SUMMARY

Innate immunity is the first line of defense against invading microorganisms and provides clues to adaptive immunity for the development of memory for subsequent infections. Insects, similar to other invertebrates, do not have adaptive immunity and thus rely on their innate immune system to combat infections. We have analyzed the role of the peptidoglycan (PGN) receptor protein (PGRP) family and other components of innate immune signaling pathways in the immune defense of the mosquito *Anopheles gambiae*, the main vector of human malaria in sub-Saharan Africa. The *PGRP* gene family consists of seven genes with ten PGRP domains. We have shown that from all *PGRP* genes only *PGRPLC* has a role in the resistance to bacterial infections of both Gram types. In our experiments we have used the Gram-positive bacterium *Staphylococcus aureus* and the Gram-negative bacterium *Escherichia coli*. The *PGRPLC* gene encodes at least three isoforms that derive from infection-driven alternative splicing of a pool of immature transcripts. Each isoform contains a different PGRP domain, encoded by three exons that all contribute equally to the PGN binding pocket. Structural modeling of the PGRPLC isoforms revealed a potential for all isoforms to bind both types of PGN, the Lys-type, which is mostly found in Gram-positive bacteria and the DAP-type, mostly found in Gram-negative bacteria. The isoform PGRPLC3 seems to be the most important in the defense against both bacteria species, although PGRPLC1 also has a crucial role in the defense against *S. aureus*.

Bacterial defense is mediated by the NF-κB transcription factor REL2, which is the ortholog of the *Drosophila* Relish. The *REL2* gene also encodes two protein isoforms: REL2-S, which only has the NF-κB domain, and REL2-F, which carries an I-κB inhibitory domain and a death domain in addition. REL2-F functions together with the receptor adaptor protein IMD to deal with *S. aureus* infections, whereas REL2-S has a role in the defense against *E. coli*. The *PGRPLC/IMD/REL2*-F pathway (*IMD*) is also partly responsible for the losses of *Plasmodium berghei*, which can be observed during the first stages of malaria infection of *A. gambiae*. *P. berghei* is a rodent malaria parasite, which has been used as a model in our studies. Whether the pathway is able to recognize the malaria parasite through PGRPLC or another associated receptor is still unclear. Another possibility is that the pathway is activated by the proliferation of commensal bacteria in the mosquito gut following a blood meal. However, we have shown that more than one of the three main isoforms of PGRPLC are required for the reaction to *P. berghei*. Other

PGRP genes, which have been proven to play a role during infection with *P. berghei,* are *PGRPLA2,* which also mediates parasite killing, and the almost identical and thus hardly indistinguishable *PGRPS2* and *S3,* which appear to inhibit parasite killing. This is possibly achieved by negative regulation of the *IMD* pathway through sequestration of PGN, which derives from the commensal bacteria and constitutively activates the pathway.

We have shown that REL1, the second NF-κB transcription factor of *A. gambiae*, which is orthologous to the *Drosophila* Dorsal (Dif does not exist in *Anopheles*), is not involved in the mosquito resistance to bacterial infections. This fact provides additional evidence that the *REL2*-associated pathways are of utmost importance in the *A. gambiae* defense to bacteria. In addition, REL1 has no role in the documented *P. berghei* killing. However, silencing of the REL1 inhibitor CACT (the ortholog of the *Drosophila* Cactus) during a parasite infection leads to a very strong refractoriness phenotype: most of the midgut-invading ookinetes are eliminated (presumably by lysis) and the remaining of the ookinetes are melanized. We thus assume that under wild-type infection conditions the parasite is either evading recognition by the REL1-associated receptors or actively modulating activation of REL1.

In conclusion, the data reported in this PhD thesis suggest significant divergence of immune signaling between the mosquito *A. gambiae* and the fruit fly *D. melanogaster*. The observed differences most likely reflect their different lifestyles and, consequently, different infectious agents, which the two insects encountered during their evolutionary lifetimes. In mosquitoes one of these agents is the malaria parasite.

ZUSAMMENFASSUNG:

Angeborene Immunität ist die primäre Verteidigungsstrategie gegen eindringende Mikroorganismen und liefert dem adaptiven Immunsystem Signale für die Entwicklung von Gedächniszellen für nachfolgende Infektionen. Ähnlich wie andere Invertebraten, verfügen Insekten nicht über eine adaptive Immunreaktion und verlassen sich deshalb voll auf ihr angeborenes Immunsystem, um Infektionen zu bekämpfen. Wir haben analysiert, welche Rolle die Familie der Peptidoglycan (PGN) Rezeptor Proteine (PGRP) und andere Komponenten der angeborenen Immunsignalkaskaden bei der Immunantwort des Moskitos *Anopheles gambiae* spielen. *A. gambiae* ist der Hauptüberträger der menschlichen Malaria im südlich der Sahara gelegenen Teil Afrikas. Die Genfamilie der PGRPs besteht aus sieben Genen mit zehn PGRP Domänen. Wir konnten zeigen, dass von allen PGRP Genen nur *PGRPLC* eine Rolle in der Verteidigung gegen bakterielle Infektionen, egal welchen Gramtyps, spielt. In unseren Experimenten haben wir das grampositive Bakterium *Staphylococcus aureus* und das gramnegative Bakterium *Escherichia coli* benutzt. Das *PGRPLC* Gen kodiert mindestens 3 Isoformen, die – je nach Infektion – aus einem Pool von unreifen Transkripten durch alternatives Splicing gebildet werden. Jede Isoform hat eine andere PGRP Domäne, welche jeweils von drei Exons kodiert wird, die alle gleich viel zur PGN Erkennungstasche beitragen. Strukturelle Modelle von PGRPLC zeigten, dass alle Isoformen dazu in der Lage sind, beide Arten von PGN zu binden, wobei die Lys-Form hauptsächlich in grampositiven Bakterien und die DAP-Form hauptsächlich in gramnegativen Bakterien vorkommt. Die PGRPLC3 Isoform scheint die wichtigste Rolle bei der Verteidigung gegen die beiden bakteriellen Formen zu haben, obwohl PGRPLC1 auch eine wichtige Rolle bei der Verteidigung gegen *S. aureus* zu spielen scheint.

Die Verteidigung gegen Bakterien wird von dem NF-κB Transkriptionsfaktor REL2 bewerkstelligt, welcher das Ortholog des *Drosophila Relish* Gens ist. Das *REL2* Gen kodiert zwei Protein Isoformen: REL2-S, das lediglich die NF-κB Domäne hat, und REL2-F, das zusätzlich noch eine inhibierende I-κB und eine Death Domäne hat. REL2-F arbeitet mit dem Rezeptor-Adaptor Protein IMD zusammen, um *S. aureus* Infektionen zu bekämpfen, wogegen REL2-S eine Rolle in der Verteidigung gegen *E. coli* spielt. Die *PGRPLC/IMD/REL2*-F Signalkaskade (*IMD*) ist auch teilweise verantwortlich für die Verluste, die *Plasmodium berghei*, ein Nager Malariaparasit, während der ersten Stadien der Malaria Infektion in *A.gambiae* erleidet. Dieser Malariaparasit ist als Modelsystem für

unsere Studien verwendet worden. Ob die *IMD* Signalkaskade auch fähig ist, den Malariaparasiten direkt durch PGRPLC oder einen anderen beteiligten Rezeptor zu erkennen, ist immer noch unklar. Nach einer Blutmahlzeit vermehren sich die residenten Bakterien im Moskitodarm. Es ist durchaus möglich, dass die Signalkaskade dadurch aktiviert wird. Unabhängig davon, konnten wir zeigen, dass mehr als eine der drei PGRPLC Isoformen benötigt werden, um eine Reaktion auf *P. berghei* hervorzurufen. Weitere PGRP Gene, die nachweislich eine Rolle während der Infektion mit *P. berghei* spielen, sind *PGRPLA2* und *PGRPS2* und *S3*. PGRPLA2 begünstigt den Kampf gegen den Parasiten, während PGRPS2 und S3, die nahezu identisch und deshalb kaum zu unterscheiden sind, den Kampf gegen den Parasiten zu behindern scheinen. Letzteres könnte durch Sequestration des PGN der residenten Bakterien erreicht werden, welches zu einer Inhibierung der *IMD* Signalkaskade führen würde.

Wir haben gezeigt, dass der zweite NF-κB Transkriptionsfaktor in *A. gambiae*, REL1, der ein Ortholog des *Drosophila* Dorsal ist (Dif existiert nicht in *Anopheles*), nicht in der Moskito Verteidigung gegen bakterielle Infektionen involviert ist. Dieser Umstand beweist noch einmal mehr, welche große Bedeutung den *REL2* Signalkaskaden in *A. gambiae* im Kampf gegen die Bakterien zukommt. Zudem scheint REL1 keine Rolle in der Verteidigung gegen *P. berghei* zu spielen. Setzt man jedoch den REL1 Inhibitor CACT (das Ortholog des *Drosophila* Cactus) während einer Malaria Infektion außer Kraft, so kommt es zu einem sehr stark refraktorischen Phenotypus: Die meisten der Ookineten, die versuchen in den Moskito Mitteldarm einzudringen, werden eliminiert (vermutlich lysiert) und die restlichen Ookineten werden melanisiert. Wir vermuten, dass unter den Bedingungen einer natürlich hervorgerufenen Infektion der Parasit entweder die Erkennung durch die Rezeptoren der REL1 Signalkaskade vermeidet oder aber gezielt die Aktivierung von REL1 verhindert.

Zusammenfassend lässt sich sagen, dass die Daten, die in dieser Doktorarbeit angeführt werden, eine signifikante Abweichung der Signalkaskaden des Moskito *A. gambiae* von denen der Fruchtfliege *D. melanogaster* beschreiben. Die beobachteten Unterschiede lassen sich wahrscheinlich auf die unterschiedlichen Lebensumstände und infektiösen Organismen zurückführen, denen diese zwei Insekten während ihrer Evolution ausgesetzt waren. Für den Moskito war einer dieser Organismen der Malaria Parasit.

The Role of PGRP Proteins in Innate Immunity Pathways in the Malaria Vector *Anopheles gambiae*

Stephan Meister

2006

Dissertation

submitted to the

Combined Faculties for the Natural Sciences and for Mathematics

of the Ruperto-Carola University of Heidelberg, Germany

for the degree of

Doctor of Natural Sciences

presented by

Diplom-Biologist Stephan Meister

born in Heilbronn-Neckargartach

Oral examination:

The Role of PGRP Proteins in Innate Immunity Pathways in the Malaria Vector *Anopheles gambiae*

Group/Team Leader (chair of TAC):	**Prof. Fotis C. Kafatos**
Co-Supervisor:	**George K. Christophides, PhD**
Second advisor from same unit:	**Dr. Iain Mattaj**
EMBL advisor from different unit:	**Dr. Eileen Furlong**
University advisors:	**Dr. Renato Paro (University of Heidelberg)**
	Dr. Christine Clayton (University of Heidelberg)

European Molecular Biology Laboratory EMBL Heidelberg &
Ruprecht-Karls-Universität Heidelberg

2006

DEDICATED TO MY LITTLE SISTER BÄRBEL

(my parents got my Diploma Thesis already)

ACKNOWLEDGEMENTS

My first and foremost thanks go to my supervisor Giorgos K. Christophides, PhD and my advisor, Prof. Fotis C. Kafatos for the opportunity to work in their joint lab and for their outstanding continuous support during this work. Their guidance was critical in that they gave me the freedom to develop my own research ideas in critical discussions while making sure my work stayed focused – most of the time… and if there is someone that this dissertation would not have been possible without – it's Giorgos. Thank you for also being a friend and listening to many a crazy story.

I would also like to thank my PhD thesis committee members, Dr. Eileen Furlong, Dr. Iain Mattaj and Dr. Renato Paro for their fruitful discussions and suggestions during the 3.5 years. I would also like to thank the Dr. Christine Clayton for sacrificing her valuable time for my PhD defense.

Other than that I do not believe in pages filled with flowery gratitude – there were basically three categories of people during my time as a PhD student::

1. People that were relevant to the work as well as good friends – thanks for putting up with me…
Bogos (Pavlos) Agianian, Claudia Blass, Dolores Doherty, Tibebu Habtewold, Anastasios Koutsos (in alphabetical order)

…and all the past and present members of the Kafatos' lab that overlapped with my being there.

