARD, S. NOBBS *et al.*, 2001 Increased yield and activity of soluble single-chain antibody fragments by combining high-level expression and the Skp periplasmic chaperonin. *Protein Expr Purif* **23**: 289-295.
- MCMENIMAN, C. J., R. V. LANE, B. N. CASS, A. W. FONG, M. SIDHU *et al.*, 2009 Stable introduction of a life-shortening *Wolbachia* infection into the mosquito *Aedes aegypti*. *Science* **323**: 141-144.
- MENGE, D. M., T. GUDA, D. ZHONG, A. PAI, G. ZHOU *et al.*, 2005 Fitness consequences of *Anopheles gambiae* population hybridization. *Malar J* **4**: 44.
- MERGULHAO, F. J., D. K. SUMMERS and G. A. MONTEIRO, 2005 Recombinant protein secretion in *Escherichia coli*. *Biotechnol Adv* **23**: 177-202.
- MEYER, R. R., and P. S. LAINE, 1990 The single-stranded DNA-binding protein of *Escherichia coli*. *Microbiol Rev* **54**: 342-380.

- MIN, K. T., and S. BENZER, 1997 *Wolbachia*, normally a symbiont of *Drosophila*, can be virulent, causing degeneration and early death. *Proc Natl Acad Sci U S A* **94**: 10792-10796.
- MINAMINO, T., and K. NAMBA, 2004 Self-assembly and type III protein export of the bacterial flagellum. *J Mol Microbiol Biotechnol* **7**: 5-17.
- MOREIRA, L. A., A. K. GHOSH, E. G. ABRAHAM and M. JACOBS-LORENA, 2002a Genetic transformation of mosquitoes: a quest for malaria control. *Int J Parasitol* **32**: 1599-1605.
- MOREIRA, L. A., J. ITO, A. GHOSH, M. DEVENPORT, H. ZIELER *et al.*, 2002b Bee venom phospholipase inhibits malaria parasite development in transgenic mosquitoes. *J Biol Chem* **277**: 40839-40843.
- MOREIRA, L. A., J. WANG, F. H. COLLINS and M. JACOBS-LORENA, 2004 Fitness of anopheline mosquitoes expressing transgenes that inhibit *Plasmodium* development. *Genetics* **166**: 1337-1341.
- MORI, H., and K. ITO, 2001 The Sec protein-translocation pathway. *Trends Microbiol* **9**: 494-500.
- MORTIMER, P., and C. ARNOLD, 2001 FAFLP: last word in microbial genotyping? *J Med Microbiol* **50**: 393-395.
- NAKAMOTO, H., and J. C. BARDWELL, 2004 Catalysis of disulfide bond formation and isomerization in the *Escherichia coli* periplasm. *Biochim Biophys Acta* **1694**: 111-119.
- NAUM, M., E. W. BROWN and R. J. MASON-GAMER, 2008 Is 16S rDNA a reliable phylogenetic marker to characterize relationships below the family level in the enterobacteriaceae? *J Mol Evol* **66**: 630-642.
- NIARE, O., K. MARKIANOS, J. VOLZ, F. ODUOL, A. TOURE *et al.*, 2002 Genetic loci affecting resistance to human malaria parasites in a West African mosquito vector population. *Science* **298**: 213-216.
- NISHIHARA, K., M. KANEMORI, H. YANAGI and T. YURA, 2000 Overexpression of trigger factor prevents aggregation of recombinant proteins in *Escherichia coli*. *Appl Environ Microbiol* **66**: 884-889.
- NOUWEN, N., N. RANSON, H. SAIBIL, B. WOLPENSINGER, A. ENGEL *et al.*, 1999 Secretin PulD: association with pilot PulS, structure, and ion-conducting channel formation. *Proc Natl Acad Sci U S A* **96**: 8173-8177.
- PAGET, M. S., L. CHAMBERLIN, A. ATRIH, S. J. FOSTER and M. J. BUTTNER, 1999 Evidence that the extracytoplasmic function sigma factor sigmaE is required for normal cell wall structure in *Streptomyces coelicolor* A3(2). *J Bacteriol* **181**: 204-211.
- PALMER, T., F. SARGENT and B. C. BERKS, 2005 Export of complex cofactor-containing proteins by the bacterial Tat pathway. *Trends Microbiol* **13**: 175-180.
- PEABODY, C. R., Y. J. CHUNG, M. R. YEN, D. VIDAL-INGIGLIARDI, A. P. PUGSLEY *et al.*, 2003 Type II protein secretion and its relationship to bacterial type IV pili and archaeal flagella. *Microbiology* **149**: 3051-3072.
- PELOQUIN, J. J., C. R. LAUZON, S. POTTER and T. A. MILLER, 2002 Transformed bacterial symbionts re-introduced to and detected in host gut. *Curr Microbiol* **45**: 41-45.

