

BIOCHEMICAL CHARACTERIZATION OF SERPINS IN THE MALARIA VECTOR,  
*ANOPHELES GAMBIAE*

by

MELISSA M GULLEY

B.S., Kansas State University, 2007

A THESIS

Submitted in partial fulfillment of the requirements for the degree

MASTER OF SCIENCE

Division of Biology  
College of Arts and Sciences

KANSAS STATE UNIVERSITY  
Manhattan, Kansas

2013

Approved by:

Major Professor  
Dr. Kristin Michel

## Abstract

To date malaria is the most important tropical disease, which is caused by *Plasmodium sp.* and vectored by anopheline mosquitoes. The mosquito's immune system is one of the limiting factors of malaria transmission. Immune reactions, such as the prophenoloxidase (PPO) pathway result in the melanization of pathogens, and are effective at limiting parasite numbers. Novel strategies for malaria control aim to exploit the immune system to interrupt parasite transmission by boosting the immune responses in the mosquito vector.

Serpins play a crucial role in regulating protease cascades involved in immunity of arthropods. In *Anopheles gambiae*, the major malaria vector in Sub-Saharan Africa, 18 SRPN genes encoding 23 distinct proteins have been identified. So far, two are identified as active inhibitors, and both affect parasite survival. This research aims to identify additional inhibitory serpins in *An. gambiae* and elucidate their potential function. Identification of such serpins will enhance our understanding of the immune system of this important vector species and may identify immunoregulators to be used in malaria control.

SRPN7, 9, and 18 were tested for their ability to inhibit commercial proteases *in vitro*. Recombinant SRPN18 had no inhibitory activity, while SRPN7 and 9 inhibited several serine proteases. SRPN7, 9 and 18 were tested against two recombinant *An. gambiae* clip serine proteases (CLIPBs) that are required for activation of phenoloxidase and thus regulate melanization. Only SRPN9 strongly inhibited CLIPB9 *in vitro*, suggesting that this serpin is a potential negative regulator of melanization. This hypothesis is further supported by the finding that SRPN9 can inhibit PO activity in insect hemolymph, *ex vivo*.

Taken together, this research identifies SRPN18 as the first non-inhibitory serpin described in mosquitoes. Additionally, this study describes the larval-specific SRPN7 as a functional inhibitor. Future studies on these proteins will elucidate their precise physiological functions. Finally, this thesis provides strong evidence that SRPN9 is a negative regulator of melanization in *An. gambiae* and may therefore affect pathogen survival within this important vector species.

# Table of Contents

|   |      |
|---|------|
| List of Figures .....   | vii  |
| List of Tables .....  | viii |
| Acknowledgements.....   | ix   |
| Dedication .....  | x    |
| Chapter 1 - Introduction.....                                     | 1    |
| 1.1 Malaria .....   | 1    |
| 1.2. Current and novel malaria control strategies .....           | 3    |
| 1.3 Immunity.....   | 5    |
| 1.3.1 Recognition .....   | 5    |
| 1.3.2 The Toll Pathway.....                                       | 7    |
| 1.3.3 The Phenoloxidase Activation Pathway .....                  | 9    |
| 1.3.4 Protease Cascades .....                                     | 11   |
| 1.3.5 Serpins.....  | 11   |
| 1.3.5.1 Structure .....   | 11   |
| 1.3.5.2 Function .....  | 12   |
| 1.3.5.3 SRPNs Encoded in <i>An. gambiae</i> Genome.....           | 13   |
| 1.4 References.....   | 15   |
| 1.5 Figures Legends .....   | 22   |
| 1.6 Figures .....   | 24   |
| Chapter 2 - Biochemical characterization of SRPN7 and SRPN18..... | 30   |
| 2.1 Abstract.....   | 30   |
| 2.2 Introduction.....   | 30   |
| 2.3 Materials and methods .....                                   | 33   |
| 2.3.1 Mosquito Rearing.....                                       | 33   |
| 2.3.2 SRPN7 and 18 Cloning .....                                  | 33   |
| 2.3.3 Recombinant Protein Purification.....                       | 34   |
| 2.3.4 SDS-PAGE electrophoresis and Western Blotting .....         | 35   |
| 2.3.5 Amidase Activity assays (Commercial).....                   | 36   |

|   |    |
|---|----|
| 2.3.6 Blood Feeding .....   | 36 |
| 2.3.7 qRT-PCR.....  | 37 |
| 2.3.9 Activation of CLIPB8, B9 .....  | 37 |
| 2.3.10 Florescence assays. (B8/B9) .....  | 37 |
| 2.4 Results.....  | 38 |
| SRPN7 and 18 RCL implicate SRPN7 as an inhibitory serpin with trypsin-like specify<br>while SRPN18 is unlikely to be inhibitory. .... | 38 |
| Recombinant protein and purification yield pure proteins. ....  | 38 |
| SRPN18 transcripts do not increase after blood meal .....   | 39 |
| SRPN7 inhibits Trypsin like serine proteases.....   | 39 |
| SRPN7 and 18 do not inhibit CLIPB8 or CLIPB9 .....  | 40 |
| 2.5 Discussion.....   | 40 |
| 2.6 References.....   | 43 |
| 2.7 Tables.....   | 45 |
| 2.8 Figure Legends .....  | 47 |
| 2.9 Figures .....   | 49 |
| 2.10 Acknowledgements.....  | 56 |
| Chapter 3 - Biochemical Characterization of SRPN9.....  | 57 |
| 3.1 Abstract.....   | 57 |
| 3.2 Introduction.....   | 58 |
| 3.3 Materials and methods .....   | 59 |
| 3.3.1 Mosquito Rearing.....   | 59 |
| 3.3.2 Mosquito Injection .....  | 60 |
| 3.3.3 SRPN9 Cloning.....  | 60 |
| 3.3.4 Recombinant Proteins .....  | 60 |
| 3.3.5 Recombinant Protein Purification.....   | 61 |
| 3.3.6 SRPN9 Polyclonal Antibody Production.....   | 62 |
| 3.3.7 SDS-PAGE electrophoresis and Western Blotting .....   | 62 |
| 3.3.8 Amidase Activity assays (Commercial).....   | 63 |
| 3.3.9 Blood Feeding .....   | 63 |
| 3.3.10 RNA Isolation .....  | 64 |

|   |    |
|---|----|
| 3.3.11 cDNA Synthesis.....  | 64 |
| 3.3.12 qRT-PCR.....   | 64 |
| 3.3.13 dsRNA synthesis.....   | 65 |
| 3.3.14 PO activity assays.....  | 65 |
| 3.3.15 Activation of CLIPB8, B9.....  | 66 |
| 3.3.16 Florescence assays. (B8/B9).....   | 66 |
| 3.3.17 Stoichiometry of Inhibition.....   | 66 |
| 3.3.18 Association Constant.....  | 67 |
| 3.4 Results.....  | 67 |
| 3.4.1 SRPN9 has a canonical reactive center loop and is predicted to be inhibitory.....                   | 67 |
| 3.4.2 Recombinant SRPN9 expression and purification yielded pure protein.....                             | 67 |
| 3.4.3 SRPN9 depletion does not affect mosquito survival.....  | 68 |
| 3.4.4 SRPN9 increases after blood meal.....   | 68 |
| 3.4.5 SRPN9 inhibits trypsin-like serine proteases.....   | 68 |
| 3.4.6 rSRPN9 decreases M. sexta phenoloxidase activity.....   | 69 |
| 3.4.7 rSRPN9 forms SDS stable complex with rCLIPB9 but not rCLIPB8.....                                   | 69 |
| 3.4.8 rSRPN9 to rCLIPB9 has low SI and similar $k_a$ as other arthropod serpin protease combinations..... | 70 |
| 3.5 Discussion.....   | 71 |
| 3.6 References.....   | 74 |
| 3.7 Tables.....   | 76 |
| 3.8 Figure Legends.....   | 80 |
| 3.9 Figures.....  | 83 |
| 3.10 Acknowledgements.....  | 94 |
| Appendix A.....   | 95 |

## List of Figures

|   |    |
|---|----|
| Figure 1-1: Global map of malaria transmission and use of IRS or LLAN in areas of high transmission. .... | 24 |
| Figure 1-2: <i>Plasmodium sp.</i> Lifecycle .....   | 25 |
| Figure 1-3: Global map of malaria vectors .....   | 26 |
| Figure 1-4: Overview of <i>An. gambiae</i> immune response .....  | 27 |
| Figure 1-5: Melanogenesis pathway in insects .....  | 28 |
| Figure 1-6: Serpin inhibition mechanism .....   | 29 |
| Figure 2-1: Amino acid sequence alignment of reactive center loops from AgSRPNs and homologues .....      | 49 |
| Figure 2-2: SRPN 7 Recombinant Protein Expression and Purification .....                                  | 50 |
| Figure 2-3: SRPN 18 Recombinant Protein Expression and Purification .....                                 | 51 |
| Figure 2-4: Blood meal has no effect on SRPN18 transcript abundance. ....                                 | 52 |
| Figure 2-5: SRPN inhibition of commercial proteases .....   | 53 |
| Figure 2-6: SRPN7 and 18 do not inhibit CLIPB8 or B9 .....  | 54 |
| Figure 2-7: SRPN18 does not form complex with CLIPB8 or B9 .....  | 55 |
| Figure 3-1: Amino acid sequence alignment of reactive center loops from AgSRPN9 and orthologs .....       | 83 |
| Figure 3-2: Nucleic acid and deduced amino acid sequence of SRPN9.....                                    | 84 |
| Figure 3-3: Protein purification of recombinant SRPN9 .....   | 85 |
| Figure 3-4: SRPN9 depletion does not affect survival .....  | 86 |
| Figure 3-5: SRPN9 expression after blood meal time course.....  | 87 |
| Figure 3-6: Recombinant AgSRPN9 inhibits trypsin-like serine proteases.....                               | 88 |
| Figure 3-7: SRPN9 inhibition of PO activity in <i>M. sexta</i> hemolymph.....                             | 89 |
| Figure 3-8: SRPN9 inhibits and forms complex with CLIPB9 .....  | 90 |
| Figure 3-9: High molecular weight band is complex of SRPN9 and CLIPB9 .....                               | 91 |
| Figure 3-10: Stoichiometry of inhibition of recombinant SRPN9 to CLIPB9 .....                             | 92 |
| Figure 3-11: $K_a$ of recombinant SRPN9 to CLIPB9 .....   | 93 |

## List of Tables

|   |    |
|---|----|
| Table 2.1: <i>SRPN7 and 18 inhibition of commercially available proteases</i> ..... | 45 |
| Table 2.2: <i>Primer sequences</i> .....  | 46 |
| Table 3.1: <i>SRPN9 commercial protease assays</i> .....                            | 76 |
| Table 3.2: <i>Inhibition of protease activity by recombinant SRPN9</i> .....        | 77 |
| Table 3.3: <i>MS/MS Confirms SRPN9 forms complex with CLIPB9</i> .....              | 78 |
| Table 3.4: <i>Primer sequences</i> .....  | 79 |

## **Acknowledgements**

I would like to thank my committee members, Dr Kristin Michel, Dr Maureen Gorman, and Dr. Rollie Clem for their time, expertise, and suggestion for my research and help in writing my thesis. I would especially like to thank Dr. Kristin Michel, my PI and mentor through this process. Without her knowledge and guidance I would have been lost. She epitomizes what a great mentor and scientist should be. Thank you Kristin!

I would also like to thank the members of the Michel lab for their support, kindness and help. I would especially like to thank Mrs. Tinea Graves, Dr. Chunju An, Dr. Xin Zhang, and Mrs. Karajo Sprigg. Tinea, Chunju and Xin went out of their way on numerous occasions to teach me molecular and biochemical techniques. Additionally, they were all very helpful in troubleshooting when my experiments went awry. Karajo and Tinea were always volunteering to aid in experiments when I was called away due to sick kids. Finally, I would like to thank each of them for providing friendship and camaraderie in the lab.

I would like to acknowledge Insect Biochemistry seminar members for their insightful suggestions and comments.

I would also like to thank my family for bearing with me when I was stressed and for supporting me so that I could pursue this degree. Thank you for your understanding, support and sacrifices that you have made!

I would also like to thank my sister, Tamera Christensen who thankfully watched my children for me whenever she could so that I could get into the lab and work. Thank you!

## **Dedication**

I would like to dedicate this work to my family. To my wonderful husband Justin and my son's: Jayden, Emery and Caleb.

# Chapter 1 - Introduction

## List of Abbreviations

|   |   |
|---|---|
| EIP- Extrinsic incubation period            | PAP- Prophenoloxidase activating protease |
| LLIN- Long lasting insecticide treated nets | SRPN- Serine Protease inhibitor           |
| IRS- Indoor residual spraying               | CLIP-SP- Clip domain serine protease      |
| PPO- Prophenoloxidase                       | AMP- Antimicrobial peptides               |
| PO- Phenoloxidase                           | MS- Mass Spectrometry                     |
|   | SPH- Serine Protease Homologue            |

## 1.1 Malaria

Malaria is a devastating disease that in 2010 alone killed 660,000 people. The largest numbers of these deaths are in Africa and mostly children under five (WHO 2012). It is especially a problem in Sub-Saharan Africa where high mortality rates are seen. Malaria parasites, including *Plasmodium falciparum*, the most lethal human parasite, are transmitted via female *Anopheles* mosquitoes (Ross, 1899). To date, vector control methods rely primarily on long-lasting insecticide treated bednets (LLIN) and indoor residual spraying (IRS) (Fig 1.1). Currently, there are no effective sterilizing vaccines available and larval insecticide targets are difficult to utilize because female Anopheline mosquitoes may lay their eggs in small pools of water (i.e. hoofmarks and old tires filled with rainwater). Coupled with increasing insecticide resistance, it is crucial that novel malarial control methods are implemented to control this disease. The Malaria Eradication Research Agenda (malERA) has acknowledged the need for new vector control strategies aimed at preventing completion of *Plasmodium sp.* lifecycle within the vector (Control, 2011).

Since its discovery in 1880 by Charles Laveran (Laveran, 1982), research has continued to uncover facts about the human malaria parasite. There are five main malarial *Plasmodium sp.* species infecting humans, *P. falciparum*, *P. vivax*, *P. ovale*, *P. malariae* and *P. knowlesi* (Oddoux et al., 2011). To be infective, each of these must fulfill a complex life cycle in both the definitive as well as intermediate hosts.

The definitive host is defined as the organism in which parasitic sexual maturity is reached. In *Plasmodium sp.*, the definitive host is the mosquito, which also acts as a vector to human populations. Some mosquitoes are anthropophilic and preferentially bite certain mammals. For example, *An. gambiae* and *Anopheles dirus*, malaria vectors of Africa and Southeast Asia, respectively, are anthropophilic.

The lifecycle of the malaria parasite is complex. A female mosquito takes a blood meal from an infected individual and ingests gametocytes. In the lumen of the midgut the gametocytes then develop into male and female gametes. The gametes are fertilized and form zygotes, which develop into motile ookinetes that breach the peritrophic matrix and midgut epithelial cells approximately one day after ingestion of infective blood meal (Fig 1.2). This breaching of the midgut is a major immune stimulating event causing midgut epithelial cells to slough off in what is described in a “time bomb” model (Han et al., 2000). The immune response mounted by the mosquito significantly decreases parasitic load. Proteins which are present in the mosquito hemolymph may cause parasite encapsulation and lysis. Ookinetes that survive the immune response continue to develop into young then mature oocysts, which undergo sporogony. Sporogony results in the production of sporozoites and ultimately leads to the release of thousands of sporozoites into the mosquito hemocoel. They then travel to the salivary glands where they penetrate a second epithelial layer. From there, they are mixed with saliva and can

be transmitted to their human hosts for completion of the lifecycle. After ingestion of infective blood meal it takes approximately two weeks until the female mosquito can infect an individual. The time between the mosquito infections to the point where they are infectious is called the extrinsic incubation period (EIP) and depends on the successful completion of the lifecycle within the mosquito (Baton and Ranford-Cartwright, 2005; Vlachou et al., 2006).

## **1.2. Current and novel malaria control strategies**

Competent *Plasmodium sp.* vectors can be found throughout the world but it is clear that *An. gambiae* is by far the greatest threat to human population as this species cohabitates with humans in the highest transmission regions (Fig 1.3). Current vector control strategies rely mainly on the use of chemical insecticides. DDT, the first chemical insecticide, was introduced in the 1940s and was a resounding success decreasing the number of competent vectors (Enayati and Hemingway, 2010). After questions arose about DDT's long lasting human health and ecological effects became a concern, its use was limited (but not abolished) and alternate insecticides were pursued. Since then an effort to use alternate chemicals in IRS and the use of LLIN has taken precedence. Currently, these methods are the most successful tools for limiting parasite transmission (World Malaria Report, 2012). IRS, or the spraying of potential mosquito resting spots with organophosphates, carbamates, and synthetic pyrethroids is widely effective, yet time and cost inefficient. Decreasing the mosquito density, it relies heavily on repeated spraying as the average IRS lasts only 4 months (Enayati and Hemingway, 2010). LLIN have had major, recent effects on malaria transmission by providing a physical barrier to infectious bites as well as exposure of the mosquito to a chemical insecticide. Currently the WHO has only

approved the use of pyrethroids as safe for the use in LLIN. The push for IRS and LLINs has been largely successful. LLIN alone decreased child mortality by (at least) 14% and reduced the number of malaria cases by at least 39% in areas where malaria is stable (Lengeler, 2004). Despite this success, the continued use of singular insecticides in bednets will increase the selection for resistance in the mosquito populations (Read et al., 2009; Raghavendra et al., 2011). The need for novel insecticides has never been higher as resistance to available insecticides mount.

In the last decade, it has been proposed that Late life acting (LLA) insecticides, that kill the mosquitoes after several gonadotrophic cycles, will decrease the selective pressure that current insecticides, which kill nearly on contact, are driving (Read et al., 2009). One such idea is the use of entomopathogenic fungus or lower concentrations of current insecticides (permethrin) that would kill older less robust mosquitoes (Blanford et al., 2005; Glunt et al., 2011). Another approach utilizes the intracellular endosymbiont *Wolbachia*. *Wolbachia* has been recently studied as a potential novel vector control agent with reduced fitness costs. *Wolbachia* infection decreases the lifespan of infected insects and it is also driven into the populations by cytotoxic incompatibility (CI) and increased pathogen interference (PI) (Yen and Barr, 1971; Hughes et al., 2011; Osei-Poku et al., 2012). Currently, stable *Wolbachia* infections have been successful in insect species such as *Aedes aegypti*, *Culex quinquefasciatus* and *Drosophila melanogaster*; yet no stable transgeneration infections been successful in *An. gambiae* (Hoffmann et al., 2011). Additionally, inconsistent results of different *Wolbachia* strains and their effect on infection of different *Plasmodium* species have recently been published. *An. gambiae* somatically infected with *Wolbachia* showed inhibition of *Plasmodium falciparum* infection, whereas it increased infection with the murine parasite, *Plasmodium bergeri* (Hughes et al., 2011; Hughes et al., 2012).

Furthermore, chemicals targeting proteins involved in negatively regulating immune responses have been proposed. One such target, is a serine protease inhibitor (SRPN) SRPN2, which functions as a negative regulator of the melanization response in *An. gambiae*. Knockdown of this immunity related protein increases the number of encapsulated parasites as well as decreases the mosquito's life span and is an interesting new insecticide target. (An et al., 2011).

## **1.3 Immunity**

Mosquitoes are exposed to numerous potentially pathogenic bacteria, virus, parasites and fungi that they must defend themselves against. Because of the close proximity in which they live with potential pathogens, and their hematophagous nature, mosquitoes have many routes of protection against these microbes. Defenses such as chitin dense exoskeleton and the peritrophic matrix provide physical barriers between pathogen and insect. If these are breached, a host of cellular and humoral responses work to protect the insect. First, the pathogen is recognized and then, typically, an immune signal is amplified by extracellular protease cascades. These cascades may lead to the activation of effector mechanisms such as humoral immune cascades i.e. TOLL, melanization, or complement-like TEP mediated killing. These immune pathways are summarized in Fig 1.4.

### ***1.3.1 Recognition***

Arthropod immunity against parasites and parasitoids (and all pathogens) is initiated upon recognition of foreign invaders. Components of microbes, called microbial associated molecular patterns (PAMP) are non-self motifs foreign to the insect such as LPS,  $\beta$ -1,3 glucans and peptidoglycans. PAMPs are recognized by pattern recognition receptors (PRR) such as Gram

Negative Binding Protein (GNBP), Peptidoglycan Recognition Proteins (PGRP) and thioester-containing proteins (TEP). TEP proteins circulate as soluble hemolymph factors secreted by hemocytes, activated by yet unknown mechanisms (Levashina et al., 2001). Other proteins involved in parasite clearance, Leucine-rich repeat proteins (LRIM) LRIM1 and APL1C (Riehle et al., 2008) form a hetero-dimeric complex (Povelones et al., 2011). Interaction between TEP1 and this hemolymph complex may cause C-terminal domain cleavage and activation of TEP1. The exact mechanism behind this is unclear, and the role of LRIM1 and APL1C in different mosquito/parasite combinations can vary (Osta et al., 2004; Dong et al., 2006; Riehle et al., 2006; Mitri et al., 2009). Ultimately, activated TEP1 binds to invading pathogens (bacteria (Levashina et al., 2001) and *Plasmodium* parasites (Blandin et al., 2004; Povelones et al., 2009) after interacting with the APL1C/LRIM1 complex (Fraiture et al., 2009; Povelones et al., 2009). TEP proteins increase parasite melanization and lysis (Povelones et al., 2009).

Toll and immune deficiency (IMD) pathways are arguably the best studied branches of insect immunity and culminate in the production of antimicrobial peptides as well as effector proteins. The Toll pathway has been most studied in the model organism *D. melanogaster* where it was first described (Hashimoto et al., 1991). The Toll pathway is activated in response to fungi and Gram positive bacteria. Humoral PRRs such as short (S) secreted PGRP-SA and PGRP-SD (Michel et al., 2001; Bischoff et al., 2004) or GNBP-1 (Gobert et al., 2003) bind to peptidoglycan on the surface of the bacteria. Activation of another branch of the Toll pathway is stimulated when GNBP3 binds  $\beta$ -glucans on fungi (Gottar et al., 2006). Pathogen secreted virulence factors can also activate the Toll pathway, which stimulates serine proteases such as Persephone (Fullaondo et al., 2011). This initiates extracellular protease cascades that ultimately lead to the cleavage and activation of the Toll ligand, Spätzle (DeLotto and DeLotto, 1998).

IMD is similarly activated by a different class of PRRs, membrane bound PGRPLC. In *An. gambiae* this trans-membrane protein recognizes and binds to peptidoglycan from Gram positive and Gram negative bacteria (Meister et al., 2009). In *D. melanogaster* this causes formation of a large receptor-adaptor complex (Georgel et al., 2001; Leulier et al., 2002). In *An. gambiae* this complex associates through death domains and cause the proteolytic activation of transcription factor, REL2 (Frolet et al., 2006) and antimicrobial peptide and effector molecule synthesis.

### ***1.3.2 The Toll Pathway***

To date, most functional studies of immunity-related pathways focus on genetic model organisms such as *D. melanogaster* or larger insects, such as *M. sexta*, which make biochemical analysis easier. Toll is a key component in insect immunity as well as *D. melanogaster* development where it was first described (Anderson et al., 1985; Lemaitre et al., 1995; Lemaitre et al., 1996). In immunity, it is currently thought that there are three branches of the Toll pathway. These branches are serine protease cascades that propagate the danger signal. Circulating as inactive zymogens, these serine proteases sequentially activate downstream proteases and converge upon Spätzle processing. Two PRR pathways are activated by PGRP-SA and PGRP-SD (Michel et al., 2001; Bischoff et al., 2004) or GGBP-1 (Gobert et al., 2003) in response to Gram positive bacteria and fungi (see above). These signals are integrated by a modular serine protease ModSP (Buchon et al., 2009), which initiates serine protease cascades. The third branch of Toll extracellular cascades is activated by virulence factors (proteases) that cleave Toll proteases such as Persephone (Gottar et al., 2006). Further experiments using RNAi identified an additional three serine proteases upstream of Toll ligand activation. The current Toll pathway model for these two branches in *D. melanogaster* immunity is that ModSp organizes messages from the PRRs and activates Grass. Spirit is then activated downstream and

subsequently spätzle processing enzyme (SPE) (Kambris et al., 2006). SPE cleaves and activates the Toll ligand Spätzle (DeLotto and DeLotto, 1998), which binds to and activates the transmembrane receptor Toll (Weber et al., 2003). After Toll activation intracellular proteins Tube and Pelle are recruited to the cell membrane (Towb et al., 1998), and MyD88 associates with the cytoplasmic side of TOLL (Tauszig-Delamasure et al., 2002) via a TIR domain. Tube and Pelle associate with MyD88 through the formation of death domains forming a heterotrimeric complex (Sun et al., 2002). Pelle may then activate kinases, which phosphorylate the IκB homologue Cactus. Cactus circulates in the cytoplasm as an inhibitory Cactus-Dorsal/Dif complex. Its phosphorylation initiates cactus degradation and NF-κB/REL family transcription factors Dorsal or Dif are translocated into the nucleus. Here, they regulate expression of antifungal AMP *drosomycin* (Lemaitre et al., 1996) as well as other AMPs and other effector proteins in the functional equivalent to the mammalian liver, the fat body. A microarray study in 2002 showed that in addition to AMP production, Toll and IMD regulate most genes induced by microbial infection (De Gregorio and Lemaitre, 2002).

The end result of Toll, AMP and effector protein transcription and translation, uses up precious resources and energy. Because of this, the Toll pathway is tightly regulated. In *D. melanogaster* as well in *An. gambiae*, serine protease inhibitors (serpins) are key regulators of the extracellular Toll cascades. There have been three serpin genes shown to have immune functions in *D. melanogaster*. Necrotic (Spn43Ac) inhibits the gram positive/fungal arm of Toll, likely through Persephone (Levashina et al., 1999; Robertson et al., 2003). Spn1 inhibits a trypsin-like serine protease, Grass, upstream in the fungal recognition branch of the Toll pathway (Fullaondo et al., 2011). Spn5 (Spn88Ea) was also shown to regulate Toll, though its location in the pathway is unclear (Ahmad et al., 2009).

Sequencing of the *An. gambiae* genome in 2002 (Holt et al., 2002) drastically aided the understanding of the *An. gambiae* immune responses. A comparative analysis of the known *D. melanogaster* genome to the newly sequenced *An. gambiae* genome identified 18 gene families, containing 242 *An. gambiae* gene homologous of 185 immunity related genes in *D. melanogaster* (Christophides et al., 2002). 41 clip-domain serine proteases (CLIPs) were identified in *An. gambiae* (compared to 35 in *D. melanogaster*), and six were up-regulated in response to bacterial and/or malaria infection (Christophides et al., 2002). Low 1:1 orthology of CLIP serine proteases between the species exists and there is little functional data indicating which proteases are involved in the extracellular cascade. Despite this, it is likely that the mosquito extracellular Toll activation cascade is similar to that in *D. melanogaster*. Additionally, 1:1 orthologues of intracellular components of the pathway have been identified for MyD88, Tube, Pelle, and Cactus (Christophides et al., 2002). *An. gambiae* orthologs to *D. melanogaster* Dorsal, and Rel1 (previously Gambif1) were identified and shown to translocate to the nucleus and increase AMP production (Barillas-Mury et al., 1996). Functional data are needed to confirm that these genes in *An. gambiae* are operating as their *D. melanogaster* orthologs. (Fig 1.5)

### ***1.3.3 The Phenoloxidase Activation Pathway***

Prophenoloxidase (PPO), the inactive zymogen of phenoloxidase (PO) is a large metalloprotein enzyme that exists as a dimer of approximately 160 kDa (Hall et al., 1995) containing 3 copper binding sites (Decker et al., 2007). Synthesized in hemocytes, the zymogen enters the hemolymph apparently via cell lysis and is either transported to the cuticle where it has an unknown function or it is activated in response to pathogens or wound healing (Asano and Ashida, 2001).

The phenoloxidase cascade has been studied in detail in large insect species such as *B. mori* and *M. sexta* where biochemical analysis is much easier due to the large volume of hemolymph available from a single insect. Ultimately leading to melanization (sequestration) of invading pathogens, this pathway plays an integral role in arthropod immunity. Phenoloxidase activity is stimulated by injury and/or recognition of pathogen moieties such as  $\beta$ -1-3-glucans and LPS (Lee et al., 1998). This stimulates a cascade of CLIP domain serine proteases (CLIPs) which are modulated by non-catalytic CLIPs called serine protease homologues (SPH)(Yu et al., 2003). CLIPs are serine proteases that have a unique N-terminal paperclip-like domain and a catalytic serine protease domain at the C-terminus. These proteases circulate as zymogens until they are proteolytically activated by cleavage at their linker region found between the CLIP and catalytic domains (Jiang and Kanost, 2000). Activation of terminal CLIPs, known as phenoloxidase activating proteases (PAPs) result in the proteolytic cleavage and activation of zymogen prophenoloxidase to phenoloxidase and ultimately, melanization.

Active PO enzyme catalyzes a very complex and multifaceted reaction that results in eumelanin synthesis. PO causes hydroxylation of tyrosine to DOPA and the oxidation of DOPA and dopamine to dopaquinone. DOPA and dopaquinones can undergo redox cycling yielding semiquinones and quinones. Additional enzymatic and non-enzymatic reactions lead to the formation of eumelanin (Fig 1.5). Byproducts of this pathway include toxic free radicals which may help in pathogen clearance, but must be strictly regulated by the insect to limit self-damage (Nappi et al., 2009). The melanization pathway is negatively regulated by a family of proteins known as serine protease inhibitors, or serpins.

### ***1.3.4 Protease Cascades***

Extracellular protease cascades function in a wide range of biological processes from development to coagulation and immunity. In arthropods, most of these are CLIP serine proteases (Muta et al., 1990). The first CLIP identified was proclotting enzyme from the horseshoe crab (Tai and Liu, 1977) (Nakamura et al., 1985) (Muta et al., 1990) and was described to contain a unique disulfide knotted domain. Subsequently, a similar domain was described in *D. melanogaster* snake and easter, serine proteases required for dorsal-ventral patterning in developing *D. melanogaster* larvae (DeLotto and Spierer, 1986; Chasan and Anderson, 1989; Wichmann et al., 2001). In insect species that have been studied serine protease cascades are responsible for acute phase defenses against infection. Many are secreted into the hemolymph, allowing them to be quickly activated to aid in coagulation, wound healing, TOLL activation, and pathogen sequestration in the hemocele of the mosquito (Gorman and Paskewitz, 2001).

### ***1.3.5 Serpins***

#### ***1.3.5.1 Structure***

In 1980 Hunt and Dayhoff recognized similarity in sequence between antitrypsin, ovalbumin and antithrombin and suggested a novel family of proteins (Hunt and Dayhoff, 1980) that were later named Serpins, or **serine protease inhibitors** (Carrell and Owen, 1985). These proteins make up a large family of metastable proteins that maintain low sequence similarity but are structurally conserved. Serpins contain 350-400 amino acids and are single domain proteins, made of up to three  $\beta$ -sheets which are surrounded by up to 9  $\alpha$ -helices (Huntington, 2011). The most functionally important region of the protein is the so-called reactive center loop (RCL), which

lends specificity and inhibitory efficacy to the molecule. Located at the C-terminus, the exposed RCL sits acts as bait for the target protease. Elongating or shortening the RCL by even 1 amino acid can limit inhibitory potential as well as decrease complex stability (Zhou et al., 2001).

Serpins are known as suicide inhibitors because the serpin as well as the protease are effectively irreversibly inhibited. In 2000, the crystal structure of the first complexed serpin was resolved to 2.6Å and verified that after cleavage and formation of an acyl intermediate, a covalent ester bond is formed between the P1 of the serpin and the active Ser of the target protease, ultimately destroying the protease's active site. The RCL completely inserts into  $\beta$ -sheet A and the target protease is translocated to the opposite pole of the serpin (Huntington et al., 2000) (Fig. 1.6).

This insertion allows the serpin to reach its lowest energy state, known as the relaxed state and is irreversible (Loebermann et al., 1984). Important regions of the RCL that are essential for inhibition are the hinge region and sessile bond denoted P1-P1' (Schechter and Berger, 1967).

To accommodate the large conformational change, the hinge region is typically made up of small amino acids. Additionally, the P1 amino acid determines the specificity of the inhibitor as its identity determines what type of serine protease may cleave it. A positively charged amino acid (such as Arg or Lys) at the P1 position indicates a trypsin-like serine protease target (Olsen et al., 2004), whereas an aromatic amino acid (such as Phe, Tyr or Trp) indicates chymotrypsin specificity (Schellenberger et al., 1991).

### ***1.3.5.2 Function***

Serpins are the largest family of serine protease inhibitors and are ubiquitously found from humans to viruses and plants (Kotwal and Moss, 1989; Brandt et al., 1990). These inhibitors have been shown to be involved in a myriad of biological function, yet the best studied serpins are involved in hemostasis regulating cascades such as blood coagulation, fibrinolysis, protein C

systems and complement activation (Hunt and Dayhoff, 1980; Patston et al., 1991).

Additionally, some serpins are no longer inhibitory and have evolved to have roles in physiological processes such as molecular transport (Thyroxin Binding Globulin) (Flink et al., 1986), blood pressure (Angiotensinogen) (Doolittle, 1983), and storage (Ovalbumin) (Hunt and Dayhoff, 1980). In insects, serpins play similar roles to their mammalian counterparts and have been shown to be involved in mating (Wolfner et al., 1997), development (Ligoxygakis et al., 2003) and extracellular, complement-like innate immune cascades such as Toll (Levashina et al., 1999) and PPO pathways (De Gregorio et al., 2002; An et al., 2012).

#### ***1.3.5.3 SRPNs Encoded in An. gambiae Genome***

Presently, in *Anopheles*, 18 SRPN genes have been identified (Christophides et al., 2002; Waterhouse et al., 2007) (Suwanchaichinda and Kanost, 2009). Alternate splicing of SRPN4 (3 isoforms) (Suwanchaichinda and Kanost, 2009) and SRPN10 (4 isoforms) (Danielli et al., 2003) contributes to a total of 23 distinct SRPN proteins encoded in the *An. gambiae* genome. The majority of the SRPNs form four clusters on two chromosomes (Christophides et al., 2002) (Suwanchaichinda and Kanost, 2009). There are two clusters on chromosome arm 2L, one containing SRPN1, 2, 3 and another encompassing SRPN7, 14, 18. Clusters on two chromosome arms 2R and 3R encode SRPN11, 12, 17 and SRPN5, 6, 16 respectively (Suwanchaichinda and Kanost, 2009). Additionally, *An.gambiae* SRPNs form phylogenetic clusters that may be mirrored by their physiological relatedness where SRPN1, 2, 3 are most similar, and SRPNs 6, 4, 5, 16 and 7, 14, 18 cluster together (Michel et al., 2005) (Suwanchaichinda and Kanost, 2009). These serpins, like in other insect species have variable expression under different physiological conditions and tissue tropism. All are presumed to be secreted as they contain a putative signal peptide (Gulley et al., 2013) with the exception of

SRPN10, which has been shown to be nucleocytoplasmic in hemocytes and pericardial cells. In midgut cells, SRPN10 is nuclear until infection when it is exported to the cytoplasm (Danielli et al., 2003).

#### **1.3.5.4 *An. gambiae* SRPNs Involved in Immunity**

In *An. gambiae*, only a few SRPNs have been shown to be involved in immunological processes. *SRPN10* encodes four different isoforms, of which *SRPN10B* (KRAL) is most highly expressed (Danielli et al., 2003). Additionally, *SRPN10* (KRAL and RCM) expression is increased in *Plasmodium bergeri* infected midgut cells and has been proposed as a midgut invasion marker (Danielli et al., 2005). *SRPN6* is also expressed in hemocytes. In addition, *SRPN6* is expressed in the midgut and salivary gland epithelia after human and murine *Plasmodium* infection. Depletion of *SRPN6* in *An. stephensi* increases the number of *P. bergeri* oocysts (Abraham et al., 2005; Pinto et al., 2008). Additionally, *SRPN6* knockdown in *An. gambiae* increases the number of salivary gland invading parasites (Pinto et al., 2008). *SRPN2* depletion markedly reduced the number of successful oocyst numbers in *Plasmodium bergeri* infected mosquitoes due to increased lysis/melanization (Michel et al., 2005). However, *SRPN2* depletion did not have an effect on the human malaria parasite, *Plasmodium falciparum* (Michel et al., 2006). Further analysis showed that *SRPN2* directly inhibited a known PAP, CLIPB9 and double knockdown reversed the *SRPN2* melanization phenotype (An et al., 2011). To date, no functional data have shown SRPN's involvement in regulating TOLL in *An. gambiae*. This data further substantiates the complex nature of serpin's role in *An. gambiae* immunity and the need for further functional studies.