2. People that were less important to the actual work, but never the less good friends, accompanying me parts of the way and/or making some sort of impact – thank you, for you made this time worthwhile…
Bärbel Meister, Sebastian Jacobi, Asli Kocaömer, Fanta Kocaömer, Benjamin Becker, Stephanie Blandin, Jennifer Volz, Gareth Lycett, Denise Sofia, Janus Jacobsen, Gerd Görres, Pit Jäger, Björn Pospiech, Klaus Dieter Horlacher, Susanne Ortelt, Kate McElroy, Yvonne Winter, Fabian Felipp, Ina Mack, Candace Zaiden, Shifu Shi Yanzi, Gabriella Silvestri, Gianmaria Liccardi (in chronological order)

…and of course my parents Elisabeth von Lübtow and Walter Meister.

3. People I would have preferred to see less of – thanks for nothing…
You know who you are, but I doubt you will ever read this!

TABLE OF CONTENTS

TABLE OF CONTENTS

INTRODUCTION

1.1 MALARIA

Historical perspective.

Alphonse Laveran (1845-1922) discovered the malaria parasite over a hundred years ago in the 1880s, and by 1897 Sir Ronald Ross (1857-1932) had elucidated the complex development of the malarial parasite in the mosquito. Already at that time, there was an effective drug to fight malaria: Quinine. However, over a hundred years later, we still do not know enough about the disease to be able to defeat it permanently.

Among human diseases, malaria still is one of the top killers - it is the leading cause of pediatric morbidity and mortality in Africa. It looms virtually everywhere in the tropics, but also occurs in many temperate regions. It threatens travelers and immigrants with no previous exposure to the malady as the immunological status of a person has a strong bearing on the severity of the disease. Since the advent and rise of global tourism imported non-endemic cases have been on the rise (estimates calculate about 12,000 annual cases in Europe).

Concerted malaria control programs were developed in the late 1940s and during the 1950s and 1960s an ambitious worldwide campaign to eradicate malaria was initiated. The control measures put in place in the WHO member countries had tremendous initial success and malaria incidence came down considerably. Indoor residual spraying with DDT was the principle method by which malaria transmission was eradicated or greatly reduced on the periphery of its transmission range (Hemingway, Field et al. 2002). However, at the end of the 1960s, the concept of malaria eradication was formally given up and the countries switched to sustainable control, largely because insecticide resistance swept through

mosquito vector species, crippling the eradication efforts. In the following decades a resurgence of malaria ensued due to the spread of resistance to malaria drugs and insecticides.

In 1998, the 'Roll Back Malaria' (www.rbm.who.int/) global partnership was founded to implement and coordinate malaria counter measures on a global scale, yet again. This partnership was initiated by UNICEF, the UN Development Program, the World Bank and the WHO. One of its main goals is to halve the burden of malaria by 2010. It has grown to include governments of countries affected by the disease, non-governmental organizations, academic institutions, international private sector representatives and research groups.

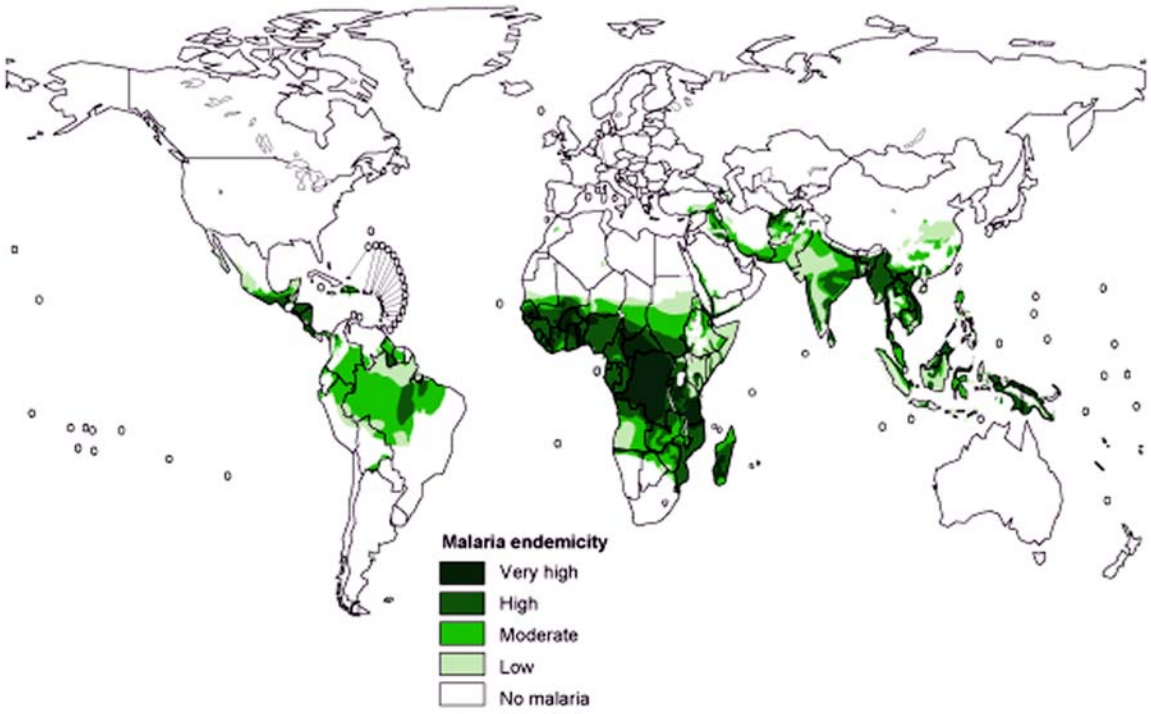


Fig. 1.1.1 Global distribution of malaria transmission risk, 2003 (source: The World Malaria Report 2005 by the Roll Back Malaria Partnership)

At the end of 2004, 107 countries and dependent territories had areas at risk of malaria transmission (Fig. 1.1.1), and according to the World Malaria Report 2005 by the Roll Back Malaria Partnership, some 3.2 billion people, half the world's population, lived in such areas.

Malaria affects the standard of education and the economy in endemic countries. Possible counter measures that are supported by the Roll Back Malaria initiative and have been proven effective are the improvement of health education; a better case management with prompt access to effective treatment (waiting even six hours for treatment can mean the difference between life and death to a child with cerebral malaria); the extensive use of pyrethroid-impregnated bed nets for mosquito control; early detection of and response to malaria epidemics; and a good coordination of all of the above.

Impact and significance on Global Health.

Malaria has always been a global problem (Fig. 1.1.2) – endemic in 107 countries with more than 300 to 500 million clinical cases and 1.5 to 2.7 million deaths worldwide each year. Most of these malaria infections can be attributed to *Plasmodium falciparum* and *Plasmodium vivax*. This makes malaria the second most frequent infectious disease in the world, second only to tuberculosis. Ninety percent of fatal malaria cases are in Africa, where malaria accounts for twenty percent of all childhood deaths. The region hit hardest by malaria is sub-Saharan Africa. About 60 percent of worldwide malaria cases, 75 percent of *falciparum* malaria cases and more than 80 percent of malaria deaths occur there. Outside Africa, some two-thirds of the cases occur in just three countries: Brazil, India and Sri Lanka.

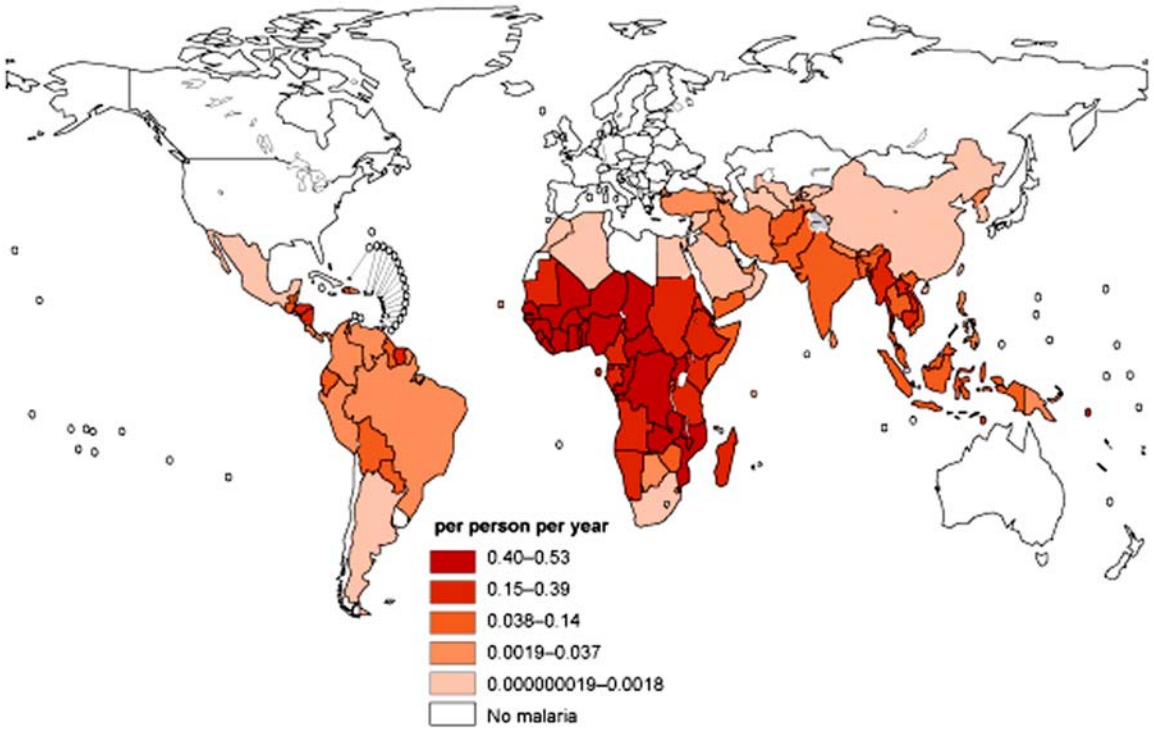


Fig. 1.1.2 Estimated incidence of clinical malaria episodes (caused by any species) country level averages, 2004 (source: The World Malaria Report 2005 by the Roll Back Malaria Partnership)

As mentioned above, children are the part of the population that is hit the hardest by malaria; even if they survive, the disease can cripple their intellectual and physical development. The disease contributes greatly to anemia among children, hampering their growth and development. During pregnancy, malaria results in maternal illness and severe anemia and contributes to low birth weight among newborns - one of the leading risk factors for infant mortality. All this results in a severe slowing of the development and growth of national economies, perpetuating poverty. Malaria truly is a disease of the poor.

Clinical features of malaria.

After a passage through liver cells, the time depending on the parasite species, *Plasmodium* infects red blood cells (RBC) and multiplies within them in an asexual cycle. Upon parasite maturation, RBCs rupture and the metabolic byproducts are released into the blood stream with the new plasmodia that cause a fever bout which subsides when the parasite invade new RBCs (generally within 3-5 hours). The cycle of fever fits is very strenuous and the time between them can vary from a few to 72 hours (Table 1.1.1).

Table 1.1.1: Different types of malaria

		Malaria tertiana	Malaria quartana	Malaria tropica
Parasite		*P. vivax & P. ovale*	*P. malariae*	*P. falciparum*
Regions		temperate zones	tropics	tropics & subtropics
Fever bout	start	fast	quickly	very fast
	length	3-4 h	4-5 h	irregular
	recurrence	every 48 h	every 72 h	irregular
recovery without treatment		after 12 fever fits (3 weeks)	after 20 fever fits (8 weeks)	mostly deadly within short time
recurrence if not treated		often within first 5 years	often within first year	often in case of survival – with same gravity and danger

The clinical features of malaria can vary from person to person depending on their immunological history and the type of malaria. Generally, the incubation period from the time of infection to full-blown malaria lasts one to four weeks. The encountered classic symptoms include a persistent fever, shivering, joint pains, headaches and vomiting. Severe and complicated malaria causes renal failure, hypoglycemia, anemia, pulmonary edema, shock and coma with potentially fatal consequences.

Malaria can be cured if diagnosed on time and treated adequately. However, only 60% of those suffering from malaria have prompt access to appropriate treatment within 24 hours of the onset of symptoms.

P. falciparum may persist in the blood at low non-clinical levels (due to partially effective immunity or incomplete drug treatment) and then increase to cause obvious illness.

If untreated, the disease can progress causing a variety of serious complications such as blockage of blood vessels leading to cerebral malaria, coma and death. For sufferers without partial immunity (e.g. Western travelers, migrant workers), it is possible for death to occur within 24 hours from the first appearance of symptoms.