- POSSOT, O. M., L. LETELLIER and A. P. PUGSLEY, 1997 Energy requirement for pullulanase secretion by the main terminal branch of the general secretory pathway. *Mol Microbiol* **24**: 457-464.
- PUSEY, P. L., 2002 Biological control agents for fire blight of apple compared under conditions limiting natural dispersal. *Plant Disease* **86**: 639-644.
- QUAKYI, I. A., R. CARTER, J. RENER, N. KUMAR, M. F. GOOD *et al.*, 1987 The 230-kDa gamete surface protein of *Plasmodium falciparum* is also a target for transmission-blocking antibodies. *J Immunol* **139**: 4213-4217.
- RAO, S., S. HU, L. MCHUGH, K. LUEDERS, K. HENRY *et al.*, 2005 Toward a live microbial microbicide for HIV: commensal bacteria secreting an HIV fusion inhibitor peptide. *Proc Natl Acad Sci U S A* **102**: 11993-11998.
- REDAK, R. A., A. H. PURCELL, J. R. LOPES, M. J. BLUA, R. F. MIZELL, 3RD *et al.*, 2004 The biology of xylem fluid-feeding insect vectors of *Xylella fastidiosa* and their relation to disease epidemiology. *Annu Rev Entomol* **49**: 243-270.
- RENER, J., P. M. GRAVES, R. CARTER, J. L. WILLIAMS and T. R. BURKOT, 1983 Target antigens of transmission-blocking immunity on gametes of *Plasmodium falciparum*. *J Exp Med* **158**: 976-981.
- RIEHLE, M. A., and M. JACOBS-LORENA, 2005 Using bacteria to express and display anti-parasite molecules in mosquitoes: current and future strategies. *Insect Biochem Mol Biol* **35**: 699-707.
- RIEHLE, M. A., C. K. MOREIRA, D. LAMPE, C. LAUZON and M. JACOBS-LORENA, 2007 Using bacteria to express and display anti-*Plasmodium* molecules in the mosquito midgut. *Int J Parasitol* **37**: 595-603.
- RIEHLE, M. A., P. SRINIVASAN, C. K. MOREIRA and M. JACOBS-LORENA, 2003 Towards genetic manipulation of wild mosquito populations to combat malaria: advances and challenges. *J Exp Biol* **206**: 3809-3816.
- ROSE, M. D., F. WINSTON and P. HIETER, 1990 *Methods in Yeast Genetics: A Laboratory Course Manual*. Cold Spring Harbor Laboratory Press.
- ROYCE, R. A., A. SENA, W. CATES, JR. and M. S. COHEN, 1997 Sexual transmission of HIV. *N Engl J Med* **336**: 1072-1078.
- SAMBROOK, J., and D. W. RUSSELL, 2001 *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press.
- SANDKVIST, M., 2001 Biology of type II secretion. *Mol Microbiol* **40**: 271-283.
- SAUER, F. G., K. FUTTERER, J. S. PINKNER, K. W. DODSON, S. J. HULTGREN *et al.*, 1999 Structural basis of chaperone function and pilus biogenesis. *Science* **285**: 1058-1061.
- SCHEIRLINCK, I., R. VAN DER MEULEN, A. VAN SCHOOR, M. VANCANNEYT, L. DE VUYST *et al.*, 2008 Taxonomic structure and stability of the bacterial community in belgian sourdough ecosystems as assessed by culture and population fingerprinting. *Appl Environ Microbiol* **74**: 2414-2423.
- SCHLAPSCHY, M., S. GRIMM and A. SKERRA, 2006 A system for concomitant overexpression of four periplasmic folding catalysts to improve secretory protein production in *Escherichia coli*. *Protein Eng Des Sel* **19**: 385-390.
- SHANKS, R. M., N. C. CAIAZZA, S. M. HINSA, C. M. TOUTAIN and G. A. O'TOOLE, 2006 *Saccharomyces cerevisiae*-based molecular tool kit for manipulation of genes from gram-negative bacteria. *Appl Environ Microbiol* **72**: 5027-5036.

- SINKINS, S. P., 2004 *Wolbachia* and cytoplasmic incompatibility in mosquitoes. *Insect Biochem Mol Biol* **34**: 723-729.
- SLETTA, H., A. NEDAL, T. E. AUNE, H. HELLEBUST, S. HAKVAG *et al.*, 2004 Broad-host-range plasmid pJB658 can be used for industrial-level production of a secreted host-toxic single-chain antibody fragment in *Escherichia coli*. *Appl Environ Microbiol* **70**: 7033-7039.
- SLETTA, H., A. TONDERVIK, S. HAKVAG, T. E. AUNE, A. NEDAL *et al.*, 2007 The presence of N-terminal secretion signal sequences leads to strong stimulation of the total expression levels of three tested medically important proteins during high-cell-density cultivations of *Escherichia coli*. *Appl Environ Microbiol* **73**: 906-912.
- SMITH, S. G., V. MAHON, M. A. LAMBERT and R. P. FAGAN, 2007 A molecular Swiss army knife: OmpA structure, function and expression. *FEMS Microbiol Lett* **273**: 1-11.
- SNOW, R. W., C. A. GUERRA, A. M. NOOR, H. Y. MYINT and S. I. HAY, 2005 The global distribution of clinical episodes of *Plasmodium falciparum* malaria. *Nature* **434**: 214-217.
- SPRENG, S., G. DIETRICH, W. GOEBEL and I. GENTSCHKEV, 1999 The *Escherichia coli* haemolysin secretion apparatus: a potential universal antigen delivery system in gram-negative bacterial vaccine carriers. *Mol Microbiol* **31**: 1596-1598.
- ST GEME, J. W., 3RD, and S. GRASS, 1998 Secretion of the *Haemophilus influenzae* HMW1 and HMW2 adhesins involves a periplasmic intermediate and requires the HMWB and HMWC proteins. *Mol Microbiol* **27**: 617-630.
- STEEN, H., and M. MANN, 2004 The ABC's (and XYZ's) of peptide sequencing. *Nat Rev Mol Cell Biol* **5**: 699-711.
- STEINER, D., P. FORRER, M. T. STUMPP and A. PLUCKTHUN, 2006 Signal sequences directing cotranslational translocation expand the range of proteins amenable to phage display. *Nat Biotechnol* **24**: 823-831.
- STRAIF, S. C., C. N. MBOGO, A. M. TOURE, E. D. WALKER, M. KAUFMAN *et al.*, 1998 Midgut bacteria in *Anopheles gambiae* and *An. funestus* (Diptera: Culicidae) from Kenya and Mali. *J Med Entomol* **35**: 222-226.
- SULTAN, A. A., V. THATHY, U. FREVERT, K. J. ROBSON, A. CRISANTI *et al.*, 1997 TRAP is necessary for gliding motility and infectivity of *Plasmodium* sporozoites. *Cell* **90**: 511-522.
- THIE, H., T. SCHIRRMANN, M. PASCHKE, S. DUBEL and M. HUST, 2008 SRP and Sec pathway leader peptides for antibody phage display and antibody fragment production in *E. coli*. *N Biotechnol* **25**: 49-54.
- TOLEIKIS, L., O. BRODERS and S. DUBEL, 2004 Cloning single-chain antibody fragments (scFv) from hybridoma cells. *Methods Mol Med* **94**: 447-458.
- TURELLI, M., and A. A. HOFFMANN, 1991 Rapid spread of an inherited incompatibility factor in California *Drosophila*. *Nature* **353**: 440-442.
- TZSCHASCHEL, B. D., C. A. GUZMAN, K. N. TIMMIS and V. DE LORENZO, 1996 An *Escherichia coli* hemolysin transport system-based vector for the export of polypeptides: export of Shiga-like toxin IIeB subunit by *Salmonella typhimurium aroA*. *Nat Biotechnol* **14**: 765-769.
- VALENT, Q. A., 2001 Signal recognition particle mediated protein targeting in *Escherichia coli*. *Antonie Van Leeuwenhoek* **79**: 17-31.