## 1.4 References

- Abraham, E.G., Pinto, S.B., Ghosh, A., Vanlandingham, D.L., Budd, A., Higgs, S., Kafatos, F.C., Jacobs-Lorena, M., Michel, K., 2005. An immune-responsive serpin, SRPN6, mediates mosquito defense against malaria parasites. *Proc Natl Acad Sci U S A* 102, 16327-16332.
- Ahmad, S.T., Sweeney, S.T., Lee, J.A., Sweeney, N.T., Gao, F.B., 2009. Genetic screen identifies serpin5 as a regulator of the toll pathway and CHMP2B toxicity associated with frontotemporal dementia. *Proc Natl Acad Sci U S A* 106, 12168-12173.
- An, C., Budd, A., Kanost, M.R., Michel, K., 2011. Characterization of a regulatory unit that controls melanization and affects longevity of mosquitoes. *Cell Mol Life Sci* 68, 1929-1939.
- An, C., Hiromasa, Y., Zhang, X., Lovell, S., Zolkiewski, M., Tomich, J.M., Michel, K., 2012. Biochemical characterization of *Anopheles gambiae* SRPN6, a malaria parasite invasion marker in mosquitoes. *PLoS One* 7, e48689.
- Anderson, K.V., Bokla, L., Nusslein-Volhard, C., 1985. Establishment of dorsal-ventral polarity in the *Drosophila* embryo: the induction of polarity by the Toll gene product. *Cell* 42, 791-798.
- Anonymous, 2011. A research agenda for malaria eradication: vector control. *PLoS Med* 8, e1000401.
- Asano, T., Ashida, M., 2001. Cuticular pro-phenoloxidase of the silkworm, *Bombyx mori*. Purification and demonstration of its transport from hemolymph. *J Biol Chem* 276, 11100-11112.
- Barillas-Mury, C., Charlesworth, A., Gross, I., Richman, A., Hoffmann, J.A., Kafatos, F.C., 1996. Immune factor Gambif1, a new rel family member from the human malaria vector, *Anopheles gambiae*. *EMBO J* 15, 4691-4701.
- Baton, L.A., Ranford-Cartwright, L.C., 2005. How do malaria ookinetes cross the mosquito midgut wall? *Trends in Parasitology* 21, 22-28.
- Bischoff, V., Vignat, C., Boneca, I.G., Michel, T., Hoffmann, J.A., Royet, J., 2004. Function of the *drosophila* pattern-recognition receptor PGRP-SD in the detection of Gram-positive bacteria. *Nat Immunol* 5, 1175-1180.
- Blandin, S., Shiao, S.H., Moita, L.F., Janse, C.J., Waters, A.P., Kafatos, F.C., Levashina, E.A., 2004. Complement-like protein TEPI is a determinant of vectorial capacity in the malaria vector *Anopheles gambiae*. *Cell* 116, 661-670.
- Blanford, S., Chan, B.H., Jenkins, N., Sim, D., Turner, R.J., Read, A.F., Thomas, M.B., 2005. Fungal pathogen reduces potential for malaria transmission. *Science* 308, 1638-1641.
- Brandt, A., Svendsen, I., Hejgaard, J., 1990. A plant serpin gene. Structure, organization and expression of the gene encoding barley protein Z4. *Eur J Biochem* 194, 499-505.
- Buchon, N., Poidevin, M., Kwon, H.M., Guillou, A., Sottas, V., Lee, B.L., Lemaitre, B., 2009. A single modular serine protease integrates signals from pattern-recognition receptors upstream of the *Drosophila* Toll pathway. *Proc Natl Acad Sci U S A* 106, 12442-12447.
- Carrell, R.W., Owen, M.C., 1985. Plakalbumin, alpha 1-antitrypsin, antithrombin and the mechanism of inflammatory thrombosis. *Nature* 317, 730-732.

- Chasan, R., Anderson, K.V., 1989. The role of easter, an apparent serine protease, in organizing the dorsal-ventral pattern of the *Drosophila* embryo. *Cell* 56, 391-400.
- Christophides, G.K., Zdobnov, E., Barillas-Mury, C., Birney, E., Blandin, S., Blass, C., Brey, P.T., Collins, F.H., Danielli, A., Dimopoulos, G., Hetru, C., Hoa, N.T., Hoffmann, J.A., Kanzok, S.M., Letunic, I., Levashina, E.A., Loukeris, T.G., Lycett, G., Meister, S., Michel, K., Moita, L.F., Muller, H.M., Osta, M.A., Paskewitz, S.M., Reichhart, J.M., Rzhetsky, A., Troxler, L., Vernick, K.D., Vlachou, D., Volz, J., von Mering, C., Xu, J., Zheng, L., Bork, P., Kafatos, F.C., 2002. Immunity-related genes and gene families in *Anopheles gambiae*. *Science* 298, 159-165.
- Control, T.m.C.G.o.V., 2011. A research agenda for malaria eradication: vector control. *PLoS Med* 8, e1000401.
- Danielli, A., Kafatos, F.C., Loukeris, T.G., 2003. Cloning and characterization of four *Anopheles gambiae* serpin isoforms, differentially induced in the midgut by *Plasmodium berghei* invasion. *J Biol Chem* 278, 4184-4193.
- Danielli, A., Barillas-Mury, C., Kumar, S., Kafatos, F.C., Loukeris, T.G., 2005. Overexpression and altered nucleocytoplasmic distribution of *Anopheles* ovalbumin-like SRPN10 serpins in *Plasmodium*-infected midgut cells. *Cell Microbiol* 7, 181-190.
- De Gregorio, E., Han, S.J., Lee, W.J., Baek, M.J., Osaki, T., Kawabata, S., Lee, B.L., Iwanaga, S., Lemaitre, B., Brey, P.T., 2002. An immune-responsive Serpin regulates the melanization cascade in *Drosophila*. *Dev Cell* 3, 581-592.
- De Gregorio, E., Lemaitre, B., 2002. The mosquito genome: the post-genomic era opens. *Nature* 419, 496-497.
- Decker, H., Schweikardt, T., Nillius, D., Salzbrunn, U., Jaenicke, E., Tucek, F., 2007. Similar enzyme activation and catalysis in hemocyanins and tyrosinases. *Gene* 398, 183-191.
- DeLotto, R., Spierer, P., 1986. A gene required for the specification of dorsal-ventral pattern in *Drosophila* appears to encode a serine protease. *Nature* 323, 688-692.
- DeLotto, Y., DeLotto, R., 1998. Proteolytic processing of the *Drosophila* Spatzle protein by easter generates a dimeric NGF-like molecule with ventralising activity. *Mech Dev* 72, 141-148.
- Dong, Y., Aguilar, R., Xi, Z., Warr, E., Mongin, E., Dimopoulos, G., 2006. *Anopheles gambiae* immune responses to human and rodent *Plasmodium* parasite species. *PLoS Pathog* 2, e52.
- Doolittle, R.F., 1983. Angiotensinogen is related to the antitrypsin-antithrombin-ovalbumin family. *Science* 222, 417-419.
- Enayati, A., Hemingway, J., 2010. Malaria management: past, present, and future. *Annu Rev Entomol* 55, 569-591.
- Flink, I.L., Bailey, T.J., Gustafson, T.A., Markham, B.E., Morkin, E., 1986. Complete amino acid sequence of human thyroxine-binding globulin deduced from cloned DNA: close homology to the serine antiproteases. *Proc Natl Acad Sci U S A* 83, 7708-7712.
- Fraiture, M., Baxter, R.H., Steinert, S., Chelliah, Y., Frolet, C., Quispe-Tintaya, W., Hoffmann, J.A., Blandin, S.A., Levashina, E.A., 2009. Two mosquito LRR proteins function as complement control factors in the TEPI-mediated killing of *Plasmodium*. *Cell Host Microbe* 5, 273-284.
- Frolet, C., Thoma, M., Blandin, S., Hoffmann, J.A., Levashina, E.A., 2006. Boosting NF-kappaB-dependent basal immunity of *Anopheles gambiae* aborts development of *Plasmodium berghei*. *Immunity* 25, 677-685.

- Fullaondo, A., Garcia-Sanchez, S., Sanz-Parra, A., Recio, E., Lee, S.Y., Gubb, D., 2011. Spn1 regulates the GGBP3-dependent Toll signaling pathway in *Drosophila melanogaster*. *Mol Cell Biol* 31, 2960-2972.
- Georgel, P., Naitza, S., Kappler, C., Ferrandon, D., Zachary, D., Swimmer, C., Kopczynski, C., Duyk, G., Reichhart, J.M., Hoffmann, J.A., 2001. *Drosophila* immune deficiency (IMD) is a death domain protein that activates antibacterial defense and can promote apoptosis. *Dev Cell* 1, 503-514.
- Glunt, K.D., Thomas, M.B., Read, A.F., 2011. The effects of age, exposure history and malaria infection on the susceptibility of *Anopheles* mosquitoes to low concentrations of pyrethroid. *PLoS One* 6, e24968.
- Gobert, V., Gottar, M., Matskevich, A.A., Rutschmann, S., Royet, J., Belvin, M., Hoffmann, J.A., Ferrandon, D., 2003. Dual activation of the *Drosophila* toll pathway by two pattern recognition receptors. *Science* 302, 2126-2130.
- Gorman, M.J., Paskewitz, S.M., 2001. Serine proteases as mediators of mosquito immune responses. *Insect Biochem Mol Biol* 31, 257-262.
- Gottar, M., Gobert, V., Matskevich, A.A., Reichhart, J.M., Wang, C., Butt, T.M., Belvin, M., Hoffmann, J.A., Ferrandon, D., 2006. Dual detection of fungal infections in *Drosophila* via recognition of glucans and sensing of virulence factors. *Cell* 127, 1425-1437.
- Gulley, M.M., Zhang, X., Michel, K., 2013. The roles of serpins in mosquito immunology and physiology. *J Insect Physiol* 59, 138-147.
- Hall, M., Scott, T., Sugumaran, M., Soderhall, K., Law, J.H., 1995. Proenzyme of *Manduca sexta* phenol oxidase: purification, activation, substrate specificity of the active enzyme, and molecular cloning. *Proc Natl Acad Sci U S A* 92, 7764-7768.
- Han, Y.S., Thompson, J., Kafatos, F.C., Barillas-Mury, C., 2000. Molecular interactions between *Anopheles stephensi* midgut cells and *Plasmodium berghei*: the time bomb theory of ookinete invasion of mosquitoes. *Embo Journal* 19, 6030-6040.
- Hashimoto, C., Gerttula, S., Anderson, K.V., 1991. Plasma membrane localization of the Toll protein in the syncytial *Drosophila* embryo: importance of transmembrane signaling for dorsal-ventral pattern formation. *Development* 111, 1021-1028.
- Hoffmann, A.A., Montgomery, B.L., Popovici, J., Iturbe-Ormaetxe, I., Johnson, P.H., Muzzi, F., Greenfield, M., Durkan, M., Leong, Y.S., Dong, Y., Cook, H., Axford, J., Callahan, A.G., Kenny, N., Omodei, C., McGraw, E.A., Ryan, P.A., Ritchie, S.A., Turelli, M., O'Neill, S.L., 2011. Successful establishment of *Wolbachia* in *Aedes* populations to suppress dengue transmission. *Nature* 476, 454-457.
- Holt, R.A., Subramanian, G.M., Halpern, A., Sutton, G.G., Charlab, R., Nusskern, D.R., Wincker, P., Clark, A.G., Ribeiro, J.M., Wides, R., Salzberg, S.L., Loftus, B., Yandell, M., Majoros, W.H., Rusch, D.B., Lai, Z., Kraft, C.L., Abril, J.F., Anthouard, V., Arensburger, P., Atkinson, P.W., Baden, H., de Berardinis, V., Baldwin, D., Benes, V., Biedler, J., Blass, C., Bolanos, R., Boscus, D., Barnstead, M., Cai, S., Center, A., Chaturverdi, K., Christophides, G.K., Chrystal, M.A., Clamp, M., Cravchik, A., Curwen, V., Dana, A., Delcher, A., Dew, I., Evans, C.A., Flanigan, M., Grundschober-Freimoser, A., Friedli, L., Gu, Z., Guan, P., Guigo, R., Hillenmeyer, M.E., Hladun, S.L., Hogan, J.R., Hong, Y.S., Hoover, J., Jaillon, O., Ke, Z., Kodira, C., Kokoza, E., Koutsos, A., Letunic, I., Levitsky, A., Liang, Y., Lin, J.J., Lobo, N.F., Lopez, J.R., Malek, J.A., McIntosh, T.C., Meister, S., Miller, J., Mobarry, C., Mongin, E., Murphy, S.D., O'Brochta, D.A., Pfannkoch, C., Qi, R., Regier, M.A., Remington, K., Shao, H.,

- Sharakhova, M.V., Sitter, C.D., Shetty, J., Smith, T.J., Strong, R., Sun, J., Thomasova, D., Ton, L.Q., Topalis, P., Tu, Z., Unger, M.F., Walenz, B., Wang, A., Wang, J., Wang, M., Wang, X., Woodford, K.J., Wortman, J.R., Wu, M., Yao, A., Zdobnov, E.M., Zhang, H., Zhao, Q., Zhao, S., Zhu, S.C., Zhimulev, I., Coluzzi, M., della Torre, A., Roth, C.W., Louis, C., Kalush, F., Mural, R.J., Myers, E.W., Adams, M.D., Smith, H.O., Broder, S., Gardner, M.J., Fraser, C.M., Birney, E., Bork, P., Brey, P.T., Venter, J.C., Weissenbach, J., Kafatos, F.C., Collins, F.H., Hoffman, S.L., 2002. The genome sequence of the malaria mosquito *Anopheles gambiae*. *Science* 298, 129-149.
- Hughes, G.L., Koga, R., Xue, P., Fukatsu, T., Rasgon, J.L., 2011. Wolbachia infections are virulent and inhibit the human malaria parasite *Plasmodium falciparum* in *Anopheles gambiae*. *PLoS Pathog* 7, e1002043.
- Hughes, G.L., Vega-Rodriguez, J., Xue, P., Rasgon, J.L., 2012. Wolbachia strain wAlbB enhances infection by the rodent malaria parasite *Plasmodium berghei* in *Anopheles gambiae* mosquitoes. *Appl Environ Microbiol* 78, 1491-1495.
- Hunt, L.T., Dayhoff, M.O., 1980. A surprising new protein superfamily containing ovalbumin, antithrombin-III, and alpha 1-proteinase inhibitor. *Biochem Biophys Res Commun* 95, 864-871.
- Huntington, J.A., Read, R.J., Carrell, R.W., 2000. Structure of a serpin-protease complex shows inhibition by deformation. *Nature* 407, 923-926.
- Huntington, J.A., 2011. Serpin structure, function and dysfunction. *J Thromb Haemost* 9 Suppl 1, 26-34.
- Jiang, H., Kanost, M.R., 2000. The clip-domain family of serine proteinases in arthropods. *Insect Biochem Mol Biol* 30, 95-105.
- Kambris, Z., Brun, S., Jang, I.H., Nam, H.J., Romeo, Y., Takahashi, K., Lee, W.J., Ueda, R., Lemaitre, B., 2006. *Drosophila* immunity: a large-scale in vivo RNAi screen identifies five serine proteases required for Toll activation. *Curr Biol* 16, 808-813.
- Kotwal, G.J., Moss, B., 1989. Vaccinia virus encodes two proteins that are structurally related to members of the plasma serine protease inhibitor superfamily. *J Virol* 63, 600-606.
- Laveran, C.L., 1982. Classics in infectious diseases: A newly discovered parasite in the blood of patients suffering from malaria. Parasitic etiology of attacks of malaria: Charles Louis Alphonse Laveran (1845-1922). *Rev Infect Dis* 4, 908-911.
- Lee, S.Y., Cho, M.Y., Hyun, J.H., Lee, K.M., Homma, K.I., Natori, S., Kawabata, S.I., Iwanaga, S., Lee, B.L., 1998. Molecular cloning of cDNA for pro-phenol-oxidase-activating factor I, a serine protease is induced by lipopolysaccharide or 1,3-beta-glucan in coleopteran insect, *Holotrichia diomphalia* larvae. *Eur J Biochem* 257, 615-621.
- Lemaitre, B., Meister, M., Govind, S., Georgel, P., Steward, R., Reichhart, J.M., Hoffmann, J.A., 1995. Functional analysis and regulation of nuclear import of dorsal during the immune response in *Drosophila*. *EMBO J* 14, 536-545.
- Lemaitre, B., Nicolas, E., Michaut, L., Reichhart, J.M., Hoffmann, J.A., 1996. The dorsoventral regulatory gene cassette *spatzle/Toll/cactus* controls the potent antifungal response in *Drosophila* adults. *Cell* 86, 973-983.
- Lengeler, C., 2004. Insecticide-treated bed nets and curtains for preventing malaria. *Cochrane Database Syst Rev*, CD000363.
- Leulier, F., Vidal, S., Saigo, K., Ueda, R., Lemaitre, B., 2002. Inducible expression of double-stranded RNA reveals a role for dFADD in the regulation of the antibacterial response in *Drosophila* adults. *Curr Biol* 12, 996-1000.

- Levashina, E.A., Langley, E., Green, C., Gubb, D., Ashburner, M., Hoffmann, J.A., Reichhart, J.M., 1999. Constitutive activation of toll-mediated antifungal defense in serpin-deficient *Drosophila*. *Science* 285, 1917-1919.
- Levashina, E.A., Moita, L.F., Blandin, S., Vriend, G., Lagueux, M., Kafatos, F.C., 2001. Conserved role of a complement-like protein in phagocytosis revealed by dsRNA knockout in cultured cells of the mosquito, *Anopheles gambiae*. *Cell* 104, 709-718.
- Ligoxygakis, P., Roth, S., Reichhart, J.M., 2003. A serpin regulates dorsal-ventral axis formation in the *Drosophila* embryo. *Curr Biol* 13, 2097-2102.
- Loebermann, H., Tokuoka, R., Deisenhofer, J., Huber, R., 1984. Human alpha 1-proteinase inhibitor. Crystal structure analysis of two crystal modifications, molecular model and preliminary analysis of the implications for function. *J Mol Biol* 177, 531-557.
- Meister, S., Agianian, B., Turlure, F., Relogio, A., Morlais, I., Kafatos, F.C., Christophides, G.K., 2009. *Anopheles gambiae* PGRPLC-mediated defense against bacteria modulates infections with malaria parasites. *PLoS Pathog* 5, e1000542.
- Michel, K., Budd, A., Pinto, S., Gibson, T.J., Kafatos, F.C., 2005. *Anopheles gambiae* SRPN2 facilitates midgut invasion by the malaria parasite *Plasmodium berghei*. *EMBO Rep* 6, 891-897.
- Michel, K., Suwanchaichinda, C., Morlais, I., Lambrechts, L., Cohuet, A., Awono-Ambene, P.H., Simard, F., Fontenille, D., Kanost, M.R., Kafatos, F.C., 2006. Increased melanizing activity in *Anopheles gambiae* does not affect development of *Plasmodium falciparum*. *Proc Natl Acad Sci U S A* 103, 16858-16863.
- Michel, T., Reichhart, J.M., Hoffmann, J.A., Royet, J., 2001. *Drosophila* Toll is activated by Gram-positive bacteria through a circulating peptidoglycan recognition protein. *Nature* 414, 756-759.
- Mitri, C., Jacques, J.C., Thiery, I., Riehle, M.M., Xu, J., Bischoff, E., Morlais, I., Nsango, S.E., Vernick, K.D., Bourgouin, C., 2009. Fine pathogen discrimination within the APL1 gene family protects *Anopheles gambiae* against human and rodent malaria species. *PLoS Pathog* 5, e1000576.
- Muta, T., Hashimoto, R., Miyata, T., Nishimura, H., Toh, Y., Iwanaga, S., 1990. Proclotting enzyme from horseshoe crab hemocytes. cDNA cloning, disulfide locations, and subcellular localization. *J Biol Chem* 265, 22426-22433.
- Nakamura, T., Morita, T., Iwanaga, S., 1985. Intracellular proclotting enzyme in limulus (*Tachypleus tridentatus*) hemocytes: its purification and properties. *J Biochem* 97, 1561-1574.
- Nappi, A., Poirie, M., Carton, Y., 2009. The role of melanization and cytotoxic by-products in the cellular immune responses of *Drosophila* against parasitic wasps. *Adv Parasitol* 70, 99-121.
- Oddoux, O., Debourgogne, A., Kantele, A., Kocken, C.H., Jokiranta, T.S., Vedy, S., Puyhardy, J.M., Machouart, M., 2011. Identification of the five human *Plasmodium* species including *P. knowlesi* by real-time polymerase chain reaction. *European Journal of Clinical Microbiology & Infectious Diseases* 30, 597-601.
- Olsen, J.V., Ong, S.E., Mann, M., 2004. Trypsin cleaves exclusively C-terminal to arginine and lysine residues. *Mol Cell Proteomics* 3, 608-614.
- Osei-Poku, J., Han, C., Mbogo, C.M., Jiggins, F.M., 2012. Identification of *Wolbachia* strains in mosquito disease vectors. *PLoS One* 7, e49922.

- Osta, M.A., Christophides, G.K., Kafatos, F.C., 2004. Effects of mosquito genes on Plasmodium development. *Science* 303, 2030-2032.
- Patston, P.A., Gettins, P., Beechem, J., Schapira, M., 1991. Mechanism of serpin action: evidence that C1 inhibitor functions as a suicide substrate. *Biochemistry* 30, 8876-8882.
- Pinto, S.B., Kafatos, F.C., Michel, K., 2008. The parasite invasion marker SRPN6 reduces sporozoite numbers in salivary glands of *Anopheles gambiae*. *Cell Microbiol* 10, 891-898.
- Povelones, M., Waterhouse, R.M., Kafatos, F.C., Christophides, G.K., 2009. Leucine-rich repeat protein complex activates mosquito complement in defense against Plasmodium parasites. *Science* 324, 258-261.
- Povelones, M., Upton, L.M., Sala, K.A., Christophides, G.K., 2011. Structure-function analysis of the *Anopheles gambiae* LRIM1/APL1C complex and its interaction with complement C3-like protein TEP1. *PLoS Pathog* 7, e1002023.
- Raghavendra, K., Barik, T.K., Reddy, B.P., Sharma, P., Dash, A.P., 2011. Malaria vector control: from past to future. *Parasitol Res* 108, 757-779.
- Read, A.F., Lynch, P.A., Thomas, M.B., 2009. How to make evolution-proof insecticides for malaria control. *PLoS Biol* 7, e1000058.
- Riehle, M.M., Markianos, K., Niare, O., Xu, J., Li, J., Toure, A.M., Podiougou, B., Oduol, F., Diawara, S., Diallo, M., Coulibaly, B., Ouatarara, A., Kruglyak, L., Traore, S.F., Vernick, K.D., 2006. Natural malaria infection in *Anopheles gambiae* is regulated by a single genomic control region. *Science* 312, 577-579.
- Riehle, M.M., Xu, J., Lazzaro, B.P., Rottschaefer, S.M., Coulibaly, B., Sacko, M., Niare, O., Morlais, I., Traore, S.F., Vernick, K.D., 2008. *Anopheles gambiae* APL1 is a family of variable LRR proteins required for Rel1-mediated protection from the malaria parasite, *Plasmodium berghei*. *PLoS One* 3, e3672.
- Robertson, A.S., Belorgey, D., Lilley, K.S., Lomas, D.A., Gubb, D., Dafforn, T.R., 2003. Characterization of the necrotic protein that regulates the Toll-mediated immune response in *Drosophila*. *J Biol Chem* 278, 6175-6180.
- Ross, R., 1899. Life-History of the Parasites of Malaria. *Nature* Volume 60, pp. 322-324.
- Schechter, I., Berger, A., 1967. On the size of the active site in proteases. I. Papain. 1967. *Biochem Biophys Res Commun* 425, 497-502.
- Schellenberger, V., Braune, K., Hofmann, H.J., Jakubke, H.D., 1991. The specificity of chymotrypsin. A statistical analysis of hydrolysis data. *Eur J Biochem* 199, 623-636.
- Sun, H., Bristow, B.N., Qu, G., Wasserman, S.A., 2002. A heterotrimeric death domain complex in Toll signaling. *Proc Natl Acad Sci U S A* 99, 12871-12876.
- Suwanchaichinda, C., Kanost, M.R., 2009. The serpin gene family in *Anopheles gambiae*. *Gene* 442, 47-54.
- Tai, J.Y., Liu, T.Y., 1977. Studies on *Limulus* amoebocyte lysate. Isolation of pro-clotting enzyme. *J Biol Chem* 252, 2178-2181.
- Tauszig-Delamasure, S., Bilak, H., Capovilla, M., Hoffmann, J.A., Imler, J.L., 2002. *Drosophila* MyD88 is required for the response to fungal and Gram-positive bacterial infections. *Nat Immunol* 3, 91-97.
- Towb, P., Galindo, R.L., Wasserman, S.A., 1998. Recruitment of Tube and Pelle to signaling sites at the surface of the *Drosophila* embryo. *Development* 125, 2443-2450.
- Vlachou, D., Schlegelmilch, T., Runn, E., Mendes, A., Kafatos, F.C., 2006. The developmental migration of Plasmodium in mosquitoes. *Curr Opin Genet Dev* 16, 384-391.

- Waterhouse, R.M., Kriventseva, E.V., Meister, S., Xi, Z., Alvarez, K.S., Bartholomay, L.C., Barillas-Mury, C., Bian, G., Blandin, S., Christensen, B.M., Dong, Y., Jiang, H., Kanost, M.R., Koutsos, A.C., Levashina, E.A., Li, J., Ligoxygakis, P., Maccallum, R.M., Mayhew, G.F., Mendes, A., Michel, K., Osta, M.A., Paskewitz, S., Shin, S.W., Vlachou, D., Wang, L., Wei, W., Zheng, L., Zou, Z., Severson, D.W., Raikhel, A.S., Kafatos, F.C., Dimopoulos, G., Zdobnov, E.M., Christophides, G.K., 2007. Evolutionary dynamics of immune-related genes and pathways in disease-vector mosquitoes. *Science* 316, 1738-1743.
- Weber, A.N., Tauszig-Delamasure, S., Hoffmann, J.A., Lelievre, E., Gascan, H., Ray, K.P., Morse, M.A., Imler, J.L., Gay, N.J., 2003. Binding of the *Drosophila* cytokine Spatzle to Toll is direct and establishes signaling. *Nat Immunol* 4, 794-800.
- Wichmann, M.W., Muller, C., Hornung, H.M., Lau-Werner, U., Schildberg, F.W., 2001. Gender differences in long-term survival of patients with colorectal cancer. *Br J Surg* 88, 1092-1098.
- Wolfner, M.F., Harada, H.A., Bertram, M.J., Stelick, T.J., Kraus, K.W., Kalb, J.M., Lung, Y.O., Neubaum, D.M., Park, M., Tram, U., 1997. New genes for male accessory gland proteins in *Drosophila melanogaster*. *Insect Biochem Mol Biol* 27, 825-834.
- Yen, J.H., Barr, A.R., 1971. New hypothesis of the cause of cytoplasmic incompatibility in *Culex pipiens* L. *Nature* 232, 657-658.
- Yu, X.Q., Jiang, H., Wang, Y., Kanost, M.R., 2003. Nonproteolytic serine proteinase homologs are involved in prophenoloxidase activation in the tobacco hornworm, *Manduca sexta*. *Insect Biochem Mol Biol* 33, 197-208.
- Zhou, A., Carrell, R.W., Huntington, J.A., 2001. The serpin inhibitory mechanism is critically dependent on the length of the reactive center loop. *J Biol Chem* 276, 27541-27547.

## 1.5 Figures Legends

### **Figure 1-1: Global map of malaria transmission and use of IRS or LLIN in areas of high transmission**

**A.** Global map of countries with ongoing transmission of malaria. Areas with persistent transmission are tan. Countries with continued transmission of malaria which have also reported cases of resistance to insecticides are brown. **B.** Percentage of populations from countries in Sub-Saharan Africa that are reporting usage of LLIN and/or IRS. Countries reporting less than 25% of the population with protection are colored beige. Countries in tan report 50-80% of their populations have protection against malaria. Brown colored countries report over 80% coverage of their populations using IRS and/or LLIN. Images taken from World Malaria Report 2012.

### **Figure 1-2: *Plasmodium sp.* Lifecycle**

The complex life cycle of *Plasmodium sp.* in its intermediate and definitive hosts. See text for further details. Image taken from (Su et al., 2007).

### **Figure1-3: Global map of malaria vectors**

Global map of the dominant or potentially important malaria vectors in regions which malaria is or has been endemic. Dominant vectors were identified by selecting the species that fed most frequently and survived the longest. Image taken from (Kiszewski et al., 2004).

### **Figure1-4: Overview of *An. gambiae* immune response**

Graphical representation of the primary immune responses of *An. gambiae*: Toll, melanization and TEP mediated pathogen clearance. Extracellular protease cascades lead the activation of these main immune responses and culminate in the production of (1) AMP, (2) Melanized pathogens and (3) TEP mediated pathogen lysis and/or phagocytosis. For detailed description of these pathways please see text. Figure adapted from (Gulley et al., 2013).

**Figure1-5: Melanogenesis pathway in insects**

Overview of the phenoloxidase pathway in insects. This figure shows the synthesis of eumelanin as well as alternate byproducts of the pathway, pheomelanin, quinones and free radicals. The figure was taken from (Nappi et al., 2009). Further details in text.

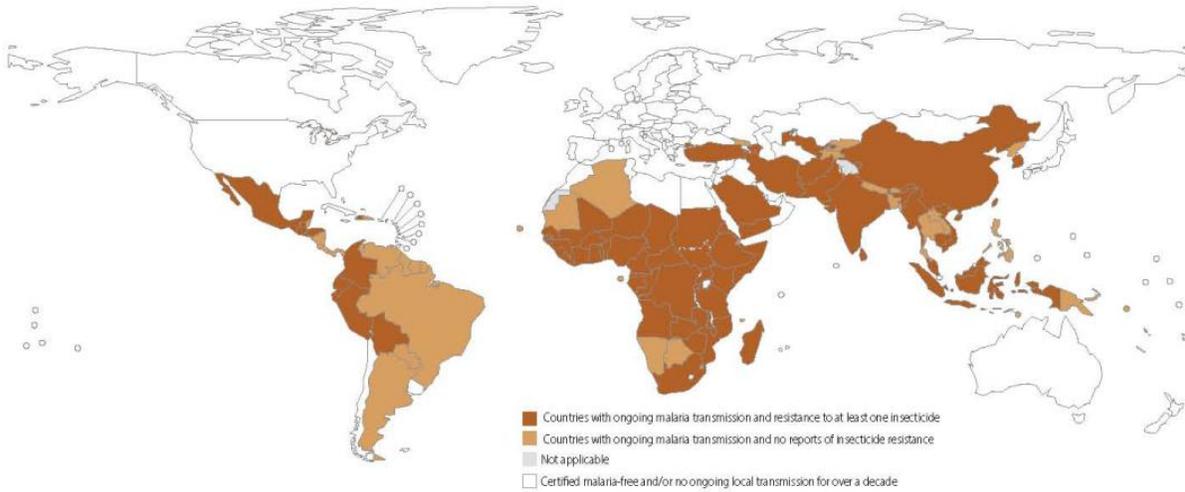
**Figure1-6: Serpin suicide inhibition mechanism**

Serpin inhibition mechanism. Target protease in green approaches and cleaves the bond between the P1-P1' of the RCL. This initiates a conformation change in which the RCL swings around and inserts into  $\beta$ -Sheet A of the serpin. In this reaction, the protease is dragged to the opposite pole of the serpin and is covalently bond to the serpin. Figure adapted from (Huntington, 2011)