In contrast, *P. vivax* can lie dormant in the liver and relapse up to several years after the initial illness. The nature of the relapse 'trigger' is unknown. *P. vivax* does not adhere to blood vessels and therefore does not cause the associated complications.

1.2 *ANOPHELES GAMBIAE*

The malaria vector.

There are more than 2,500 known mosquito species worldwide. However, only around 50 to 60 species of *Anopheles* mosquitoes are capable of transmitting malaria parasites between humans.

Life cycle.

Anophelines, like all other mosquitoes, undergo complete metamorphosis with four distinct developmental stages in the course of their lifetime: egg, larvae, pupae and adult (Fig. 1.2.1). The pre-adult stages complete their life cycle in water. The length of the mosquito's life cycle depends on temperature and species characteristics. Some species have adapted to go through their life cycle in as little as four days whereas other species take as long as one month.

Eggs: It is only the adult females that bite man and other animals for a blood meal after which they lay their eggs in standing or slow-moving water that can be found in a variety of sources ranging from small containers to vast expanses of marshland. In her lifespan a female mosquito will lay eggs approximately one to three times.

Larval stage: Within 24 to 48 hours, the laid eggs hatch into larvae (also termed wrigglers). The larval stage is always aquatic and larvae feed on organic matter, micro-organisms and each other. They only shuttle to the surface to obtain oxygen through a snorkel-like breathing apparatus. There are four larval stages (instars), at the end of each one the larvae molt; after the fourth molt, they become pupae. Larval development usually takes from 7 to 10 days.

Pupal stage: The pupal stage does not feed but unlike most insect pupae it is extremely active. Pupae also live near the surface of the water, breathing air through two horn-like tubes (called siphons) that are on their back.

Adult (Imago) stage: The adult hatches from the pupal case using air pressure after a few days. The life span of an adult mosquito depends on several factors: temperature,

humidity, sex of the mosquito and time of year. Most males live for a very short time, about a week; females can live for about a month, depending on the aforementioned factors.

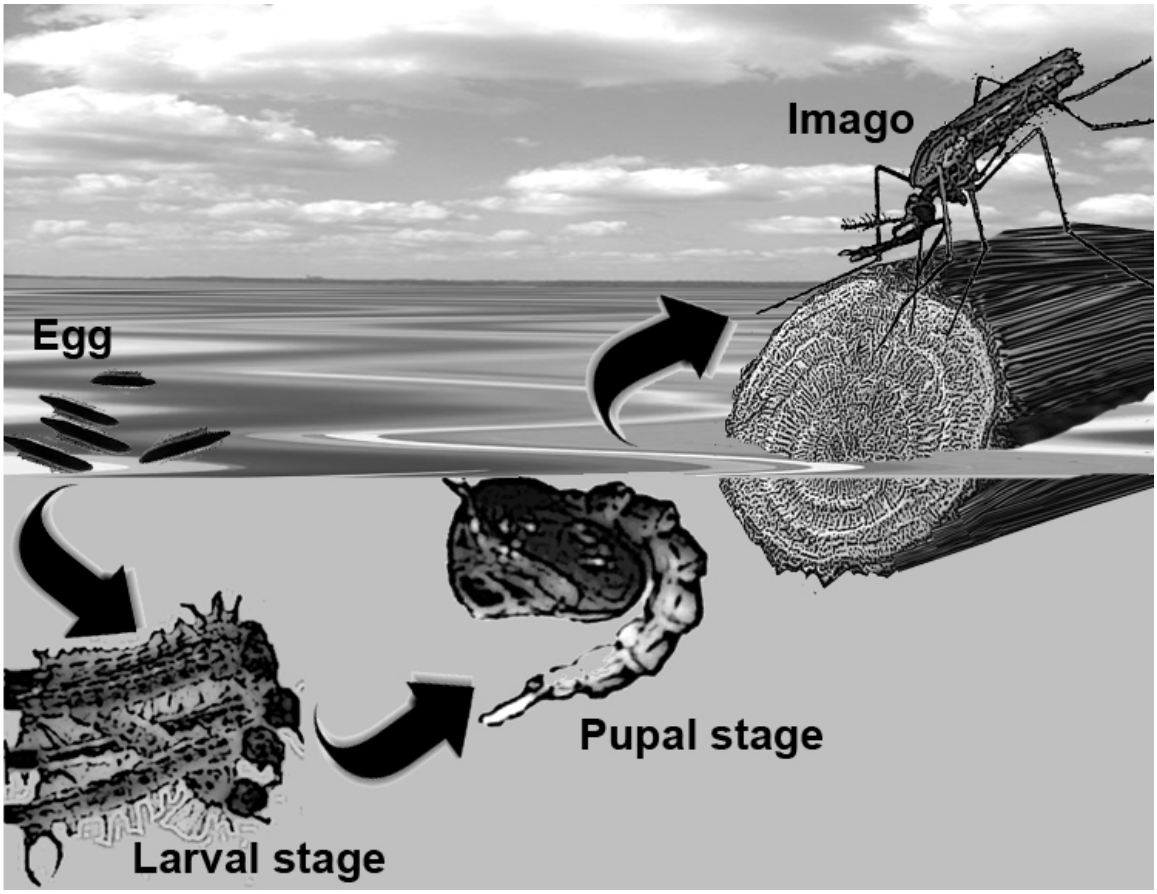


Fig. 1.2.1: The four stages of the mosquito life cycle are: egg, pupa, larva, and adult (imago). © S. Meister 2005

1.3 *PLASMODIUM* SPECIES

The malaria parasite.

There are four species of the apicomplexan, protozoan parasite genus *Plasmodium* that can cause malaria in humans: *P. falciparum*, *P. vivax*, *P. ovale* and *P. malariae*. Only two are highly prevalent, *P. falciparum* and *P. vivax*. Though *P. vivax* malaria is the most common, the malaria brought on by *Plasmodium falciparum* is the most lethal and accounts for around 90% of malaria deaths in Africa and about 50% in South East Asia/Latin America.

All human malaria is spread by female anopheline mosquitoes which have piercing mouthparts and need a supply of blood in order to produce and lay eggs. Malaria can also be transmitted, more rarely, by blood transfusion, contaminated needles and syringes and in rare cases from mother to child before and/or during birth.

Life cycle.

Stage I (sporozoite and merozoite stages)**:** When a parasite-infected mosquito bites, the sporozoite form of the parasite gets injected into the human bloodstream and moves into the liver - within ~30 minutes – where it reproduces in an asexual proliferation cycle for 5 days or more, depending on the species (*P. falciparum* or *P. vivax*). During this phase hundreds of thousands of merozoite stage parasites are produced.

Stage II (merozoite and trophozoite stages)**:** Merozoites break out from the liver, enter the bloodstream, and within minutes invade the RBC to grow and multiply asexually via the trophozoite stage, producing more merozoites. Every 36-72 hours (depending on the species) the infected RBCs rupture, releasing metabolic byproducts and more parasites into

the blood stream causing fever, chills and anemia - all the clinical symptoms associated with malaria. The released parasites then invade other RBC, beginning a new cycle. Eventually, up to 10% of all RBC become infected.

Stage III (gametocyte stage)**:** At some point during the course of the infection, most notably when the asexual proliferation slows down, the merozoites grow but do not divide, and produce the sexual parasite stages in a subset of infected RBC. These are the gametocytes, which are the gamete precursors and are distinguishable as males (Microgametocyte) and females (Macrogametocyte). These gametocytes rest inert in the vertebrate host blood and only upon being ingested by a biting mosquito will they initiate the mosquito part of the parasite life cycle. A population of mature *P. falciparum* gametocytes has a half-life of 2.5 days and, depending on its size, may persist at levels infective to mosquitoes for periods as long as 22 days (Smalley and Sinden 1977).

Stage IV (gamete, zygote and ookinete stages): Upon being taken up with a bloodmeal by a biting female mosquito, the gametocytes transform into gametes. The male gametocyte undergoes exflagellation and produces up to 8 male gametes (microgametes). Female gametocytes produce only one female gamete (macrogamete) each. Gametogenesis occurs within 10−15 min following the uptake of the bloodmeal and is triggered by the drop in temperature and change in pH. Within 30 min, microgametes must actively find and fertilize the macrogametes. Gamete pairs fuse to form a zygote which then develops into the motile ookinete and sequentially traverses the mosquito peritrophic membrane and midgut epithelial mono-layer to encyst on the opposite, basal side.

Stage V (oocyst and sporozoite stages): On the hemolymph side of the mosquito midgut epithelium, underneath the basal lamina, the ookinete will develop to a sporozoite-

filled oocyst which will take 8−15 days (depending on the *Plasmodium* species) to mature. During this maturation process a meiotic cycle is followed by several rounds of mitosis to produce several thousand haploid sporozoites. Upon maturation, the oocyst ruptures and releases the sporozoites which travel through the mosquito hemolymph to reach the salivary glands.

Stage VI (sporozoite stage): At the salivary glands, the sporozoite stage parasites traverse the second mosquito epithelium, this time in the basal to apical direction, and accumulate within the gland ducts. From there, the parasites get injected into the next human host with the mosquito saliva, which contains a complex mixture of anti-hemostatic, anti-inflammatory and immunomodulatory compounds. The dual host cycle of the parasite then starts anew.

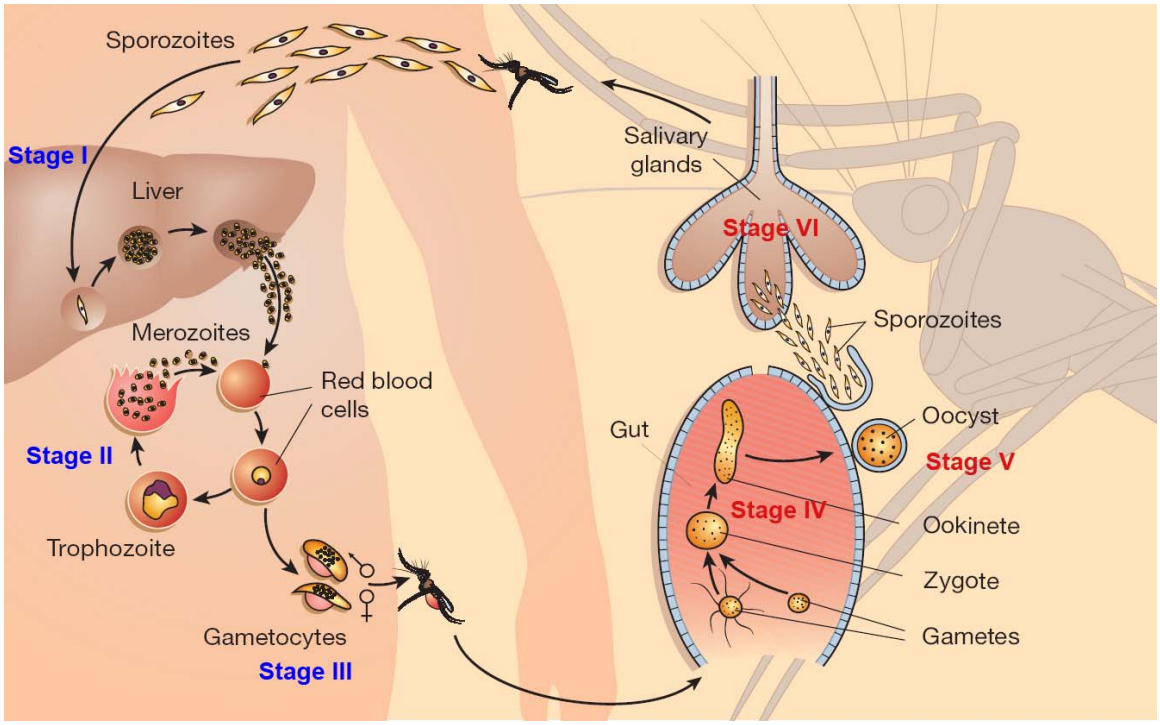


Fig. 1.3.1 Life cycle of the parasite *Plasmodium falciparum*. The developmental stages of the parasite cycle in the human host (I - III) are labeled in blue and the developmental stages in the mosquito vector (IV – VI) are labeled in red. (modified from (Wirth 2002))

1.4 MOSQUITO-PARASITE INTERACTIONS

Hematophagy or blood feeding is a behavior shared between various arthropod taxa. In mosquitoes it is exhibited by nonautogenous females that require a blood meal for egg production. Pathogens such as the human malaria parasite profit from this circumstance and use the mosquito as a vehicle to travel between vertebrate hosts.