- VEGH, B. M., P. GAL, J. DOBO, P. ZAVODSZKY and F. VONDERVISZT, 2006 Localization of the flagellum-specific secretion signal in *Salmonella* flagellin. *Biochem Biophys Res Commun* **345**: 93-98.
- WANDERSMAN, C., and P. DELEPELAIRE, 1990 TolC, an *Escherichia coli* outer membrane protein required for hemolysin secretion. *Proc Natl Acad Sci U S A* **87**: 4776-4780.
- WEINER, J. H., P. T. BILOUS, G. M. SHAW, S. P. LUBITZ, L. FROST *et al.*, 1998 A novel and ubiquitous system for membrane targeting and secretion of cofactor-containing proteins. *Cell* **93**: 93-101.
- WHITTY, C. J., D. LALLOO and A. USTIANOWSKI, 2006 Malaria: an update on treatment of adults in non-endemic countries. *BMJ* **333**: 241-245.
- WINTER, G., A. D. GRIFFITHS, R. E. HAWKINS and H. R. HOOGENBOOM, 1994 Making antibodies by phage display technology. *Annu Rev Immunol* **12**: 433-455.
- WONG, T., A. AMIDI, A. DODDS, S. SIDDIQI, J. WANG *et al.*, 2007 Evolution of the Bacterial Flagellum, pp. in *ASM Microbe*.
- WULFING, C., and R. RAPPUOLI, 1997 Efficient production of heat-labile enterotoxin mutant proteins by overexpression of *dsbA* in a *degP*-deficient *Escherichia coli* strain. *Arch Microbiol* **167**: 280-283.
- YAMADA, Y., K. KATSURA, H. KAWASAKI, Y. WIDYASTUTI, S. SAONO *et al.*, 2000 *Asaia bogorensis* gen. nov., sp. nov., an unusual acetic acid bacterium in the alpha-Proteobacteria. *Int J Syst Evol Microbiol* **50 Pt 2**: 823-829.
- YONEKURA, K., S. MAKI-YONEKURA and K. NAMBA, 2003 Complete atomic model of the bacterial flagellar filament by electron cryomicroscopy. *Nature* **424**: 643-650.
- YOSHIDA, S., D. IOKA, H. MATSUOKA, H. ENDO and A. ISHII, 2001 Bacteria expressing single-chain immunotoxin inhibit malaria parasite development in mosquitoes. *Mol Biochem Parasitol* **113**: 89-96.
- YOSHIDA, S., H. MATSUOKA, E. LUO, K. IWAI, M. ARAI *et al.*, 1999 A single-chain antibody fragment specific for the *Plasmodium berghei* ookinete protein Pbs21 confers transmission blockade in the mosquito midgut. *Mol Biochem Parasitol* **104**: 195-204.
- YOSHIDA, S., Y. SHIMADA, D. KONDOH, Y. KOUZUMA, A. K. GHOSH *et al.*, 2007 Hemolytic C-type lectin CEL-III from sea cucumber expressed in transgenic mosquitoes impairs malaria parasite development. *PLoS Pathog* **3**: e192.
- YOUNG, G. M., D. H. SCHMIEL and V. L. MILLER, 1999 A new pathway for the secretion of virulence factors by bacteria: the flagellar export apparatus functions as a protein-secretion system. *Proc Natl Acad Sci U S A* **96**: 6456-6461.
- YOUNG, J. M., and D. C. PARK, 2007 Relationships of plant pathogenic enterobacteria based on partial *atpD*, *carA*, and *recA* as individual and concatenated nucleotide and peptide sequences. *Syst Appl Microbiol* **30**: 343-354.
- YUAN, S., H. DUAN, C. LIU, X. LIU, T. LIU *et al.*, 2004 The role of thioredoxin and disulfide isomerase in the expression of the snake venom thrombin-like enzyme calobin in *Escherichia coli* BL21 (DE3). *Protein Expr Purif* **38**: 51-60.
- ZENGEL, J. M., D. MUECKL and L. LINDAHL, 1980 Protein L4 of the *E. coli* ribosome regulates an eleven gene r protein operon. *Cell* **21**: 523-535.

- ZHANG, G., S. BROKX and J. H. WEINER, 2006 Extracellular accumulation of recombinant proteins fused to the carrier protein YebF in *Escherichia coli*. *Nat Biotechnol* **24**: 100-104.
- ZIELER, H., D. B. KEISTER, J. A. DVORAK and J. M. RIBEIRO, 2001 A snake venom phospholipase A(2) blocks malaria parasite development in the mosquito midgut by inhibiting ookinete association with the midgut surface. *J Exp Biol* **204**: 4157-4167.

APPENDIX 1

Isolation of Secondary Secreted Protein Candidates: *P. agglomerans flgL* and *ssb*

Objective and Summary

In order for a paratransgenic strain of *P. agglomerans* to effectively target *Plasmodium* development in the mosquito gut, the effector proteins that it expresses must be secreted from the bacterial cell. To this end, a search for native secreted proteins was performed using spent growth medium, 2D-PAGE resolution of the proteins present, and MALDI-TOF analysis to identify the proteins. The primary secreted protein candidate found in this search was FliC/Flagellin and information regarding the isolation of the *fliC* gene and its utilization are found in Chapter 2. Among the remaining proteins identified were FlgL (flagellar hook-associated protein 3) and Ssb (single-stranded DNA-binding protein). Experimental steps similar to those in Chapter 2 were taken to isolate the genes for these secondary secreted protein candidates. These steps involved genome walking PCR using degenerate primers designed from the peptide fragments identified with MALDI-TOF. In the case of *flgL*, the gene that was isolated was believed to be *E. coli* contamination because the sequence similarity to *E. coli flgL* was 99.7%. This contamination must have arisen at some point in the preparation of the genomic DNA template. As a result, the *flgL* gene that was isolated was not used for the creation of any

secretion constructs. The *P. agglomerans ssb* gene was successfully isolated and sequenced, and as a result, two secretion constructs were made using the *ssb* ORF. Ultimately, it is doubtful that Ssb is a true secreted protein as it is mainly involved in intracellular DNA processes. Ssb most likely appeared in the secreted protein profile due to cell lysis during growth of the culture or collection of the cells prior to protein precipitation.

Materials and Methods

MALDI-TOF analysis. This analysis is explained in Ch. 2 Materials and Methods.

Genome Walking PCR. This PCR technique is explained in Ch. 2 Materials and Methods.