## 1.6 Figures

**Figure 1-1: Global map of malaria transmission and use of IRS or LLAN in areas of high transmission.**

A



B

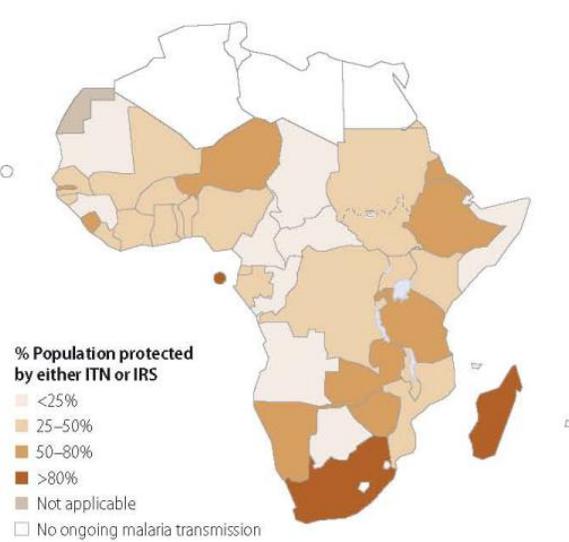


Figure 1-2: *Plasmodium sp.* Lifecycle

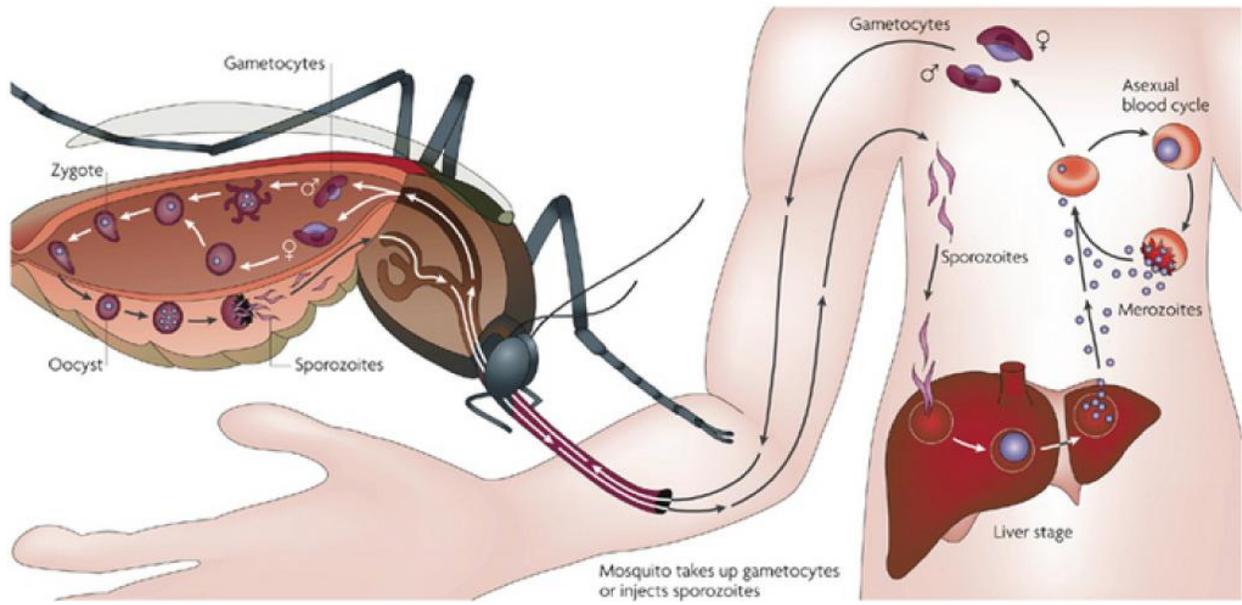




Figure 1-4: Overview of *An. gambiae* immune response

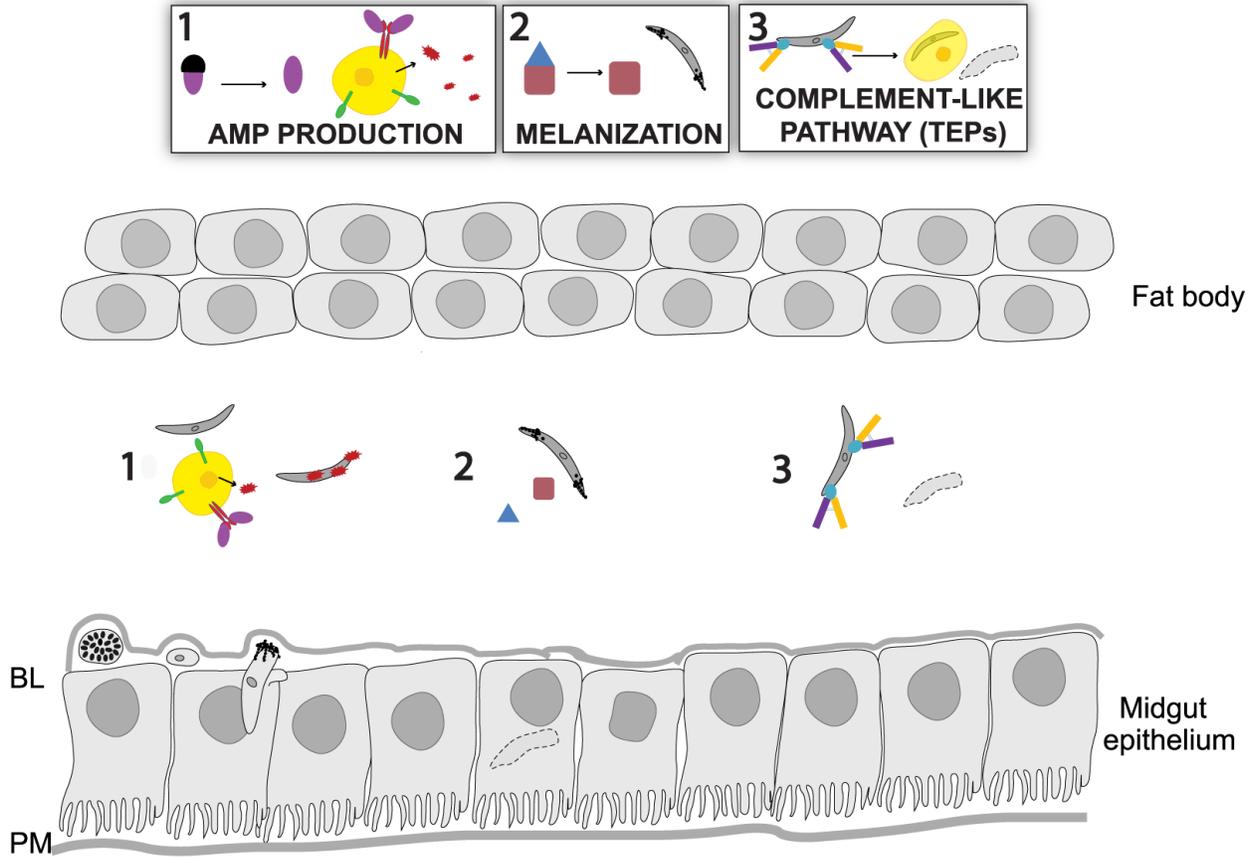


Figure 1-5: Melanogenesis pathway in insects

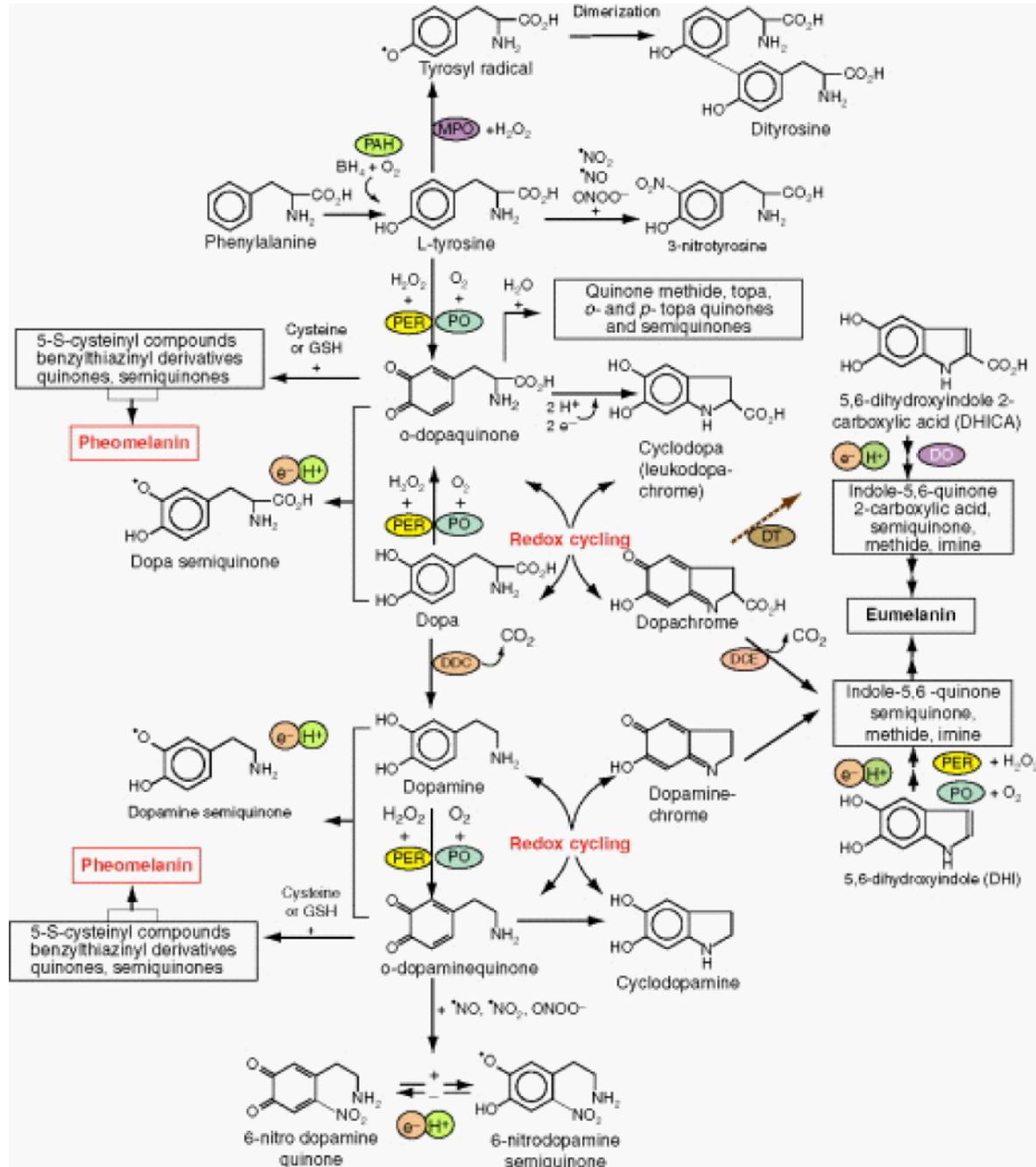
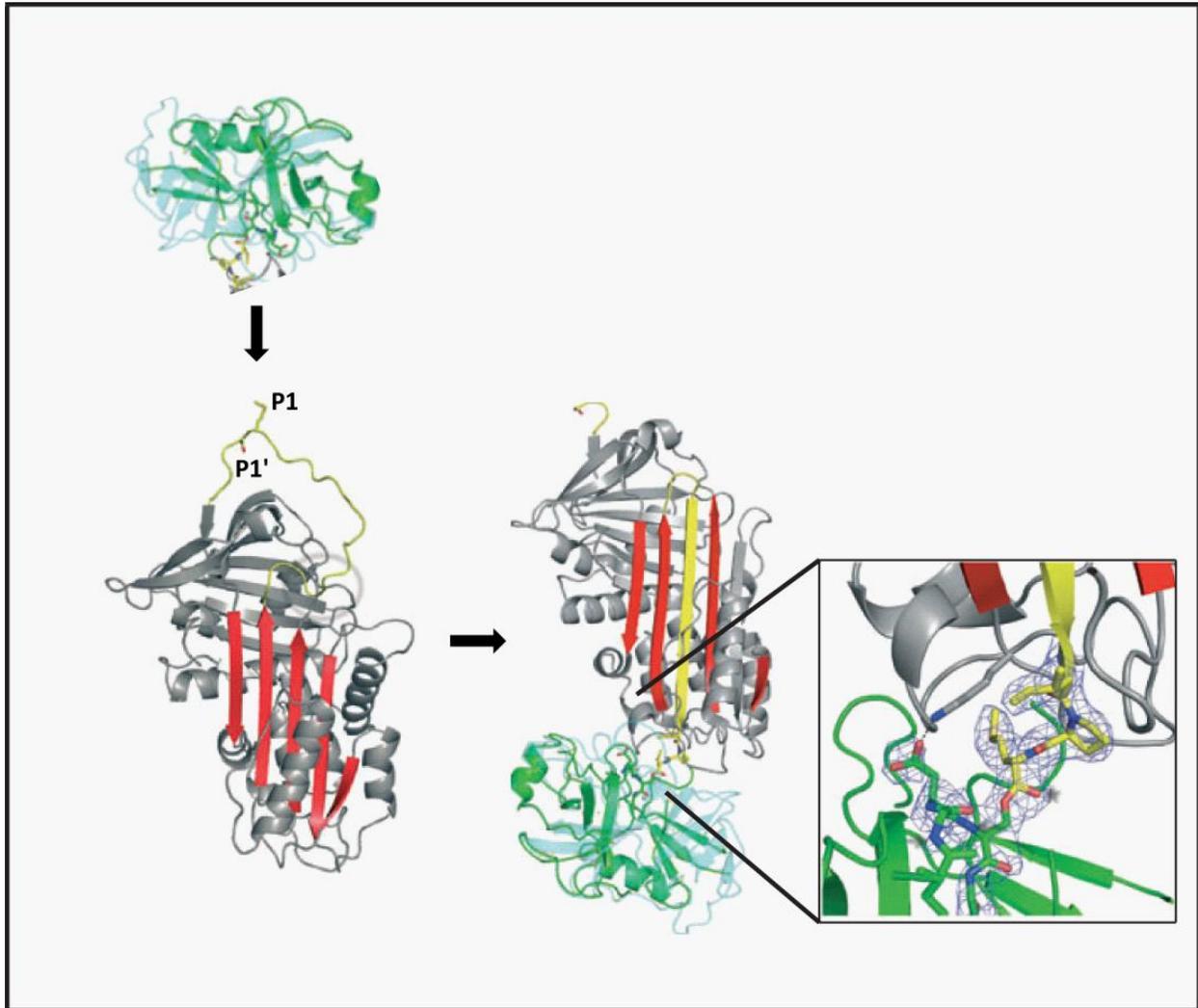


Figure 1-6: Serpin inhibition mechanism



## **Chapter 2 - Biochemical characterization of SRPN7 and SRPN18**

### **2.1 Abstract**

In insects, serine protease inhibitors (Serpins, SRPNs) are involved in numerous physiologic events such as development, mating, digestion, clotting and immunity. They are found in all tissues and throughout all life stages. These proteins form covalent complexes with target proteases which can be seen using western blot analysis by the appearance of higher weight molecular bands. In *Anopheles gambiae*, 18 serpin genes encoding 23 distinct proteins have been identified; though few have been biochemically characterized and even fewer native targets have been identified. In this study *An. gambiae* SRPN7 and 18 were characterized biochemically. One, canonical SRPN7, is predicted to be an inhibitory SRPN with trypsin like specificity though was not able to inhibit any of the native proteases tested. Published expression data shows that it is most highly expressed in larval and early pupal stages. This may indicate a possible role in development as transcript levels severely decline in late pupae and adults (Suwanchaichinda and Kanost, 2009). Another, SRPN18, which clusters on chromosome arm 2L with SRPN7, has been classified as non-inhibitory. This was confirmed after testing against an array of native and commercially available serine proteases. This study aims to aid to the current knowledge of serpin biology in the major malaria vector, *An gambiae*. Here, I show that SRPN7 is an inhibitory serpin which inhibits trypsin-like proteases. Additionally, SRPN18 is the first non-inhibitory serpin described in mosquitoes.

### **2.2 Introduction**

Insects do not have an adaptive arm of their immune system; instead they rely solely on innate defenses. Extracellular protease cascades regulate many immune responses similar to that of the

mammalian complement system. These cascades are largely made up of serine proteases such as in the Toll (Levashina et al., 1999; Ahmad et al., 2009; Fullaondo et al., 2011) and prophenoloxidase (PO) pathways (Gorman et al., 2007; Wang and Jiang, 2007; An et al., 2011). Stimulation of the Toll pathway increases the production of antimicrobial peptides (AMPs) and effector proteins (Barillas-Mury et al., 1996; De Gregorio and Lemaitre, 2002). PO activation results in melanization, or encapsulation of pathogens. If activated, these pathways are energy consuming generally culminating in protein production which requires the use of scarce amino acids. Alternately, these pathways can be toxic to the organism itself by mechanisms such as the production of toxic quinones in melanization (Nappi et al., 2009). Because it is costly for the organism to maintain these defenses as constitutively “on” many of these proteases are expressed at a basal level and circulate throughout the hemolymph in inactive zymogen forms. Zymogens may then be quickly activated upon stimulation.

In serine proteases the catalytically important residues His57, Ser195 and Asp102, known as the catalytic triad (Matthews et al., 1967; Blow et al., 1969) are located in a region known as the binding pocket and are responsible for cleaving the target protein. Serine protease specificity is determined by the amino acids that make up the binding pocket. Elastase-like serine proteases (e.g. elastase and Subtilisin Carlsberg) have a binding pocket made up of large side chain containing amino acids, and thus prefer to cleave after small neutral amino acids. Conversely, chymotrypsin’s binding pocket is made up of small side chain amino acids and can accommodate larger amino acids and prefers to cleave after aromatic amino acids. Plasmin, kallikrein, thrombin and trypsin are trypsin-like serine proteases which tend to cleave after positively charged amino acids such as Lys or Arg. These proteases have a multitude of functions in mammals including, but not limited to, blood coagulation, fibrinolysis, protein C systems and

complement activation. They are negatively regulated by a large family of proteins called serine protease inhibitors or serpins (Carrell and Owen, 1985).

Serpins are largely secreted and circulate as metastable proteins. After cleavage of the sessile bond (P1-P1') by a target protease, the serpin undergoes a dramatic conformational change where their RCL inserts into the core of the serpin and reaches its lowest energy state. As this occurs, the target proteases' active site is distorted and becomes covalently linked to the serpin. Both the serpin as well as the protease are irreversibly inhibited and thus they are referred to as suicide inhibitors. Due to the huge conformational change associated with the inhibitory mechanism, predictions can be made about serpins based on key regions within the reactive center loop (RCL).

The RCL is a portion of the protein that sits away from the body of the serpin and acts as bait for the target protease. The RCL is integral for the functionality of serpins and inhibitory potential can be predicted by analyzing this stretch of amino acids. The most important region of the RCL is the sessile bond denoted P1-P1', which is the bond cleaved by the target protease. Serpins tend to inhibit certain classes of serine proteases (e.g. trypsin-like, chymotrypsin-like or elastase-like), and this specificity is determined by the amino acid in the P1 position. In insects they are responsible for many biologically important events from mating to immune regulation (Coleman et al., 1995) (Levashina et al., 1999). Currently, only a few serpins have been shown to be involved in regulating immune responses in anophelines. *SRPN6* is expressed in hemocytes and salivary glands in *An. gambiae* and *An. stephensis* until the insect is stimulated by the murine malaria parasite, *Plasmodium bergeri*. After infection the gene is up-regulated in midgut epithelium where it may aid in preventing the parasite from exiting the midgut. This hypothesis is supported as *SRPN6* depletion in *An. stephensi* increases the number of successful oocysts

(Abraham et al., 2005). *SRPN10* (KRAL) expression is increased in *Plasmodium bergeri* infected midgut cells and has been proposed as a midgut invasion marker (Danielli et al., 2005) Finally, *SRPN2* has been characterized as a negative regulator of the PO pathway as it can directly inhibit a terminal PO activating protease (PAP), CLIPB9. Furthermore, *SRPN2* depletion causes the spontaneous appearance of melanotic pseudotumors (An et al., 2011). In this study two *An. gambiae* SRPNs have been biochemically characterized, though native targets have not yet been identified.

## **2.3 Materials and methods**

### ***2.3.1 Mosquito Rearing***

*An. gambiae* G3 line was used in these experiments. Females were allowed to oviposit on wet filter paper then floated in water to allow hatching. For 48 hours larvae were fed on 2% baker's yeast (Red Star, Active Dry Yeast) then switched to 2% w/v fish food (TetraMin Tropical Flakes, Tetra) and baker's yeast 2:1 ratio. Adult mosquitoes were fed on an 8% fructose + 2.5mM PABA (Sigma) solution ad libitum, and heparinized horse blood (Plasvacc) was fed through an artificial glass feeder using parafilm as a membrane. All stages were reared at 80% humidity, 27°C and a 12:12 light:dark cycle

### ***2.3.2 SRPN7 and 18 Cloning***

cDNA fragments encoding the full-length mature proteins were amplified using the gene-specific primers listed in Table 2.2. The forward primers included an NcoI site and the reverse primers contained an EcoR1 restriction site. A second round of PCR was performed and the forward primers contained six codons for histidine residues. The reverse primers are the same as round 1 PCR. The PCR products were digested with NcoI and EcoR1, and then inserted into the same

restriction sites into a pET-28a vector (Novagen). The plasmids were transformed into BL21 (DE3) competent *E. coli* cells and stored at -80°.

To biochemically characterize SRPN7 and 18, recombinant SRPN18 protein was expressed using an *Escherichia coli* expression system. The full coding region, minus the predicted signal peptide, was amplified using gene specific primers. SRPN18 protein was expressed using BL-21 DE3 competent *E. coli* cells. Cells containing plasmid were grown overnight at 37°C from bacterial stocks on LB agar plates containing 50µg/ml kanamycin. A single colony was inoculated into 250 ml flask containing 50ml LB containing and 50µg/ml Kanamycin then shook overnight at 37°C 150rpm. 15ml of the overnight culture was used to inoculate 2, 2 liter flasks of 500ml LB with 50µg/ml Kanamycin. Inoculated culture was incubated at 37.0°C and shook at 225rpm for approximately 2 hours to an OD<sub>600</sub> between 0.6-0.8. Protein expression was induced using 0.1mM IPTG and incubated at least 8 hours at 20 °C, 150rpm. The culture underwent centrifugation 4000rpm 20 minutes and the pellet was kept at -80°C. Cloning and purification of SRPN7 was performed by Mrs. Tinea Canady.

### ***2.3.3 Recombinant Protein Purification***

Cell pellets from expression was re-suspended in lysis buffer (300mM NaCl + 10mM imidazole) 4 ml buffer/gram pellet. Cells were lysed by sonication (Vibra Cell High Intensity Ultrasonic Processor 750 Watt model), and soluble and insoluble fractions were separated by centrifugation at 10,000g for 40min at 4.0°C. Soluble portions were retained and purified by running the sample through a column containing 2 ml Ni-NTA (Qiagen) beads. Bound proteins were washed twice with wash buffer (20ml 300µM NaCl + 20mM imidazole) to remove unbound proteins. Bound proteins were eluted in 4, 5ml fractions using elution buffer (300µM NaCl + 250mM imidazole). Salts were removed by dialysis, three times for eight hours in dialysis buffer (20mM

Tris + 20mM NaCl, pH = 8.) The sample was further purified by running the sample through column containing 2 ml Q Sepharose Fast Flow beads (GE Healthcare) equilibrated with 20mM Tris + 20mM NaCl. Unbound proteins were removed by 10 ml wash buffer (20mM Tris + 20mM NaCl). Protein was eluted off the column by a low to high salt gradient (50-200mM NaCl+20mM Tris). Fractions were analyzed by 10% SDS-PAGE and pooled. Protein concentrations were determined by Bradford assay using Comassie Plus Protein Assay Reagent (Pierce) and bovine serum albumin as a standard (Sigma). Purification of SRPN7 was done by Mrs. Tinea Canady.

#### ***2.3.4 SDS-PAGE electrophoresis and Western Blotting***

Recombinant proteins were denatured with 6X SDS buffer containing  $\beta$ -mercaptoethanol for 5 minutes at 95 °C. The samples were then separated using 10% SDS-PAGE electrophoresis for 20 min at 100 volts followed by 40 min at 180 volts. The gel was stained with EZ-Run Protein Gel Staining solution (Fisher). For immunoblot analysis proteins were transferred to PVDF membrane (GE Healthcare) at 0.110 Amps for 1 hour. Membranes were blocked in 5% skim milk for 1 hour and washed 3 times for 10 minutes in 1X Tris Buffered saline with Tween 20 then incubated with either mouse anti-His antibody (1:1000)(Genescript) or rabbit anti-CLIPB9 antibody (1:500) for at least 1 hour on a rocking shaker (Model 55, MidSci). The antibody solution was removed and the membrane was washed 3 times in 1X Tris Buffered saline with Tween 20, 10 minutes. Anti-rabbit or Anti-mouse IgG HRP conjugated antibody (Promega) was applied (1:20,000) for 1 hour and placed on rocking shaker. Antibodies were visualized by Western Lightning Chemiluminescence Reagent Plus Kit (Perkin Elmer).

### **2.3.5 Amidase Activity assays (Commercial)**

Purified rSRPN7 and 18 were tested for inhibitory activity against the following commercially available serine proteases: bacterial Subtilisin Carlsberg, fungal protease K, porcine elastase-1, human plasmin, bovine trypsin, porcine kallikrein, human thrombin, and bovine chymotrypsin. Activity of each protease was tested with different p-nitroanalide (pNa) linked colorimetric substrates (See Table 2.1). Bovine serum albumin (BSA) (Thermo Scientific) was added (3-4  $\mu$ l of 1  $\mu$ g/ $\mu$ l) in each assay to protect the protease from self-cleavage. SRPN inhibitory activity was assessed in triplicate at a 1:1 and 10:1 (SRPN:Protease) ratio by incubating BSA, SRPN and protease at room temperature for 10 min followed by addition of 200  $\mu$ l of 50  $\mu$ M pNA linked substrate to reaction and plate was immediately placed in plate reader (BioTek PowerWave Xs). Absorbance was measured every 30s, 41 times. Activity was determined by calculating average slope of reads. One unit of amidase activity was defined as  $\Delta A_{405}/\text{min}=0.001$ . Statistical analysis was done using Graphpad Prism 4. Tests included a one way ANOVA followed by Newman-Kuels post test.

Percent inhibition was determined by the equation 2.1:

$$1 - (\text{Average activity with Serpin} / \text{Average Initial activity of Protease only}) \times 100$$

Amount of SRPN was calculated by the equation 2.2:

$$\text{ng SRPN} = \text{Molar Ratio} \times (\text{ng protease} / \text{MW protease}) \times \text{MW SRPN}$$

### **2.3.6 Blood Feeding**

Single cages of 2-3 day old *An. gambiae* G3 strain mosquitoes were split into two treatment groups and placed in separate population cages. Ten females were removed from each cage and frozen in liquid nitrogen (Non-Blood fed sample). One population cage was offered a

heparinized horse blood meal for 30 min then another 10 females were removed from each cage (OH sample) and frozen in liquid nitrogen. 100 females from each population cage were collected (only blood fed for experimental cage) and separated into new population cages. Samples were then collected from each at 1, 6, 12, and 24 hours after the OH time point.

### **2.3.7 qRT-PCR**

iQ SYBR GREEN Supermix (Bio-Rad) kit was used as per manufacturer's specifications. 1  $\mu$ l cDNA template was used in each reaction. Primers were designed using the Primer Express Version 3.0.1 and are listed on Table 2.2. Relative gene expression was determined using the  $\Delta\Delta$ ct method and StepOnePlus RT-PCR machine. RPS7 was used as housekeeping gene. Statistical analysis was performed in Graphpad Prism 4. A two-way ANOVA followed by two-tailed t test was used to compare treatments and time.

### **2.3.9 Activation of CLIPB8, B9**

CLIPs were activated using 1  $\mu$ g Factor Xa (New England BioLabs) per 1.5  $\mu$ g B8 or B9 in 0.1M Tris + 2mM CaCl<sub>2</sub> (final concentration of CLIP 50ng/ $\mu$ l). 1.2  $\mu$ l .05% Tween 20 (Fisher) was added per 1  $\mu$ g Factor Xa in CLIPB8 reactions to increase the fraction of activated protease. Reactions were incubated at 37° C for 12-36 hours.

### **2.3.10 Florescence assays. (B8/B9)**

SRPN7 and 18 protein was tested against recombinant CLIP Serine proteases at 10:1 SRPN:protease ratio. rSRPNs were incubated with BSA and protease for ten min at room temperature then 100  $\mu$ l of 200  $\mu$ M AMC linked substrate (FVR) was added. Reactions were immediately placed in florescence reader (Wallac Victor<sup>3</sup> 1420 Multilabel counter, PerkinElmer) and reads were taken 41 times at 30 second intervals and the slope was calculated.

## 2.4 Results

***SRPN7 and 18 RCL implicate SRPN7 as an inhibitory serpin with trypsin-like specificity while SRPN18 is unlikely to be inhibitory.***

The RCL of known mosquito serpins range from 12 to 47 amino acids, and very few contain less than 28. This is most likely because this number is required for the RCL to stretch over the serpin core (Gulley et. al.). *An. gambiae* SRPN7 and its orthologs in *C. quinquefasciatus* and *Ae. aegypti* are canonical serpins with the average RCL length of 28 amino acids and Arg residue at the predicted P1 site. Furthermore, the hinge is conserved and is made up of amino acids containing small side chains. Based on these observations, I hypothesized that *An. gambiae* SRPN7 is an inhibitory serine protease inhibitor with trypsin like specificity. *An. gambiae* SRPN18, on the other hand has a very short RCL containing only 23 amino acids. Additionally, the hinge region contains an Isoleucine, which has a large side chain as well as a Proline. The presence of proline at even numbers of amino acids away from the sessile bond has been shown to decrease inhibitory ability of serpins (Perry et al., 1989; Stein and Carrell, 1995). The decrease may be caused firstly by reducing the flexibility in hinge region and secondly because these even numbered amino acids are the ones that get buried into the core of the serpin after RCL insertion (Gettins, 2002). Based on this, I predict that SRPN18 is not an inhibitory SRPN (Fig. 2.1).

***Recombinant protein and purification yield pure proteins.***

To investigate the validity of inhibitory activity of these SRPNs, 6-His tagged proteins were expressed in an *Escherichia coli* expression system. Soluble protein was purified using Nickel and Q-sepharose chromatography. The calculated MW of mature SRPN7 and 18 is 42.9 and

42.6 kDa, respectively. Under reducing conditions SRPN7 and 18 migrated to approximately, 40kDa on a 10% SDS gel and did not contain any significant contaminating protein shown in Figure 2-2 and 2-3. Protein concentration was determined by the Bradford assay using Bovine serum albumin (Sigma) as the standard and verified via Nano Drop2000 (Thermo Scientific). SRPN7 fraction 13 was used in experiments and its concentration was determined to be 185ng/ $\mu$ l. SRPN18 fractions 14-17 were pooled and the final concentration was determined to be 1.7  $\mu$ g/ $\mu$ l.

### ***SRPN18 transcripts do not increase after blood meal***

Female anophelines are obligatory hematophagous insects that require protein from the blood meal for egg production and expression of 33% of all *An. gambiae* genes are altered after the injection of a blood meal (Marinotti et al., 2005). To determine if *SRPN18* expression levels change in response to the blood meal, *SRPN18* was monitored at different times after blood meal and compared to sugarfed mosquitoes. *SRPN18* transcripts did not significantly increase or decrease at any timepoint after blood meal (Fig 2.4).

### ***SRPN7 inhibits Trypsin like serine proteases***

To test for activity of SRPN7 and 18, each serpin was incubated with an array of commercially available serine proteases with trypsin (plasmin, trypsin, kallikrein and thrombin), chymotrypsin (protease K and chymotrypsin) or elastase (elastase and subtilisin Carlsberg) like activities. The predicted P1 for SRPN7 is Arg which is the preferred amino acid after which trypsin-like serine proteases cleave. SRPN7 inhibited three of the trypsin like serine proteases tested: trypsin, plasmin and kallikrein. SRPN7 inhibited trypsin and plasmin at a 10:1 SRPN:protease ratio and kallikrein at a 1:1 and 10:1 ratio. SRPN18, which has a predicted P1 of Phe, was able to marginally inhibit trypsin-like serine proteases, kallikrein, trypsin and plasmin at a 10:1 ratio

26.3%, 32.1%, and 20.1% respectively (Fig 2.5). Assays for SRPN7 and 18 are summarized in Table 2.1. See Appendix for detailed graphs for all assays

### ***SRPN7 and 18 do not inhibit CLIPB8 or CLIPB9***

In *An. gambiae*, only CLIPB9 has been shown to be a phenoloxidase activating protease (PAP) (An et al., 2011). Additionally, another CLIP, CLIPB8 has been shown to be involved in the PPO pathway, though its biochemical function within the pathway is still unclear (Paskewitz et al., 2006). In an attempt to identify native targets, both SRPNs were incubated with recombinant CLIPB8 and CLIPB9. FVRase activity was measured and compared to the FVRase activity of CLIPB8 and B9 alone. Neither SRPN7 nor 18 were able to decrease the FVRase activity *in vitro* (Fig2.6). SRPN18 was further analyzed by western blot to confirm that no higher molecular weight band appeared (Fig 2.7).

## **2.5 Discussion**

Currently, 18 SRPN genes encoding 23 proteins have been identified in *An. gambiae* (Christophides et al., 2002; Waterhouse et al., 2007; Suwanchaichinda and Kanost, 2009). Of these few have been characterized. SRPN6 and 10 are both upregulated in response to parasite infection and have been proposed as midgut invasion markers (Abraham et al., 2005; Danielli et al., 2005). Additionally, SRPN2 has been shown to be a negative regulator of the melanization response and can directly inhibit *An. gambiae* PAP, CLIPB9 (An et al., 2011). SRPN1 has been implicated in the PPO pathway as it can inhibit *M sexta* PAP3, and prevent spontaneous melanization of *M sexta* plasma *in vitro* (Michel et al., 2006). The function of all other SRPNs remains to be shown. In this study two additional SRPNs, SRPN7 and 18 were biochemically characterized.

SRPN7 is a mosquito specific SRPN, and previous expression profiles have shown that it is most highly expressed in larvae but it is expressed at low levels in adults, mainly in female heads (Michel et al., 2005; Suwanchaichinda and Kanost, 2009; Baker et al., 2011). It maintains a canonical RCL with conserved residues in key areas necessary for the dramatic conformational change associated with serpin inhibition. SRPN7 was predicted to be an inhibitory serpin of serine proteases with trypsin-like specificity. This was verified by the inhibitory effects on the trypsin like serine proteases tested (trypsin, plasmin and kallikrein) as well as its lack of inhibition of chymotrypsin and elastase like serine proteases.

One to one orthologs of all *An. gambiae* serpins can be found in *Ae. aegypti* and *C. quinquefasciatus* with the notable exception of SRPN13 and 18. SRPN18 has only recently been annotated (Waterhouse et al., 2007), and is expressed in most tissues (Marinotti et al., 2005) and throughout all life stages (Suwanchaichinda and Kanost, 2009). *SRPN18* clusters with *SRPN7* and *14* close to the *SRPN2* cluster on chromosome arm 2L. Based on the conservation of key amino acids making up the RCL and identity of the predicted P1, SRPN18 was predicted to be a non-inhibitory SRPN. When SRPN18 was tested against an array of commercially available serine proteases marginal inhibition of trypsin like serine proteases (trypsin, plasmin, and kallikrein) at high SRPN:protease ratios was observed. This could be explained by erroneous P1 prediction for SRPN18 of Phe. Alternately, SRPN18 may have acted as a substrate for these proteases. The protease could cleave SRPN18 at alternate sites besides the P1. Alternately, if the protease cleaves at the P1 quickly enough the serpin will not have sufficient time to insert into the  $\beta$  sheet and will instead function as a substrate. This would seem to be more likely as we only see marginal decrease in protease activity at the higher SRPN:protease ratio and can be tested by N-terminal sequencing. SRPN18 did not inhibit chymotrypsin-like serine proteases

which further substantiates the prediction that SRPN18 is not functioning as a typical serine protease inhibitor. Further analysis showed that SRPN7 and 18 were unable to inhibit two CLIPs, CLIPB8 and CLIPB9, known to be involved in the melanization response in *An. gambiae*.

## 2.6 References

- Abraham, E.G., Pinto, S.B., Ghosh, A., Vanlandingham, D.L., Budd, A., Higgs, S., Kafatos, F.C., Jacobs-Lorena, M., Michel, K., 2005. An immune-responsive serpin, SRPN6, mediates mosquito defense against malaria parasites. *Proc Natl Acad Sci U S A* 102, 16327-16332.
- Ahmad, S.T., Sweeney, S.T., Lee, J.A., Sweeney, N.T., Gao, F.B., 2009. Genetic screen identifies serpin5 as a regulator of the toll pathway and CHMP2B toxicity associated with frontotemporal dementia. *Proc Natl Acad Sci U S A* 106, 12168-12173.
- An, C., Budd, A., Kanost, M.R., Michel, K., 2011. Characterization of a regulatory unit that controls melanization and affects longevity of mosquitoes. *Cell Mol Life Sci* 68, 1929-1939.
- Baker, D.A., Nolan, T., Fischer, B., Pinder, A., Crisanti, A., Russell, S., 2011. A comprehensive gene expression atlas of sex- and tissue-specificity in the malaria vector, *Anopheles gambiae*. *BMC Genomics* 12, 296.
- Barillas-Mury, C., Charlesworth, A., Gross, I., Richman, A., Hoffmann, J.A., Kafatos, F.C., 1996. Immune factor Gambif1, a new rel family member from the human malaria vector, *Anopheles gambiae*. *EMBO J* 15, 4691-4701.
- Blow, D.M., Birktoft, J.J., Hartley, B.S., 1969. Role of a buried acid group in the mechanism of action of chymotrypsin. *Nature* 221, 337-340.
- Carrell, R.W., Owen, M.C., 1985. Plakalbumin, alpha 1-antitrypsin, antithrombin and the mechanism of inflammatory thrombosis. *Nature* 317, 730-732.
- Christophides, G.K., Zdobnov, E., Barillas-Mury, C., Birney, E., Blandin, S., Blass, C., Brey, P.T., Collins, F.H., Danielli, A., Dimopoulos, G., Hetru, C., Hoa, N.T., Hoffmann, J.A., Kanzok, S.M., Letunic, I., Levashina, E.A., Loukeris, T.G., Lycett, G., Meister, S., Michel, K., Moita, L.F., Muller, H.M., Osta, M.A., Paskewitz, S.M., Reichhart, J.M., Rzhetsky, A., Troxler, L., Vernick, K.D., Vlachou, D., Volz, J., von Mering, C., Xu, J., Zheng, L., Bork, P., Kafatos, F.C., 2002. Immunity-related genes and gene families in *Anopheles gambiae*. *Science* 298, 159-165.
- Coleman, S., Drahn, B., Petersen, G., Stolorov, J., Kraus, K., 1995. A *Drosophila* male accessory gland protein that is a member of the serpin superfamily of proteinase inhibitors is transferred to females during mating. *Insect Biochem Mol Biol* 25, 203-207.
- Danielli, A., Barillas-Mury, C., Kumar, S., Kafatos, F.C., Loukeris, T.G., 2005. Overexpression and altered nucleocytoplasmic distribution of *Anopheles* ovalbumin-like SRPN10 serpins in *Plasmodium*-infected midgut cells. *Cell Microbiol* 7, 181-190.
- De Gregorio, E., Lemaitre, B., 2002. The mosquito genome: the post-genomic era opens. *Nature* 419, 496-497.
- Fullaondo, A., Garcia-Sanchez, S., Sanz-Parra, A., Recio, E., Lee, S.Y., Gubb, D., 2011. Spn1 regulates the GGBP3-dependent Toll signaling pathway in *Drosophila melanogaster*. *Mol Cell Biol* 31, 2960-2972.
- Gettins, P.G., 2002. Serpin structure, mechanism, and function. *Chem Rev* 102, 4751-4804.
- Gorman, M.J., Wang, Y., Jiang, H., Kanost, M.R., 2007. *Manduca sexta* hemolymph proteinase 21 activates prophenoloxidase-activating proteinase 3 in an insect innate immune response proteinase cascade. *J Biol Chem* 282, 11742-11749.

- Levashina, E.A., Langley, E., Green, C., Gubb, D., Ashburner, M., Hoffmann, J.A., Reichhart, J.M., 1999. Constitutive activation of toll-mediated antifungal defense in serpin-deficient *Drosophila*. *Science* 285, 1917-1919.
- Marinotti, O., Nguyen, Q.K., Calvo, E., James, A.A., Ribeiro, J.M., 2005. Microarray analysis of genes showing variable expression following a blood meal in *Anopheles gambiae*. *Insect Mol Biol* 14, 365-373.
- Matthews, B.W., Sigler, P.B., Henderson, R., Blow, D.M., 1967. Three-dimensional structure of tosyl-alpha-chymotrypsin. *Nature* 214, 652-656.
- Michel, K., Budd, A., Pinto, S., Gibson, T.J., Kafatos, F.C., 2005. *Anopheles gambiae* SRPN2 facilitates midgut invasion by the malaria parasite *Plasmodium berghei*. *EMBO Rep* 6, 891-897.
- Michel, K., Suwanchaichinda, C., Morlais, I., Lambrechts, L., Cohuet, A., Awono-Ambene, P.H., Simard, F., Fontenille, D., Kanost, M.R., Kafatos, F.C., 2006. Increased melanizing activity in *Anopheles gambiae* does not affect development of *Plasmodium falciparum*. *Proc Natl Acad Sci U S A* 103, 16858-16863.
- Nappi, A., Poirie, M., Carton, Y., 2009. The role of melanization and cytotoxic by-products in the cellular immune responses of *Drosophila* against parasitic wasps. *Adv Parasitol* 70, 99-121.
- Paskewitz, S.M., Andreev, O., Shi, L., 2006. Gene silencing of serine proteases affects melanization of Sephadex beads in *Anopheles gambiae*. *Insect Biochem Mol Biol* 36, 701-711.
- Perry, D.J., Harper, P.L., Fairham, S., Daly, M., Carrell, R.W., 1989. Antithrombin Cambridge, 384 Ala to Pro: a new variant identified using the polymerase chain reaction. *FEBS Lett* 254, 174-176.
- Sievers, F., Wilm, A., Dineen, D., Gibson, T.J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Soding, J., Thompson, J.D., Higgins, D.G., 2011. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* 7, 539.
- Stein, P.E., Carrell, R.W., 1995. What do dysfunctional serpins tell us about molecular mobility and disease? *Nat Struct Biol* 2, 96-113.
- Suwanchaichinda, C., Kanost, M.R., 2009. The serpin gene family in *Anopheles gambiae*. *Gene* 442, 47-54.
- Wang, Y., Jiang, H., 2007. Reconstitution of a branch of the *Manduca sexta* prophenoloxidase activation cascade in vitro: snake-like hemolymph proteinase 21 (HP21) cleaved by HP14 activates prophenoloxidase-activating proteinase-2 precursor. *Insect Biochem Mol Biol* 37, 1015-1025.
- Waterhouse, R.M., Kriventseva, E.V., Meister, S., Xi, Z., Alvarez, K.S., Bartholomay, L.C., Barillas-Mury, C., Bian, G., Blandin, S., Christensen, B.M., Dong, Y., Jiang, H., Kanost, M.R., Koutsos, A.C., Levashina, E.A., Li, J., Ligoxygakis, P., Maccallum, R.M., Mayhew, G.F., Mendes, A., Michel, K., Osta, M.A., Paskewitz, S., Shin, S.W., Vlachou, D., Wang, L., Wei, W., Zheng, L., Zou, Z., Severson, D.W., Raikhel, A.S., Kafatos, F.C., Dimopoulos, G., Zdobnov, E.M., Christophides, G.K., 2007. Evolutionary dynamics of immune-related genes and pathways in disease-vector mosquitoes. *Science* 316, 1738-1743.