The interactions between *Plasmodium* and *Anopheles* are rather complicated and not well understood. The mosquito is far from being a willing vehicle to the parasite, contrary to the initial notion of the parasite just passing swiftly and without damage through the mosquito. The developmental steps the parasite has to undergo are many and complex, and can fail on a number of levels as indicated by the huge amount of parasite losses in the mosquito (Fig. 1.4.1).

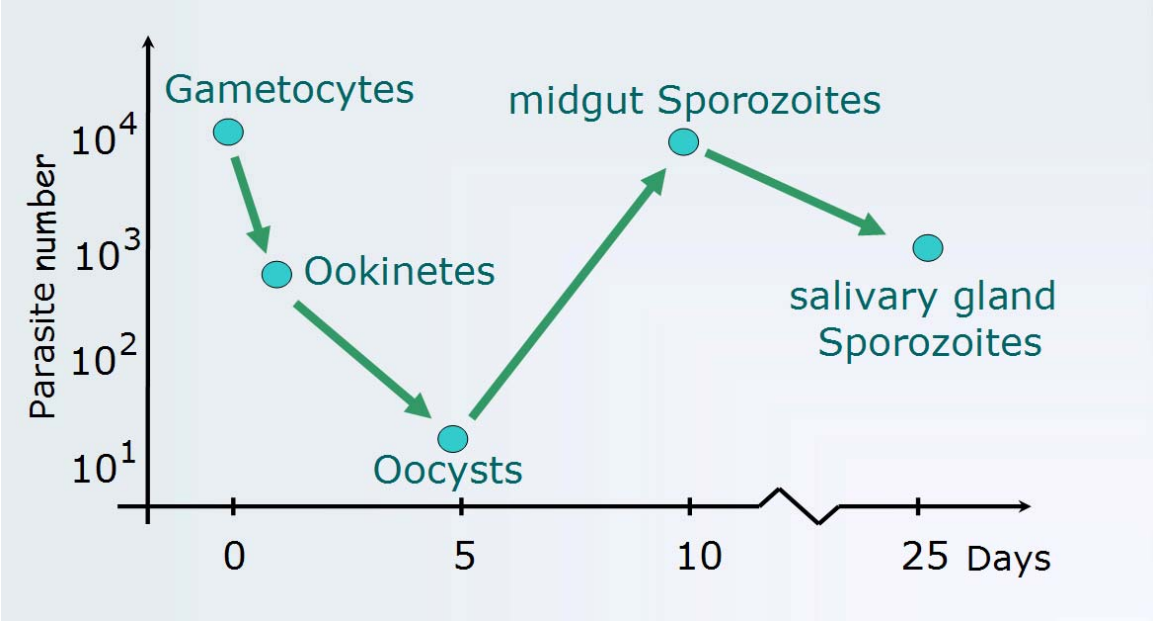


Fig. 1.4.1 Parasite losses in the mosquito vector (modified from (Sinden 1999)).

Some researchers would claim that the anthropocentric view of humans as hosts and mosquitoes as vectors is incorrect. Firstly, the adaptation of the parasite to humans has in all likelihood only occurred recently in evolutionary time. Secondly, as a consequence, the parasite is much better adapted and causes less damage to the mosquito, whereas the so-called human host suffers immensely. Lastly, the all important sexual part of the parasite life-cycle occurs in the mosquito and the proliferation in the human is purely asexual.

Susceptibility and refractoriness of mosquitoes to parasites.

Because development of the parasite in the mosquito is delicate, most malaria parasites develop successfully only in a few mosquito species and therefore, only a limited number of *Plasmodium* – *Anopheles* combinations can cause malaria in humans or other animals. This number of effective combinations is further narrowed by mosquito biting preferences and vectorial capacity (the varying ability of individual mosquitoes of the same species to sustain parasite development). For example, a mosquito might passively lack some requirement for successful parasite development, such as Xanthurenic acid to activate *Plasmodium berghei* (a rodent parasite) gametogenesis in the midgut (Billker, Lindo et al. 1998), and/or it might actively mount a strong immune response against the parasite. The hypothesis of mosquitoes mounting immune responses against the malaria parasite is supported by laboratory selection of mosquito strains that are refractory to the parasites. In these strains the parasites are killed while traversing the midgut epithelium, e.g. by melanization in the L3-5 strain (Collins, Sakai et al. 1986) (Fig. 1.4.2c) or by lysis in the SUAF2 strain (Vernick, Fujioka et al. 1995) (Fig. 1.4.2b).

Refractoriness thus appears to have a genetic basis and occurs in the wild. It is thought to be conveyed by natural resistance alleles (Niare, Markianos et al. 2002) that limit parasite development in the vector to relatively small numbers of oocysts, typically fewer than ten (Pringle 1966; Billingsley, Medley et al. 1994). However, the genetic control of refractoriness appears to be complex. Not only does it involve several quantitative trait loci (QTL), but the relative contribution of each QTL varies with the parasite species (Zheng, Cornel et al. 1997; Zheng, Wang et al. 2003). However, these aspects make it quite clear that the mosquito-parasite interactions are a promising target for malaria control efforts.

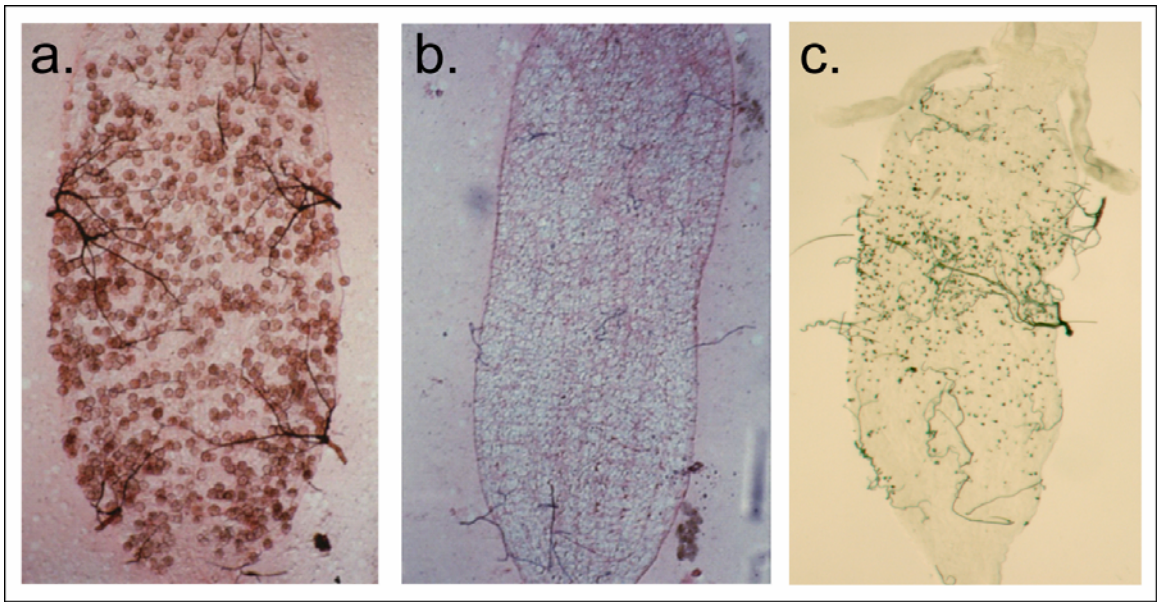


Fig. 1.4.2 Modes of parasite killing in the mosquito midgut. **a.** Infected wild-type midgut **b.** Lysis of the *Plasmodium* parasite in a infected midgut of the SUAF2 strain (Vernick, Fujioka et al. 1995) **c.** Melanized parasites in the midgut of the L3-5 strain (Collins, Sakai et al. 1986)

Mosquito midgut invasion.

As shown in Fig. 1.4.1, parasite losses in the mosquito are documented at three decisive developmental transitional stages: the gamete-to-ookinete, the ookinete-to-oocyst

and the midgut sporozoite-to-salivary gland sporozoite transitions (Sinden 1999; Sinden 2002; Alavi, Arai et al. 2003). However, the most serious reduction occurs at the ookinete-to-oocyst transition at which point the parasite numbers often drop to single digits, necessitating the key amplification step that follows (Fig. 1.4.1; Fig. 1.4.3).

Different routes of midgut invasion have been reported for various mosquito–parasite species combinations, *P. berghei* ookinetes penetrate the midgut epithelium of *A. gambiae* and *A. stephensi* using a combination of inter and intracellular routes (Han, Thompson et al. 2000; Vlachou, Zimmermann et al. 2004) whereas *P. falciparum* has been reported to invade *A. stephensi* mosquito midguts solely via the intracellular route (Baton and Ranford-Cartwright 2004; Baton and Ranford-Cartwright 2005).

Recent work has shown that many ookinetes are eliminated inside the midgut epithelium through the action of mosquito immune factors such as TEP1 (Thioester containing protein 1) (Blandin, Shiao et al. 2004) and LRIM1 (Leucine-rich repeat protein 1) (Osta, Christophides et al. 2004). TEP1, which is produced by hemocytes and secreted in the hemolymph, binds to the surface of the parasites mediating their lysis. LRIM1 is thought to act in the same fashion, however, CTL4 (and to a lesser extend CTLMA2) protect the remaining parasites from the action of LRIM1, assuring their development to oocysts that some days later will release sporozoites in the mosquito hemolymph (Fig. 1.4.3a). However, in the absence of CTL4 by genetic knockdown, LRIM1 promotes killing of the remaining parasites, which are subsequently melanized in a reaction thought to be mediated by phenoloxidases (Fig. 1.4.3b).

Thus, it is reasonably established now, that the mosquito immune reactions are indeed in part responsible for this decrease (Blandin, Shiao et al. 2004; Meister, Koutsos et al. 2004;

Osta, Christophides et al. 2004; Meister, Kanzok et al. 2005). However, how some parasites manage to escape these immune reactions remains vaguely elusive and is a central questions among the research community.

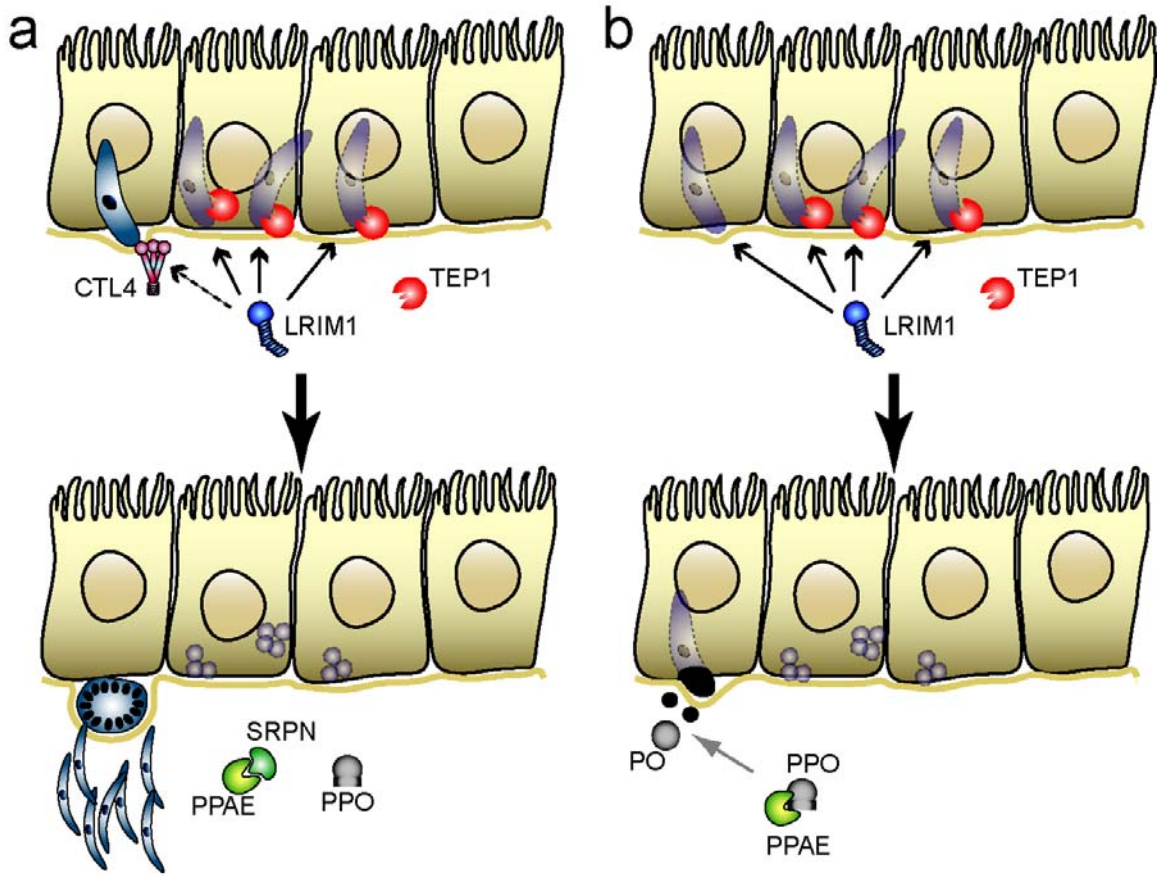


Fig. 1.4.3 *Anopheles-Plasmodium* interactions in the mosquito midgut **(a)** TEP1 and LRIM1 eliminates many ookinetes inside the midgut epithelium. CTL4 protects the remaining parasites from the action of LRIM1. **(b)** After CTL4 knockdown by RNAi, LRIM1 promotes killing of the remaining parasites, which are subsequently melanized (from (Meister, Koutsos et al. 2004)).