Plasmid construction. The plasmids in this study were constructed using the yeast gap repair method as previously described by Shanks *et al* (SHANKS *et al.* 2006). Briefly, the vector to be modified was digested with restriction enzymes and treated with a phosphatase. DNA inserts were amplified with oligos that contained 40 bp of homology to the digested vector (up to 2000 bases from the digestion site). Agarose gel-purified vector (~20-200 ng) and inserts (50-500 ng), along with herring sperm DNA (100 µg) were co-transformed into 8-10 colonies of *S. cerevisiae* INVSc-1 cells (Invitrogen) resuspended in 500 µl of PLATE solution (50% PEG, 100 mM LiAc, 10 mM Tris, pH 7.4, 1 mM EDTA). The cells were incubated on the benchtop for ~2 d. After heat shock at 42°C for 30 min, the yeast cells were resuspended in 150 µl of deionized H₂O and plated on uracil drop-out medium (6.7 g/L yeast nitrogen base minus amino acids, 1.92 g/L yeast synthetic drop-out media supplement without uracil, 2% glucose (w/v) and 20 g/L bacteriological agar) and incubated at 30°C for up to 2 d. Total yeast DNA was

purified from the colonies using the “Yeast Smash and Grab DNA Miniprep” protocol (ROSE *et al.* 1990). Fifty nanograms of total yeast DNA was transformed into an appropriate strain of *E. coli* and bacterial clones were verified for the resultant plasmid by restriction enzyme digestion and DNA sequencing.

The yeast-replicating vector, pDB14, was built as follows. pACYC184 (New England Biolabs) was digested with *XmnI* for 1.5 h followed by treatment with calf intestinal phosphatase for 1 h at 37°C. The 2 μ m *ori* and *URA3* gene from the yeast cloning vector pMQ64 were amplified as a single PCR product with 40 bp of homology to pACYC184 and recombined into the *XmnI* cut site.

pDB24 was made by digesting pDB14 with *HindIII-EagI* and recombining in the *P. agglomerans fliC* 5'UTR (amplified from the *P. agglomerans* chromosome) with an *AscI* site at the 3' end, and the 6His and myc epitope sequences and stop codon from pIT2-*scFv*. Then, pDB27 was constructed by digesting pDB24 with *AscI* and recombining in the *malE* gene from pMALTM-c2X and *scFv* gene from pIT2-*scFv* (see Table A1.2). pDB27 allows for the expression of the fusion protein Male-Anti-BSA *scFv* under the control of the *P. agglomerans fliC* promoter. Finally, the *fliC* 5'UTR in pDB27 was replaced by digesting the vector with *EcoNI* and recombining the P_{*tac*} promoter from pMALTM-cRI to make pDB28.

To make pDB29, pDB28 was digested with *BglII* and the *P. agglomerans ssb* ORF (amplified from the *P. agglomerans* chromosome) was recombined in frame behind the anti-BSA *scFv* ORF by yeast recombination. To make pDB31, pDB27 was digested with *EcoNI* and the *P. agglomerans ssb* ORF was recombined in frame after the P_{*tac*} promoter and before the *malE* ORF using yeast recombination.

Results

Isolating flgL and ssb based on peptide sequence.

The *flgL* and *ssb* genes were isolated from the *P. agglomerans* chromosome using a genome walking PCR technique (GUO and XIONG 2006). Figure A1.5 and A1.6 show the *P. agglomerans* FlgL and Ssb peptide fragments that were identified and their relative placement in the *E. coli flgL* and *ssb* sequences (shown in green), as well as the site of degenerate PCR primer design (red arrows). The primers were based on Gram-negative codon usage tables and allowed for as much degeneracy as possible to ensure a successful amplification of the corresponding DNA sequence (SAMBROOK and RUSSELL 2001). The sequence of these primers and the series of universal “walking primers” are given in Table A1.3. PCR products ranging in size from 200-600 bp were cloned using TOPO technology (Invitrogen Corporation). The DNA sequence from these PCR products was used to generate a *P. agglomerans flgL* or *ssb* contig from which additional specific primers could be designed and sequential rounds of genome walking PCR could be performed. This procedure was repeated until the entire *P. agglomerans* gene and the flanking DNA up to and including some of the neighboring genes was isolated and sequenced (Figures A1.5 and A1.6).

TABLE A1.1 Strains Used in This Study

Strain	Relevant characteristics ^a	Source or Reference
<i>E. coli</i> Top10	F <i>mcrA</i> Δ(<i>mrr-hsdRMS-mcrBC</i>) ø80 <i>lacZ</i> ΔM15 Δ <i>lacX74 recA1 ara</i> Δ139 D(<i>ara-leu</i>)7697 <i>galU galK</i> <i>rpsL</i> (Str ^R) <i>endA1 nupG</i>	Invitrogen Corporation
<i>E. coli</i> Top10 F'	Top 10 with F'[<i>lacI</i> ^f Tn10 (Tc ^R)]	Invitrogen Corporation
<i>S. cerevisiae</i> INVSc-1	Sc1: <i>MATα his3D1 leu2 trp1-289 ura3-52</i> <i>MATα his3D1 leu2 trp1-289 ura3-52</i>	Invitrogen Corporation

^a Str^R, streptomycin resistance; Tc^R, tetracycline resistance; Nal^R, nalidixic acid resistance

TABLE A1.2 Plasmids Used in This Study

Plasmid	Relevant Characteristics ^a	Source or Reference
pACYC184	Cam ^R , Tc ^R ; cloning vector with p15A low copy <i>ori</i>	(CHANG and COHEN 1978)
pMQ64	Gent ^R ; yeast recombination vector and source of yeast <i>ori</i> (2μm) and yeast <i>URA3</i> gene	(SHANKS <i>et al.</i> 2006)
pMAL TM -c2X	Ap ^R ; source of <i>malE</i> (maltose binding protein) gene	New England Biolabs
pMAL TM -cRI	Ap ^R ; source of P _{tac}	New England Biolabs
pIT2- <i>scFv</i>	Ap ^R ; source of anti BSA <i>scFv</i> gene with 6His and myc epitopes	(DE WILDT <i>et al.</i> 2000)
pDB14	Cam ^R ; pACYC184/2μm <i>ori</i> and <i>URA3</i> gene	This study
pDB24	Cam ^R ; pDB14/ <i>P. agg fliC</i> 5'UTR- <i>AscI</i> -6His-myc- STOP	This study
pDB27	Cam ^R ; pDB24/ <i>P. agg fliC</i> 5'UTR- <i>malE</i> -anti BSA <i>scFv</i>	This study
pDB28	Cam ^R ; pDB14/P _{tac} - <i>malE</i> -anti-BSA <i>scFv</i>	This study
pDB29	Cam ^R ; pDB28/P _{tac} - <i>malE</i> -anti-BSA <i>scFv-ssb</i>	This study
pDB31	Cam ^R ; pDB27/ P _{tac} - <i>ssb-malE</i> -anti BSA <i>scFv</i>	This study

^a Ap^R, ampicillin resistance; Cam^R, chloramphenicol resistance; Gent^R, gentamycin resistance; Tc^R, tetracycline resistance