## 2.7 Tables

**Table 2.1: SRPN7 and 18 inhibition of commercially available proteases**

| Protease  | Subtilisin<br>Carlsberg | Fungal<br>Protease K | Porcine<br>Elastase | Human<br>Plasmin | Bovine<br>Trypsin<br>Gold | Porcine<br>Kallikrein | Human<br>Thrombin | Bovine<br>Chymotrypsin |            |
|-----------|-------------------------|----------------------|---------------------|------------------|---------------------------|-----------------------|-------------------|------------------------|------------|
| Substrate | FVR-pNA                 | AAPF-pNA             | AAPF-pNA            | PFR-pNA          | IEAR-pNA                  | IEGR-pNA              | FVR-pNA           | AAPF-pNA               |            |
| SRPN7     | 1:1                     | 4.6%± 4.8            | 4.0%± 7.0           | -                | 6.7%± 11.6                | 3.0%± 5.2             | 54.0%± 47.8*      | 5.2%± 7.1              | 1.7%± 2.9  |
|           | 10:1                    | 9.1%± 9.1            | 9.3%± 8.2           | -                | 78.9%± 9.2*               | 36.3%± 26.2*          | 99.0%± 0.6*       | 1.8%± 3.1              | 8.1%± 14.0 |
| SRPN18    | 1:1                     | 4%                   | 5.8%±5.2            | 5.8%±0.4         | 12.9%±8.4                 | 8.2%±3.7              | 8.9%±3.1          | 0.0%±0.0               | 0.0%± 0.1  |
|           | 10:1                    | 7%                   | 8.2%±8.4            | 9.8%±2.3         | 20.1%±2.6**               | 32.1%±4.7***          | 26.3%±8.5*        | 0.0%±0.0               | 0.2%±0.4   |

One-way ANOVA, post test Newman-Keuls as compared to protease and substrate alone. \* $p < 0.001$ , \*\* $p < 0.01$ , \*\*\* $p < 0.05$

**Table 2.2: Primer sequences**

| Name                       | Sequence   |
|----------------------------|--|
| <u>rProtein production</u> |  |
| SRPN7-1F                   | 5'-TCATCACGGCCAAGGAACCACC-3'                               |
| SRPN7-2F                   | 5'-AT{ <b>CCATGG</b> }GCC <b>CATCATCATCATCAT</b> CACGGC-3' |
| SRPN7-1R                   | 5'-TT[ <b>GAATTC</b> ]CTAGTTGTGGTTAGAATAGTG-3'             |
| SRPN7-2R                   | 5'-TT[ <b>GAATTC</b> ]TCAAAACTGTTCATCGG-3'                 |
| SRPN18-1F                  | 5'-TCATCACGGCGATCCTACGACAG-3'                              |
| SRPN18-2F                  | 5'-AT{ <b>CCATGG</b> }GCC <b>CATCATCATCATCAT</b> CACGGC-3' |
| SRPN18-1R                  | 5'-TT[ <b>GAATTC</b> ]TCAAAACTGTTCATCGG-3'                 |
| SRPN18-2R                  | 5'-TT[ <b>GAATTC</b> ]TCAAAACTGTTCATCGG-3'                 |
| <u>qRT-PCR</u>             |  |
| SRPN18-qF                  | 5'-GCCAACACCGAGATTGAAGATT-3'                               |
| SRPN18-qR                  | 5'-GGACACACCGCTGAACAG-3'                                   |
| <u>dsRNA synthesis</u>     |  |
| SRPN18R-F                  | 5'- <b>TAATACGACTCACTATAGGG</b> AGCTACGCCAGATCGTGAAC-3'    |
| SRPN18R-R                  | 5'- <b>TAATACGACTCACTATAGGG</b> AAAAGTTCATCGGGATCG-3'      |
| T7                         | 5'- <b>TAATACGACTCACTATAGGG</b> -3'                        |

**BOLD** = T7 sequence, **BOLD** = His tail, [**BOLD**] = ECOR1 site, {**BOLD**} = NCO1 site  
q = quantitative

## 2.8 Figure Legends

### **Figure 2-1: Amino acid sequence alignment of reactive center loops from AgSRPNs and homologues**

Alignment of *An. gambiae* SRPN7 and 18 reactive center loop with orthologous genes in *Aedes aegypti* and *Culex quinquefasciatus*. Alignment was performed in “Clustal Omega”(Sievers et al., 2011). Predicted scissile bond is indicated by red line. Consensus sequence is on bottom. Hinge region is in blue.

### **Figure 2-2: SRPN 7 Recombinant Protein Expression and Purification**

Coomassie Blue <sup>TM</sup> stained 10% SDS gel after (A) Nickel (Ni-NTA) affinity chromatography or (B) Q-Sepharose chromatography. (A) From left to right: Pre-induction (PI), Hours post protein induction (hpi) 1, 4, 7, Insoluble protein (IS), Soluble protein(S), Marker (M), Flow Thru (FT), Wash 1, 2, Elutions 1-4, Marker (M). Protein was eluted out in 300mM NaCl + 250mM imidazole pH= 8.0. (B) Protein was run through column containing Q-Sepharose beads washed, and eluted in a 50-500mM pH= 8.0 linear NaCl gradient. From right to left Flow Thru (FT), Wash (W), and elutions 2-27.

### **Figure 2-3: SRPN18 Recombinant Protein Expression and Purification**

Coomassie Blue <sup>TM</sup> stained 10% SDS gel after (A) Nickel (Ni-NTA) affinity chromatography or (B) Q-Sepharose chromatography. (A) From left to right: Marker (M), Pre-induction (PI), Hours post protein induction (hpi) 1, 6, 18, Insoluble protein (IS), Soluble protein(S), Flow Thru (FT) 1, 2, Wash 1, 2, Elutions 1-4. Protein was eluted out in 300mM NaCl + 250mM imidazole pH= 8.0. (B) Protein was run through column containing Q-Sepharose beads washed, and eluted in a 50-500mM pH= 8.0 linear NaCl gradient. From right to left Marker (M), and elutions 2-29.

### **Figure 2-4: SRPN18 transcripts do not change after blood meal**

qRT-PCR analysis of SRPN18 transcript abundance . Relative expression of SRPN18 after blood meal time course. Expression was normalized to RPS7. Nonblood fed control samples were used

as calibrator. Statistical tests performed is 2 way ANOVA. Error bars are mean and SEM of three replicates

**Figure 2-5: SRPN7 and 18 inhibition of commercial proteases**

(A) SRPN7 protein was incubated with plasmin or (B) SRPN18 incubated with kallikrein for ten minutes at room temperature at different SRPN:protease molar ratios. 200  $\mu$ l of 50 $\mu$ M substrate (PFR-pNA or IEGR-pNA respectively) was added and absorbance was measured immediately every 30 s for 41 times at 405nm. Bar graphs represent residual activity and dot graphs represent individual reads during the assay. Error bars are mean and SEM of three biological replicates.

**Figure 2-6: SRPN7 and 18 do not inhibit CLIPB8 or B9**

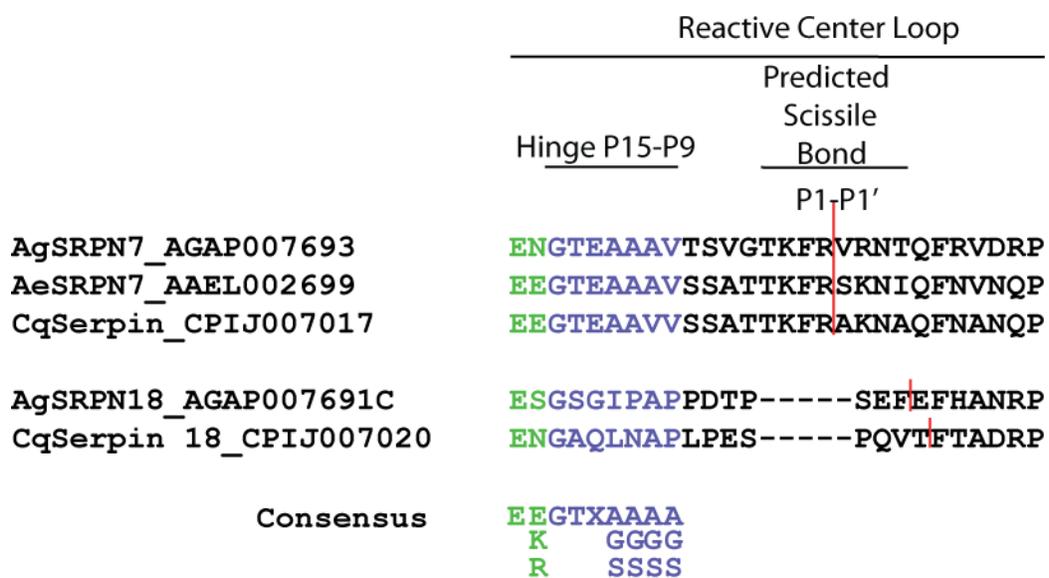
SRPN7 or 18 was incubated with Factor Xa activated (A) CLIPB9 or (B) CLIPB8 at room temperature for Ten minutes. FVR-AMC was used as substrate and FVRase activity was measured immediately every 30 s for 41 times. Error bars are mean and SEM of three biological replicates.

**Figure 2-7: SRPN18 does not form complex with CLIPB8 or B9**

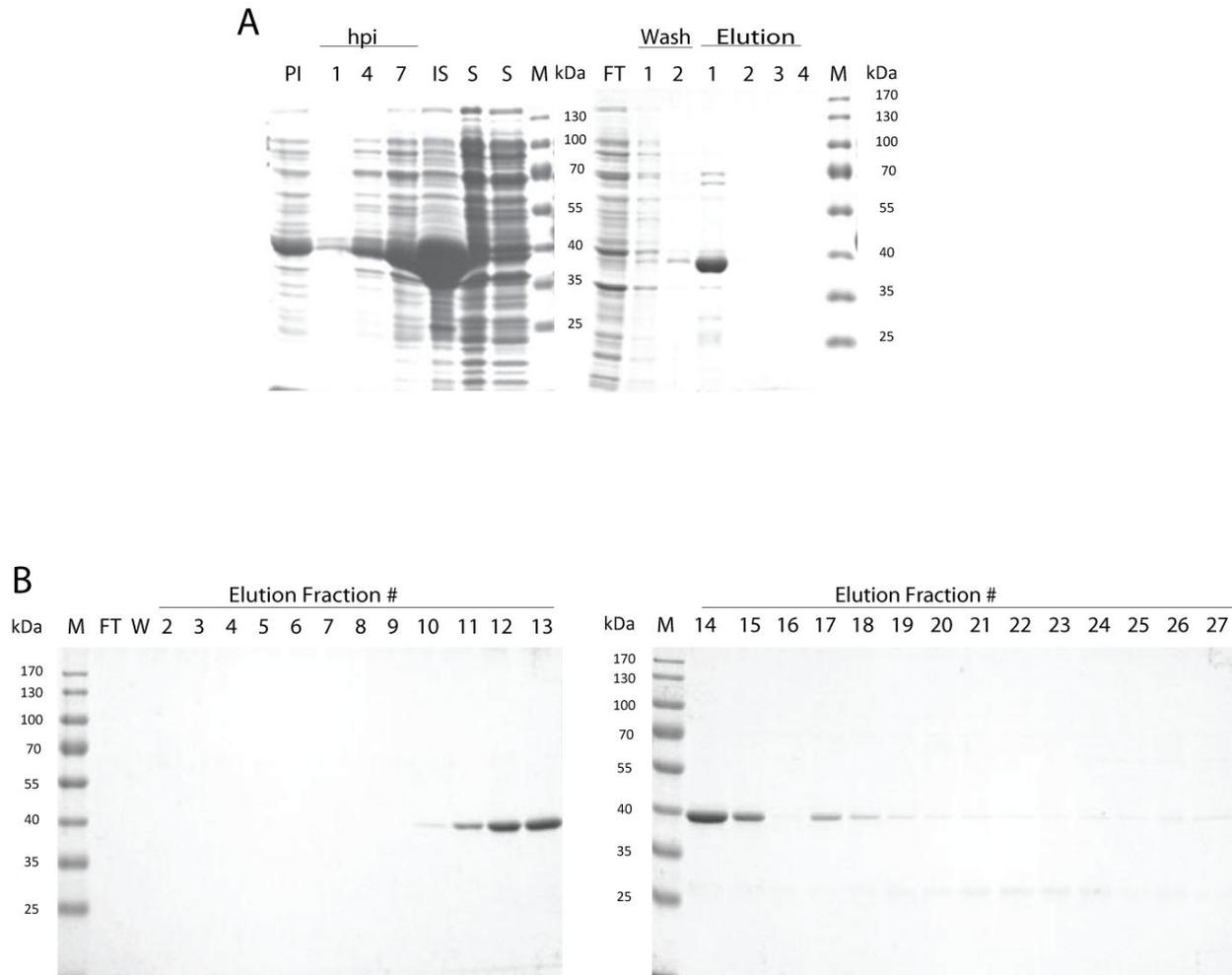
Activation of recombinant CLIPB8 and CLIPB9 and no complex formation via Western Blot analysis using anti-His antibodies

## 2.9 Figures

**Figure 2-1:** *Amino acid sequence alignment of reactive center loops from AgSRPNs and homologues*



**Figure 2-2: *SRPN 7 Recombinant Protein Expression and Purification***



**Figure 2-3: SRPN 18 Recombinant Protein Expression and Purification**

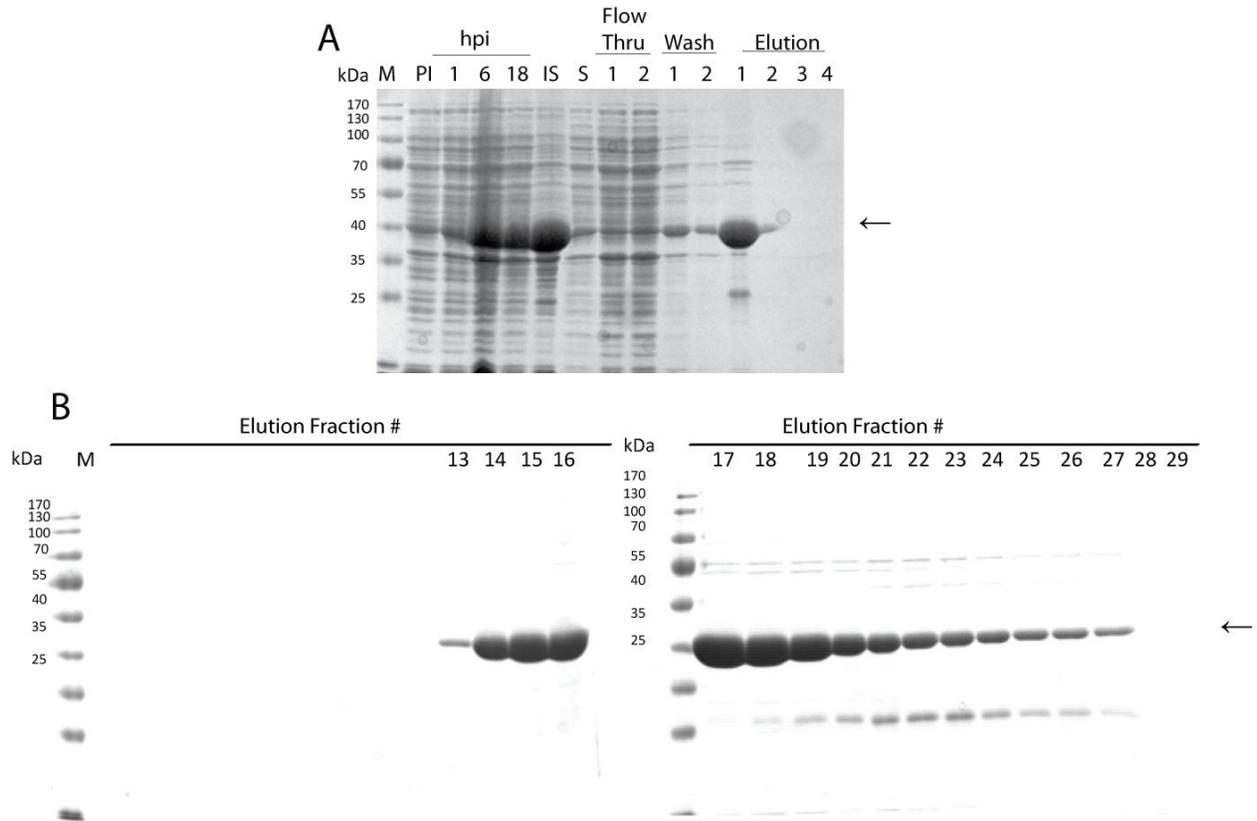


Figure 2-4: *Blood meal has no effect on SRPN18 transcript abundance.*

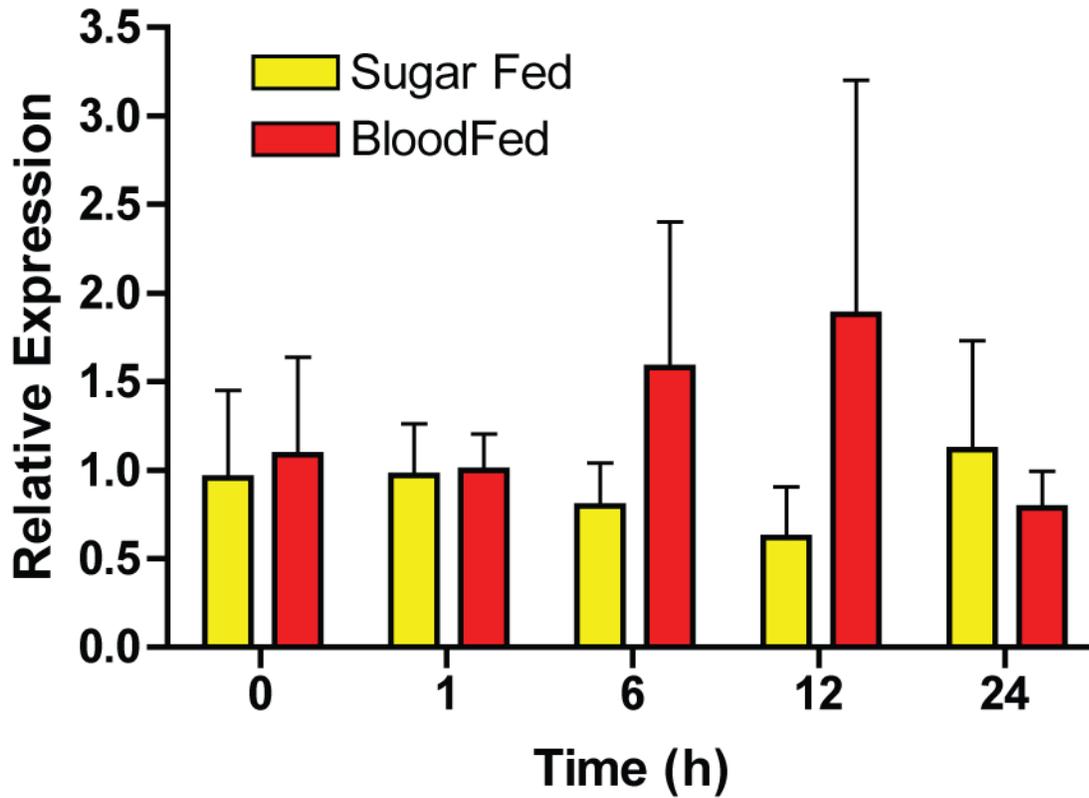


Figure 2-5: *SRPN inhibition of commercial proteases*

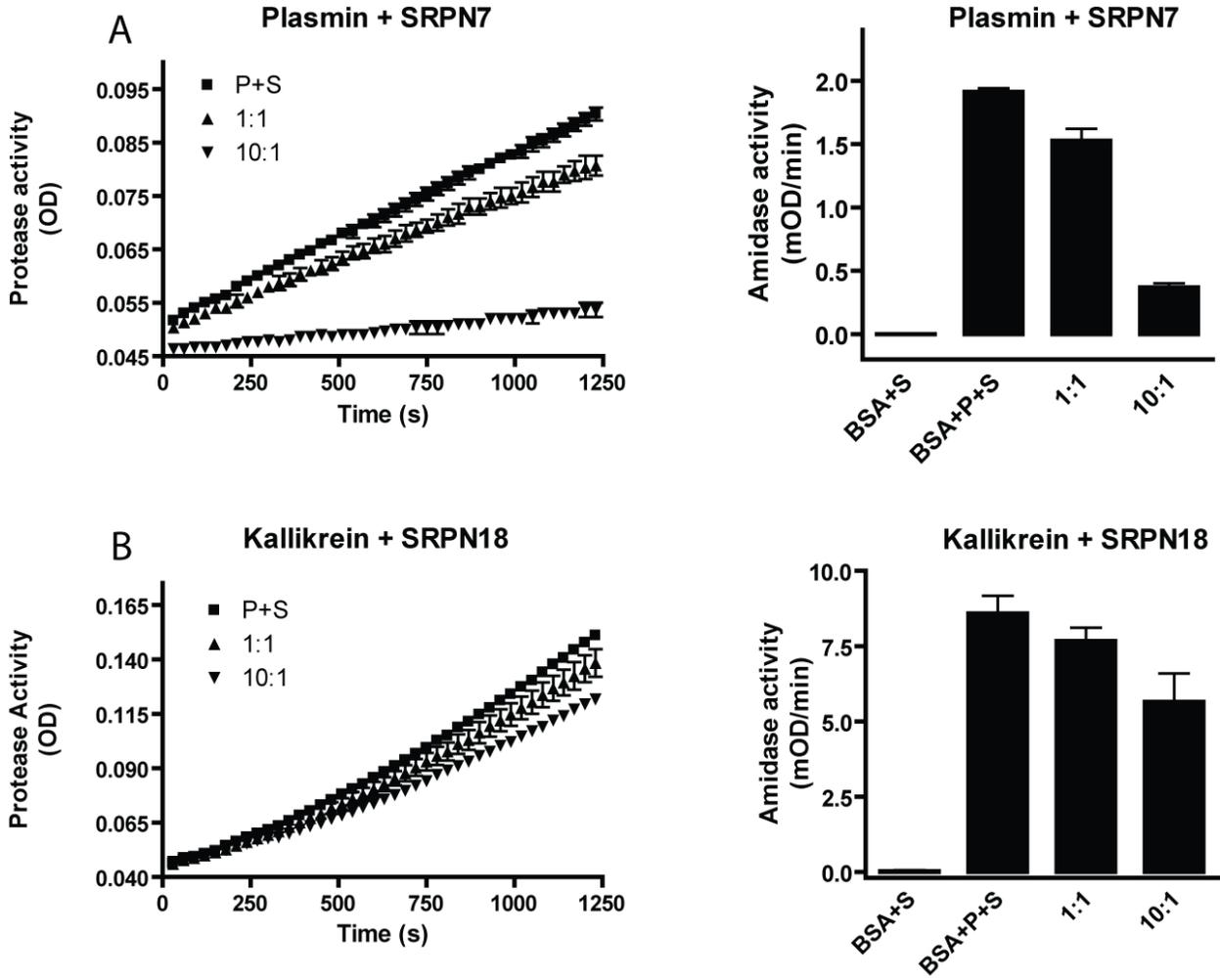
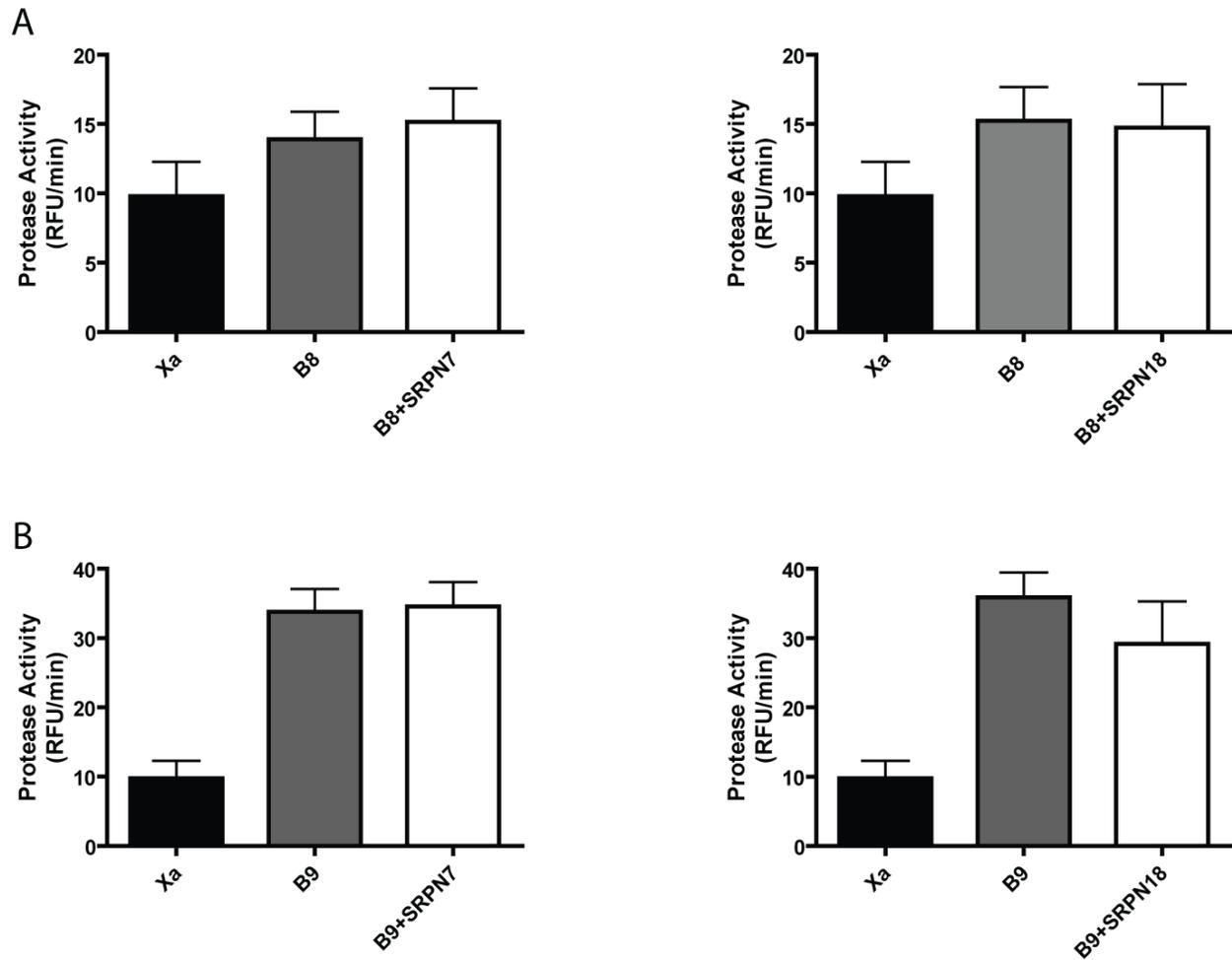
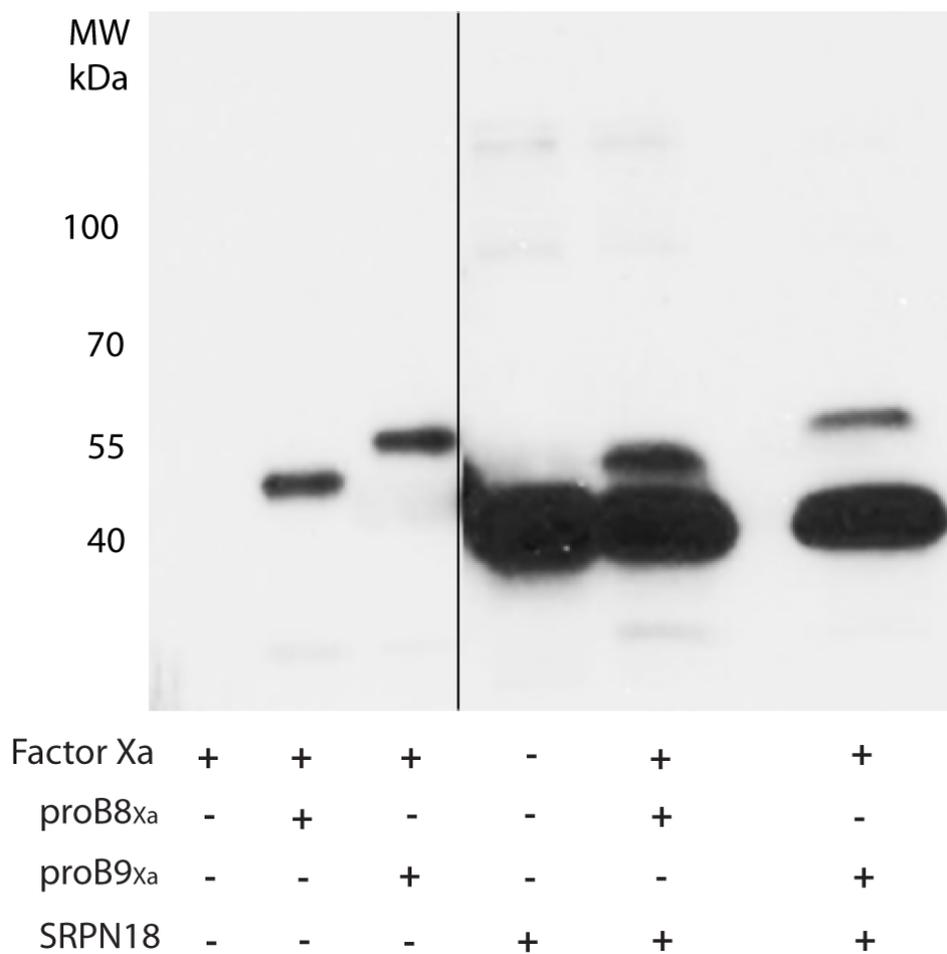


Figure 2-6: *SRPN7* and *18* do not inhibit *CLIPB8* or *B9*



**Figure 2-7:** *SRPN18* does not form complex with *CLIPB8* or *B9*



## **2.10 Acknowledgements**

I would like to thank Mrs. Tinea Canady for SRPN7 cloning, expression and purification.. I would like to thank KaraJo Srigg, John Cladwell, Amanda Osarczuk, and Sarah Smith for mosquito rearing. I would like to thank Dr Chunju An and Dr. Xin Zhang for kindly providing CLIPB8 and B9 for experiments. Thanks to Dr. Kristin Michel for collecting M. sexta plasma. I would also like to thank Dr. Michael Kanost and his lab for allowing me the use of their equipment. I would also like to thank Dr. Yasuaki Hiromasa for performing the Mass spectrometry analysis.

## Chapter 3 - Biochemical Characterization of SRPN9

### 3.1 Abstract

One important aspect of Anopheline immunity is melanization, the encapsulation of pathogens in the insoluble pigment eumelanin. This immune response is partly responsible for certain strains of mosquitoes being refractive to *Plasmodium sp.* infection. Melanization occurs after pathogen recognition stimulates extracellular CLIP domain serine protease (CLIPs) cascades. Clips circulate as zymogens and are sequentially activated. In the melanization pathway, terminal CLIPs known as prophenoloxidase activating proteases (PAPs), ultimately cleave inactive prophenoloxidase (PPO) to Phenoloxidase (PO). PO converts mono and diphenols to quinones which polymerize to eumelanin. This process is thought to limit oxygen accessibility, essentially suffocating the pathogen. Due to harmful byproducts of eumelanin synthesis, such as superoxides, semi-quinones and reactive oxygen species, these cascades are tightly regulated by one class of serine protease inhibitors, called serpins (SRPNs). Although reverse genetic approaches have identified several CLIPs that act as agonists/antagonists in the melanization response, the extracellular protease cascade is largely unknown. To date in *An. gambiae*, only SRPN2 has been shown to be a negative regulator of a terminal phenoloxidase activating protease (PAP), CLIPB9. Using biochemical techniques I characterized SRPN9 as a suicide inhibitor. I found that SRPN9 was not able to inhibit CLIPB8 but did inhibit known PAP, CLIPB9. Kinetic analysis of SRPN9 and CLIPB9 showed that that Stoichiometry of Inhibition is 1.33 and the rate of association ( $k_a$ ) for this SRPN to protease is  $9.8 \times 10^4 \text{M}^{-1} \text{s}^{-1}$ . Overall these data show *An. gambiae* SRPN9 is a functional serine protease inhibitor of trypsin-like serine proteases and of CLIPB9 in particular.

### 3.2 Introduction

Serpins are the largest family of serine protease inhibitors found ubiquitously in prokaryotes as well as eukaryotes. These inhibitors tend to inhibit serine proteases with trypsin like specificity while others have evolved to perform different biological functions such as storage proteins and molecular chaperones (Hunt and Dayhoff, 1980; Flink et al., 1986). These metastable proteins contain 350-400 amino acids, are chiefly secreted and circulate in a stressed (S) state (Carrell and Owen, 1985). They share limited sequence similarity, yet are structurally conserved. The body of the serpin contains up to nine  $\beta$ -sheets surrounded by  $\alpha$ -helices. A so-called reactive center loop (RCL) sits out from the body of the protein and acts as a lure for target proteases. The catalytic serine of the target protease performs nucleophilic attack on the sessile bond (denoted P1-P1') of the serpin. Upon cleavage of the RCL, the SRPN undergoes a massive conformational change in which the RCL flips around and inserts into  $\beta$ -sheet A (Huntington et al., 2000). As this insertion occurs an acyl intermediate (Lawrence et al., 1995) is formed followed by covalent bond formation between the serpin and protease. This traps the protease and destroys its active site as the protease is dragged to the opposite pole of the serpin (Huntington et al., 2000).

Serpins are known as suicide inhibitors because both the protease and the SRPN are irreversibly inhibited in this interaction (Gettins, 2002).

Extracellular immune cascades in insects, such as the PO activation pathway, consist of sequentially activating serine proteases. These serine proteases contain an N-terminal disulfide knotted domain (called a paperclip like domain) and a C-terminal trypsin-like protease domain held together by a linker region (Muta et al., 1990). CLIPs are a large family of proteins. Fifty-six CLIPs have been identified in *An. gambiae* (Waterhouse et al., 2007) yet few show 1:1

orthology indicating species specific gene expansion(Christophides et al., 2002). Many CLIPs are continually expressed at a basal level which allows them to be quickly activated in an acute phase response to pathogens. CLIPs circulate in arthropod hemolymph in an inactive zymogen form and are activated by cleavage between the CLIP domain and the catalytic domain by limited proteolysis (Muta et al., 1990). Additionally, several CLIPs (CLIPAs and Es) are non-catalytic proteases called serine protease homologues (SPH) and instead function as co factors needed for the propagation of the PO signal (Nappi et al., 2009). Ten CLIPAs have been identified in the *An. gambiae* genome (Christophides et al., 2002). Three, CLIPA2, A5, and A7 are putative SPH and are involved in melanization of parasites (Volz et al., 2006). Previous studies using reverse genetics and biochemical techniques have shown several CLIPBs (CLIPB3, B4, B8, B9, B14, B17) to be involved in the melanization response(Volz et al., 2005; Paskewitz et al., 2006; Volz et al., 2006; An et al., 2011). Of the 18 genes encoding 23 distinct SRPNs in *An. gambiae* only SRPN2 has been shown to inhibit a phenoloxidase activating protease (PAP), CLIPB9(An et al., 2012). *An. gambiae* SRPN9 orthologs, *D. melanogaster* Spn5 and *M. sexta* Serpin-6, negatively regulate Toll and PPO activity (Zou and Jiang, 2005; Ahmad et al., 2009). Here, we test whether SRPN9 is a functional serine protease inhibitor that can prevent hemolymph melanization in insects.

### **3.3 Materials and methods**

#### **3.3.1 Mosquito Rearing**

*An. gambiae* G3 line was used in these experiments. Females were allowed to oviposit on wet filter paper then floated in water to allow hatching. For 48 hours larvae were fed on 2% baker's yeast (Red Star, Active Dry Yeast) then switched to 2% w/v fish food (TetraMin Tropical

Flakes, Tetra) and baker's yeast 2:1 ratio. Adult mosquitoes were fed on an 8% fructose + 2.5mM PABA (Sigma) solution ad libitum, and heparinized horse blood (Plasvacc) was fed through an artificial glass feeder using parafilm as a membrane. All stages were reared at 80% humidity, 27°C and a 12:12 light:dark cycle.

### ***3.3.2 Mosquito Injection***

Female mosquitoes were collected 2-3 days post emergence, placed in paper cups and provided sugar water. Females were anesthetized by exposure to CO<sub>2</sub> and injected (Drummond Nanoject™ II) with 69nl of dsRNA (2.89µg/µl). Mosquitoes were allowed to recover in paper cups and sugar water was supplied. Survival was recorded daily.

### ***3.3.3 SRPN9 Cloning***

A cDNA fragment encoding the full-length mature SRPN9 was amplified using the gene-specific primers listed in Table 3-4. The forward primer included an NcoI site and the reverse primer contained an ECOR1 restriction site. A second round of PCR was performed and the forward primer contained six codons for histidine residues. The reverse primer is the same as round 1 PCR. The PCR product was digested with NcoI and EcoR1, and then inserted into the same restriction sites into a pET-28a vector (Novagen). The plasmid was transformed into BL21 (DE3) competent *E. coli* cells and stored at -80°. Cloning was done previously by Mrs. Tinea Canady.

### ***3.3.4 Recombinant Proteins***

To biochemically characterize SRPN9, recombinant SRPN9 protein was expressed using an *Escherichia coli* expression system. The full coding region, minus the predicted signal peptide, was amplified using gene specific primers. SRPN9 protein was expressed using BL-21 DE3 competent *E. coli* cells. Cells containing plasmid were grown overnight at 37°C from bacterial

stocks on LB agar plates containing 50µg/ml kanamycin. A single colony was inoculated into 250 ml flask containing 50ml LB containing and 50µg/ml Kanamycin then shook overnight at 37°C 150rpm. 15ml of the overnight culture was used to inoculate 2, 2 liter flasks of 500ml LB with 50µg/ml Kanamycin. Inoculated culture was incubated at 37.0°C and shook at 225rpm for approximately 2 hours to an OD<sub>600</sub> between 0.6-0.8. Protein expression was induced using 0.1mM IPTG and incubated at least 8 hours at 20 °C, 150rpm. The culture underwent centrifugation 4000rpm 20 minutes and the pellet was kept at -80°C.