1.5 INSECT INNATE IMMUNITY

Adaptive vs. innate immunity.

Innate Immunity has developed as a first line of efficient defense against microbial invaders and as such it is common to all metazoans. Characteristic of it is the detection of surface patterns of microorganisms by "pattern recognition receptors" (PRR) (Janeway 1989) that have conserved the memory of motives (or patterns) associated with the microorganisms such as i.e. peptidoglycan, lipopolysaccharide, lipoteichoic acids, lipoproteins, bacterial CpG DNA, and flagellin. These receptors rapidly activate signaling pathways and effector mechanisms such as phagocytosis, proteolytic cascades and of course the synthesis of antimicrobial peptides (AMPs) that quickly limit the infection (Medzhitov and Janeway 2000; Janeway and Medzhitov 2002).

Adaptive immunity was acquired later in evolution (it is shared by ~45,000 vertebrate species) by the ancestors of cartilaginous fish and introduced the advantage of individual immune memory. The adaptive immunity is able to create an almost limitless variety of cells and molecules to recognize and eliminate an equally vast variety of foreign invaders. It achieves this feat by: (1) somatic gene rearrangement to create a large repertoire of receptors, (2) the clonal multiplication of antigen-specific effector cells targeting the invading pathogen and (3) the generation of memory cells to prevent re-infection.

To sum it up: adaptive immunity is built on innate immunity as its foundation. It depends on innate immunity for its activation, and consequently it is possible for the majority of species on earth to do just fine without adaptive immunity. However, a hypothetical species with adaptive immunity but no innate immunity seems rather unlikely.

Innate immunity.

Drosophila has served as the model system for innate immune responses in insects for many years (Hoffmann 2003). It resists challenges from various microorganisms remarkably well, relying solely on its innate immunity. For this reason the basics of insect innate immune systems shall first be discussed drawing from our knowledge of *Drosophila*.

The insects' first line of defense against infections is the structural barriers of the body, mainly the hardened exoskeleton, but also the chitinous trachea and the peritrophic matrix of the midgut. To protect insect interior tissues, breaches are quickly sealed by coagulation (Theopold, Li et al. 2002) and melanization reactions (Soderhall and Cerenius 1998).

Should the pathogens however manage to penetrate these structural barriers into the open circulatory system in which the organs and tissues are bathed in the hemolymph, they encounter the insects' immune reaction by the fat body and hemocyte cells. There are constitutively present molecules in the hemolymph(prophenoloxidase (PPO), coagulation factors, pattern recognition molecules and opsonins) which sometimes require post-translational activation like proteolytic cleavage.

Upon binding to pathogens, the recognition proteins (or opsonins) immediately trigger mainly two types of innate immune responses: (1) cellular responses: phagocytosis and encapsulation, and (2) humoral responses: secretion of AMPs, melanization and coagulation. Depending on the intensity of infection, these responses may be sufficient to clear the microorganisms from the insects' hemolymph within a few minutes.

Humoral immunity: AMPs.

As mentioned before, the binding of PRRs to pathogen-specific molecules activates signaling cascades that, within hours after infection, cause the transcriptional activation of hundreds of immune-inducible molecules, including AMPs (Dimopoulos, Christophides et al. 2002). AMPs represent the best-characterized insect humoral reaction and are mostly produced by the fat body and the hemocytes, although various barrier epithelia, such as tracheae, anterior midgut, the genital tract and the Malpighian tubules are also capable of producing these peptides (Lehane, Wu et al. 1997; Richman, Dimopoulos et al. 1997; Tzou, Ohresser et al. 2000). They are typically small, cationic and structurally diverse peptides (Bulet, Hetru et al. 1999) and can reach micromolar concentrations in the insect hemolymph (Lowenberger 2001).

The effector mechanisms vary between different AMP families; however, in general, AMPs are believed to be attracted to the negatively charged surface of microbes killing them by damaging the structure of their cytoplasmic membrane, for example by permeabilization or by forming voltage-dependent channels (Bulet, Hetru et al. 1999; Shai 2002). Once induced, AMPs can persist in the hemolymph for at least 3 weeks, and may provide protection against reoccurring infections for the rest of the insect life.

Eight distinct classes of AMPs have been discovered mainly in the fruit fly *D. melanogaster* by biochemical analysis (Imler and Bulet 2005). They can be grouped into three families based on their main biological targets in *Drosophila*:

1. gram-positive (Gram+) bacteria

Defensin (Dimarcq, Hoffmann et al. 1994)

2. gram-negative (Gram-) bacteria

Cecropins (Kylsten, Samakovlis et al. 1990)

Drosocin (Bulet, Dimarcq et al. 1993)

Attacins (Asling, Dushay et al. 1995)

Diptericin (Wicker, Reichhart et al. 1990)

MPAC/AttacinC (Imler and Bulet 2005)

3. fungi

Drosomycin (Fehlbaum, Bulet et al. 1994)

Metchnikowin (Levashina, Ohresser et al. 1995)

Gambicin (Vizioli, Bulet et al. 2001) of *A. gambiae* is active against both Gram+ and Gram- bacteria.

Cellular immunity: phagocytosis, encapsulation & melanization.

Hemocytes are the major cell type in the hemolymph responsible for phagocytosis (Hillyer, Schmidt et al. 2003) and encapsulation (Meister and Lagueux 2003). In phagocytosis, specialized blood cells recognize, internalize and destroy microbial invaders (Aderem and Underhill 1999). The process is paramount for clearing of invading bacteria in the early stages of infections.

Opsonic ligands have to bind to the microorganism in order for them to be recognized by phagocytic cells. One of the receptors involved in the phagocytosis of Gram- bacteria in *Drosophila* is the PRR and Imd pathway mediator PGRP-LC, pointing to a link between humoral and cellular immune reactions (Ramet, Manfruelli et al. 2002). In *Anopheles*, a member of the thioester containing protein family, TEP1, acts like an opsonizing

complement factor by binding to bacteria and promoting their phagocytosis (Levashina, Moita et al. 2001).

Bigger objects that cannot be phagocytosed such as eggs of parasitoid wasps in *Drosophila* are encapsulated by the hemocyte cells. In this reaction, the hemocytes attach to the surface and extend long lamellae to cover and inactivate the object. They thus cause death by asphyxiation and/or attack with free radicals (Nappi and Vass 1998). The next step often is the activation of the PPO cascade to crosslink and thus neutralize everything in a melanin capsule in a reaction called melanotic encapsulation (Fig 1.5.1).

In *Drosophila*, there have been three hemocyte types described: plasmatocytes, lamellocytes and crystal cells (Meister and Lagueux 2003) They are mostly attached to tissues, such as muscle, tracheae, fat body, midgut and Malpighian tubules (Hillyer and Christensen 2002). Plasmatocytes are performing the phagocytosis (Aderem and Underhill 1999) and represent the main population of hemocytes. Encapsulation is carried out by lamellocytes (Nappi and Vass 1998), and finally, PPO-producing crystal cells are thought to mediate melanization of encapsulated bodies (Meister and Lagueux 2003).

Although we are starting to understand hematopoiesis and blood cell differentiation in *Drosophila* (Evans, Hartenstein et al. 2003), attempts to establish a unified system to classify hemocytes in other insects have not been feasible to date (Brehelin, Zachary et al. 1978; Hillyer and Christensen 2002).

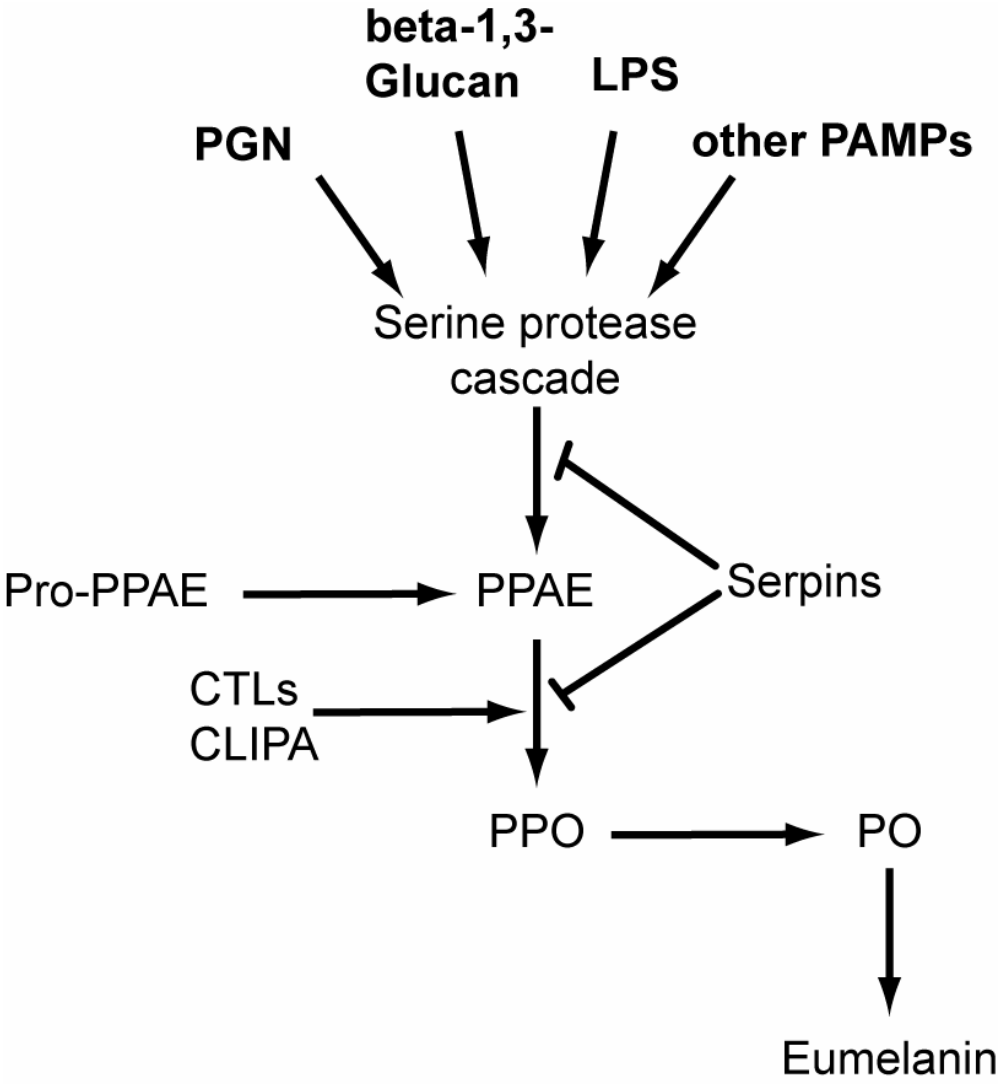


Fig. 1.5.1 Scheme of the signaling cascade leading to the activation of melanization in insects. CLIPA, clip domain serine protease A; CTL C-type Lectin; LPS, Lipopolysaccharide; PAMP, pathogen associated molecular pattern; PGN, Peptidoglycan; PPAE, PPO activating enzyme ; PPO, Prophenoloxidase; PO, Phenoloxidase.