Table A1.3: Oligonucleotides used in this Study

Oligonucleotide	Sequence 5' - 3' ^a	Purpose
M13 Forward (-20)	GTAAAACGACGGCCAG	Sequencing TOPO inserts
M13 Reverse	CAGGAAACAGCTATGAC	Sequencing TOPO inserts
Semi-1	GCCAATTCGGATNGAYKSNGGNTC	Arbitrary walking primer
Semi-2	GCCTTAAGGCCTANGARMSNCCNAG	Arbitrary walking primer
Semi-3	CGGTTAAGGCCTANYTCSKNGANGC	Arbitrary walking primer
Semi-4	GCCAATTCGGATNSAGYMNCTNCG	Arbitrary walking primer
FlgL134-140 For	TAYATYTTYGCNGGTAYAAAR	Genome Walking PCR
FlgL203-212 Rev1	NGTYTTNARNGCNGCRATNGC	Genome Walking PCR
FlgL236-247 Rev2	NACNGTNARNACRTRTTNA	Genome Walking PCR
flgL-nest1-Left	ATGCGCWTWAGYACTMRMATGATG	Nested Genome Walking PCR
flgL-nest2-Left	TACCAGCAAAAYATGCRWGGGARYAYT G	Nested Genome Walking PCR
flgL-nest1-Right	CTGGAAMAGMGACATTCCTGC	Nested Genome Walking PCR
flgL-nest2-Right	ATCGCTRAAYGCTTTATASGAWGCSTG	Nested Genome Walking PCR
flgL check Left	AATGGCACCTTGAGTGACGATG	Verify <i>P. agg flgL</i> gene
flgL check Right	GTGCAATGCCGTTTGCTGCATG	Verify <i>P. agg flgL</i> gene
Ssb5-22 For	GGNGTNAAYAARGTNATYYTNG	Genome Walking PCR
Ssb74-85 Rev1	YTGNCYTCRATRTANACYTG	Genome Walking PCR
Ssb98-116 Rev2	CATYTGATNGTNCNCCNACRTTNAC	Genome Walking PCR
ssb-nest1-Left	ATGGCCAGCAGAGGCGTWAAYAARG	Nested Genome Walking PCR
ssb-nest2-Left	ATTCTYGTYGGKAATCTGGGYCARG	Nested Genome Walking PCR
ssb-nest1-Right	GAAMGGAATRTCRTTCRAARTCCATYGG	Nested Genome Walking PCR
ssb-nest2-Right	CATYGGSGGYTCGTTAGACGG	Nested Genome Walking PCR

^a R = A or G; Y = C or T; N = any nucleotide; K = G or T; M = A or C; S = G or C; *P. agg.*, *P. agglomerans*

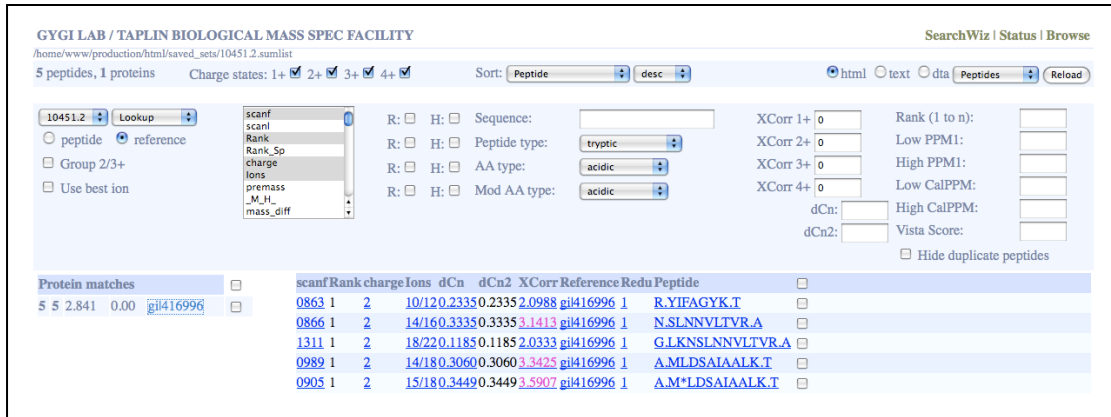


Figure A1.1 FlgL MALDI-TOF Results from Taplin Biological Mass Spec Facility.

Results from MALDI-TOF analysis of each individual spot are viewed by signing in to the Taplin Biological Mass Spec Facility website with a provided username and password. In the lower left-hand corner is a menu of protein matches that shows the number of peptide fragments isolated for this particular spot (5 total). Each ID number takes the user to a screen that shows where each of the peptide fragments align to a known protein in the NCBI Enterobacteriaceae database. In this case, the peptide fragments matched to *E. coli* FlgL.

GYGI LAB / TAPLIN BIOLOGICAL MASS SPEC FACILITY

Protein and Peptide Details for reference gi416996

Analyzing database: /facility/facility_data/Facilityparams/Ecoli.fasta

Create protein database: /lab/tmp/gi_416996_sp_P29744_FLGL.fasta

gi|416996|sp|P29744|FLGL_ECOLI Flagellar hook-associated protein 3 (HAP3) (Hook-filament junction protein)

MRFSTQMMYQ QNMRGITNSQ AEWKMYGEQN STGKRVVNS DDPIAASQAV VLSQAQAQNS QYTLARTPAT QKVSLEESVL SQVTTAIQNA QEKIVYASNG TLSDDDRASL ATDIQGLRDQ LLNLANTTDG
 NGRY**YIFAGYK** TETAPFSEER GKIVYGAESI KQVVDASRM VIGHTGDKIF DSITSNVAE PDGSAEETNL F**MLDSATAIA** LKTPVADSEA DKETAAALD KTNRGL**KNSL** **HNVLTVRAEL** GTQLNELES
 DSLGSDRALG QTQQMSDLVD VDWNATISSY IMQQTALQAS YKAFDMQGL SLFQLSK

Position	MH+	Sequence (click to highlight)
134-140	862.0160	YIFAGYK
239-247	1016.1856	SLNNVTVR
203-212	1033.2713	MLDSATAIA
236-247	1371.6229	LKNLNNVTVR

Search Links

[BLAST entire protein sequence](#)

Master Database Sources

unknown

Search on [Google](#)

Protein coverage by amino acid count: 29/317 = 9.1%

Protein coverage by mass: 3228/34282 = 9.4%

Figure A1.2: Peptide fragment matches to *E. coli* FlgL from Taplin Biological Mass Spec Facility. In the case of the spot that matched to *E. coli* FlgL, five peptide fragments were generated upon tryptic digestion. This screenshot shows where 4 of the 5 fragments matched to the *E. coli* FlgL protein sequence present in the NCBI Enterobacteriaceae protein database (2 fragments are overlapping). The amino acid positions are given as well as the amino acid sequence of the fragment.