### ***3.3.5 Recombinant Protein Purification***

Pellet from expression was re-suspended in lysis buffer (300mM NaCl + 10mM imidazole) 4 ml buffer/grams pellet. Cells were lysed by sonication (Vibra Cell High Intensity Ultrasonic Processor 750 Watt model) and soluble and insoluble fractions were separated by centrifugation at 10,000g for 40min at 4.0°C. Soluble portions were retained and purified by running the sample through a column containing 2 ml Ni-NTA (Qiagen) beads. Bound proteins were washed twice with wash buffer (20ml 300µM NaCl + 20mM imidazole) to remove unbound proteins. Bound proteins were eluted in 4, 5ml fractions using elution buffer (300µM NaCl + 250mM imidazole). Salts were removed by dialysis, three times for eight hours in dialysis buffer (20mM Tris + 20mM NaCl, pH = 8.) The sample was further purified by running the sample through column containing 2 ml Q Sepharose Fast Flow beads (GE Healthcare) equilibrated with 20mM Tris + 20mM NaCl. Unbound proteins were removed by 10 ml wash buffer (20mM Tris + 20mM NaCl). Protein was eluted off the column by a low to high salt gradient (50-200mM NaCl+20mM Tris). Fractions were analyzed by 10% SDS-PAGE and pooled. Protein concentrations were determined by Bradford assay using Comassie Plus Protein Assay Reagent (Pierce) and bovine serum albumin as a standard (Sigma).

### ***3.3.6 SRPN9 Polyclonal Antibody Production***

500µg recombinant SRPN9 protein was ran on 10% SDS-PAGE gel and stained with EZ-Run Protein Gel Staining solution (Fisher). The gel containing SRPN9 was cut in slices and sent by Mrs. Tinea Cannady to produce polyclonal rabbit antisera (Cocalico). Antibodies were purified by incubating antibodies with 400µg SRPN9 bound to a nitrocellulose membrane (GE Healthcare) overnight at 4°C. Antibodies that did not bind to SRPN9 were washed from the nitrocellulose membrane 3 times with 1% Tris-Tween Buffered Saline (TTBS) to cover membrane. Purified SRPN9 antibodies were eluted in 3, 1 ml elutions of 0.2M glycine and 1M tris base, pH= 7.5 elutions and stored at -80°C in 30 ml aliquots.

### ***3.3.7 SDS-PAGE electrophoresis and Western Blotting***

Recombinant proteins were denatured with 6X SDS buffer containing β-mercaptoethanol for 5 minutes at 95 °C. The samples were then separated using 10% SDS-PAGE electrophoresis for 20 min at 100 volts followed by 40 min at 180 volts. The gel was stained with EZ-Run Protein Gel Staining solution (Fisher). For immunoblot analysis proteins were transferred to PVDF membrane (GE Healthcare) at 0.110 Amps for 1 hour. Membranes were blocked in 5% skim milk for 1 hour and washed 3 times for 10 minutes in 1X Tris Buffered saline with Tween 20 then incubated with either rabbit anti-SRPN9 antibody (1:500) or rabbit anti-CLIPB9 antibody (1:500) for at least 1 hour on a rocking shaker (Model 55, MidSci). The antibodies were removed and the membrane was washed 3 times in 1X Tris Buffered saline with Tween 20, 10 minutes. Anti-rabbit IgG HRP conjugated antibody (Promega) was applied (1:20,000) for 1 hour and placed on rocking shaker. Antibodies were visualized by Western Lightning Chemiluminescence Reagent Plus Kit (Perkin Elmer).

### **3.3.8 Amidase Activity assays (Commercial)**

Purified rSRPN9 was tested for inhibitory activity against the following commercially available serine proteases: bacterial Subtilisin Carlsberg, fungal protease K, porcine elastase-1, human plasmin, bovine trypsin, porcine kallikrein, human thrombin, and bovine chymotrypsin. Activity of each protease was tested with a p-nitroanalide (pNa) linked colorimetric substrates (See Table 3.1). Bovine serum albumin (BSA) (Thermo Scientific) was added (3-4  $\mu\text{l}$  of  $1\mu\text{g}/\mu\text{l}$ ) in each assay to protect the protease from self-cleavage. SRPN inhibitory activity was assessed in triplicate at a 1:1 and 10:1 (SRPN:Protease) ratio by incubating BSA, SRPN and protease (See Table3-2) at room temperature for 10 min followed by addition of  $200\mu\text{l}$  of  $50\mu\text{M}$  pNA linked substrate to reaction and plate was immediately placed in plate reader (BioTek PowerWave Xs). Absorbance was measured every 30s, 41 times. Activity was determined by calculating average slope of reads. One unit of amidase activity was defined as  $\Delta A_{405}/\text{min}=0.001$ . Statistical analysis was done using Graphpad Prism 4. Tests included a one way ANOVA followed by Newman-Kuels post test.

Percent inhibition was determined by the equation 3.1:

$$1 - (\text{Average activity with Serpin} / \text{Average Initial activity of Protease only}) \times 100$$

Amount of SRPN was calculated by the equation 3.2:

$$\text{ng SRPN} = \text{Molar Ratio} \times (\text{ng protease} / \text{MW protease}) \times \text{MW SRPN}$$

### **3.3.9 Blood Feeding**

Single cages of 2-3 day old mosquitoes were split into two treatment groups and placed in separate population cages. Ten females were removed from each cage and frozen in liquid nitrogen (Non-Blood fed sample). One population cage was offered a heparinized horse blood

meal for 30 min then another 10 females were removed from each cage (0H sample) and frozen in liquid nitrogen. 100 females from each population cage were collected (only blood fed for experimental cage) and separated into new population cages. Samples were then collected from each at 1, 3, 12, and 48 hours after the 0H time point.

### ***3.3.10 RNA Isolation***

Frozen mosquitoes were homogenized in 1ml Trizol and centrifuged at 15000 rpm, 4°C for 30 min. Aqueous phase was collected and RNA was precipitated using isopropanol. RNA was pelleted by centrifugation at 15000 rpm, 4°C for 30 min and washed twice in 70% ethanol. RNA was dissolved in 100 µl RNase free water and stored at -80°C. DNA was removed by DNase digestion. Isolated RNA was mixed with 11.5µl Buffer RDD (Qiagen) and 2.88 µl DNase1 (Qiagen) (2.7unit/µl) and incubated at RT for 10 minutes. RNA CLEAN-UP was performed as per manufacturer's instructions (Quiagen RNeasy Mini-kit). RNA quality was assessed by Agilent BioAnalyzer using a nano chip as per manufacturer's instructions and concentration was determined by Nanodrop 8000.

### ***3.3.11 cDNA Synthesis***

cDNA Synthesis was performed in 20µl reaction volumes. 2 µl of 50ng/µl total RNA was used with the iSCRIPT cDNA Synthesis kit, as per manufacturer's instructions.

### ***3.3.12 qRT-PCR***

iQ SYBR GREEN Supermix (Bio-Rad) kit was used as per manufacturer's specifications. 1µl cDNA template was used in each reaction. Primers were designed using the Primer Express Version 3.0.1 and are listed on Table 3.4. Relative gene expression was determined using the  $\Delta\Delta\text{Ct}$  method and StepOnePlus RT-PCR machine. RPS7 was used as housekeeping gene.

Statistical analysis was performed in Graphpad Prism 4. A two-way ANOVA followed by two-tailed t test was used to compare treatments and time.

### **3.3.13 dsRNA synthesis**

Gene specific primers with T7 tails were designed using E-RNAi webservice (version 3.2)(Horn and Boutros, 2010), see Table 3-4. Vector plasmid Pet28a (Novagen) containing the *SRPN9* gene was amplified the gene using Promega Go taq kit. All PCR product was gel extracted using Quiagen Gel Extraction Kit and used as template for 2<sup>nd</sup> PCR using T7 primers. DNA was precipitated in isopropanol at -20°C for at least one hour, and then pelleted by 30 minutes centrifugation at 13000rpm at 4 °C. DNA was washed twice with one ml 70% ethanol. 1µg of purified PCR product was used as template for dsRNA synthesis in a 20µl reaction volume as per Ampliscribe T7-Flash Transcription kit (Epicenter) instructions. Reaction mixture was treated with RNase Free DNase1 (2.7 units/µl) for 20min at 37 °C to remove DNA template and RNase free water was added to total volume of 300 µl. RNA was extracted with equal volumes of Phenol/Chloroform/Isoamyl alcohol (Fisher) and separated from chloroform by centrifugation. The aqueous phase was transferred into a new tube and RNA was precipitated with isopropanol for 1 hour at -20°C. RNA was pelleted by centrifugation 30 minutes at 4°C. The supernatant was removed and RNA was washed twice with one ml 70% ethanol. Pellet was dissolved in 30 µl dd RNase free water and diluted to 2.89 µg/µl.

### **3.3.14 PO activity assays**

Plasma was collected by gravity from 2 day 5<sup>th</sup> instar *Manduca sexta* larvae by clipping the horn with dissecting scissors. Hemocytes were removed by centrifugation in a tabletop Eppendorf centrifuge at 8000 rpm for 10 minutes and stored at -80°C. Plasma collection was done by Dr. Kristin Michel. The plasma was incubated for 10 min with 0, 100, 200 or 400ng SRPN9 at room

temperature. Subsequently, proPO activity was activated in the reaction mixture by adding 10 $\mu$ g *Micrococcus luteus* (Sigma) and incubating the reaction at room temperature for 10 minutes. 200 $\mu$ l of 2mM dopamine-HCL was added to the reaction and immediately placed in plate reader (BioTek PowerWave Xs) Absorbance was measured 60 times at 30 second intervals. Statistical tests were performed in Graphpad Prism 4. Tests included a one way ANOVA followed by Newman-Kuels post test.

### ***3.3.15 Activation of CLIPB8, B9***

CLIPs were activated using 1 $\mu$ g Factor Xa (New England BioLabs) per 1.5  $\mu$ g B8 or B9 in 0.1M Tris + 2mM CaCl<sub>2</sub> (final concentration of CLIP 50ng/ $\mu$ l). 1.2 $\mu$ l .05% Tween 20 (Fisher) was added per 1 $\mu$ g Factor Xa in CLIPB8 reactions to increase the fraction of activated protease. Reactions were incubated at 37° C for 12-36 hours.

### ***3.3.16 Florescence assays. (B8/B9)***

SRPN9 protein was tested against recombinant CLIP Serine proteases at a 10:1 SRPN:protease ratio. SRPN9 was incubated with BSA and protease for ten min at room temperature then 100 $\mu$ l of 200 $\mu$ M AMC linked substrate (FVR) was added. Reactions were immediately place in florescence reader (Wallac Victor<sup>3</sup> 1420 Multilabel counter, PerkinElmer) and reads were taken every 41 times at 30 second intervals and the slope was calculated.

### ***3.3.17 Stoichiometry of Inhibition***

Different molar ratios of SRPN9 (6:1, 4:1, 2:1, 1.5:1, 1:1, .5:1, 0:1) was incubated with 100ng CLIPB9 and 3 $\mu$ gBSA at room temperature for ten min. 100 $\mu$ l of 300 $\mu$ M FVR-AMC substrate was added and FVRase activity was measured every 30 sec, 61 times. Percent residual activity was calculated and plotted against SRPN9:CLIPB9 molar ratio. X intercept, when Y=0 of linear regression of this data gives SI.

### 3.3.18 Association Constant

100µl of 300µM FVR-AMC was added to SRPN9 at 0, 4.3, 8.6, 13, 17.3, 25.9, 34.5 and 51.8 nM plus 213nM BSA. Activated CLIPB9 (2.37pmol) was added and FVRase activity was measured 50 times every thirty seconds.  $k_{obs}$  is calculated by nonlinear regression for each serpin:protease concentration (see equation below).  $k_{obs}$  is plotted against respective serpin:protease concentration. Linear regression of this gives a line.  $k'$  was calculated as the slope of this line.  $k_a$  was calculated (See below)

$P = V_0 / k_{obs} [1 - e^{-(1 - k_{obs}t)}]$  P=concentration of product at time (t)

$$k_a = k' (1 + [FVR]/K_{M(FVR)})SI$$

## 3.4 Results

### 3.4.1 SRPN9 has a canonical reactive center loop and is predicted to be inhibitory

Previous studies showed that SRPN9 orthologues in *D. melanogaster* and *M. sexta* are important regulators of TOLL and PO pathways (Zou and Jiang, 2005; Ahmad et al., 2009). To predict if *An. gambiae* SRPN9 could be inhibitory and whether it may play similar roles as its orthologs, the RCL was aligned with *D. melanogaster*, *M. sexta*, *A. aegypti*, *C. quinquefasciatus*, *A. mellifera* orthologs and canonical serpin, human  $\alpha$ -1-antitrypsin. (Fig 3-1) Many inhibitory serpins are cleaved by proteases with trypsin-like specificity, which tend to cleave after Arg or Lys. Arg at the predicted P' of the RCL for all SRPN9 orthologus thus suggests that these serpins will inhibit trypsin-like serine proteases. Additionally, overall sequence similarity of the SRPN9 RCL with that of its orthologes, suggests that AgSRPN9 is likely inhibitory and has similar targets and functions in *An gambiae*.

### 3.4.2 Recombinant SRPN9 expression and purification yielded pure protein

To investigate biochemical properties such as SRPN9 inhibitory activity and potential targets, 6-His tagged SRPN9 was expressed in an *E. coli* expression system. Soluble protein was purified using Nickel and Q-sepharose chromatography. The calculated molecular weight of rSRPN9 is 48.0 kDa and under reducing condition rSRPN9 migrated to 50 kDa on a 10% SDS gel and did not contain any significant contaminating protein. (Figure 3-3) Fractions 13-17 were pooled and yielded 5mg of rSRPN9 in a total volume of 5 ml. rSRPN9 concentration was determined to be 1.1µg/µl by Bradford assay using Bovine serum albumin (Sigma) as the standard and verified via Nano Drop2000 (Thermo Scientific). Nucleic acid and deduced amino acid sequences are shown in Figure 3-2.

#### **3.4.3 SRPN9 depletion does not affect mosquito survival**

Knockdown of *SRPN9* by dsRNA did not affect survival when compared to GFP injected controls. Unlike SRPN2, another CLIPB9 inhibitor, there was no visible phenotype associated with SRPN9 depletion, yet SRPN9 is indeed depleted in dsRNA injected females. (Fig. 3-4)

#### **3.4.4 SRPN9 increases after blood meal**

Female anophelines are obligatory hematophagous insects that require protein from the blood meal for egg production. The ingestion of blood causes altered expression of 33% of all *An. gambiae* genes (Marinotti et al., 2005). To determine if *SRPN9* expression levels change in response to the blood meal, *SRPN9* was monitored at different times after blood meal and compared to sugarfed mosquitoes. Similar to experiments in *A. aegypti*, *SRPN9* levels increase over time after ingestion of a blood meal. (Fig. 3-5)

#### **3.4.5 SRPN9 inhibits trypsin-like serine proteases**

Based on the sequence analysis I predicted that rSRPN9 would inhibit trypsin-like serine proteases based on the Arg at the predicted P1. To test this hypothesis and determine the

inhibitory activity and specificity, rSRPN9 was tested against an array of commercially available proteases. Percent inhibition was calculated as described in Materials and Methods and is summarized in Table 3-2. As predicted, SRPN9 was able to inhibit all tested trypsin like proteases: plasmin, trypsin, kallikrein, and thrombin at a 10:1 SRPN:protease ratio, and all but thrombin at a 1:1 ratio. Interestingly, SRPN9 was also able to inhibit chymotrypsin at both a 10:1 as well as a 1:1 ratio. (Fig 3-6) Individual replicates for each SRPN:Protease are shown in Appendix A.

#### **3.4.6 rSRPN9 decreases *M. sexta* phenoloxidase activity**

Biochemical analysis on mosquito hemolymph is difficult due to the limited volume per adult that can be collected. To determine if SRPN9 may be involved in inhibiting PO activity, the well characterized *M. sexta* system was utilized. Increasing amounts of rSRPN9 were incubated with *M. sexta* hemolymph then PO activity was stimulated by addition of *M. luteus*. Addition of 400, 200 and 100ng rSRPN9 decreased PO activity by 58.8, 49.5 and 29.4% respectively. rSRPN9 was able to significantly decrease PO activity in *M. sexta* plasma in a dose dependent manner presumably by inhibiting a protease or proteases in the PO activating cascade. (Fig 3-7)

#### **3.4.7 rSRPN9 forms SDS stable complex with rCLIPB9 but not rCLIPB8**

CLIPB8 and CLIPB9 are involved in the melanization reaction in *An. gambiae* (Paskewitz et al., 2006; An et al., 2011) and CLIPB9 is a PAP able to directly activate phenoloxidase (An et al., 2011). To study the potential inhibitory effects of SRPN9 on these proteases, rSRPN9 was incubated at room temperature with bovine Factor Xa activated rCLIPB8<sub>Xa</sub> and rCLIPB9<sub>Xa</sub>. FVRase activity was measured and complex formation was investigated by Western blot analysis using anti-CLIPB8, anti-CLIPB9 and anti-SRPN9 antibodies. Western blot analysis showed rSRPN9 was able to form complex with rCLIPB9<sub>Xa</sub>. Incubation of rSRPN9 and rCLIPB9<sub>Xa</sub>

showed the appearance of a high molecular weight band approximately 75kDa, which is the expected mass of CLIPB9/SRPN9 complex. This complex was recognized by anti-SRPN9 and anti-CLIPB9 antibodies. (Fig 3-8) To determine if the high molecular weight band was indeed a complex of rSRPN9 and rCLIPB9, rSRPN9 and rCLIPB9 were incubated and ran on a 10% SDS-PAGE gel. The gel was stained with EZ-Run Protein Gel Staining solution. The high molecular weight band was excised (Fig 3-9) and analyzed by 1D NanoLC ESI MS/MS. MS/MS confirmed that rSRPN9 and rCLIPB9 could form a complex as they were present in the high molecular weight band (approximately 75 kDa). (Table 3-3)

#### ***3.4.8 rSRPN9 to rCLIPB9 has low SI and similar $k_a$ as other arthropod serpin protease combinations***

Kinetic analysis was performed to determine the stoichiometry of inhibition (SI) and rate of association ( $k_a$ ) of rSRPN9 and rCLIPB9<sub>Xa</sub> under pseudo-second order rate conditions. The SI was determined by incubating increasing SRPN:protease molar ratios for 10 min at room temperature. Percent residual activity was plotted against SRPN:protease molar ratio. SI is extrapolated by X when Y=0, or the SRPN:protease ratio at which the protease has zero activity. The  $k_a$  was measured by mixing different molar ratios of SRPN:protease and immediately measuring FVRase activity.  $K_{obs}$  was calculated by nonlinear regression. The slope of this line equals  $k'$ , which is used to calculate the association rate  $k_a = k' (1 + [S]/K_M)$  SI. 1.3 molecules of SRPN9 was able to inhibit 1 molecule of CLIPB9 (Fig. 3-10). The calculated  $K_a$  of SRPN9 for CLIPB9 was  $9.8 \times 10^4 M^{-1} s^{-1}$  which is within the range of known serpins. (Fig. 3-11) This data shows that *in vitro*, rSRPN9 is acting as a good inhibitor of this PAP.

### 3.5 Discussion

In mammals, extracellular serine proteases are involved in a variety of processes from wound healing, blood coagulation, fibrinolysis, and complement activation (Hunt and Dayhoff, 1980; Patston et al., 1991). In arthropods, these serine protease cascades are involved in dorsal-ventral axis patterning, Toll activation, melanization, and coagulation. Serpins are the largest group of inhibitors of serine proteases in metazoans, which can inhibit multiple targets as can be seen in *M. sexta* serpin-6, which can form complexes with HP8, PAP-3 and immulectin-2. This is also observed in mammals, for example Plasminogen activator inhibitor-1 can inhibit tissue-type plasminogen activator and the urokinase-type plasminogen activator (Zou and Jiang, 2005) (Fortenberry, 2013).

The inhibitory activity of serpins can be predicted based on the composition of the RCL. Target proteases cleave the peptide bond between the P1 and P1'. Based on the distance from the hinge region, the P1 for SRPN9 in *An. gambiae* is Arg (Gulley et al., 2013) and thus SRPN9 is predicted to inhibit proteases with trypsin-like specificities. This is supported as SRPN9 had no effect on subtilisin-like proteases subtilisin Carlsberg and protease K or elastase-like serine protease porcine elastase-1 which preferentially cleave after small neutral amino acids. This is further supported as SRPN9 was able to reduce trypsin, plasmin and kallikrein activity by 38.6, 75.6, and 96.8% respectively at a 1:1 SRPN to protease ratio. At a 10:1 ratio over 90% of the amidase activity of all three proteases was inhibited. SRPN9 appears to be a promiscuous inhibitor as it was also able to significantly inhibit chymotrypsin, which prefers to cleave after aromatic amino acids, at a 1:1 and 10:1 ratio (49.9 and 81.9%, respectively). This suggests that Phe at the predicted P2 position can also act as the P1. The identity of the P1 in these different reactions will have to be confirmed experimentally, and if proven would be a novel idea.

Melanization is important in limiting successful parasite numbers within the mosquito in laboratory vector/parasite pairs (Collins et al., 1986). The first step in determining if SRPN9 was a physiologically relevant regulator of melanization in *An. gambiae* was to directly test its ability to reduce PO activity in an *ex vivo* system. SRPN9 was able to decrease PO activity in *M. sexta* hemolymph in a dose dependent manner. To further investigate the mechanism for this inhibition, SRPN9 was tested against *An. gambiae* CLIPs. Although CLIPB8 has yet to be biochemically characterized it has a predicted role in melanization (Paskewitz et al., 2006) and CLIPB9 can directly activate PPO to PO (An et al., 2011). SRPN9 had no inhibitory effect on CLIPB8 but was able to significantly decrease amidase activity of the only known *An. gambiae* PAP, CLIPB9. Furthermore, Western blot and Mass Spec analysis confirm that rSRPN9 and rCLIPB9 were able to form an SDS stable covalent complex *in vitro*. Biochemical characterization of rSRPN9 and rCLIPB9 also proves that SRPN9 is a good inhibitor of CLIPB9 with a low SI (1.3) and a  $K_a$  similar to other arthropod serpins ( $9.8 \times 10^4 \text{M}^{-1} \text{s}^{-1}$ ).

Due to its ability to inhibit many different serine proteases, across classes (trypsin like and chymotrypsin like) it is likely that SRPN9 may have different functions in various tissues. Similar to its orthologs it may be involved in Toll regulation. *SRPN9* expression increases after blood meal and is expressed in midguts and salivary glands. Taken with the data that show SRPN9 is able to inhibit blood serine proteases (plasmin and thrombin) it may also be involved in inhibiting blood proteases. SRPN9 may be functioning as the female attempts to obtain the blood meal or in the midgut after the blood meal is acquired. SRPN2 has been shown to negatively regulate melanization and is able to inhibit CLIPB9 and SRPN2 knockdown mosquitoes show spontaneous melanization and a decreased lifespan (An et al., 2011). If SRPN9 were the major regulator of CLIPB9, we would expect to see a phenotype similar to the

well characterized SRPN2, yet there was no visible phenotype associated with SRPN9 knockdown. This may be explained if SRPN2 is not the physiological regulator of CLIPB9. SRPN2 could be a better inhibitor of yet identified proteases involved in the extracellular PO activating cascade. Another explanation is that SRPN2 is expressed highly in *An. gambiae* hemolymph when compared to SRPN9. If SRPN2 is present in a much higher concentration than SRPN9, knockdown of SRPN9 would in effect be rescued by the presence of SRPN2 in the hemolymph. To tease out these specifics, more experiments will have to be done. For now, the key finding in this study is that *An. gambiae* SRPN9 is an inhibitory SRPN and is able to inhibit proteases with trypsin and chymotrypsin-like specificities including known PAP, CLIPB9 *in vitro*.

### 3.6 References

- Ahmad, S.T., Sweeney, S.T., Lee, J.A., Sweeney, N.T., Gao, F.B., 2009. Genetic screen identifies serpin5 as a regulator of the toll pathway and CHMP2B toxicity associated with frontotemporal dementia. *Proc Natl Acad Sci U S A* 106, 12168-12173.
- An, C., Budd, A., Kanost, M.R., Michel, K., 2011. Characterization of a regulatory unit that controls melanization and affects longevity of mosquitoes. *Cell Mol Life Sci* 68, 1929-1939.
- An, C., Hiromasa, Y., Zhang, X., Lovell, S., Zolkiewski, M., Tomich, J.M., Michel, K., 2012. Biochemical characterization of *Anopheles gambiae* SRPN6, a malaria parasite invasion marker in mosquitoes. *PLoS One* 7, e48689.
- Bioinformatics, C.e.a.B., 2007. Protein Family Alignment Annotation Tool. In.
- Carrell, R.W., Owen, M.C., 1985. Plakalbumin, alpha 1-antitrypsin, antithrombin and the mechanism of inflammatory thrombosis. *Nature* 317, 730-732.
- Christophides, G.K., Zdobnov, E., Barillas-Mury, C., Birney, E., Blandin, S., Blass, C., Brey, P.T., Collins, F.H., Danielli, A., Dimopoulos, G., Hetru, C., Hoa, N.T., Hoffmann, J.A., Kanzok, S.M., Letunic, I., Levashina, E.A., Loukeris, T.G., Lycett, G., Meister, S., Michel, K., Moita, L.F., Muller, H.M., Osta, M.A., Paskewitz, S.M., Reichhart, J.M., Rzhetsky, A., Troxler, L., Vernick, K.D., Vlachou, D., Volz, J., von Mering, C., Xu, J., Zheng, L., Bork, P., Kafatos, F.C., 2002. Immunity-related genes and gene families in *Anopheles gambiae*. *Science* 298, 159-165.
- Collins, F.H., Sakai, R.K., Vernick, K.D., Paskewitz, S., Seeley, D.C., Miller, L.H., Collins, W.E., Campbell, C.C., Gwadz, R.W., 1986. Genetic selection of a *Plasmodium*-refractory strain of the malaria vector *Anopheles gambiae*. *Science* 234, 607-610.
- Flink, I.L., Bailey, T.J., Gustafson, T.A., Markham, B.E., Morkin, E., 1986. Complete amino acid sequence of human thyroxine-binding globulin deduced from cloned DNA: close homology to the serine antiproteases. *Proc Natl Acad Sci U S A* 83, 7708-7712.
- Fortenberry, Y.M., 2013. Plasminogen activator inhibitor-1 inhibitors: a patent review (2006 - present). *Expert Opin Ther Pat*.
- Gettins, P.G., 2002. Serpin structure, mechanism, and function. *Chem Rev* 102, 4751-4804.
- Gulley, M.M., Zhang, X., Michel, K., 2013. The roles of serpins in mosquito immunology and physiology. *J Insect Physiol* 59, 138-147.
- Horn, T., Boutros, M., 2010. E-RNAi: a web application for the multi-species design of RNAi reagents--2010 update. *Nucleic Acids Res* 38, W332-339.
- Hunt, L.T., Dayhoff, M.O., 1980. A surprising new protein superfamily containing ovalbumin, antithrombin-III, and alpha 1-proteinase inhibitor. *Biochem Biophys Res Commun* 95, 864-871.
- Huntington, J.A., Read, R.J., Carrell, R.W., 2000. Structure of a serpin-protease complex shows inhibition by deformation. *Nature* 407, 923-926.
- Lawrence, D.A., Ginsburg, D., Day, D.E., Berkenpas, M.B., Verhamme, I.M., Kvassman, J.O., Shore, J.D., 1995. Serpin-protease complexes are trapped as stable acyl-enzyme intermediates. *J Biol Chem* 270, 25309-25312.
- Marinotti, O., Nguyen, Q.K., Calvo, E., James, A.A., Ribeiro, J.M., 2005. Microarray analysis of genes showing variable expression following a blood meal in *Anopheles gambiae*. *Insect Mol Biol* 14, 365-373.

- Muta, T., Hashimoto, R., Miyata, T., Nishimura, H., Toh, Y., Iwanaga, S., 1990. Proclotting enzyme from horseshoe crab hemocytes. cDNA cloning, disulfide locations, and subcellular localization. *J Biol Chem* 265, 22426-22433.
- Nappi, A., Poirie, M., Carton, Y., 2009. The role of melanization and cytotoxic by-products in the cellular immune responses of *Drosophila* against parasitic wasps. *Adv Parasitol* 70, 99-121.
- Paskewitz, S.M., Andreev, O., Shi, L., 2006. Gene silencing of serine proteases affects melanization of Sephadex beads in *Anopheles gambiae*. *Insect Biochem Mol Biol* 36, 701-711.
- Patston, P.A., Gettins, P., Beechem, J., Schapira, M., 1991. Mechanism of serpin action: evidence that C1 inhibitor functions as a suicide substrate. *Biochemistry* 30, 8876-8882.
- Volz, J., Osta, M.A., Kafatos, F.C., Muller, H.M., 2005. The roles of two clip domain serine proteases in innate immune responses of the malaria vector *Anopheles gambiae*. *J Biol Chem* 280, 40161-40168.
- Volz, J., Muller, H.M., Zdanowicz, A., Kafatos, F.C., Osta, M.A., 2006. A genetic module regulates the melanization response of *Anopheles* to *Plasmodium*. *Cell Microbiol* 8, 1392-1405.
- Waterhouse, R.M., Kriventseva, E.V., Meister, S., Xi, Z., Alvarez, K.S., Bartholomay, L.C., Barillas-Mury, C., Bian, G., Blandin, S., Christensen, B.M., Dong, Y., Jiang, H., Kanost, M.R., Koutsos, A.C., Levashina, E.A., Li, J., Ligoxygakis, P., Maccallum, R.M., Mayhew, G.F., Mendes, A., Michel, K., Osta, M.A., Paskewitz, S., Shin, S.W., Vlachou, D., Wang, L., Wei, W., Zheng, L., Zou, Z., Severson, D.W., Raikhel, A.S., Kafatos, F.C., Dimopoulos, G., Zdobnov, E.M., Christophides, G.K., 2007. Evolutionary dynamics of immune-related genes and pathways in disease-vector mosquitoes. *Science* 316, 1738-1743.
- Zou, Z., Jiang, H., 2005. *Manduca sexta* serpin-6 regulates immune serine proteinases PAP-3 and HP8. cDNA cloning, protein expression, inhibition kinetics, and function elucidation. *J Biol Chem* 280, 14341-14348.

### 3.7 Tables

**Table 3.1: SRPN9 commercial protease assays**

| Protease             | Amount Used         | BSA               | Amount SRPN9         |                      |
|----------------------|---------------------|-------------------|----------------------|----------------------|
|                      |                     |                   | 1:1                  | 10:1                 |
| Subtilisin Carlsberg | 2.5µl<br>(10ng/µl)  | 3µl<br>(1.0µg/µl) | 4.4µl<br>(10ng/µl)   | 4.4µl<br>(100ng/µl)  |
| Fungal Protease K    | 2.0µl<br>(10ng/µl)  | 3µl<br>(1.0µg/µl) | 6.6µl<br>(5ng/µl)    | 6.6µl<br>(50ng/µl)   |
| Porcine Elastase-1   | 2.0µl<br>(1µg/µl)   | 4µl<br>(1.0µg/µl) | 37.0µl<br>(100ng/µl) | 37.0µl<br>(1.0µg/µl) |
| Human Plasmin        | 4.0µl<br>(100ng/µl) | 3µl<br>(1.0µg/µl) | 25.5µl<br>(10ng/µl)  | 25.5µl<br>(100ng/µl) |
| Bovine Trypsin Gold  | 1.0µl<br>(10ng/µl)  | 3µl<br>(1.0µg/µl) | 4.0µl<br>(5ng/µl)    | 4.0µl<br>(50ng/µl)   |
| Porcine Kallikrein   | 4.0µl<br>(100ng/µl) | 3µl<br>(1.0µg/µl) | 5.7µl<br>(100ng/µl)  | 5.7µl<br>(1µg/µl)    |
| Human Thrombin       | 1.5µl<br>(10ng/µl)  | 3µl<br>(1.0µg/µl) | 4.0µl<br>(5ng/µl)    | 4.0µl<br>(50ng/µl)   |
| Bovine Chymotrypsin  | 4.0µl<br>(10ng/µl)  | 4µl<br>(1.0µg/µl) | 7.5µl<br>(10ng/µl)   | 7.5µl<br>(100ng/µl)  |

**Table 3.2: Inhibition of protease activity by recombinant SRPN9**

| Protease             | Substrate Specificity | Substrate | Inhibition [%] |              |
|----------------------|-----------------------|-----------|----------------|--------------|
|                      |                       |           | 1:1            | 10:1         |
| Subtilisin Carlsberg | Neutral/Acidic        | FVR-pNA   | 3.6 ± 3.1      | 5.6 ± 5.8    |
| Fungal Protease K    | Aromatic              | AAPF-pNA  | 0.5 ± 0.3      | 0.0 ± 0.0    |
| Porcine Elastase-1   | Neutral/Small         | AAPF-pNA  | 27.7 ± 33.7    | 36.4 ± 35.4  |
| Human Plasmin        | Lys/Arg               | PFR-pNA   | *75.6 ± 20.7   | *99.6 ± 0.5  |
| Bovine Trypsin Gold  | Arg/Lys               | IEAR-pNA  | *38.6 ± 12.2   | *96.2 ± 0.7  |
| Porcine Kallikrein   | Lys/Arg               | IEGR-pNA  | *96.8 ± 4.9    | *98.8 ± 1.0  |
| Human Thrombin       | Arg,Lys               | FVR-pNA   | 9.8 ± 5.2      | **81.9 ± 2.0 |
| Bovine Chymotrypsin  | Aromatic              | AAPF-pNA  | *49.9 ± 18.2   | *96.0 ± 0.3  |

One-way ANOVA, post test Newman-Keuls as compared to protease and substrate alone. \*  $p < 0.001$ , \*\*  $p < 0.01$

**Table 3.3: MS/MS Confirms SRPN9 forms complex with CLIPB9**

| Protein  | Accession #                | MS Peptides Matched | Peptide Sequences  | MASCOT Score (MS/MS) |                        |
|--|----------------------------|---------------------|--|----------------------|------------------------|
| SRPN9  | AGAP003139<br>gi 116041607 | 50                  | K. GESIFTLR. L   | 31                   | Total<br>Score<br>1942 |
|  |                            |                     | K. LAVANAAYFK. G   | 68                   |                        |
|  |                            |                     | K. DLLEPGAITR. N   | 42                   |                        |
|  |                            |                     | K. TVGMKPVLER. M   | 38                   |                        |
|  |                            |                     | R. TSLNLQWADSK. T  | 76                   |                        |
|  |                            |                     | K. TTVYEAYDTAR. K  | 45                   |                        |
|  |                            |                     | R. LYKGESIFTLR. L  | 63                   |                        |
|  |                            |                     | R. LYKGESIFTLR. L  | (38)                 |                        |
|  |                            |                     | K. EIFYVSADQQK. F  | 33                   |                        |
|  |                            |                     | R. EPILFDEVLQK. S  | 46                   |                        |
|  |                            |                     | R. SVLFNGVYRQPE. -   | 50                   |                        |
|  |                            |                     | K. TTVYEAYDTARK. S   | 41                   |                        |
|  |                            |                     | R. FSESNVGFSSADK. L  | 85                   |                        |
|  |                            |                     | R. GQIKDLLEPGAITR. N                                       | 56                   |                        |
|  |                            |                     | R. GQIKDLLEPGAITR. N                                       | (47)                 |                        |
|  |                            |                     | R. GRFSESNVGFSSADK. L                                      | 47                   |                        |
|  |                            |                     | K. LAVANAAYFKGTWQTK. F                                     | 60                   |                        |
|  |                            |                     | K. AAETNKEIFYVSADQQK. F                                    | 85                   |                        |
|  |                            |                     | K. AAETNKEIFYVSADQQK. F                                    | (41)                 |                        |
|  |                            |                     | K. AAETNKEIFYVSADQQK. F + Deamidated (NQ)                  | (23)                 |                        |
|  |                            |                     | R. LAAETDILHEVVNEGISR. K                                   | (75)                 |                        |
|  |                            |                     | R. LAAETDILHEVVNEGISR. K                                   | 93                   |                        |
|  |                            |                     | R. LAAETDILHEVVNEGISR. K                                   | (85)                 |                        |
|  |                            |                     | R. LAAETDILHEVVNEGISR. K + Deamidated (NQ)                 | (41)                 |                        |
|  |                            |                     | R. LAAETDILHEVVNEGISR. K + Deamidated (NQ)                 | (59)                 |                        |
|  |                            |                     | K. LAVANAAYFKGTWQTKFK. A + Deamidated (NQ)                 | 60                   |                        |
|  |                            |                     | K. FVDMMHVEGTFSHAANEK. L + Deamidated (NQ)                 | (51)                 |                        |
|  |                            |                     | K. FVDMMHVEGTFSHAANEK. L + Deamidated (NQ)                 | (35)                 |                        |
|  |                            |                     | R. FSESNVGFSSADKLFQGR. Q + Deamidated (NQ)                 | 90                   |                        |
|  |                            |                     | R. FSESNVGFSSADKLFQGR. Q                                   | (85)                 |                        |
|  |                            |                     | K. FVDMMHVEGTFSHAANEK. L + Deamidated (NQ); Oxidation (M)  | (46)                 |                        |
|  |                            |                     | K. FVDMMHVEGTFSHAANEK. L + Deamidated (NQ); Oxidation (M)  | 63                   |                        |
|  |                            |                     | R. LAAETDILHEVVNEGISRK. V + Deamidated (NQ)                | 53                   |                        |
|  |                            |                     | R. MGLGQLFDQGANFADFQGR. E                                  | (108)                |                        |
|  |                            |                     | R. MGLGQLFDQGANFADFQGR. E + Deamidated (NQ)                | (85)                 |                        |
|  |                            |                     | R. MGLGQLFDQGANFADFQGR. E + Oxidation (M)                  | 128                  |                        |
|  |                            |                     | R. MGLGQLFDQGANFADFQGR. E + Deamidated (NQ); Oxidation (M) | (89)                 |                        |
|  |                            |                     | K. FADTIELLDYQTPDEQR. A                                    | 85                   |                        |
|  |                            |                     | K. FADTIELLDYQTPDEQR. A                                    | (42)                 |                        |
|  |                            |                     | K. FADTIELLDYQTPDEQR. A                                    | (56)                 |                        |
|  |                            |                     | K. FKAETNKEIFYVSADQQK. F                                   | 114                  |                        |
|  |                            |                     | K. FKAETNKEIFYVSADQQK. F + Deamidated (NQ)                 | (59)                 |                        |
|  |                            |                     | R. GRFSESNVGFSSADKLFQGR. Q + Deamidated (NQ)               | 57                   |                        |
|  |                            |                     | K. SKIEVNEEGSVAASATVAFSFR. S                               | 70                   |                        |
|  |                            |                     | K. SKIEVNEEGSVAASATVAFSFR. S + Deamidated (NQ)             | (25)                 |                        |
|  |                            |                     | R. WVENATRQIKDLLEPGAITR. N + Deamidated (NQ)               | 39                   |                        |
|  |                            |                     | R. TSLNLQWADSKTTVYEAYDTAR. K                               | 132                  |                        |
| R. TSLNLQWADSKTTVYEAYDTAR. K + Deamidated (NQ)     | (61)                       |                     |  |                      |                        |
| R. TSLNLQWADSKTTVYEAYDTARK. S                      | 57                         |                     |  |                      |                        |
| K. FADTIELLDYQTPDEQRAYINR. W + Deamidated (NQ)     | 89                         |                     |  |                      |                        |
| CLIPB9   | AGAP013442<br>gi 347969169 | 12                  | K. YSCGSLINR. R  | 68                   | Total<br>Score<br>525  |
|  |                            |                     | K. LGIKVYDHAR. C   | 43                   |                        |
|  |                            |                     | K. TGEVNFVTGFGR. T   | (60)                 |                        |
|  |                            |                     | K. TGEVNFVTGFGR. T   | 62                   |                        |
|  |                            |                     | K. VWYLEGIVSYGNR. C  | 110                  |                        |
|  |                            |                     | K. LQKVWYLEGIVSYGNR. C                                     | 111                  |                        |
|  |                            |                     | K. LQKVWYLEGIVSYGNR. C + Deamidated (NQ)                   | (33)                 |                        |
|  |                            |                     | R. YVLTAAHCVIGEVERKEGK. L                                  | 34                   |                        |
|  |                            |                     | R. LAQSIYEYTSFVQPVCLPLTDFR. A                              | 67                   |                        |
|  |                            |                     | R. LAQSIYEYTSFVQPVCLPLTDFR. A + Deamidated (NQ)            | (34)                 |                        |
| R. LAQSIYEYTSFVQPVCLPLTDFR. A + 2 Deamidated (NQ)  | (28)                       |                     |  |                      |                        |
| R. LAQSIYEYTSFVQPVCLPLTDFRASK. T + Deamidated (NQ) | 46                         |                     |  |                      |                        |

**Table 3.4: Primer sequences**

| Name                       | Sequence  |
|----------------------------|---|
| <u>rProtein production</u> |   |
| SRPN9-1F                   | 5'-TCATCACGGCCATGCTCGCCGGGTGATG-3'                      |
| SRPN9-2F                   | 5'-AT{ <b>CCATGG</b> }GCC <b>CATCATCATCAT</b> CACGGC-3' |
| SRPN9-1R                   | 5'-TT[ <b>GAATTC</b> ]TTACTCCGGCTGGCGGTAGAC-3'          |
| SRPN9-2R                   | 5'-TT[ <b>GAATTC</b> ]TTACTCCGGCTGGCGGTAGAC-3'          |
| <u>qRT-PCR</u>             |   |
| SRPN9-qF                   | 5'-AACCGACATCCTGCACGA-3'                                |
| SRPN9-qR                   | 5'-ACCGTCTTCTCGATGCTGAA-3'                              |
| <u>dsRNA synthesis</u>     |   |
| SRPN9R-F                   | 5'- <b>TAATACGACTCACTATAGGG</b> GTGATGAAAAGAAGGCGGAA-3' |
| SRPN9R-R                   | 5'- <b>TAATACGACTCACTATAGGG</b> GCCACGTGCCCTTAAAGTAG-3' |
| T7                         | 5'- <b>TAATACGACTCACTATAGGG</b> -3'                     |

**BOLD** = T7 sequence, **BOLD** = His tail, [**BOLD**] = ECOR1 site, {**BOLD**} = NCO1 site  
q = quantitative

### 3.8 Figure Legends

#### **Figure 3-1: Amino acid sequence alignment of reactive center loops from AgSRPN9 and orthologs**

Alignment of *An. gambiae* SRPN9 reactive center loop with orthologous genes in *M. sexta*, *D. melanogaster*, *Ae. aegypti*, *C. quinquefasciatus*, and *Apis mellifera*. Canonical Human alpha-1-antitrypsin is also aligned. Alignment was performed in "PFAAT" (Bioinformatics, 2007). Predicted scissile bond is indicated by red line. Consensus sequence is beneath conservation score in grey.