A. gambiae*: a new model for studying innate immunity*.

Although *Drosophila* serves as the model system for innate immune response study in insects (Hoffmann 2003), results from *Drosophila* cannot necessarily be expected to also be true for *Anopheles*, as the two dipterans diverged approximately 250 million years ago (Wiegmann, Yeates et al. 2003) and have adopted quite different lifestyles. The most

important differences are hematophagy and the aquatic developmental stages in *Anopheles*. However, the use of *Anopheles* as a model system for innate immunity suffered from the complexity of rearing, and thus the lack of robust genetics, genetic manipulation techniques and biochemistry. Taken together this permitted only a limited view of the immune reactions in this important vector of human malaria.

However, advances in *Anopheles* genomics and functional genomics during the last four years have led to a rapid development in the field. Genome sequencing (Holt, Subramanian et al. 2002) allowed for a large-scale comparative genomic analysis of gene families involved in the *Drosophila* immune system (Christophides, Zdobnov et al. 2002). This analysis revealed significant divergences between the *Drosophila* and *Anopheles* genomes highlighted by a lack of orthologs and an excess of gene family expansions in the recognition and effector gene modules. This may reflect the adaptation to specific environmental requirements imposed by the different lifestyles of the respective insect species. The signal transduction gene families, however, are remarkably conserved (Christophides, Vlachou et al. 2004), suggesting that successful defense strategies remain conserved during the evolution of the two insects. In other words, the recognition and response arsenal of the immune system had to be adjusted and expanded to deal with the different encountered immune challenges in the variable ecological niches.

Defense against the parasite.

A growing number of studies have shown that the malaria parasite causes transcriptional upregulation of immunity genes during its various developmental transitions in the mosquito. This suggests that the innate immune system may account for large parts of

the documented parasite losses (Fig. 1.4.1) (Richman, Bulet et al. 1996; Dimopoulos, Richman et al. 1997; Dimopoulos, Seeley et al. 1998; Dimopoulos, Christophides et al. 2002; Tahar, Boudin et al. 2002; Christophides, Vlachou et al. 2004).

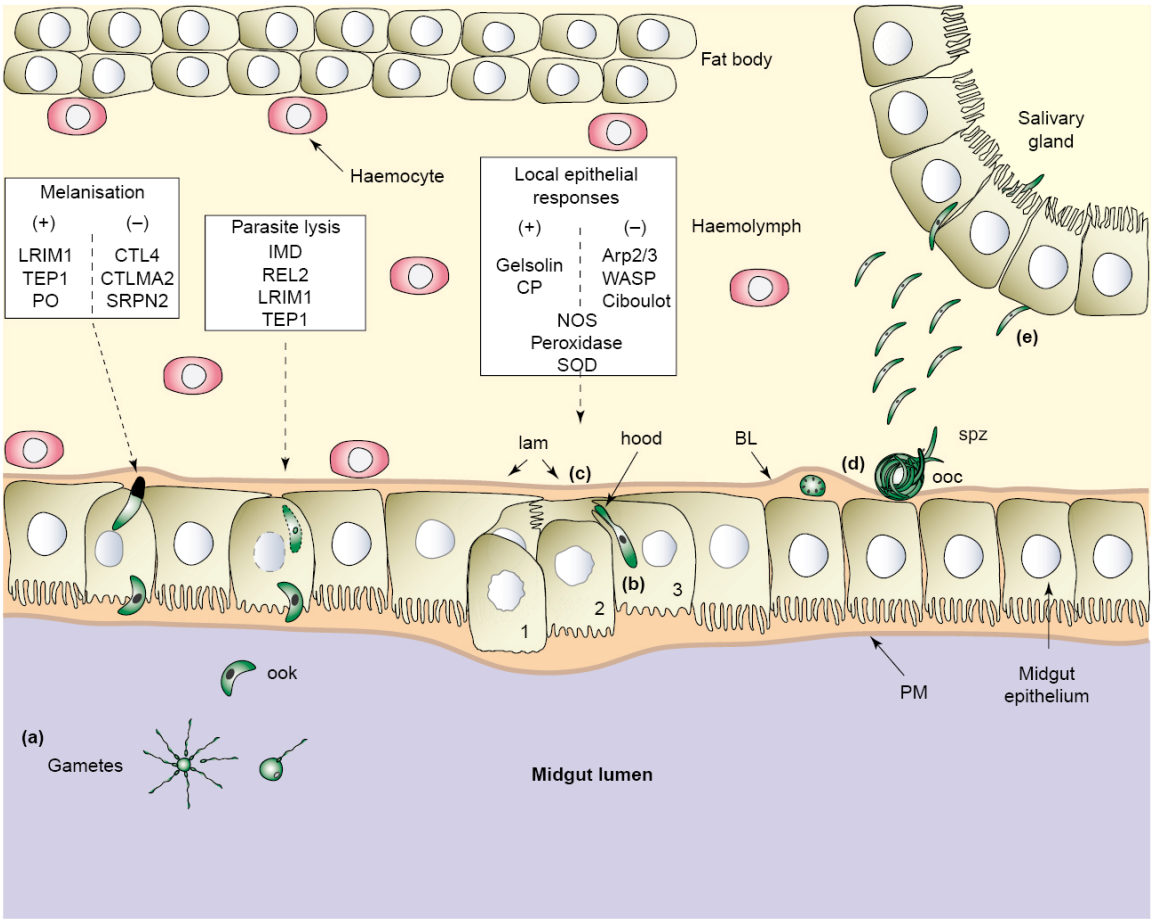


Fig. 1.5.2 *Anopheles* immune reaction against the *Plasmodium* parasite during its sexual developmental cycle in the mosquito midgut and hemolymph. Mosquito molecules acting in a positive, protagonistic (+) or negative, antagonistic (-) way towards the parasite are listed in boxes. The reactions facilitated by them include parasite melanization, lysis and local epithelial responses. (a) Following gamete development and fertilization, (b) Plasmodium ookinetes (ook) traverse the cytoplasm of several midgut cells (1, 2 and 3) before emerging basolaterally underneath the basal lamina (BL). (d) There, the parasite develops into an oocyst (ooc) to produce thousands of sporozoites (spz). (e) Upon rupture of the oocyst, the sporozoites are released and then invade the salivary gland. CP, capping protein; IMD, immune deficiency gene; PM, peritrophic matrix; PO, phenoloxidases; REL2, Relish 2; SOD, superoxide dismutase (from (Vlachou and Kafatos 2005)).

There are several mosquito immune reactions directed against the parasite (Fig. 1.5.2) – the best example probably being the *Plasmodium* ookinete melanization in the genetically

selected *A. gambiae* strain L3-5 (Collins, Sakai et al. 1986). The fact that this reaction is dependent on the mosquito-parasite combination (*P. cynomolgi*, *P. berghei* and allopatric strains of *P. falciparum* are melanized, but sympatric *P. falciparum* populations are not (Collins, Sakai et al. 1986; Paskewitz, Brown et al. 1988)) points to the possibility that parasite-specific recognition is involved and that the parasite itself has some mechanism of immune evasion.

The L3-5 strain very effectively melanizes injected Sephadex beads (Paskewitz and Riehle 1994) which indicates that melanization is not necessarily coupled or dependent on parasite recognition and killing. Another mechanism of parasite killing that has been reported for the SUAF2 strain is lysis in the cytoplasm of the midgut epithelial cells (Vernick, Fujioka et al. 1995).

More recent studies have focused on the role of innate immunity genes in parasite melanization and lysis in the mosquito midgut. Knockdown (KD) of the hemocyte-specific complement-like protein TEP1 by RNAi was found to greatly increase the number of developing oocysts as TEP1 binds to *P. berghei* ookinetes and mediates their lysis. On the other hand, silencing of *TEP1* in the L3-5 strain inhibits melanization (while also increasing the number of developing oocysts) (Blandin, Shiao et al. 2004). TEP1 was previously shown to be one of the factors involved in the phagocytosis of bacteria (Levashina, Moita et al. 2001).

Another recently identified factor in parasite killing in the mosquito midgut is LRIM1. When silenced, the oocyst numbers are substantially increased, similar to the *TEP1* KD (Osta, Christophides et al. 2004). *LRIM1* has been shown to be upregulated upon bacterial challenge (Dimopoulos, Christophides et al. 2002) and *Plasmodium* infection (Osta,

Christophides et al. 2004) but no *LRIM1* orthologs have been identified in other species to date.

Neither TEP1 nor LRIM1 are sufficient to halt the parasites progress, reinforcing the idea that parasite survival in susceptible mosquitoes may be due to specific immune evasion on the parasites part. In fact, two C-type lectins (CTLs) – *CTL4* and *CTLMA2* – seem to be recruited by *P. berghei* to protect it from *LRIM1*-mediated killing. Silencing of either of these CTLs leads to killing and melanization of almost all *P. berghei* ookinetes in the mosquito midgut (Osta, Christophides et al. 2004).

Other genes implicated in the mosquito cellular epithelial reaction to parasite invasion are serine proteases that have been shown to be transcriptionally activated and involved in ookinete killing and melanization (Dimopoulos, Christophides et al. 2002; Volz, Osta et al. 2005). The serine protease inhibitor, SRPN10, is highly and specifically induced in *Anopheles* midgut cells upon parasite invasion (Danielli, Kafatos et al. 2003) and can serve as an excellent marker of this event. At least two other serpins SRPN2 (Michel, Budd et al. 2005) and SRPN6 (Abraham, Pinto et al. 2005), are negative factors of ookinete killing and melanization.

Transcriptomic microarray analysis of the local epithelial response as revealed that as much as 7% of the surveyed mosquito transcriptome may be regulated upon parasite invasion – this includes genes involved in cytoskeletal remodeling, apoptosis, immune responses, the redox state, cell adhesion and the extracellular matrix (Vlachou, Schlegelmilch et al. 2005). An interesting gene (*RFABG*) discovered in this study encodes lipophorin a lipid transport vehicle that is a positive factor of parasite oocyst and mosquito egg development.

In summary, the picture emerging is that the development and advancement of the parasite depends on positive (e.g. CTL4, CTLMA2) and negative (e.g. TEP1, LRIM1) mosquito factors. This is consistent with the notion of parasite melanization in the L3-5 strain being complex and depending on multiple quantitative trait loci (Zheng, Cornel et al. 1997). Suspiciously, the *TEP1* gene sequence, which is in the *Pen2* region, displays high variability between refractory and susceptible mosquitoes (Blandin, Shiao et al. 2004); and the *Pen1* region shows clusters of extensive sequence polymorphisms that may relate to the refractory phenotype (Thomasova, Ton et al. 2002). Also, strong physiological differences have been shown to exist between the refractory L3-5 and the susceptible G3 strain, most notably an elevated level of reactive oxygen species contributing to parasite melanotic encapsulation in the L3-5 strain (Kumar, Christophides et al. 2003).

1.6 IMMUNE SIGNALING PATHWAYS IN *DROSOPHILA*

Studies in *D. melanogaster* studies have shown that the production of AMPs is the result of the activation of two distinct pathways: the *Toll* and the *Imd* pathway (Hoffmann 2003).

The Toll *pathway*.

Toll was originally discovered as being involved in the dorsoventral pattern formation during embryonic *Drosophila* development (Anderson, Bokla et al. 1985); it soon became evident that it was also involved in innate immune responses against fungi (Lemaitre, Nicolas et al. 1996) and Gram+ bacteria (Fig. 1.6.1).

The receptor responsible for recognition of fungi is reported to be the *GNBP3* receptor, whose null mutation *hades* is sensitive to fungal infections (Ferrandon, Imler et al. 2004). Downstream signaling of the recognition event proceeds through the Persephone protease (Psh) (Ligoxygakis, Pelte et al. 2002), which is negatively regulated by the blood serpin Necrotic (Nec) (Levashina, Langley et al. 1999), to cleave the ligand and activator of the Toll receptor, the cytokine-like polypeptide Spaetzle (Spz) (Lemaitre, Nicolas et al. 1996; Weber, Tauszig-Delamasure et al. 2003).

The activation of the other *Toll* signaling branch by Gram+ bacteria (Fig. 1.6.1) involves two extracellular proteins thought to be brought into a complex by PAMP recognition. The first is the peptidoglycan recognition protein, PGRP-SA (Michel, Reichhart et al. 2001) and the second is GNBP1 (Gram-negative binding protein 1) (Gobert, Gottar et al. 2003). This process ultimately leads to proteolytic cleavage of Spz and activation of Toll.