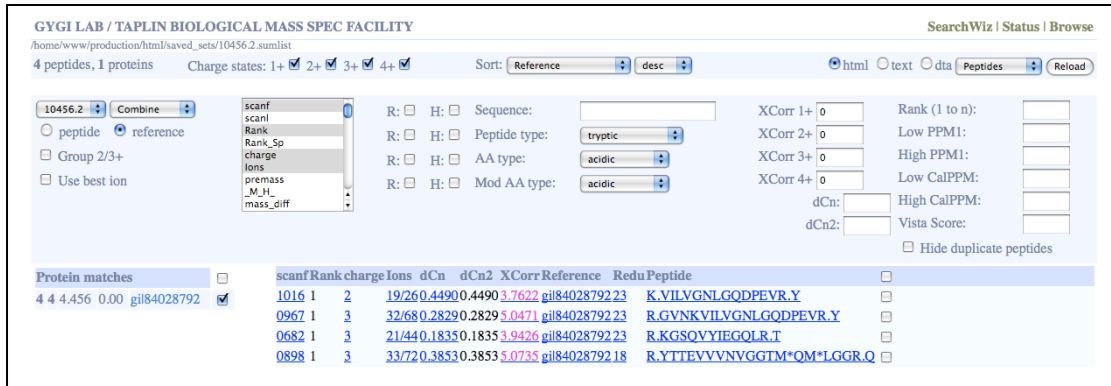


Figure A1.3: Ssb MALDI-TOF Results from Taplin Biological Mass Spec Facility.

Results from MALDI-TOF analysis of each individual spot are viewed by signing in to the Taplin Biological Mass Spec Facility website with a provided username and password. In the lower left-hand corner is a menu of protein matches that shows the number of peptide fragments isolated for this particular spot (4 total). Each ID number takes the user to a screen that shows where each of the peptide fragments align to a known protein in the NCBI Enterobacteriaceae database. In this case, the peptide fragments matched to *E. coli* Ssb.

GYGI LAB / TAPLIN BIOLOGICAL MASS SPEC FACILITY
 Protein and Peptide Details for reference gi84028792

Analyzing database: /facility/facility_data/Facilityparams/Ecoli.fasta
 Create protein database : /lab/tmp/gi_84028792_sp_POAGE0_SS.fasta

gi|84028792|sp|POAGE0|SSB_ECOLI Single-stranded DNA-binding protein (SSB) (Helix-destabilizing protein)

MASR**GVN**KV**ILVGNL**QD**PEVR**YMPNGGAV ANITLTSSES WRDKATGEMK EQTEWHRVVL FGKLAEVASE YLR**KGSQ**VY**IEQ**L**TR**TKWT DQSGQDR**YTT** **EVVV**V**GGTM** **QMLGGR**QGGG APAGGNIGGG
 QPQGGWQFQ QPQGGNQFSG GAQSRPQQA PAAPSNEPMM DFDDDIPF

Position	MH+	Sequence (click to highlight)
74-85	1378.5711	KGSQVYIEQLR
9-22	1509.7458	VILVGNLQDPEVR
5-22	1908.2082	GVNKVILVGNLQDPEVR
98-116	2013.3346	YTTTEVVVNVGGTMQLGGR

Search Links
[BLAST entire protein sequence](#)
 Master Database Sources
 unknown
 Search on [Google](#)

Protein coverage by amino acid count: 49/178 = 27.5%
 Protein coverage by mass: 5262/18976 = 27.7%

Figure A1.4: Peptide fragment matches to *E. coli* Ssb from Taplin Biological Mass Spec Facility. In the case of the spot that matched to *E. coli* Ssb, four peptide fragments were generated upon tryptic digestion. This screenshot shows where the four fragments (2 were overlapping) matched to the *E. coli* Ssb protein sequence present in the NCBI Enterobacteriaceae protein database. The amino acid positions are given as well as the amino acid sequence of the fragment.

A.

```
1 mrfstqmmq qnmgitnsq aewmkygeqm stgkrvvnps ddpiaasqav vlsqaqaqns
61 qytlartfat qkvsleesvl sqvttaiqna qekivyasng tlsdddrasl atdiqglrdq
121 llnlanttdg ngryifagyk tetapfseek gkyvgaesi kqqvdasrsm vightgdkif
181 dsitsnavae pdgsasetnl famldsaiiaa lktpvadsea dketaaaald ktnrg lknsl
241 nnvltvrael gtqlnelesl dslgsdralg qtqqmsdlvd vdwnatissy imqqtalqas
301 ykaftdmqgl slfqlsk
```

B.

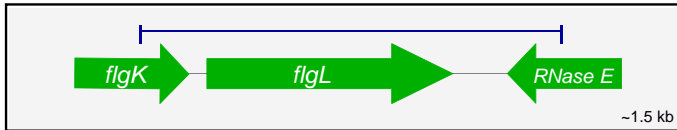


Figure A1.5: Identity and placement of sequenced *P. agglomerans* FlgL peptide fragments in the *E. coli* FlgL sequence and the *P. agglomerans* *flgL* contig assembled after genome walking PCR. Five *P. agglomerans* FlgL peptide fragments were sequenced using MALDI-TOF and identified based on homology to the *E. coli* FlgL sequence. A1.5A shows where the fragments (2 fragments are overlapping) match to the *E. coli* FlgL sequence (residues in green). The red arrows indicate where the degenerate primers were designed (Table A1.3). A1.5B shows the resulting *P. agglomerans* *flgL* contig assembled after genome walking PCR. The blue bracket roughly indicates the length of sequence obtained.

A.

```
1 masrgvnkvi lvgnlqgdpe vrympnggav anitlatses wrdkatgemk eqtewhrvvl
61 fgklaevase ylrkgsqvvi egqlrtrkwt dqsgqdrytt evvvhvvggtm qmlggrqggg
121 apaggniggg qpqggwgpq qpqggnqfsg gaqsrpqgsa paapsneppm dfdddipf
```

B.