#### **Figure 3-2: Nucleic acid and deduced amino acid sequence of SRPN9**

SRPN9 nucleic acid sequence and deduced amino acid sequence performed in Vector NTI Advance (Invitrogen). 6-His tag is underlined in green. dsRNA primer sequences are underlined in red.

#### **Figure 3-3: Protein purification of recombinant SRPN9**

EZ-Run Protein Gel Staining solution stained 10% SDS gel after (A) Nickel (Ni-NTA) affinity chromatography and (B) Q-Sepharose chromatography. (A) From left to right: Pre-induction (PI), Hours post protein induction (hpi) 2, 3, 4. Soluble protein(S), Insoluble protein (IS), Marker (M), Wash 1-3, Elutions 1-3. Protein was eluted with 300mM NaCl + 250mM imidazole pH= 8.0 at 1 ml/min. (B) Protein was run through column containing Q-Sepharose beads washed, and eluted at 1ml/min in a 50-500mM pH= 8.0 linear NaCl gradient. Numbers are elution numbers.

#### **Figure 3-4: Survival for SRPN9 depleted mosquitoes**

Percent survival of mosquitoes injected with dsRNA targeting *SRPN9* (red), *GFP* (green), or not injected (black). Each graph represents an independent, individual biological replicate. Log rank test was performed to compare the curves. A.  $X^2=0.2462$ ,  $df=1$ ,  $p=0.6197$  B.  $X^2=0.9640$ ,  $df=1$ ,  $p=0.3262$

**Figure 3-5: SRPN9 expression after blood meal time course**

Relative expression of SRPN9 after blood meal time course. Expression was normalized to RPS7 nonblood fed control samples were used as calibrator. Statistical tests performed is 2 way ANOVA. Error bars are mean and SEM of three replicates. P =0.0005.

**Figure 3-6: Recombinant AgSRPN9 inhibits trypsin-like serine proteases**

SRPN9 protein was incubated with (A) chymotrypsin or (B) trypsin for ten minutes at room temperature at 1:1 or 10:1 SRPN:protease molar ratios. 200 µl of 50µM substrate was added and absorbance was measured immediately every 30 s for 41 times at 405nm. Bar graphs represent residual activity and dot graphs are representative assay of three replicates. Error bars are mean and SEM of three biological replicates.

**Figure 3-7: SRPN9 inhibition of PO activity in *M. sexta* hemolymph**

Different amounts of SRPN9 were incubated with *M. sexta* plasma at room temperature for ten minutes. PO was activated by the addition of *M. luteus*. Dopamine HCL was added as substrate and PO activity was measured every 30 sec, 41 times at 405nm. (A). Representative assay of three replicates. Error bars are mean and SEM. One-way ANOVA, p< 0.0001 Newman Kuels post test, (B). Combined protease activity of three biological replicates. Error bars are mean and SEM. All treatments are statistically different (p < 0.05)

**Figure 3-8: SRPN9 inhibits and forms complex with CLIPB9**

rSRPN9 was incubated with Factor Xa activated (A) rCLIPB8 or (B) rCLIPB9 at room temperature for Ten minutes. FVR-AMC was used as substrate and FVRase activity was measured immediately every 30 sec, 41 times. Error bars are mean and SEM of three biological. (C) Activation of recombinant CLIPB9 and complex formation via Western Blot analysis using anti-SRPN9 and anti-CLIPB9 antibodies. Asterisk indicates complex.

**Figure 3-9: High molecular weight band is complex of SRPN9 and CLIPB9**

rSRPN9 (2µg) and rCLIPB9<sub>Xa</sub> (1µg) were incubated for 10 minutes at room temperature. Proteins were then ran on 10% SDS-PAGE as described in Materials and Methods. Proteins

were stained with EZ-Run Protein Gel Staining solution. The highest molecular band (Asterisk indicates complex) was further analyzed by 1D NanoLC ESI MS/MS (Table 3.3).

**Figure 3-10: Stoichiometry of inhibition of recombinant SRPN9 to CLIPB9**

rSRPN9 was incubated with Factor Xa activated rCLIPB9 at different molar ratios. FVR-AMC was used as substrate and FVRase activity was measured every 30 s for 61 times. Linear regression of the residual FVRase activity was plotted. Error bars are mean and SEM of three technical replicates.

**Figure 3-11:  $K_a$  of recombinant SRPN9 to CLIPB9**

(A) 300 $\mu$ M FVR-AMC was added varying concentrations of rSRPN9 (0, 4.3, 8.6, 13, 17.3, 25.9, 34.5 and 51.8 nM) plus 213nM BSA. rCLIPB9<sub>Xa</sub> (2.37pmol) was added. FVRase activity was measured every 30s, 50 times. B. Linear regression of A.



**Figure 3-2: Nucleic acid and deduced amino acid sequence of SRPN9**

H H H H H H G Q C S P G D E K K A E K A T D L E Q S R S R L Y K G  
 5077 CATCATCATC ATCATCACGG CCAATGCTCG CCGGGTGATG AAAAGAAGGC GGAAAAGGCG ACCGATCTCG AGCAGTCGCG CAGCCGCCTG TACAAGGGCG  
  
 E S I F T L R L L D A I N T A T P N E N L F F S P Y S L Y N V L L M  
 5177 AATCGATCTT TACCTGCGG CTGCTGGACG CGATCAACAC GGCCACGCGG AACGAAAACC TCTTCTTCTC ACCCTACAGC CTGTACAATG TGCTGCTGAT  
  
 M Y F G A R D T T E K L L R T S L N L Q W A D S K T T V Y E A Y D  
 5277 GATGTACTTC GGGGCGGCG ACACGACCGA GAAGTGCTC CGCACCGCC TCAACCTGCA GTGGGCGGAC AGCAAGACGA CCGTGTACGA GCGGTACGAT  
  
 T A R K S L R G R F S E S N A V G F S S A D K L F F G R Q I P V S  
 5377 ACCGCCCGGA AGTCGCTGCG GGGTCGGTTC AGCGAAAGCA ATGCCGTCGG CTCAGCTCG GCCGACAAGC TGTTCCTCGG CCGGCAGATT CCCGTGTCGA  
  
 T C V Q Q K F A D T I E L L D Y Q T Q P D E Q R A Y I N R W V E N A  
 5477 CCTGCCTGCA GCAAAAGTTT GCCGACACGA TCGAGCTGCT CGACTACCAG ACGCAGCCGG ACGAGCAGCG GCGGTACATT AACCGGTGGG TGGAGAATGC  
  
 T R G Q I K D L L E P G A I T R N T K L A V A N A A Y F K G T W Q  
 5577 GACCCGCGG CAGATTAAGG ATCTGCTCGA GCCGGGAGCG ATCACGCGCA ACACCAAGCT GCGGTCGCG AATCGCGCCT ACTTTAAGGG CACGTGGCAG  
  
 T K F K A A E T N K E I F Y V S A D Q Q K F V D M M H V E G T F S  
 5677 ACCAAGTTTA AGGCGGCCGA AACGAACAAG GAAATCTTTT ACGTGTCCG CGACCAGCAG AAGTTCGTCG ACATGATGCA CGTCGAGGGA ACGTTCAGCC  
  
 H A A N E K L G C H I L E L P Y S A G P G A D N A D D G P Y Q Q G A  
 5777 ATGCTGCCAA CGAAAAGCTC GGCTGCCACA TTCTCGAGCT CCCGTACAGT GCTGGCCCG GCGCAGACAA TGCCGACGAT GGGCCGTACC AGCAGGGCGC  
  
 A N P D N Q V S M F V F L P P A E P N A L S K L L S R L A A E T D  
 5877 CGCCAATCCG GACAATCAGG TGTGATGTT TGTCTTCTG CGCGCGGCG AACCGAACGC ACTGTCCAAG CTGCTTAGCC GGCTGGCCCG CGAAACCGAC  
  
 I L H E V V N E G I S R K V D V K L P K F S I E K T V G M K P V L  
 5977 ATCCTGCACG AGGTGGTAAA CGAGGGCATT TCGCGCAAGG TCGAGTGAA GCTGCCAAG TTCAGCATCG AGAAGACGGT CCGCATGAAG CCGGTGCTGG  
  
 E R M G L G Q L F D Q G A N F D A F T D G R E P I L F D E V L Q K S  
 6077 AGCGGATGGG CCTGGGCCAG CTGTTGACC AGGGGGCCAA CTTTGACGCG TTCACGGACG GCCGCGAACC GATCCTGTTC GACGAGTGC TGCAAAAGTC  
  
 K I E V N E E G S V A A S A T V A F S F R S S R P A D P A M F H C  
 6177 CAAGATCGAG GTCAACGAGG AAGGTTGGT GGCTGCGTCC GCTACCGTAG CGTTCTCGTT CCGGTCTGCC CGCCCGGCCG ATCCGGCCAT GTTCCACTGC  
  
 N H P F V F L I Y D Y G T R S V L F N G V Y R Q P E  
 6277 AACCATCCGT TCGTTTTCTT CATCTACGAT TACGGCACCC GGTCGGTGCT GTTCAACGGT GTCTACCGCC AGCCGGAGTA AG

**Figure 3-3: Protein purification of recombinant SRPN9**

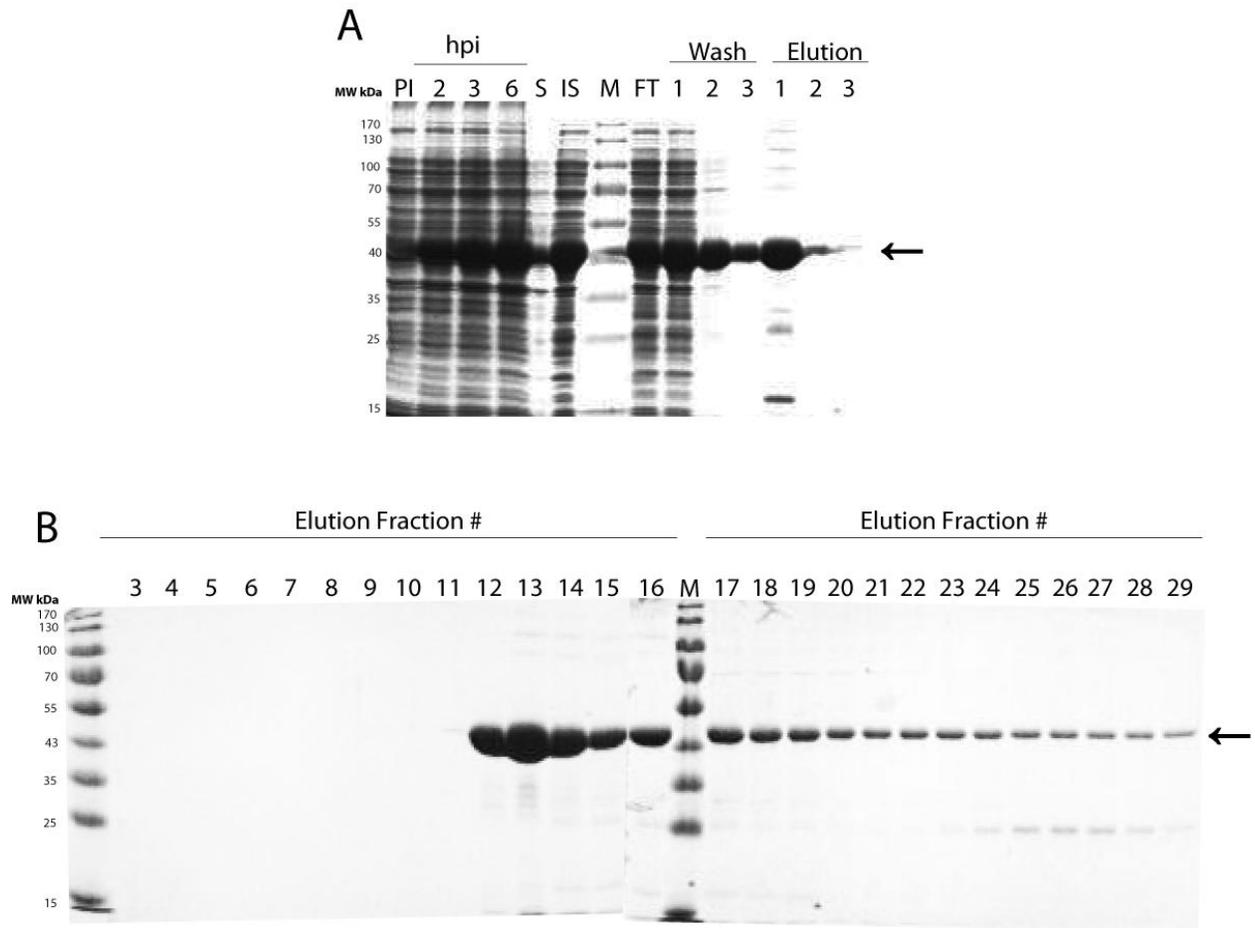


Figure 3-4: SRPN9 depletion does not affect survival

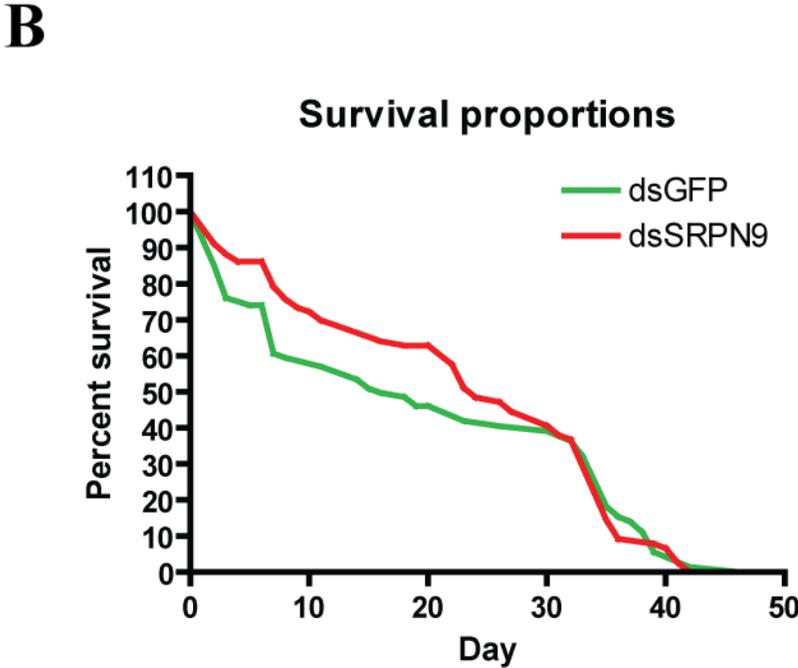
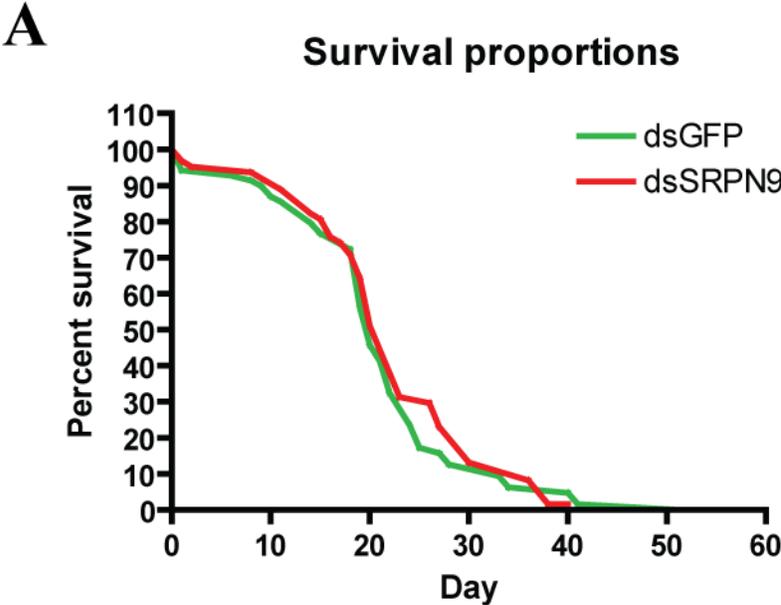


Figure 3-5: SRPN9 expression after blood meal time course

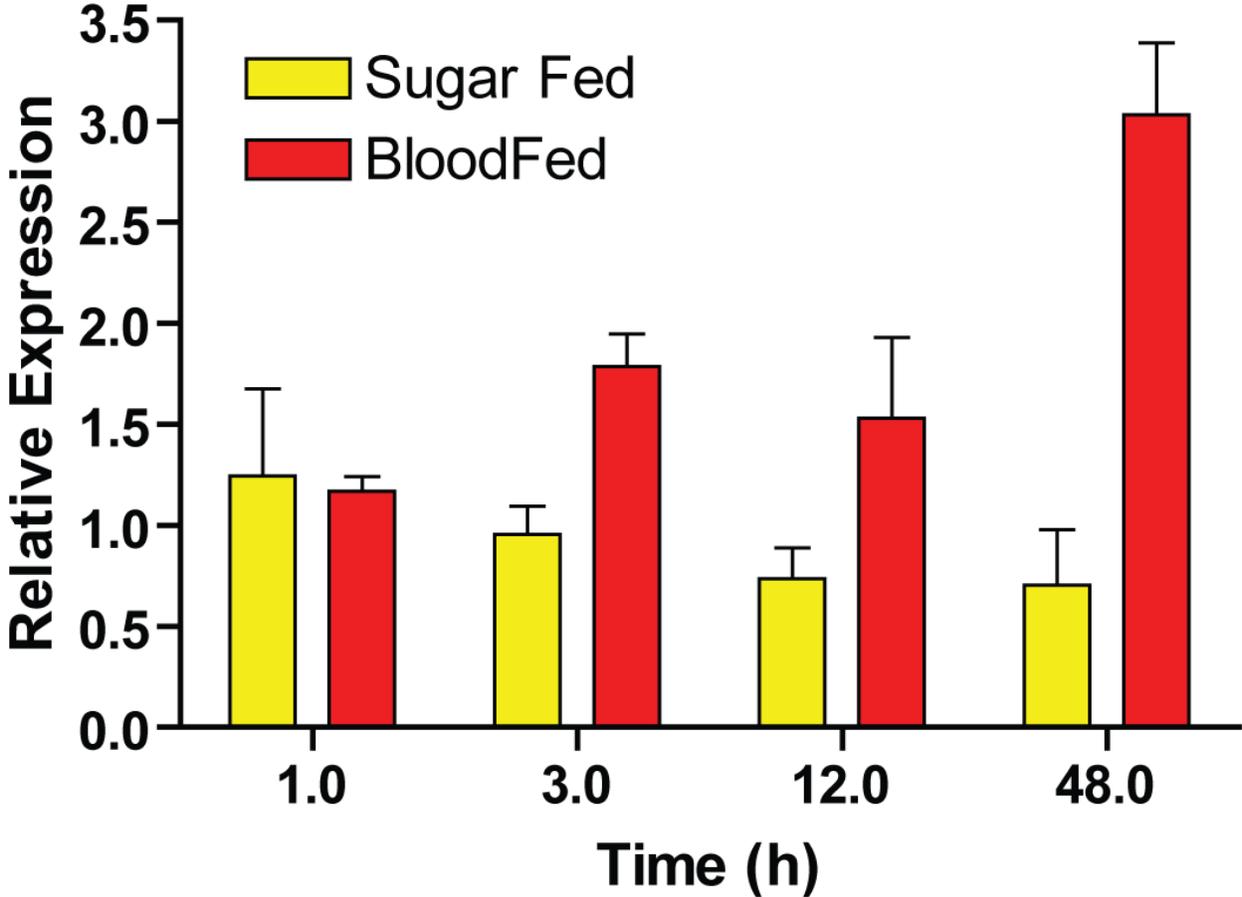


Figure 3-6: Recombinant AgSRPN9 inhibits trypsin-like serine proteases

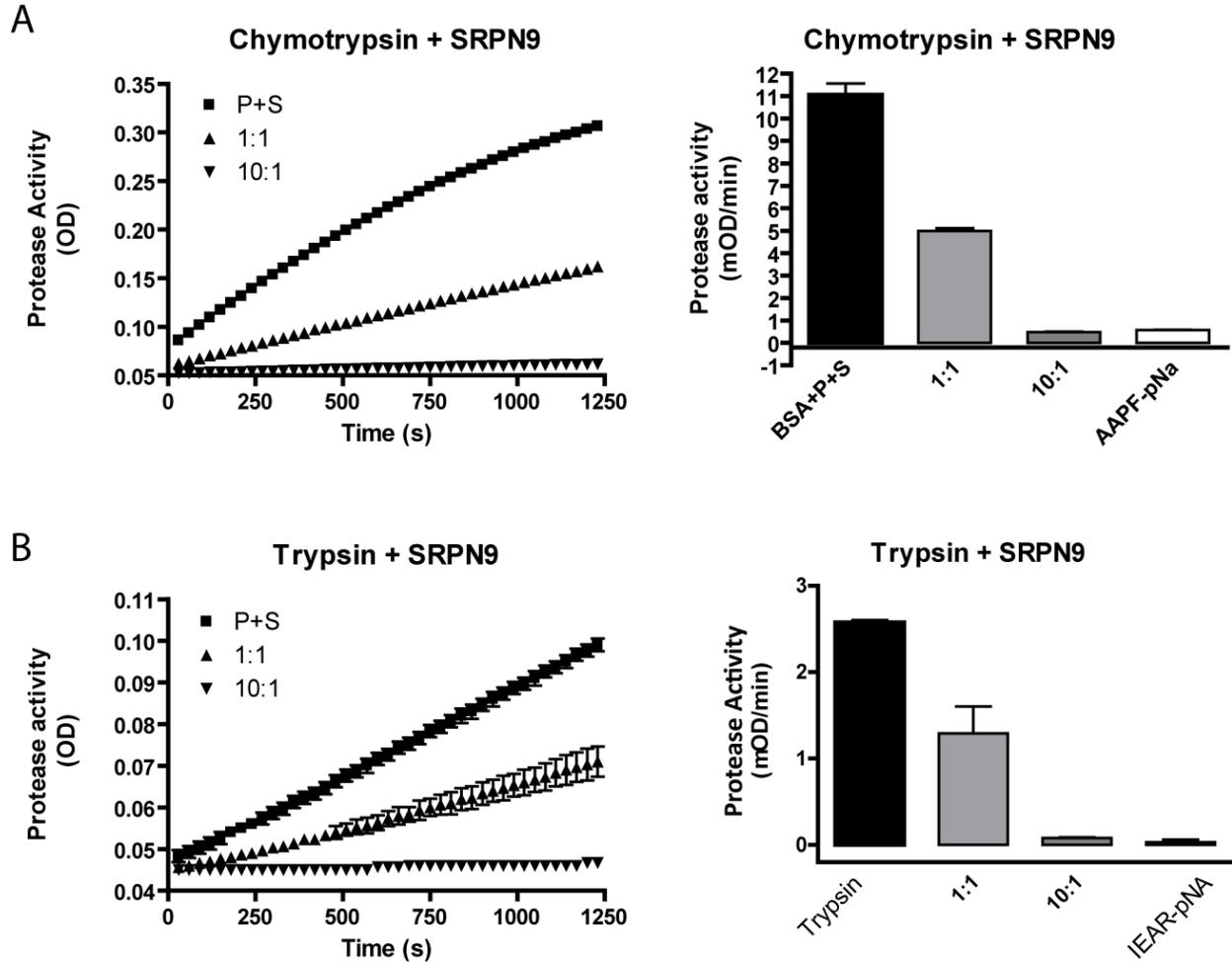


Figure 3-7: SRPN9 inhibition of PO activity in *M. sexta* hemolymph

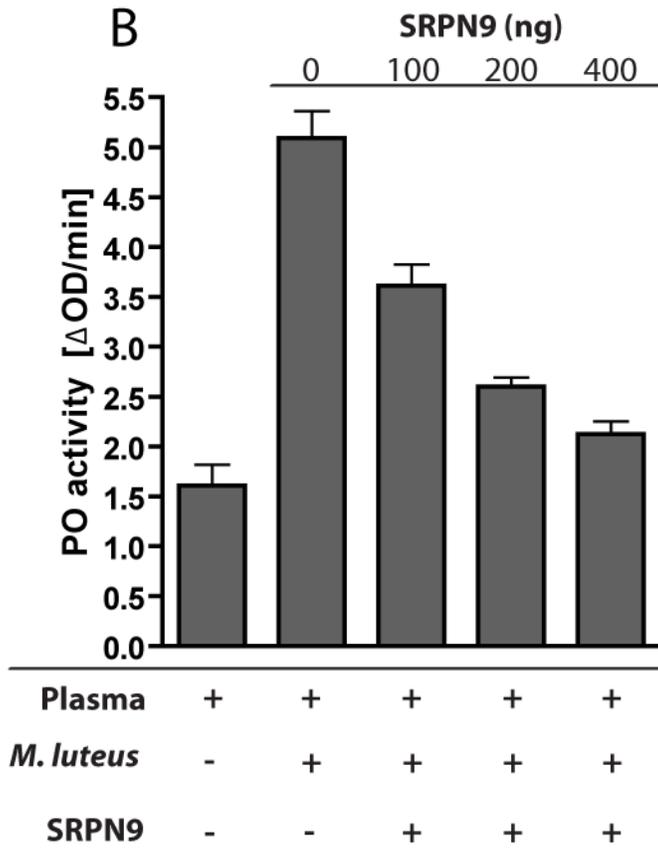
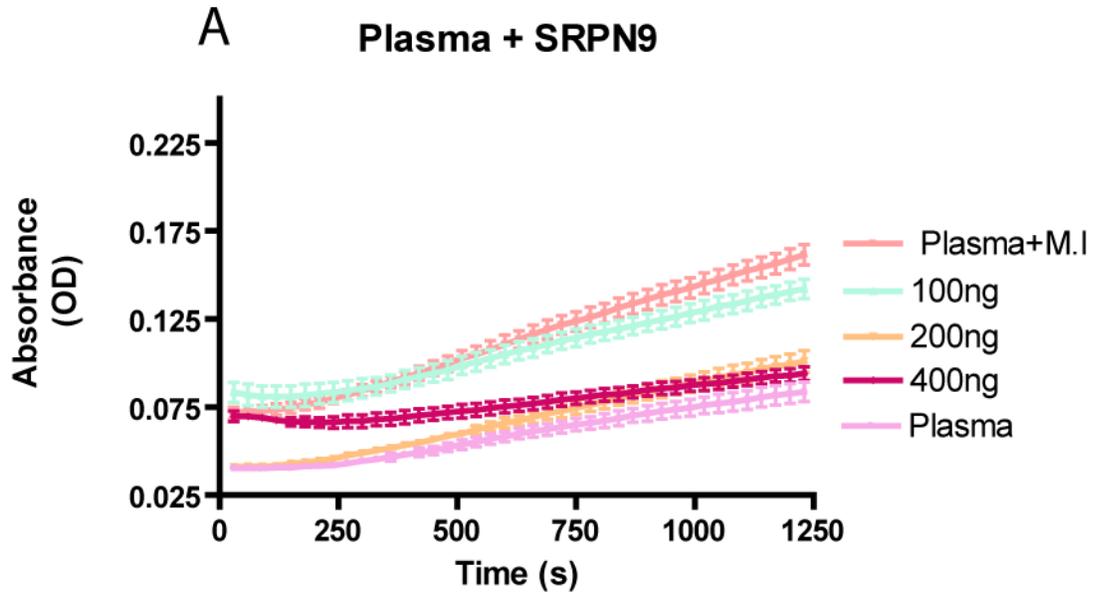
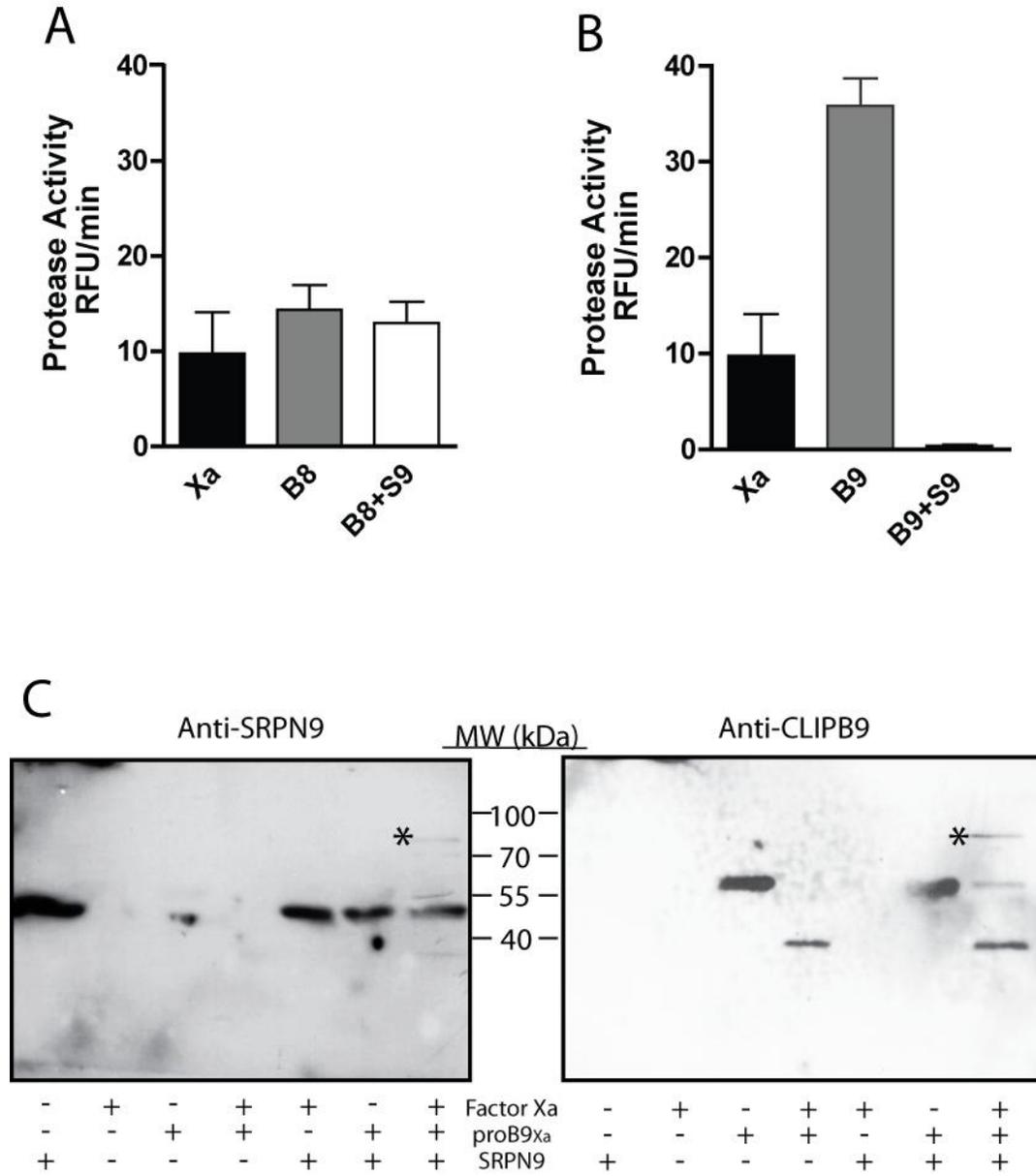


Figure 3-8: SRPN9 inhibits and forms complex with CLIPB9



**Figure 3-9: High molecular weight band is complex of SRPN9 and CLIPB9**

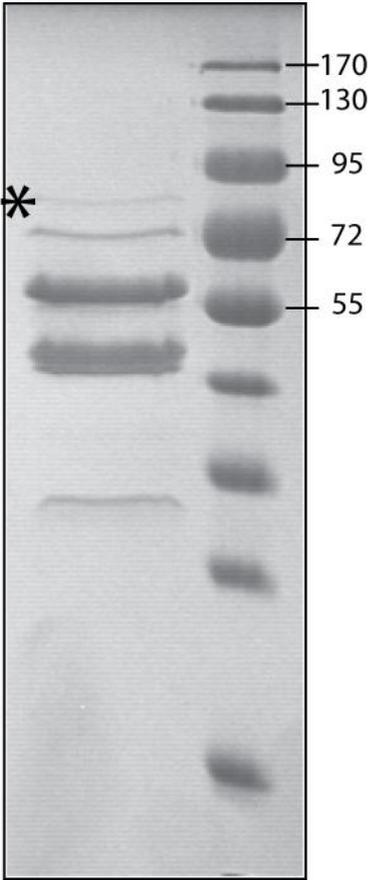


Figure 3-10: Stoichiometry of inhibition of recombinant SRPN9 to CLIPB9

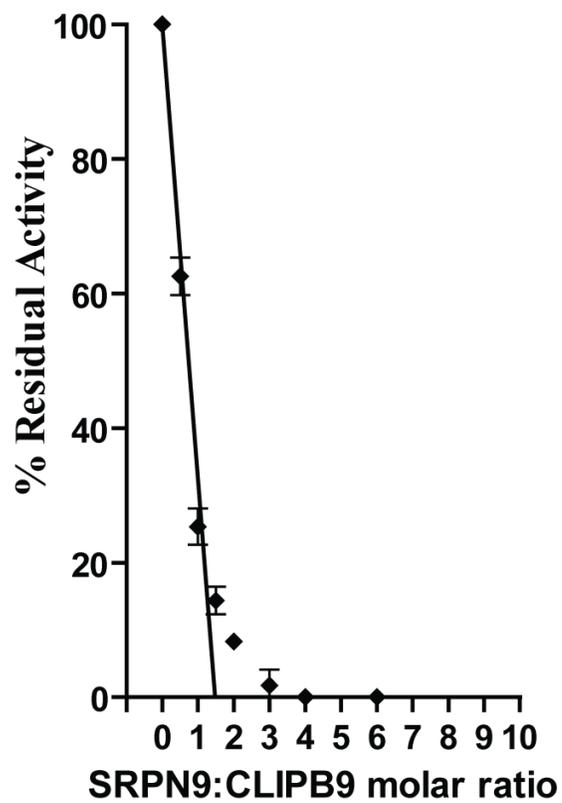
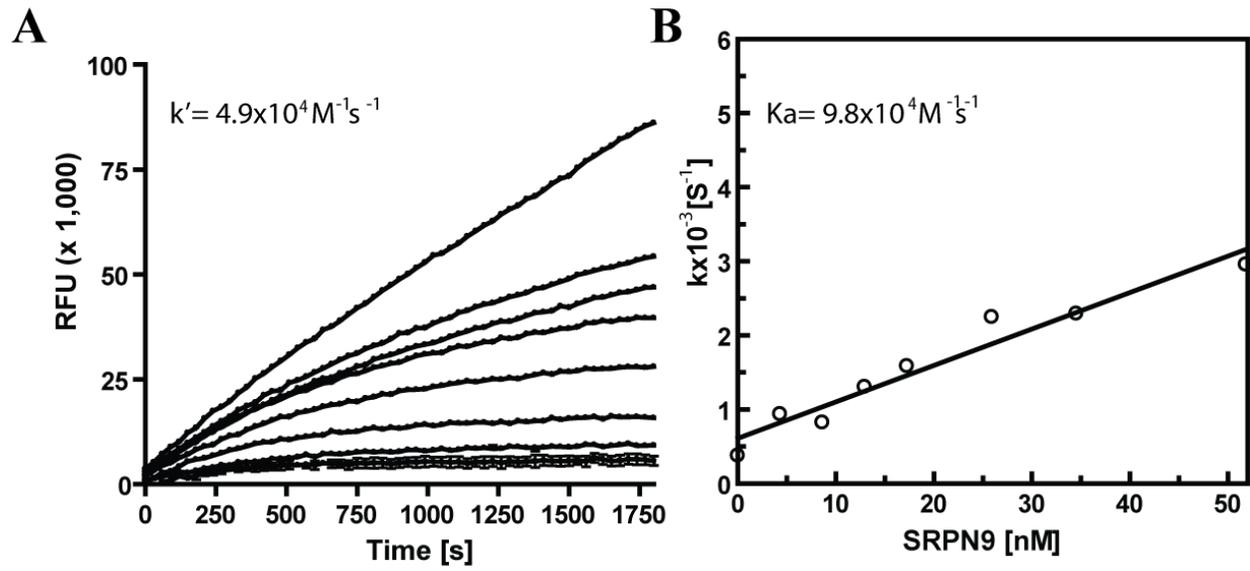


Figure 3-11:  $K_a$  of recombinant SRPN9 to CLIPB9



### **3.10 Acknowledgements**

I would like to thank Mrs. Tinea Canady for SRPN9 cloning and antibody production. I would like to thank KaraJo Sprigg, John Caldwell, Amanda Osarczuk, and Sarah Smith for mosquito rearing. I would like to thank Dr Chunju An and Dr. Xin Zhang for kindly providing CLIPB8 and B9 for experiments. Thanks to Dr. Kristin Michel for collecting *M. sexta* plasma. Thanks to Dr. Michael Kanost and his lab for allowing me the use of their equipment. I would also like to thank Dr. Yasuaki Hiromasa for performing the Mass spectrometry analysis.