At least three cytoplasmic proteins, MyD88, Tube and Pelle are recruited by the intracellular part of Toll – potentially the intracytoplasmic so-called TIR (Toll/IL-2 receptor) homology domain. Pelle is a serine-threonine kinase believed to play an indirect role in the phosphorylation and subsequent proteolytic degradation of Cactus, a member of the IκB family of proteins. The degradation of Cactus results in the nuclear translocation of NF-κB transcription factors, Dorsal and Dif, which are bound to intact Cactus and thus retained in the cytosol (Belvin and Anderson 1996).

Whereas Dorsal is essential for developmental processes, Dif is mostly implicated in the transcription of AMPs and other defense genes through specific binding to cis-acting elements (κB motifs) found in the promoter sequences of these genes.

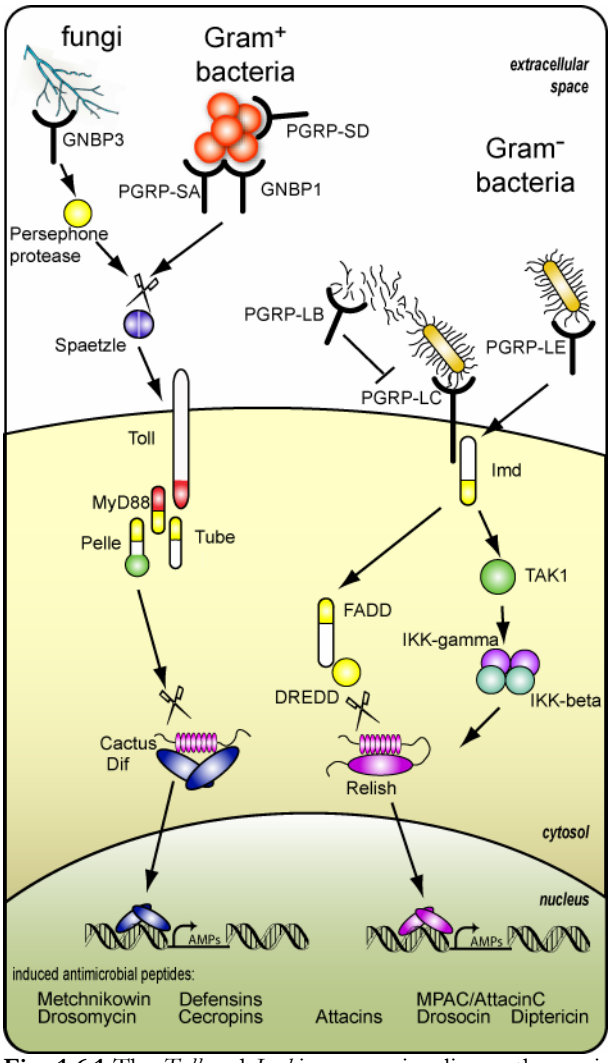


Fig. 1.6.1 The *Toll* and *Imd* immune signaling pathways in *D. melanogaster* (modified from (Meister, Koutsos et al. 2004))

The Imd *pathway.*

Gram- bacterial infections are predominantly dealt with by activation of the *Imd* (immune deficiency) pathway, named after its first characterized mutation (Lemaitre, Kromer-Metzger et al. 1995; Lemaitre, Nicolas et al. 1996). Two PRRs have been shown so far to trigger responses of the Imd pathway: the membrane spanning PGRP-LC (Choe,

Werner et al. 2002; Gottar, Gobert et al. 2002) and the extracellular PGRP-LE (Takehana, Katsuyama et al. 2002; Takehana, Yano et al. 2004). The intermediate steps of signal transduction (if any) between PGRP-LE and the Death domain carrying Imd are unknown. PGRP-LC and Imd have recently been shown to interact directly (Choe, Lee et al. 2005). Downstream of Imd the signal is transduced by two paths: one involves the *Drosophila* homologs of the mammalian IκB kinase complex (IKK) (Silverman, Zhou et al. 2000) and the other the caspase *Dredd* (Stoven, Ando et al. 2000). Downstream, they lead to the proteolytic cleavage and activation of Relish (Dushay, Asling et al. 1996), a third member of the NF-κB family of transcription factors.

Full length Relish contains an amino-terminal DNA binding domain and a carboxyterminal IκB domain, which acts similarly to the *Toll* pathway inhibitor Cactus, preventing the nuclear translocation of the transcription factor domain when the pathway is inactive (Dushay, Asling et al. 1996). After proteolytic removal of the IκB domain, Relish translocates into the nucleus and induces transcription of AMPs.

The elegant simplicity of the two different pathways responding either to Gram- or Gram+ and fungal infections is slightly misleading, however. Recent studies have highlighted that this initial clear separation of the two pathways, is to a large extent an oversimplification of the reality of complex feedback and crosstalk interactions between pathways (Hoffmann 2003; Hultmark 2003).

1.7 THE FAMILY OF PGRP RECEPTORS

Peptidoglycan structure.

Peptidoglycan (PGN) is a cell wall component of both Gram+ (90%) and Gram- (10%) bacterial cell walls. While all bacteria have PGN, the amount, location, and specific composition may vary. For example, PGN is associated with lipoteichoic acid in a thick exposed layer around the bacterial cell wall of Gram+ bacteria, whereas in Gram- bacteria it is only in a thin layer overlaid by a much thicker layer of LPS.

PGN is composed of parallel glycan strands and cross-linking peptide moieties. The glycan strands consist of two alternating sugar residues: β-1,4-linked N-Acetylglucosamine (NAG) and N-Acetyl Muramic Acid (NAM) (Fig 1.7.1). NAM is found only in bacteria and blue-green algae.

The peptide moiety consists of only 3 to 6 different, alternating L- and D-amino acids, where the D configuration is typical of PGN. Position 3 in all Gram- bacteria and in the Gram+ *bacilli* (genus *Bacillus* and *Clostridium*) is occupied by m-Diaminopimelic Acid (m-Dap) (Fig 1.7.1, right). Most other Gram-positive bacteria (e.g. Gram+ *cocci*) have L-Lysine in this position (Fig. 1.7.1, left). The Dap-type PGN peptides are usually directly cross-linked, whereas the Lys-type PGN peptides are usually linked through an "interpeptide bridge" that varies in length and amino acid composition in different bacteria (Schleifer and Kandler 1972). PGN has been classified into two Groups, six Subgroups and 16 Variations by (Schleifer and Kandler 1972).

PGN can be cleaved by several enzymes of bacterial and animal heritage. Chicken type Lysozyme has muramidase activity (N-acetylmunramide glycanohydrolase (EC 3.2.1.17)) – it

cleaves the glycosidic link between NAG and NAM in the polysaccharide chain (Chipman and Sharon 1969) (Fig. 1.7.1, blue arrows). Enzymes with N-acetylmuramoyl-L-alanine amidase activity (NAMLAA–activity (EC3.5.1.28)) are enzymes specifically cleaving the lactylamide bond between Muramic Acid and the peptide chain (Fig 1.7.1, red arrows). Examples of peptides with NAMLAA activity are the zinc dependent bacteriophage T7 Lysozyme (Cheng, Zhang et al. 1994; Dziarski 2004), and PGRP-SC1B of *D. melanogaster* the (Mellroth, Karlsson et al. 2003). Mammalian proteins with NAMLAA activity are the mouse (Gelius, Persson et al. 2003) and human PGRP-L (Gelius, Persson et al. 2003; Wang, Li et al. 2003) that are found in different tissues and body fluids (Valinger, Ladesic et al. 1982; De Pauw, Neyt et al. 1995; Hoijer, Melief et al. 1996).

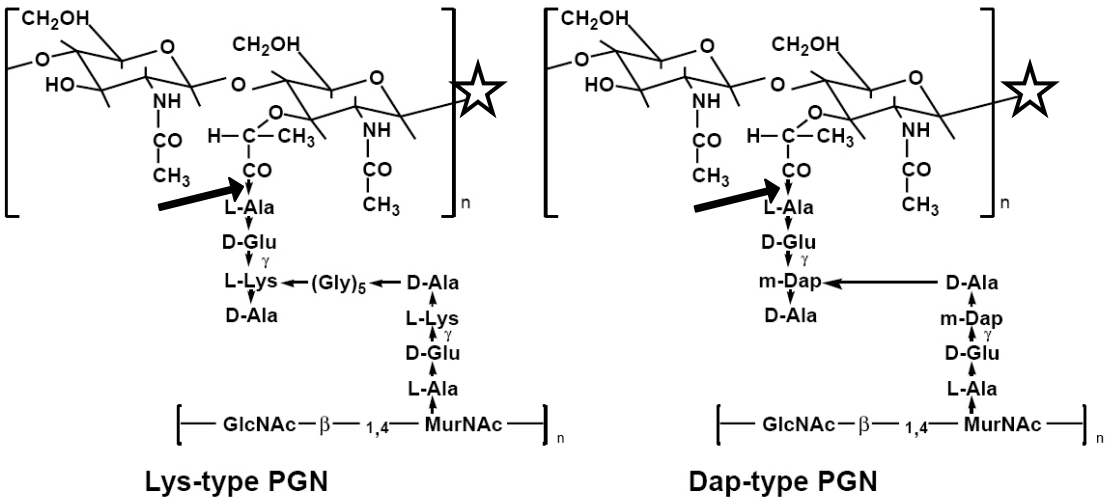


Fig. 1.7.1. Structure of Lys-type and Dap-type peptidoglycan (adapted from: (Dziarski 2004)). Amide bonds cleaved by NAMLAA enzymatic activity are indicated by arrows and glycosidic links cleaved by lysozyme muramidase activity are marked with a star.

It has been known for a long time that PGN promotes an inflammatory response. Initially was derived from intraperitoneal injection of PGN from Gram+ bacteria that

promoted inflammation. In cell culture models, PGN has been demonstrated to stimulate the production of inflammatory cytokines in monocytes, macrophages, neutrophils, and epithelial cells. Later, muramyl dipeptide (MDP), a subcomponent of PGN, was found to be the minimal chemical structure required for adjuvant activity (reviewed in (Stewart-Tull 1980)).

PGRP*s in* Drosophila *and other insects*.

The first PGRP was purified from the hemolymph of the silkworm *Bombyx mori* (Yoshida, Kinoshita et al. 1996). It was a 19 kDa protein that bound Gram+ bacteria and PGN and activated the prophenoloxidase (PPO) cascade. A few years later it was cloned (Ochiai and Ashida 1999) as well as an ortholog from the moth *Trichoplusia ni* (Kang, Liu et al. 1998). The discovery of mammalian PGRPs showed that these proteins are highly conserved from insects to mammals.

Sequencing of the fruit fly *D. melanogaster* genome facilitated the identification of a family of 13 highly diversified PGRP homologs transcribed to at least 17 PGRP proteins (Werner, Liu et al. 2000). In general, members of this PGRP gene family share one or more well-conserved PGRP domains and can be classified as short (S), which encode secreted proteins, and as long (L), which encode transmembrane and intracellular protein products. As mentioned previously, one of the short, secreted *Drosophila* PGRPs, PGRP-SA, is essential for activation of the Toll/Dif signaling pathway in response to Gram+ bacteria, such as *M. luteus*, but not fungi (Michel, Reichhart et al. 2001). A recent report also places the soluble PGRP-SD upstream of the Toll receptor in the Gram+ bacteria recognition pathway (Bischoff, Vignal et al. 2004). The PGRP-LC receptor, on the other hand, acts through the

alternative IMD/Relish pathway in response to Gram- bacteria (Choe, Werner et al. 2002; Gottar, Gobert et al. 2002).

PGRPs have been found in a wide range of organisms: silkworm (*Bombyx mori* (Yoshida, Kinoshita et al. 1996)), fruit fly (*D. melanogaster* (Werner, Liu et al. 2000)), honey bee (*Apis mellifera* (Evans 2004)), human (*Homo sapiens* (Kang, Liu et al. 1998; Liu, Xu et al. 2001)) and mouse (*Mus musculus* (Kang, Liu et al. 1998; Kiselev, Kustikova et al. 1998)). The PGRP domains share a conserved cysteine motif important for protein function. Six Cysteines, conserved in all mammalian PGRPs, have been found in bovine PGRP to form a disulfide motif of 1-6, 2-5 and 3-4 pairings (Tydell, Yount et al. 2002). This is consistent with the pattern in the moth PGRP, which has two disulfide bridges corresponding to 1-6 and 2-5 pairings (Ochiai and Ashida 1999). Importantly, in the *Drosophila* PGRP-SA loss-of-function mutation, *semmelweis*, which abolishes anti-Gram+ bacterial responses, Cys-80 (corresponding to the conserved cysteine 4) is changed to Tyr, thus disrupting the 3-4 highly conserved disulfide bond (Michel, Reichhart et al. 2001). In the rest of the *Drosophila* PGRPs, Cysteines 3 and 4 are present in all proteins with the sole exception of PGRP-LE. Notably PGRP-SA and all known Lepidopteran PGRPs have residues permitting a 1-6 cysteine pairing, similar to their mammalian homologs.