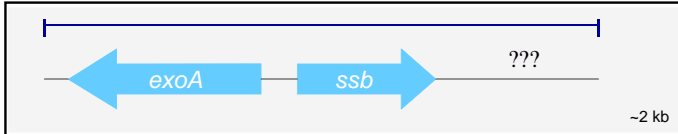


Figure A1.6: Identity and placement of sequenced *P. agglomerans* Ssb peptide fragments in the *E. coli* Ssb sequence and the *P. agglomerans* *ssb* contig assembled after genome walking PCR. Four *P. agglomerans* Ssb peptide fragments were sequenced using MALDI-TOF and identified based on homology to the *E. coli* Ssb sequence. A1.6A shows where the fragments (2 fragments are overlapping) match to the *E. coli* Ssb sequence (residues in green). The red arrows indicate where the degenerate primers were designed (Table A1.3). A1.6B shows the resulting *P. agglomerans* *ssb* contig assembled after genome walking PCR. The blue bracket roughly indicates the length of sequence obtained. Because a downstream ORF was not found, it is likely additional “walking” is needed in order to find the next gene.

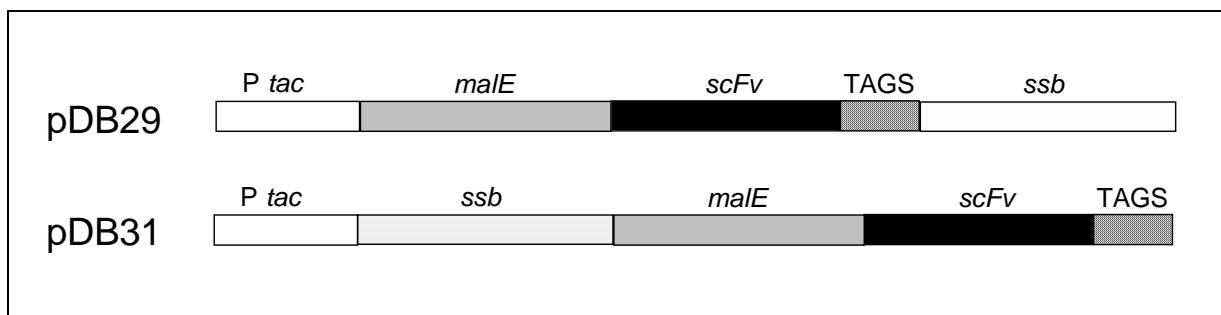


Figure A1.7: Secretion constructs using *P. agglomerans* *ssb* ORF. Two constructs were made using the *P. agglomerans* *ssb* ORF as the secretion signal for export of the MalE-anti-BSA scFv fusion protein (*malE-scFv*). These plasmids were made using yeast recombination. When expressed, the Ssb protein is fused to the C-terminal end of the passenger protein in pDB29, and also fused to the N-terminal end of the passenger protein in pDB31. The P_{tac} promoter is driving the expression of the genes. The coding sequence for these proteins contains the 6His and myc epitope tags (“TAGS”).

APPENDIX 2

Growth Curve of Wild-type *P. agglomerans*

Objective and Summary

A growth curve of wild-type *P. agglomerans* cells grown in Luria-Bertani (LB) broth was constructed. The objective was to collect spent medium for 2D-PAGE during late-*log* phase, which would allow for the highest concentration of secreted protein and the lowest concentration of lysed cells in the culture.

A 5 ml overnight culture of *P. agglomerans* in LB broth was grown at 30°C in a shaking incubator set to 250 rpm. On the next day, a 20 ml LB culture was established using a 1/100 dilution of the overnight culture. Incubation continued at 30°C in the shaking incubator and a 1 ml aliquot was removed every hour for an OD₆₀₀ reading in a Perkin Elmer MBA 2000 spectrophotometer. Deionized water was used as a blank. After four hours, a 1/10 dilution of the 1 ml aliquot of the culture was used to obtain a reading that was still within range of the spectrophotometer.

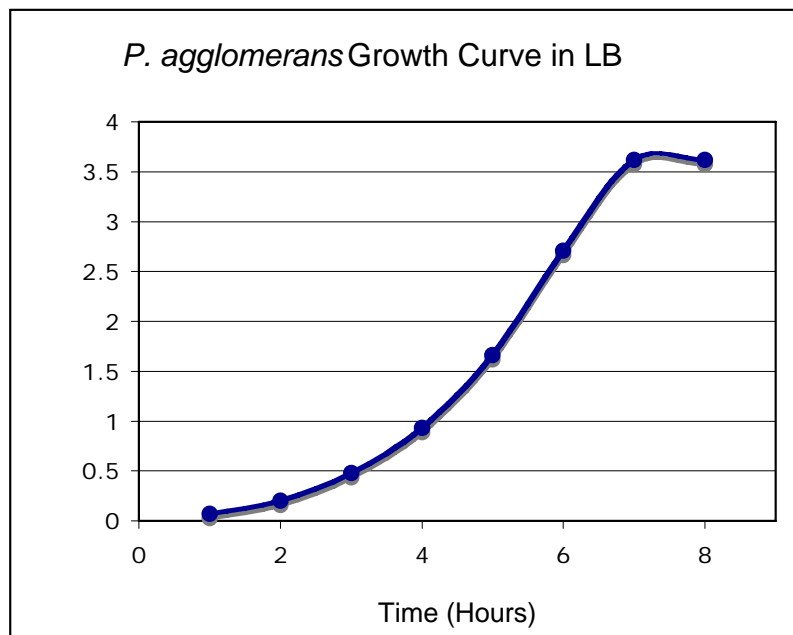


Figure A2.1: Growth curve of Wild-type *P. agglomerans* in LB Broth. This graph shows the logarithmic growth of wild-type *P. agglomerans* cells in LB broth. Samples were collected every hour for a total of 8 hours and the OD₆₀₀ reading was plotted on the y-axis.

APPENDIX 3

Parallel 2D-PAGE preparations for MALDI-TOF analysis of secreted *P. agglomerans* proteins

Objective and Summary

The Taplin Biological Mass Spec Facility at Harvard Medical School encourages researchers submitting samples for MALDI-TOF analysis to only send samples that are detectable at the Coomassie stain level. Because the majority of spots on the 2D-PAGE gels of secreted *P. agglomerans* proteins were detectable with the more sensitive reagent silver stain, in order to increase the amount of protein per spot, parallel gels were run and identical spots were cored from the gel and combined in a single tube prior to sending to the facility. The gel shown in Figure 2.1 and the six gels shown here in Appendix 3 (Figures A3.1 and A3.2) represent the source of all of the spots sent for identification at the Taplin Facility. The spots (labeled A-O) were chosen for identification based on reproducibility of the same spot from gel to gel and the relative intensity compared to more faint spots on the gels.