# Appendix A

Individual replicates of commercial activity assays for each SRPN:protease combination.

**Figure A-1: SRPN9 inhibits Chymotrypsin**

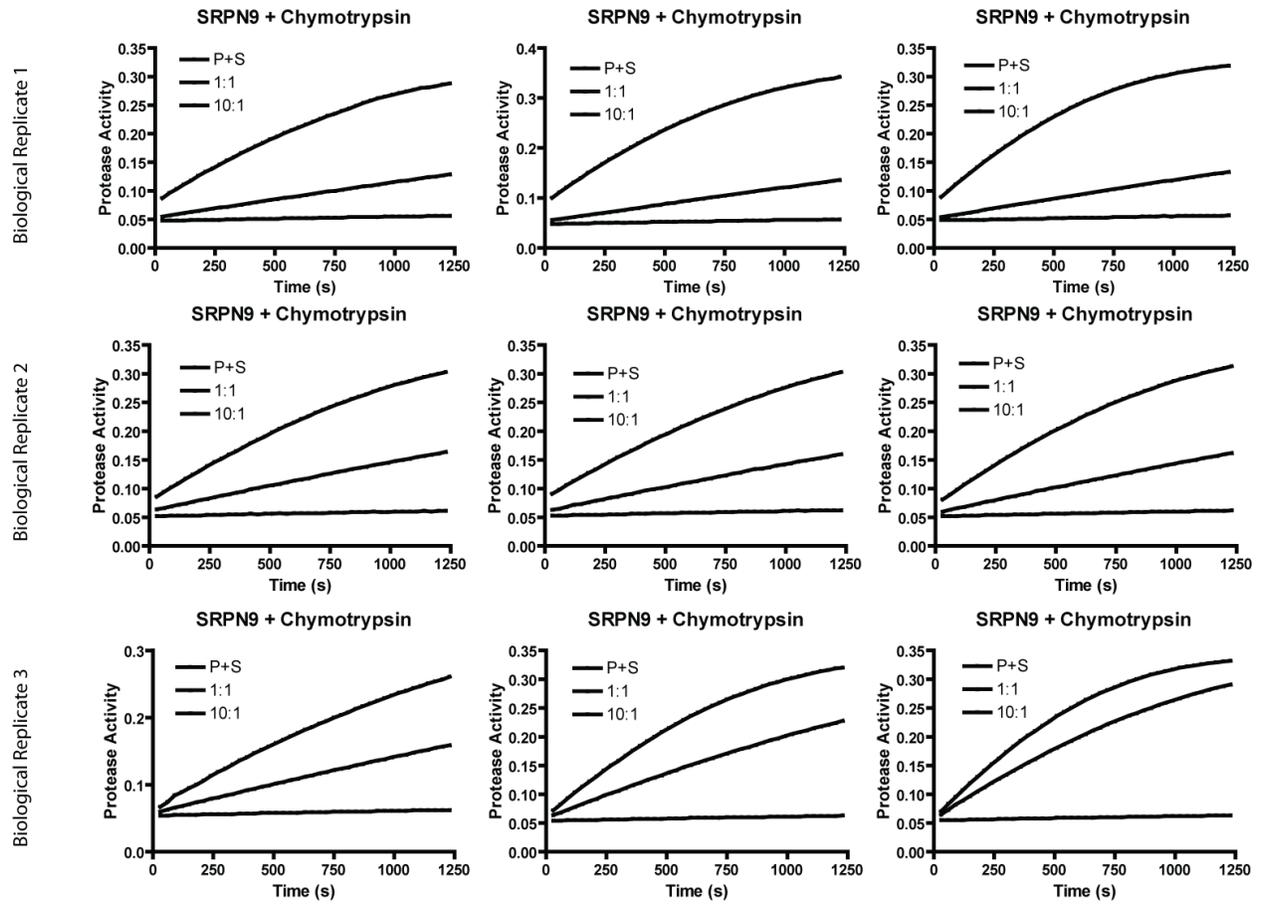
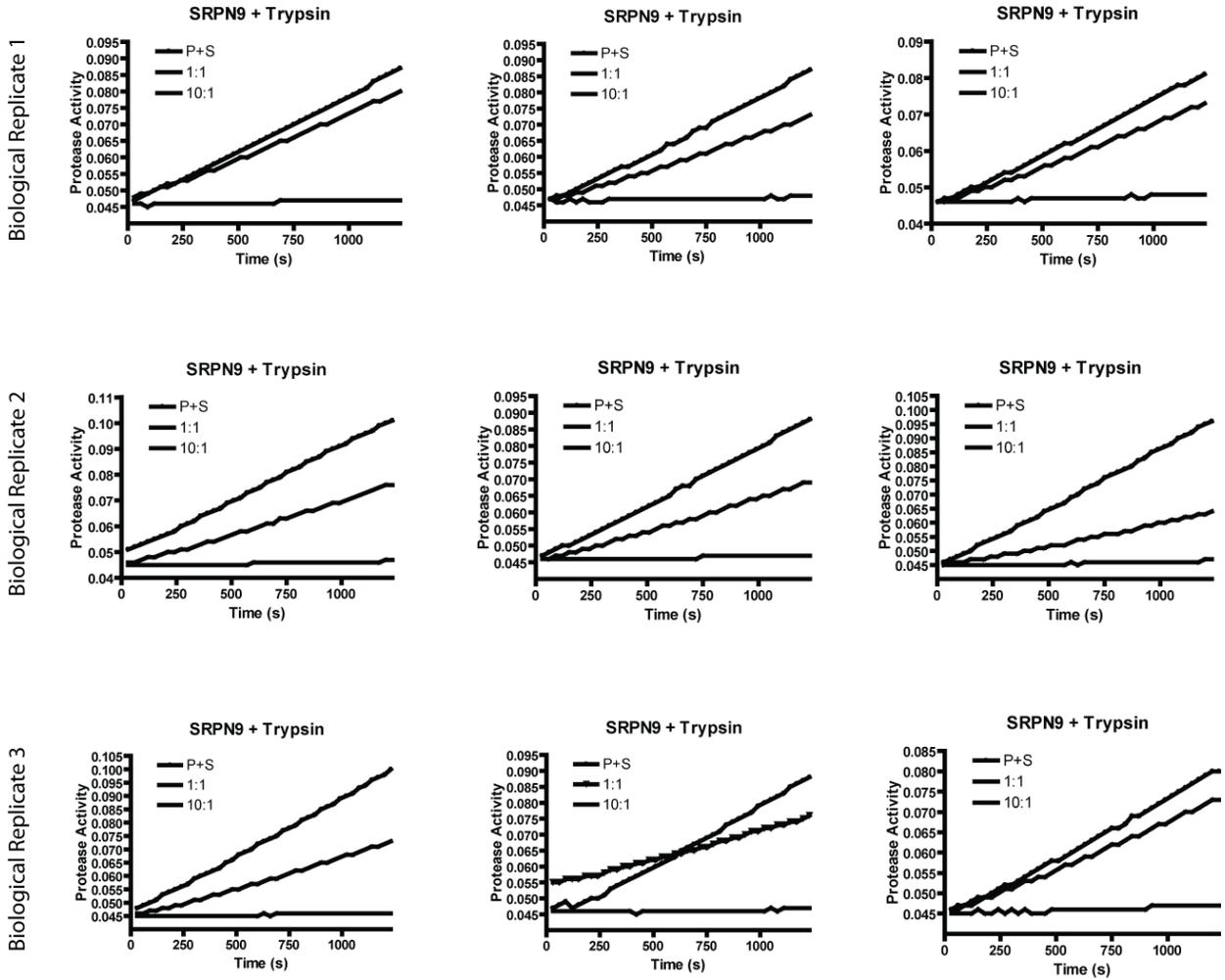
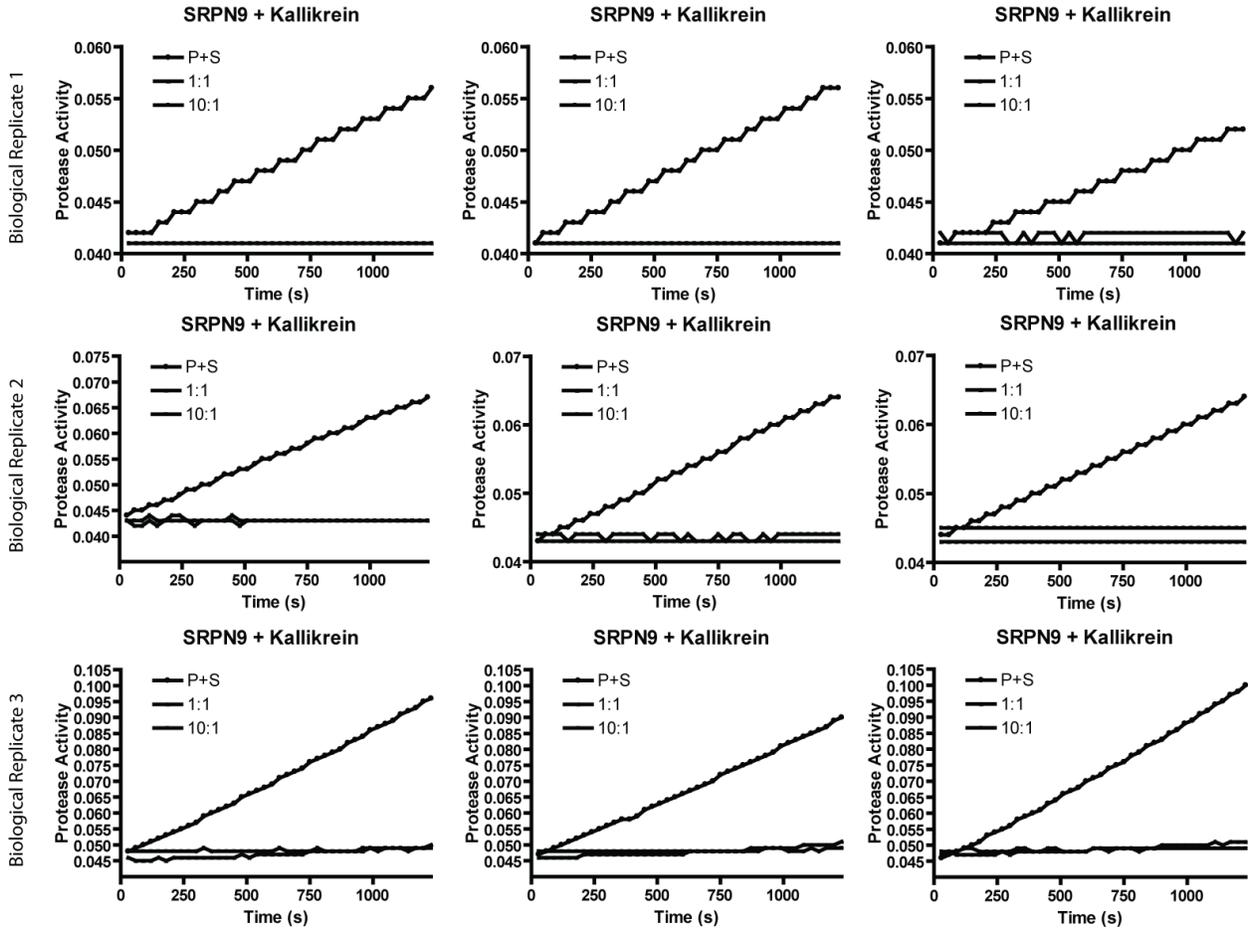


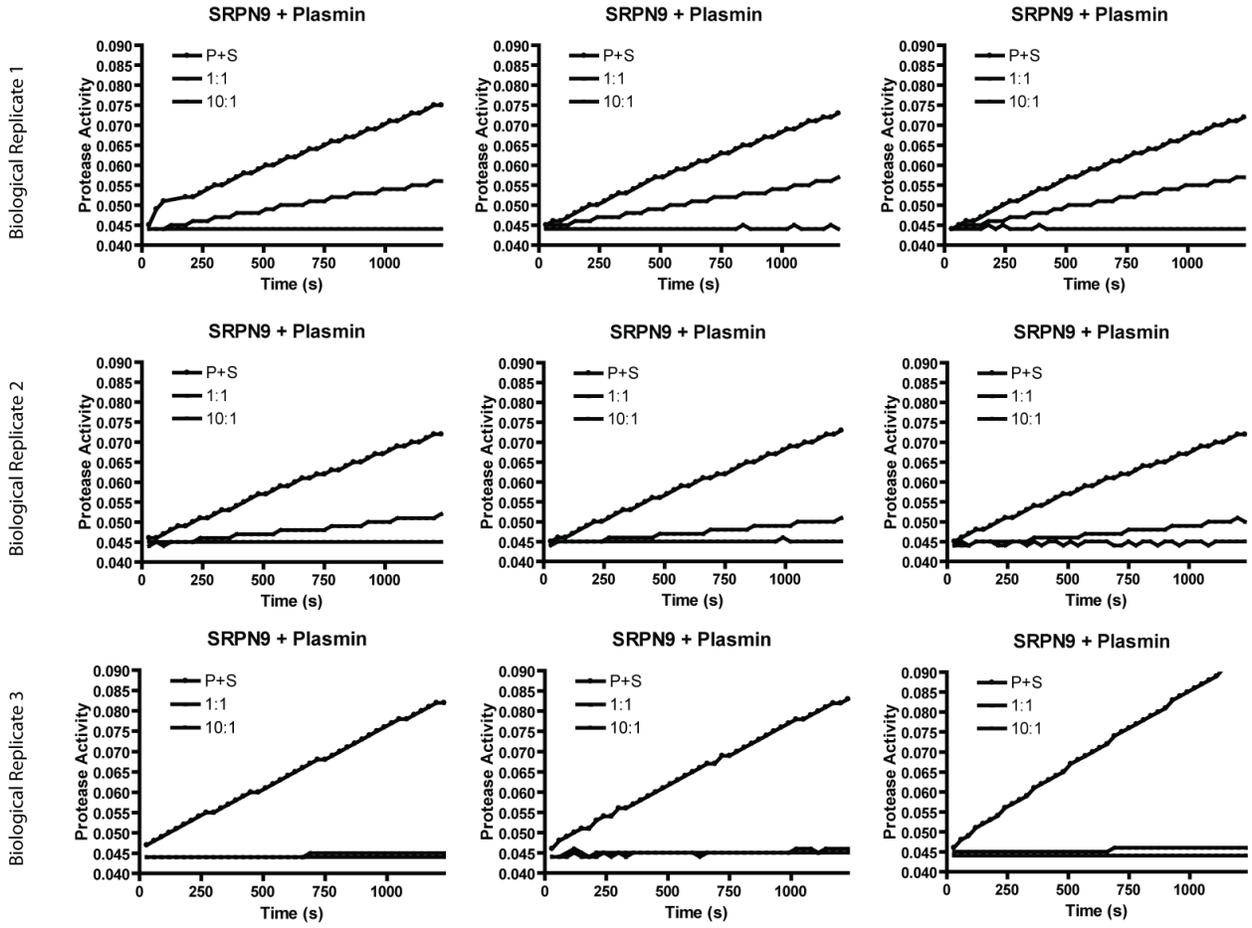
Figure A-2: SRPN9 inhibits Trypsin



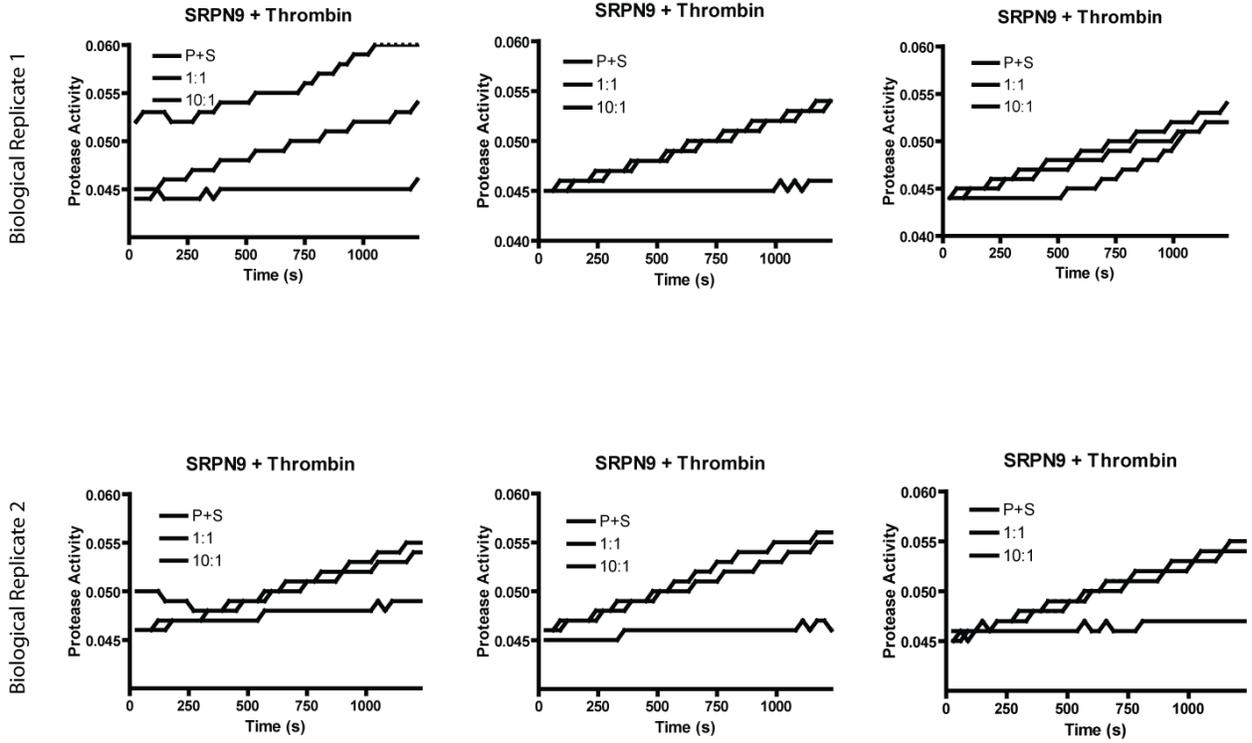
*Supplemental Figure A-3: SRPN9 inhibits Kallikrein*



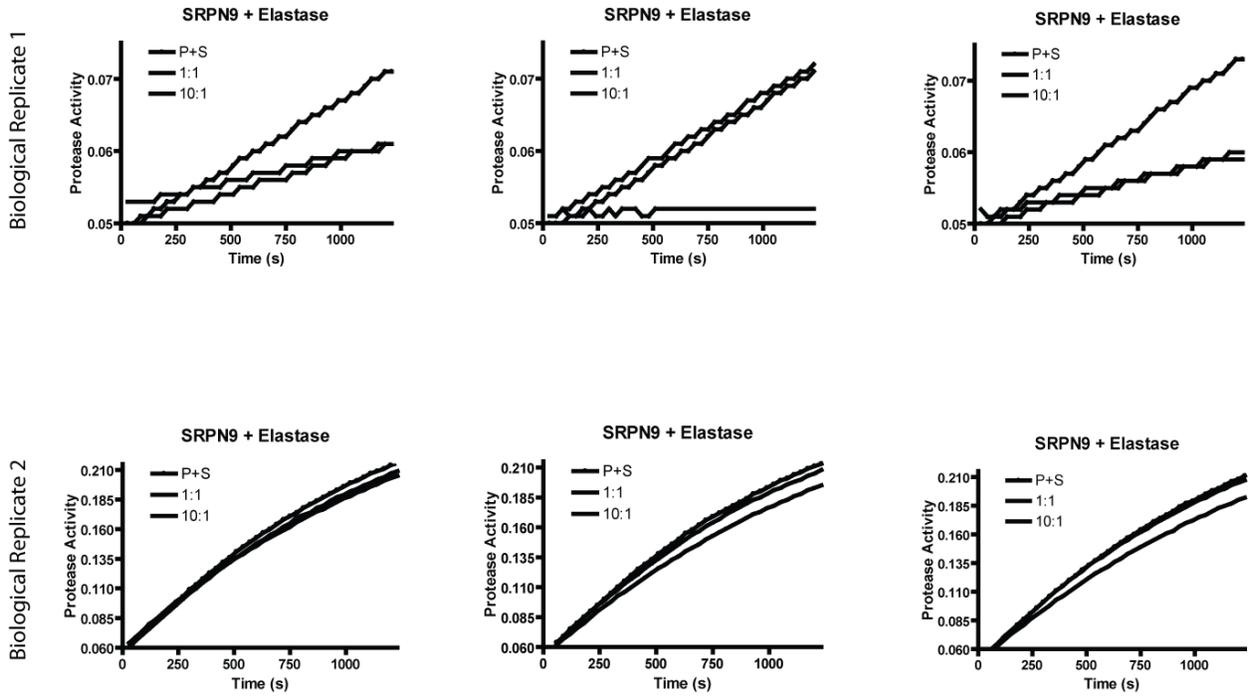
Supplemental Figure A-4: SRPN9 inhibits Plasmin



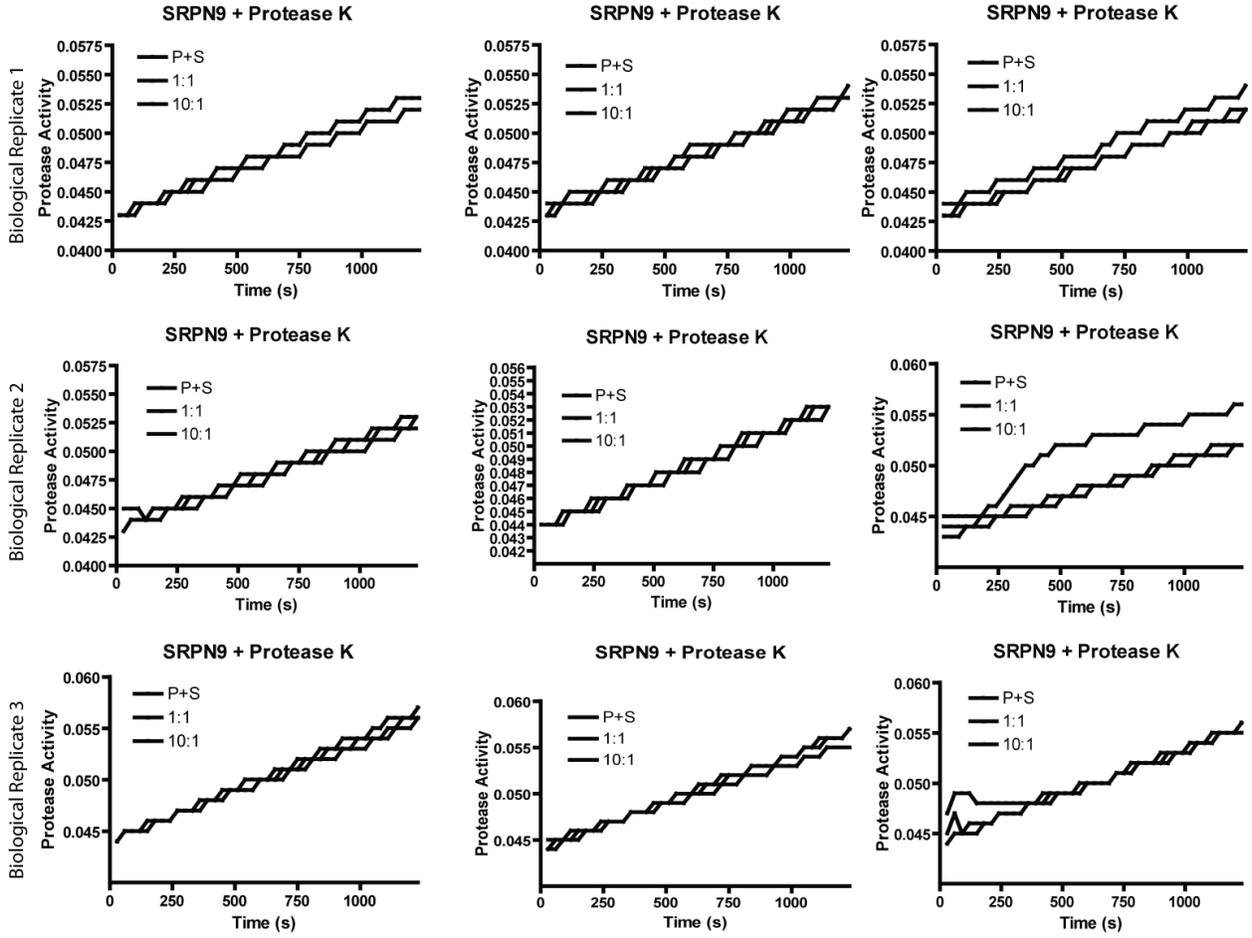
*Supplemental Figure A-5: SRPN9 inhibits Thrombin at a 10:1 ratio*



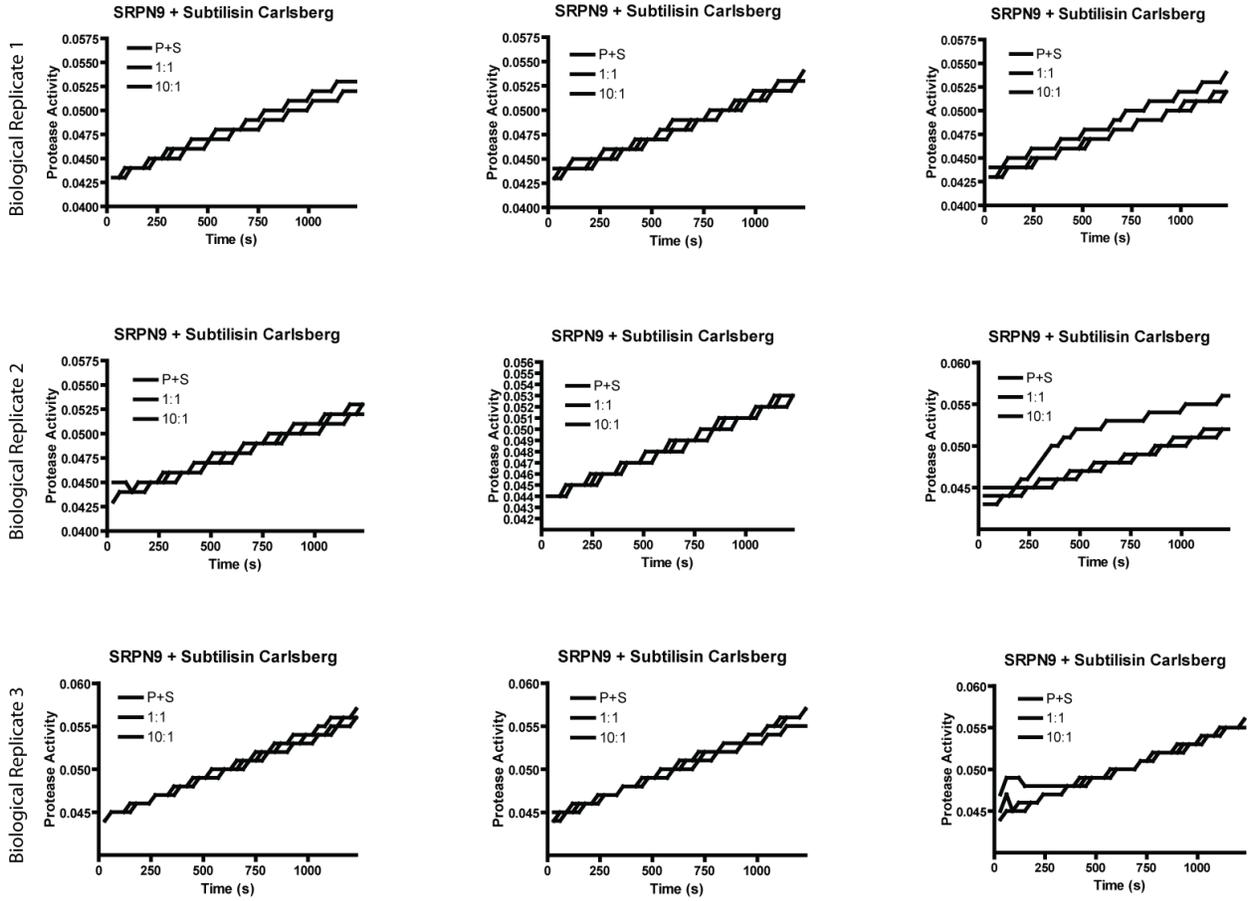
*Supplemental Figure A-6: rSRPN9 does not inhibit Elastase*



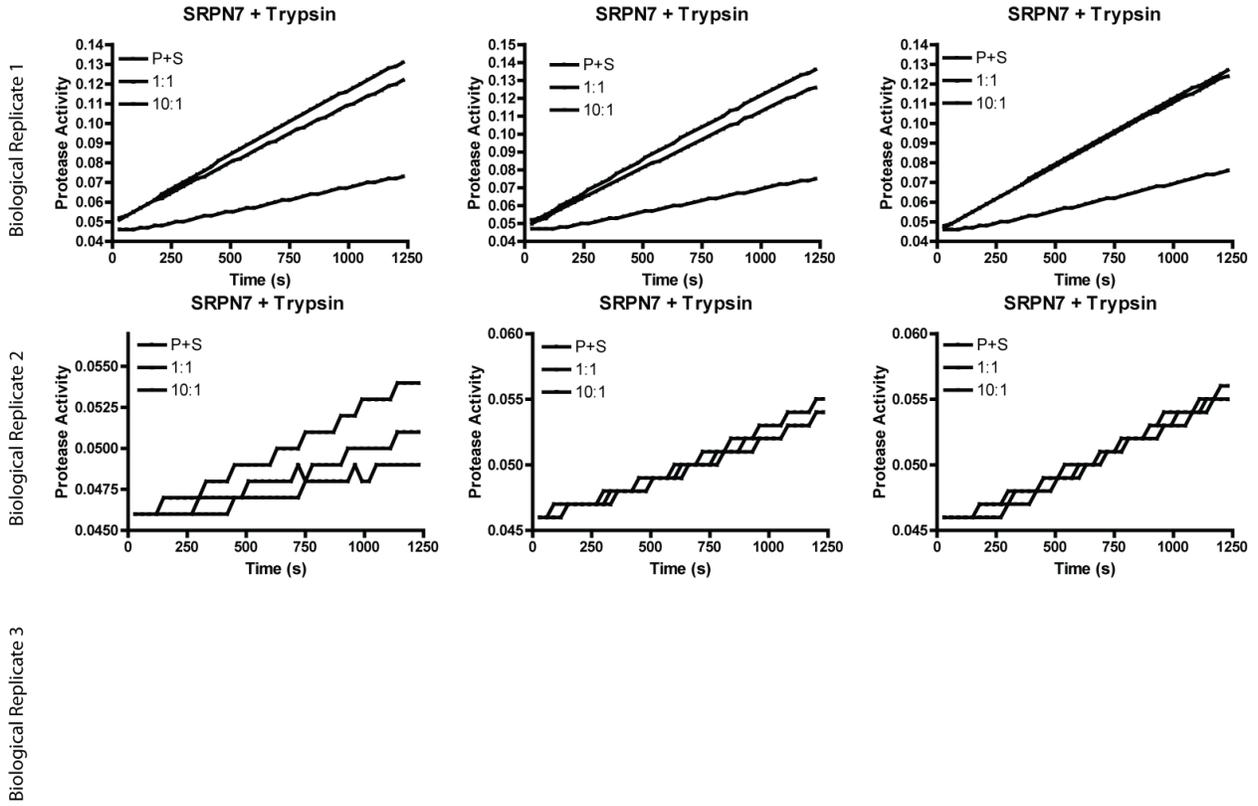
Supplemental Figure A-7: *rSRPN9* does not inhibit *Protease K*



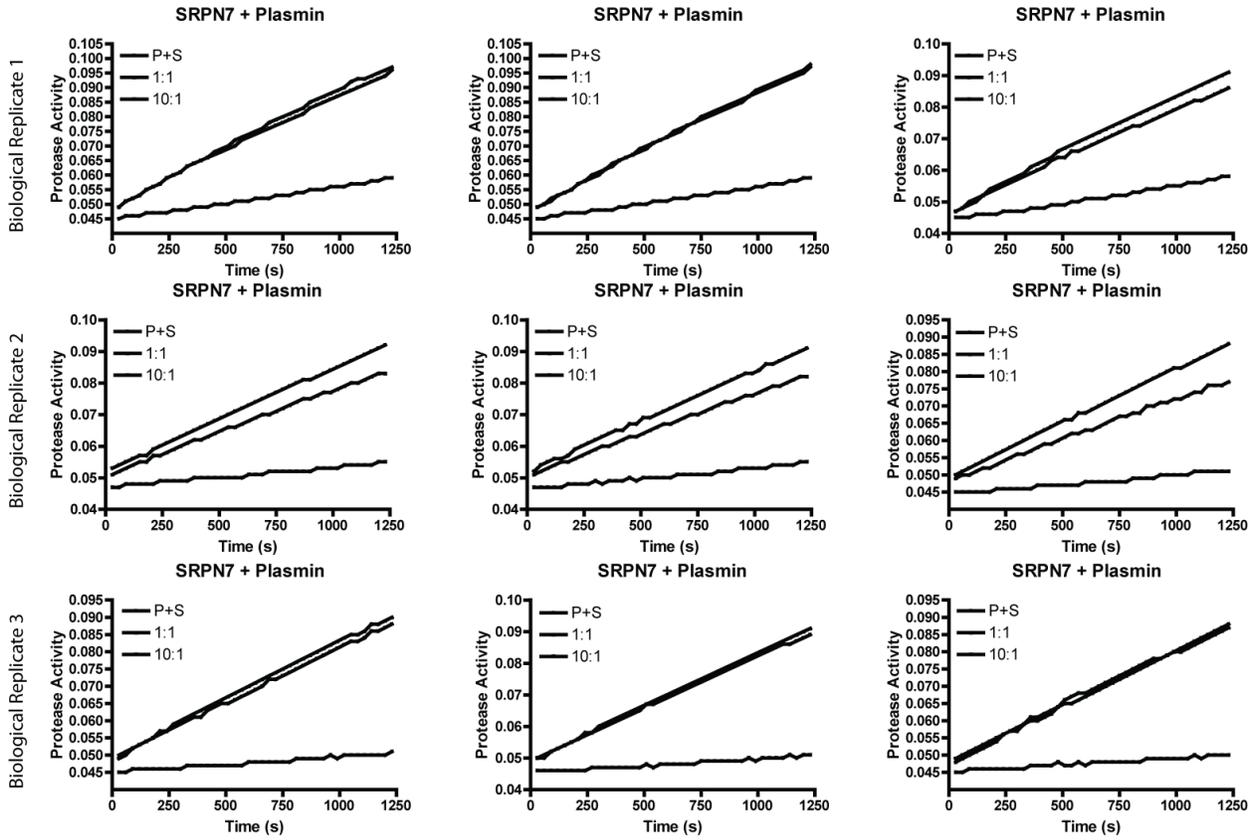
*Supplemental Figure A-8: rSRPN9 does not inhibit Subtilisin Carlsberg*