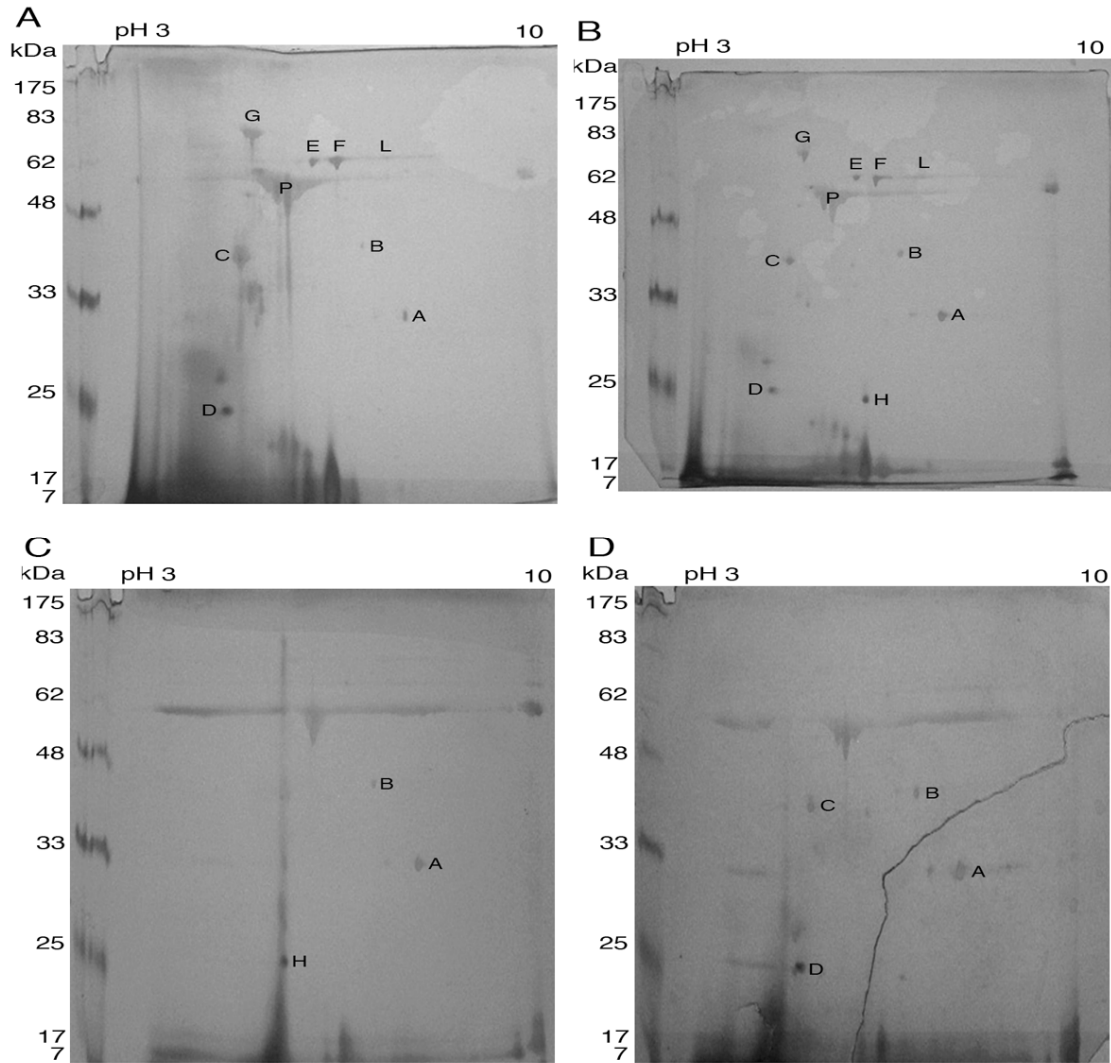


Figure A3.1: Parallel gels of late-log protein preparations of *P. agglomerans* secreted proteins. In order to increase the amount of protein per sample (spot) sent for identification, parallel 2D-PAGE gels were generated. Proteins were separated by isoelectric focusing on a pH gradient of 3-10 in the first dimension. Proteins were then separated by molecular weight on a 10% acrylamide gel in the second dimension. Proteins were visualized with silver stain and identical protein spots were cored from the gel and combined in a single tube. The letter labeling system corresponds to the labels in Figure 2.1.

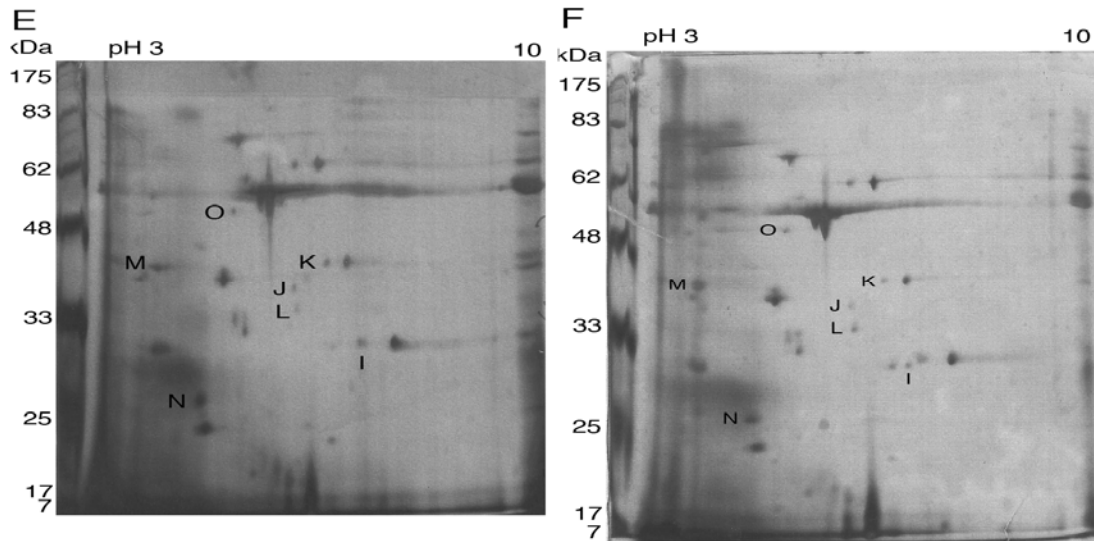


Figure A3.2: Additional parallel gels of late-*log* protein preparations of *P. agglomerans* secreted proteins.

In order to increase the amount of protein per sample (spot) sent for identification, parallel 2D-PAGE gels were generated. Proteins were separated by isoelectric focusing on a pH gradient of 3-10 in the first dimension. Proteins were then separated by molecular weight on a 10% acrylamide gel in the second dimension. Proteins were visualized with silver stain and identical protein spots were cored from the gel and combined in a single tube. The letter labeling system corresponds to the labels in Figure 2.1.