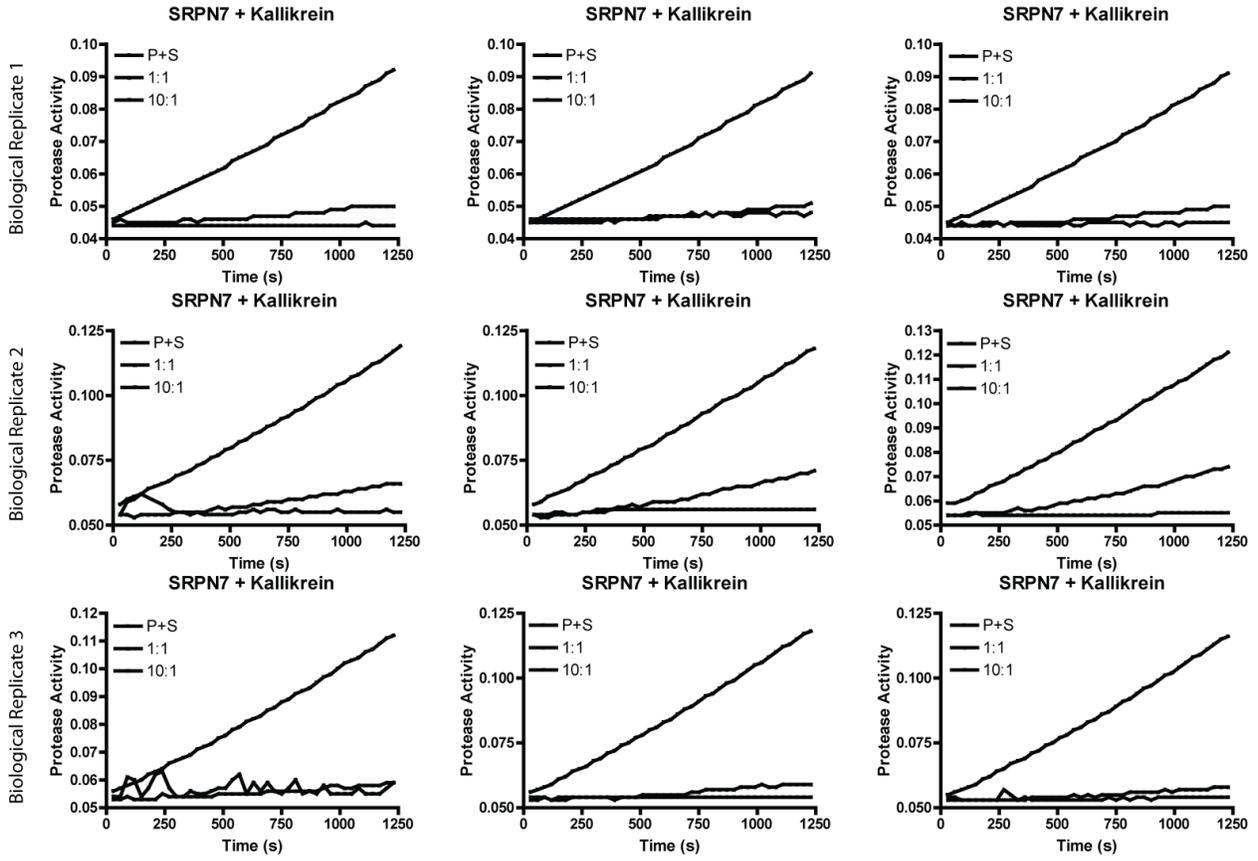
*Supplemental Figure B-1: rSRPN7 inhibits Trypsin*



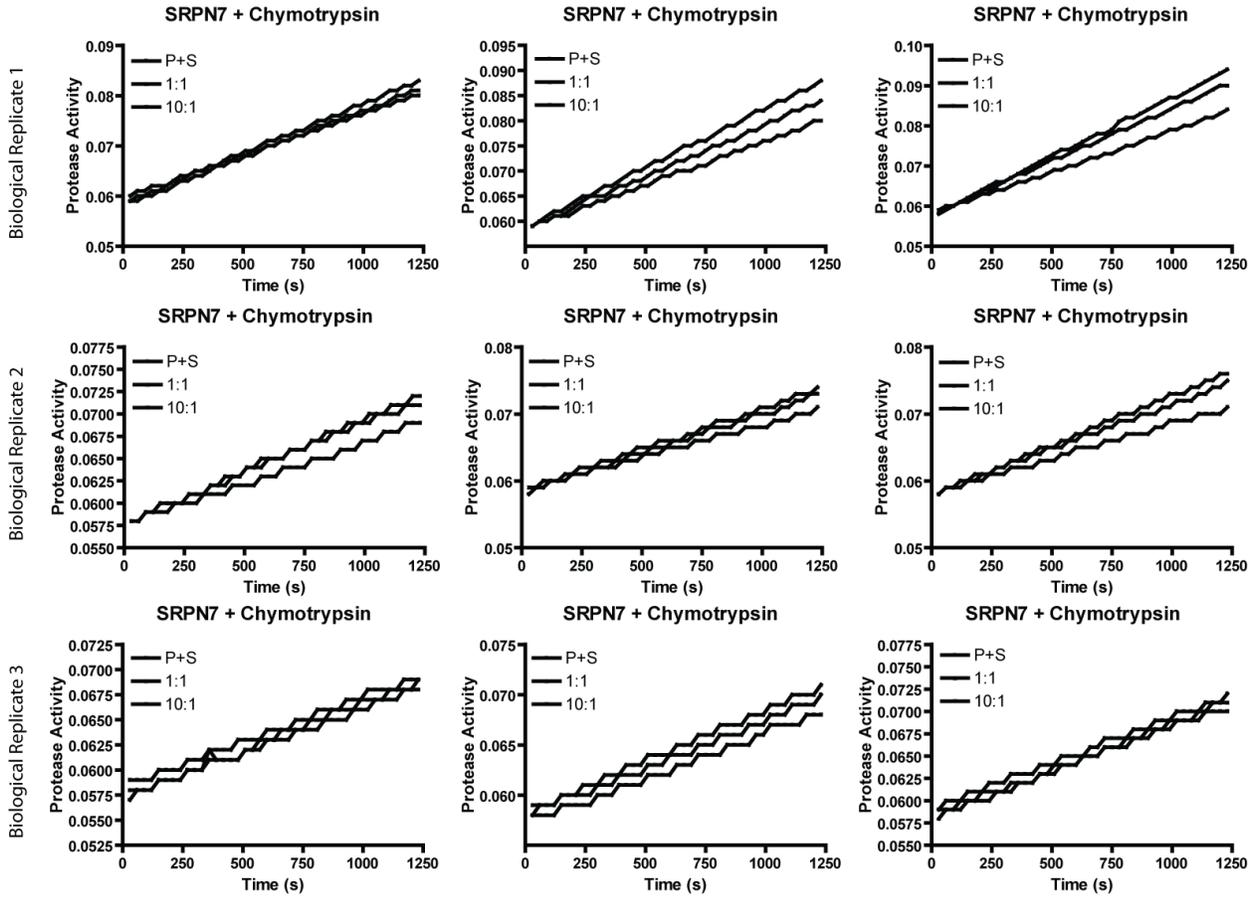
*Supplemental Figure B-2: rSRPN7 inhibits Plasmin*



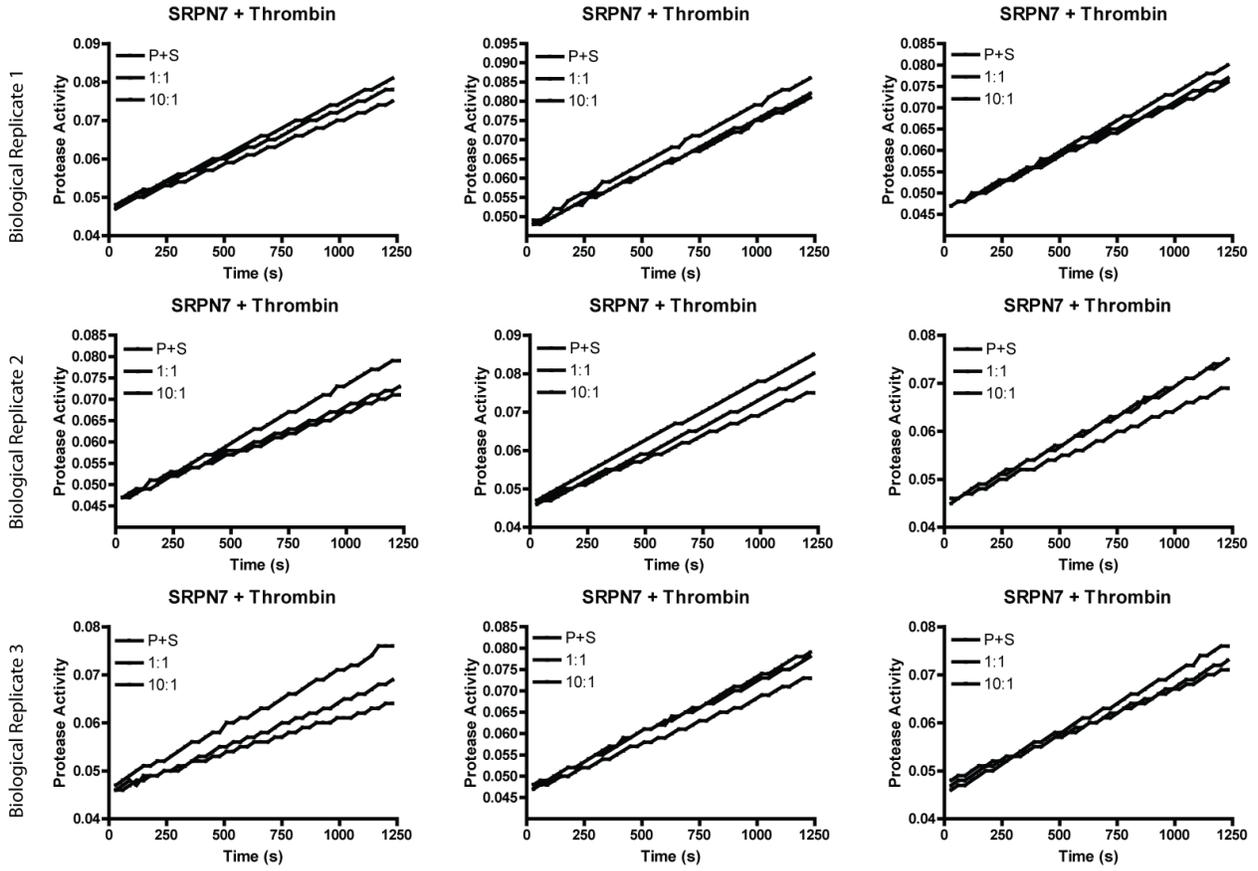
*Supplemental Figure B-3: rSRPN7 inhibits Kallikrein*



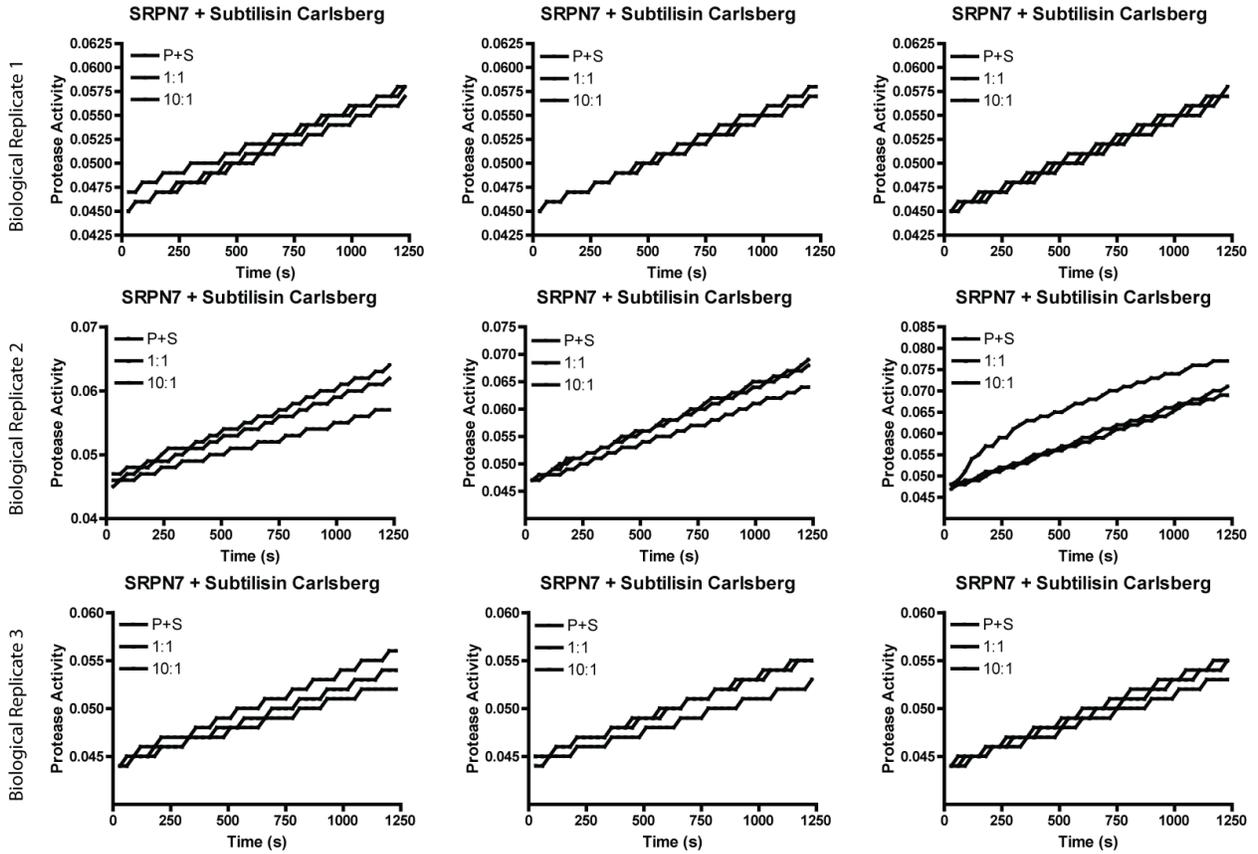
*Supplemental Figure B-4: rSRPN7 does not inhibit Chymotrypsin*



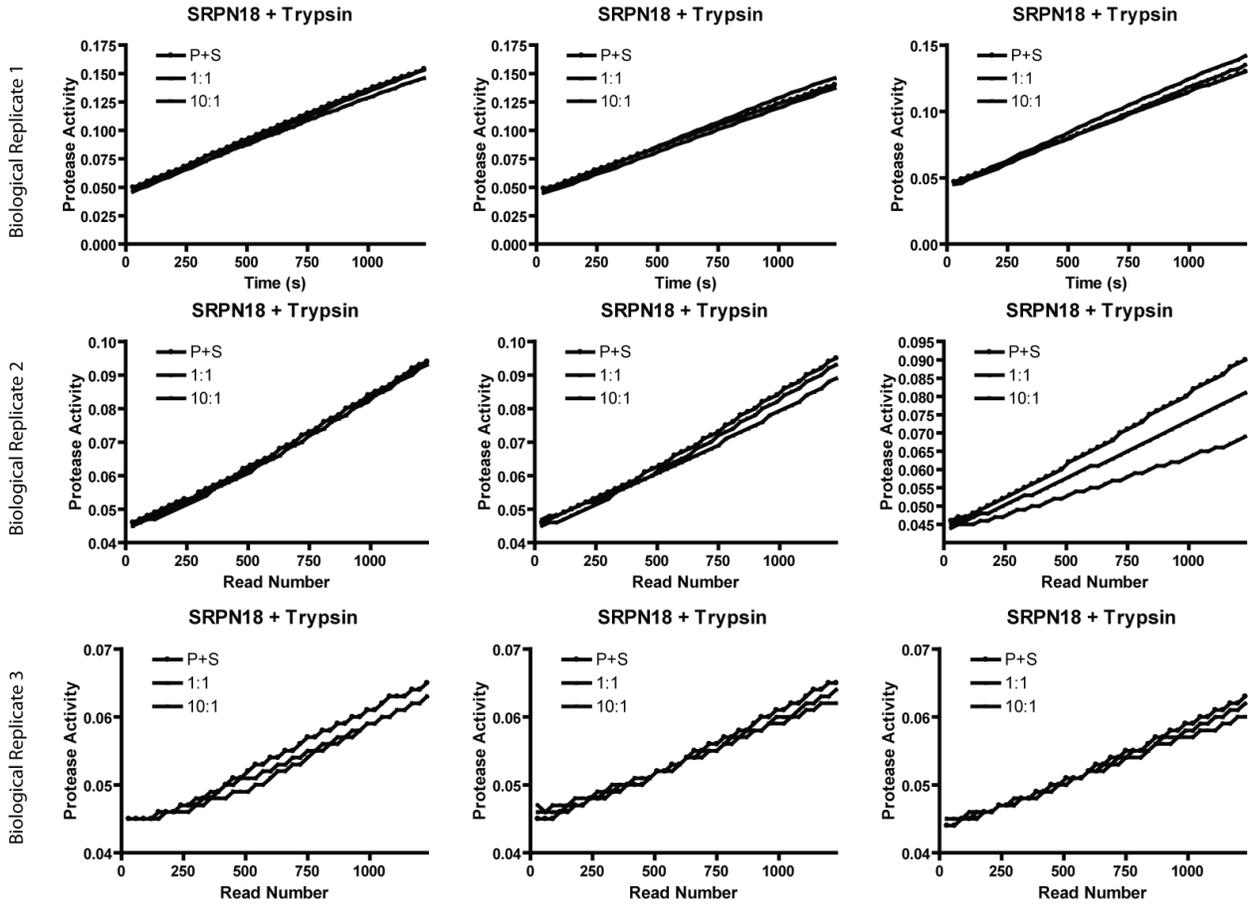
*Supplemental Figure B-5: rSRPN7 does not inhibit thrombin*



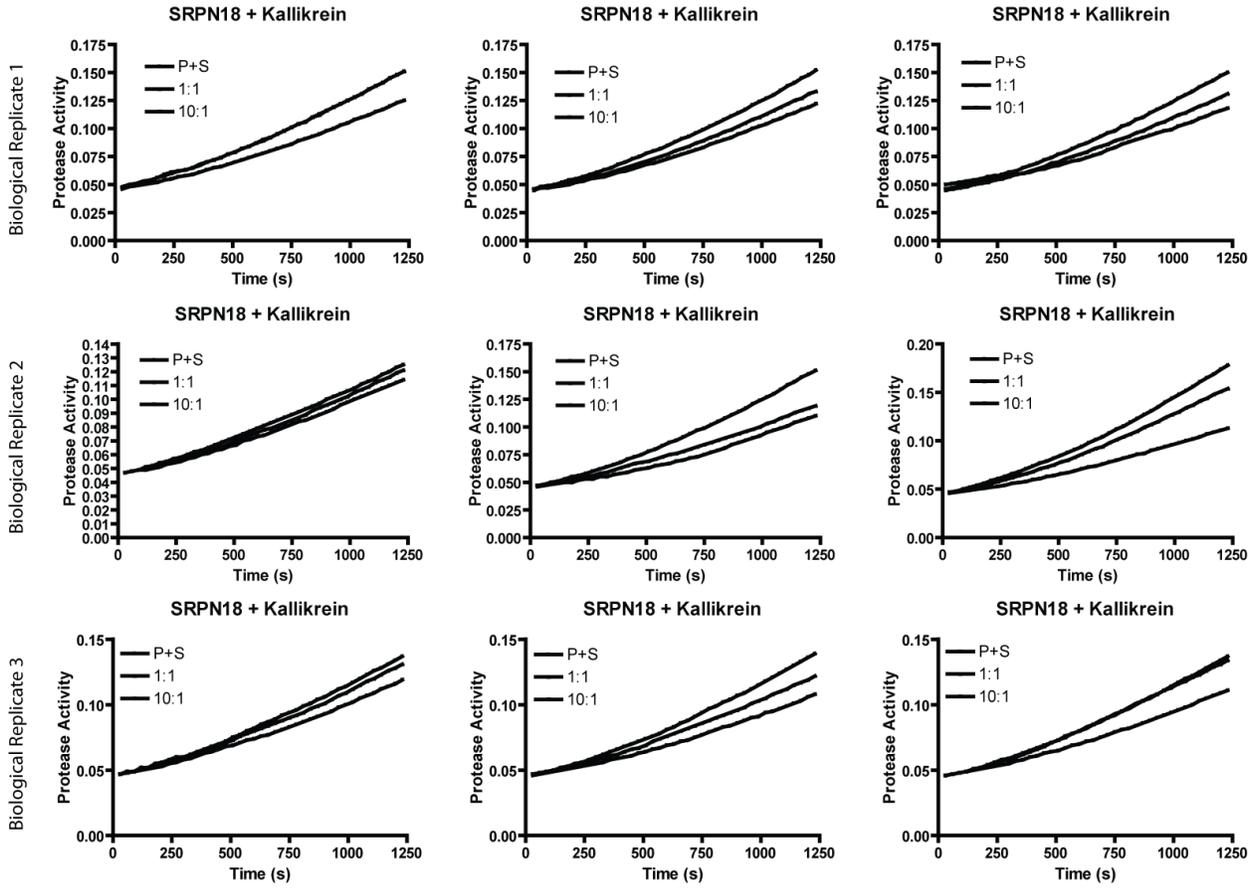
*Supplemental Figure B-6: rSRPN7 does not inhibit Subtilisin Carlsberg*



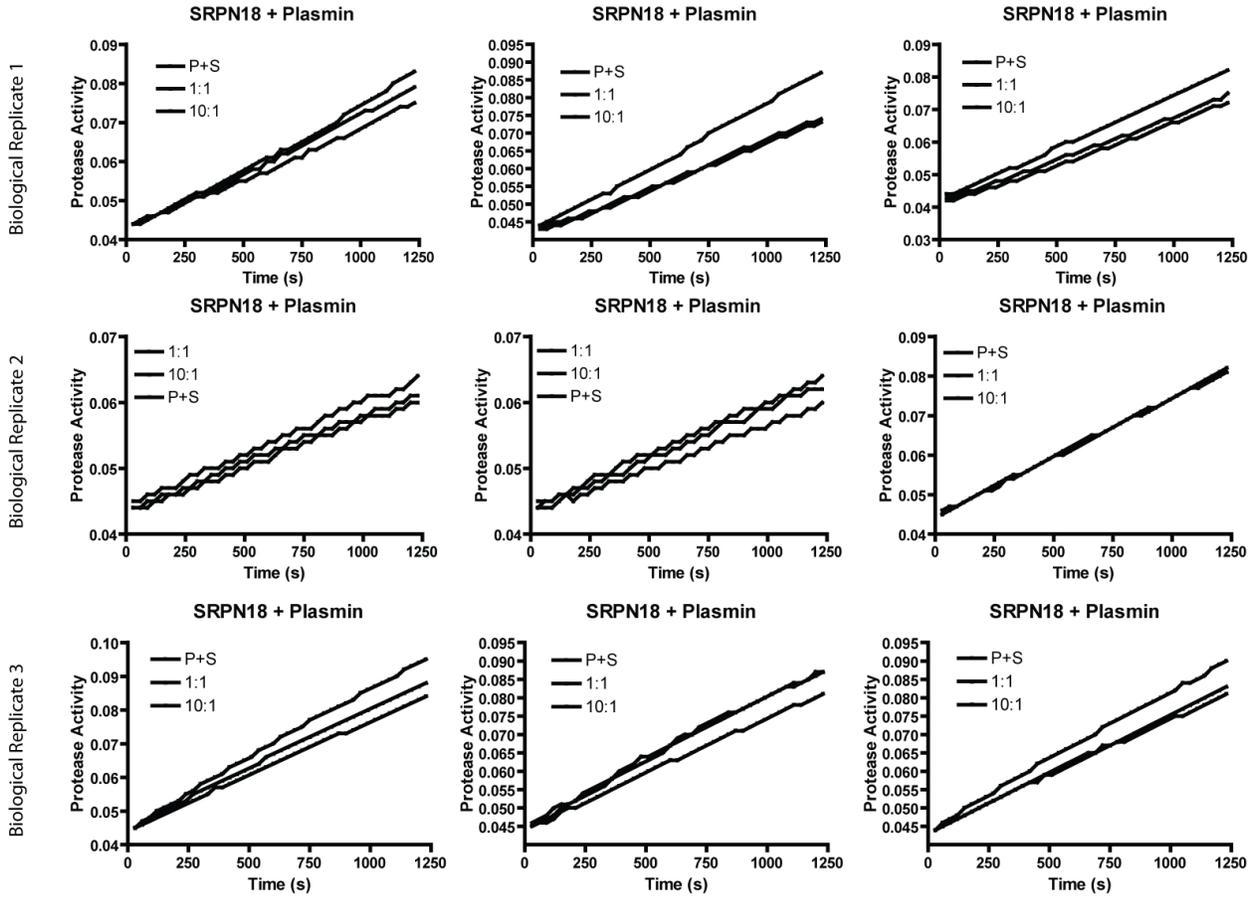
*Supplemental Figure C-1: rSRPN18 marginally inhibits trypsin*



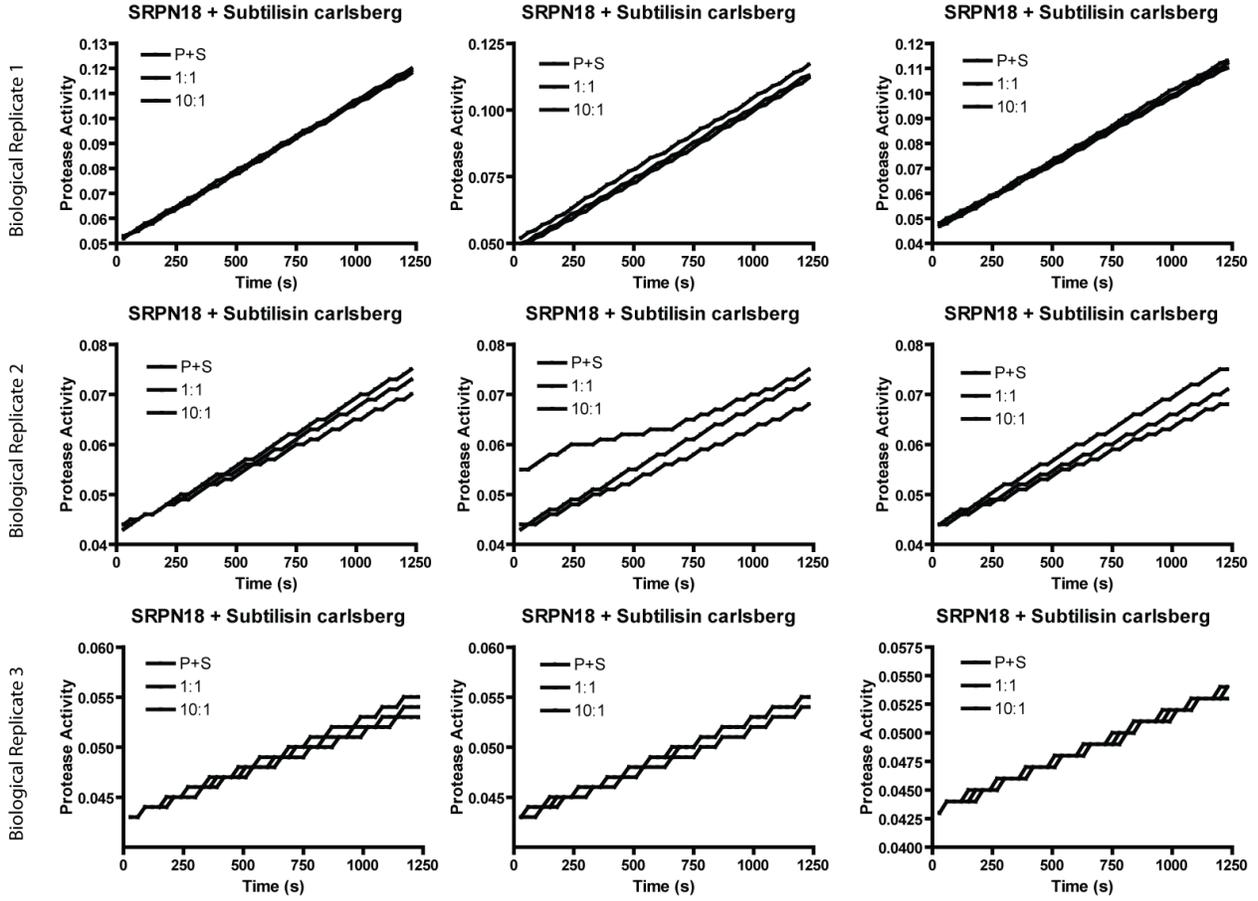
*Supplemental Figure C-2: rSRPN18 marginally inhibits kallikrein*



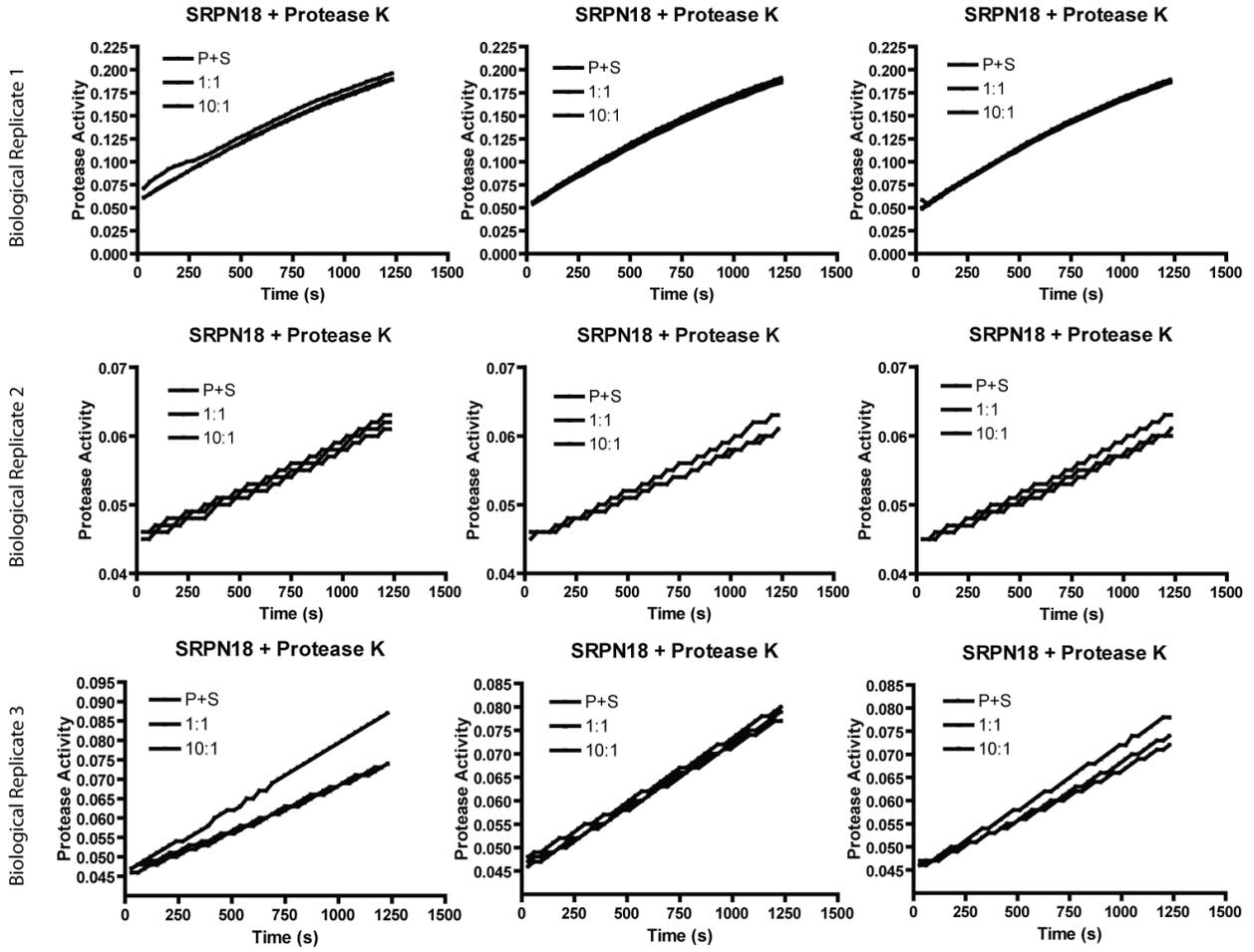
*Supplemental Figure C-3: rSRPN18 marginally inhibits plasmin*



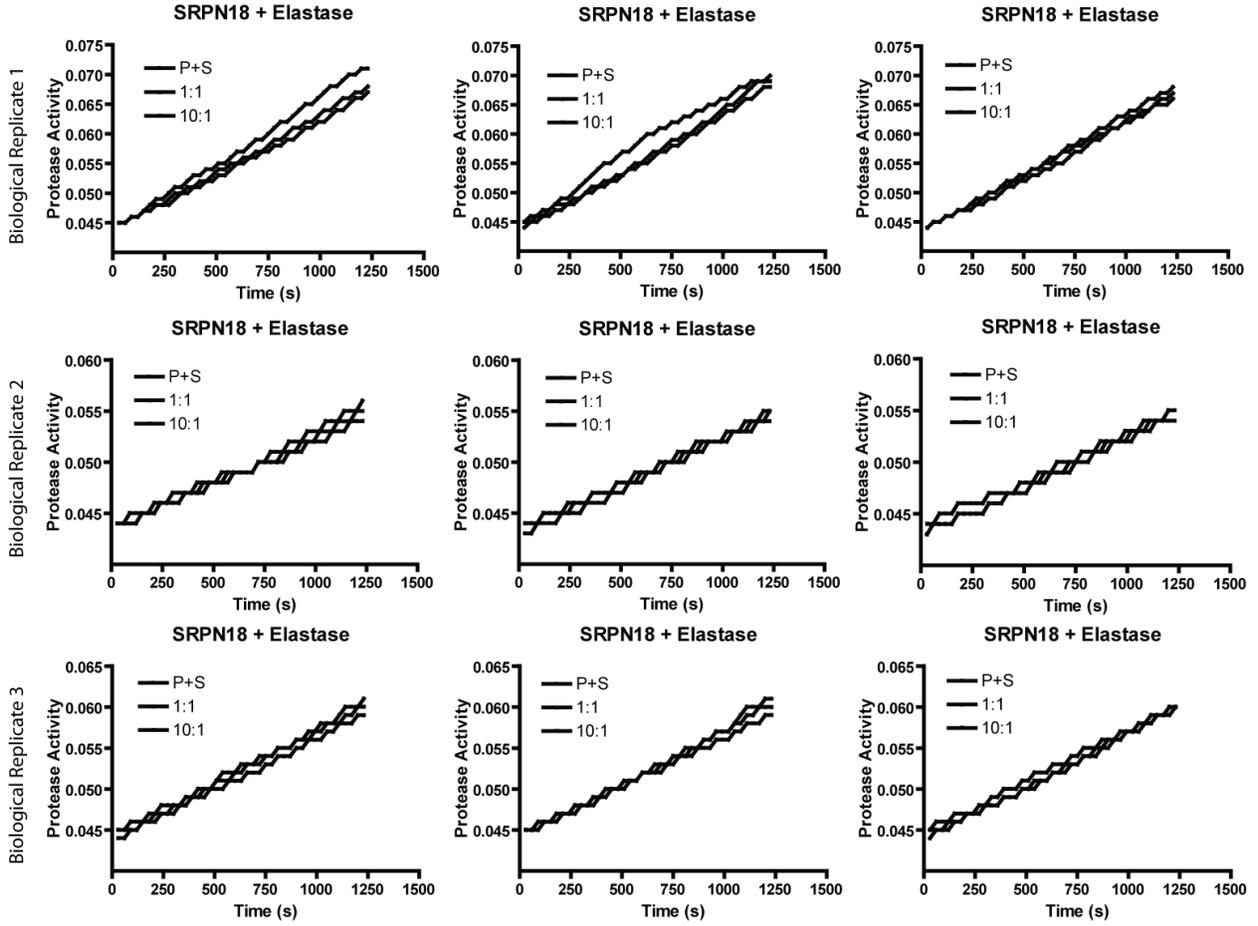
*Supplemental Figure C-4: rSRPN18 does not inhibit Subtilisin Carlsberg*



*Supplemental Figure C-5: rSRPN18 does not inhibit Protease K*



*Supplemental Figure C-6: rSRPN18 does not inhibit elastase*



*Supplemental Figure C-7: rSRPN18 does not inhibit chymotrypsin*