Other important amino acid residues in conserved relative locations in the PGRP domains are a threonine and a tyrosine that are important for amidase activity and two histidines and a cysteine that coordinate the catalytic zinc ion (Reiser, Teyton et al. 2004). Five PGRPs in *Drosophila* have all these requirements: *PGRP-SC1A, -SC2, -SB1, -SB2* and *–LB*. PGRP-LB has been also shown experimentally to possess lytic activity against sensitized *Escherichia coli* (Kim, Byun et al. 2003) and more recently to specifically degrade Gram-

bacterial PGN to down regulate the Imd pathway (Zaidman-Remy, Herve et al. 2006). PGRP-SC1 and PGRP-SC2 have been shown to have a similar function in larvae (Bischoff, Vignal et al. 2006) .

The rest of the *Drosophila* PGRPs (*PGRP-SA, -SD, -LA, -LC, -LD, -LE*) and the short human and mouse PGRP-Ss lack the zinc-binding residues, but are likely to have retained the ability to bind PGNs, as shown for short mammalians PGRPs (Liu, Xu et al. 2001) and *Drosophila*'s PGRP-SA (Werner, Liu et al. 2000)

PGRP-LC *in* D. melanogaster.

The long, transmembrane PGRP-LC receptor signaling responds to Gram- bacteria to induce production of certain antimicrobial peptides through the IMD/Relish pathway (Choe, Werner et al. 2002; Gottar, Gobert et al. 2002). A genome-wide RNAi screen in *Drosophila* cells in culture also pointed to PGRP-LC as a key player for phagocytosis of Gram- but not Gram+ bacteria (Ramet, Manfruelli et al. 2002).

The three different PGRP domains (LCa, LCx, and LCy) of the *PGRP-LC* gene are spliced to a common invariant part of the gene to produce at least three transcript and protein isoforms. The 21 to 24 kDa PGRP domains are located extracellularly, and the invariant domain is intracellular (Werner, Borge-Renberg et al. 2003). It was recently shown that the intracellular domain, though not having any significant homology to other known domains, is mediating contact with the Imd protein (Choe, Lee et al. 2005). Only the PGRP-LCx variant is absolutely required for activation of the IMD/Relish pathway (Choe, Werner et al. 2002; Werner, Borge-Renberg et al. 2003; Kaneko, Goldman et al. 2004) . The PGRP-LCa form is, in addition to LCx, involved in the responses to small monomeric

peptidoglycan subunits as shown in studies using RNAi (Kaneko, Goldman et al. 2004), possibly by forming heterodimers with the LCx form. The role of the third splice form, LCy, is still elusive because RNAi knockdown of its expression does not affect the IMD/Relish signaling upon challenge with PGN (Werner, Borge-Renberg et al. 2003; Kaneko, Goldman et al. 2004).

AIMS OF THE THESIS

Most of our knowledge about innate immunity pathways to date was gathered in *Drosophila melanogaster*, where invading pathogens activate at least two distinct signaling pathways: the *Toll* and the *Imd* pathway. The *Toll* pathway is activated by Gram+ (Gram-positive) and fungal infections whereas the *Imd* pathway responds to Gram- (Gram-negative) bacteria. The infectious non-self agents are recognized by PRRs interacting with cell wall components. Peptidoglycan Recognition Receptor Proteins (PGRPs) have been placed upstream of both the *Toll* and the *Imd* pathway.

The first objective of this thesis was to use a comparative genomics approach to investigate differences and similarities in the *Toll* and *Imd* innate immune signaling pathways between *D. melanogaster* and *Anopheles gambiae,* the main vector of malaria in sub-Saharan Africa.

The second objective was to then analyze the role of the PGRP family of pattern recognition receptors at the top of these two pathways, and to scrutinize the role of *PGRPLC* in particular. To these ends, the function of the receptor gene family in resistance to bacterial infections and in defense to malaria was to be examined by RNAi in cultured cells and whole mosquitoes.

The third objective was to uncover the tasks of the transcriptional regulators of the immune pathways, REL1 and REL2. The effector target genes of these NF-κB transcription factors and their roles in bacterial and malarial defense were to be uncovered by means of RNAi, survival assays and microarrays.

MATERIALS & METHODS

3.1 MOSQUITO / PARASITE

A. gambiae *rearing*.

The susceptible strains 4a r/r, G3 and the refractory strain L3-5 (Collins, Sakai et al. 1986), were reared at 27°C at a relative humidity of 75%, with a 12/12h light/darkness cycle. Larvae were raised in 0,1% marine salt water and fed with ground cat food. Adults were fed ad libitum on wet cotton soaked in a 15% sucrose solution. For the production of eggs, female mosquitoes were allowed to take a blood meal on anaesthetized Balb/c or CD1 mice. This method was described in (Richman, Bulet et al. 1996).

***Gene silencing in adult* A. gambiae.**

dsRNA was produced as described in this chapter. A nano-injector (Nanoject, Drummond) was used to introduce 69nl of dsRNAs (207ng at a concentration of 3 mg/ml) in water in the thorax of one to two-day old CO2-anesthetized adult female mosquitoes, which were then allowed to recover for four days. Only after this recovery time they were subjected to subsequent assays. This method was first described in (Blandin, Moita et al. 2002).

Bacterial infection & survival of the mosquito.

GFP-expressing *E. coli* OP-50 was a gift from J.J. Ewbank (INSERM, Marseille-Luminy, France). *Bacillus subtilis, M. luteus* and *S. aureus* were kind gifts from Philippe Bulet (IBMC, Strasbourg, France). Bacteria were cultured to OD_{600} = 0.7, pelleted, washed and resuspended in phosphate-buffered saline (PBS) to final concentrations (*E. coli* at

OD_{600}=0.01, *S. aureus* at OD_{600}=0.4). 69 nl of the bacterial suspension or PBS for controls was injected into the mosquito thorax with a nano-injector (Nanoject, Drummond) while the mosquitoes were anesthetized with CO_2 and allowed to recover. Dead mosquitoes were counted daily and removed over a period of 7 days. The results generally are representative of at least three independent experiments, each carried out with around 50 mosquitoes per tested group.

A. gambiae *infection with the malaria parasite.*

For malaria infections, mosquitoes were fed for 30 min at 21°C on CD1 mice infected with transgenic GFP-fluorescing *P. berghei* strains expressing GFP at the ookinete and early oocyst stages (Vlachou, Zimmermann et al. 2004) and all developmental stages (sexual and asexual) (Franke-Fayard, Trueman et al. 2004). After 5-7 days, midguts of the mosquitoes were dissected, fixed for 30min in 4% formaldehyde in 1xPBS, washed 3 times with 1xPBS and mounted on glass slides in Vectashield mounting medium. The number of GFP fluorescing oocysts and black melanized ookinetes per midgut were counted on a fluorescence microscope.

A. gambiae *challenge with fungi.*

Beauveria bassiana was a kind gift of Dominique Ferrandon, CNRS, IBMC, Strasbourg, France. The dark green fungus (Fig. 3.1.1 right) was extracted from a mosquito naturally overcome with it. For lack of better background knowledge, it was treated exactly like *B. bassiana* with regards to procedures and protocols. That seemed to work just fine.

note of caution: *B. bassiana* is extremely detrimental to most insects. Special care has to be taken to do all work under sterile conditions to avoid contamination of laboratory insect lines!

Infection of mosquitoes with B. bassiana

Genes of interest were knocked down by RNAi in adult female mosquitoes as described above. On the fourth day after KD, mosquitoes were knocked down on ice and put onto a malt-agar plate (Fig. 3.1.1) on which 8 days earlier *B. bassiana* had been plated out. The plate can be kept up to 3 weeks at room temperature. Gently shake the plate with the mosquitoes on it and then return them to their container. Keep the mosquitoes in a room dedicated to *B. bassiana* infections separate from the other mosquito strains you are maintaining.

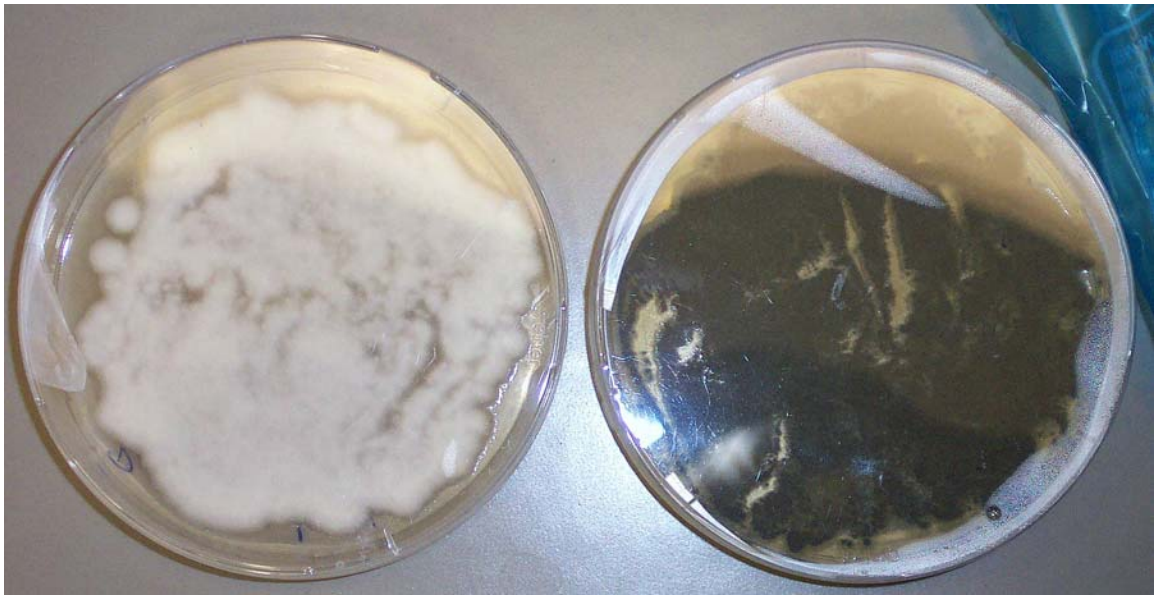


Fig 3.1.1 B. bassiana (left) and unknown fungi extracted from mosquitoes (right) grown on malt-agar plates.

Extraction of spores

Several malt-agar plates of >3 weeks old *B. bassiana* are harvested by adding 10-15ml of ddH_2O on to the plates and scratching the surface with a sterile spatula. Pour the juice over

autoclaved glass wool (in a sterile glass funnel) into a centrifuge bottle. Wash the plates twice with 10-15 ml of ddH_2O and centrifugate 10 min at 4000 rpm. Discard the supernatant and wash the pellet twice with ddH_2O. Dissolve the pellet in 100-250 µl of ddH_2O and count the number of spores in a Thoma (or Neubauer) counter by diluting an aliquot x100, x1000 or more. Add sterile 50% glycerol to have a final concentration of 10^7 spores/ml in 25% glycerol and store the spores at -80°C.

Plating out of spores

Plate out 50 µl of 10^7 spores/ml on a malt-agar plate and grow at room temperature (Fig. 3.1.1).

Malt-agar plates
(1l medium, for ~25 plates of 10 cm dish)
Peptone: 1g (Select peptone GIBCO BRL Cat No. 30392-021)
Glucose: 20g (α-D(+)glucose monohydrate ROTH Art 6780)
Malt: 20g (Sigma M-0383)
Select agar: 15g (Invitrogen Cat No 30392-023)

3.2 FRUIT FLY *DROSOPHILA MELANOGASTER*

Fruit fly rearing.

Drosophila melanogaster was reared on "*Drosophila* Quick Mix Medium" in standard *Drosophila* vials (both from Blades Biological Ltd (http://www.blades-bio.co.uk/home.htm)) at 21°C (RT). The white mutant strain was ordered from Blades Biological Ltd and the *PGRP-SA* (semmelweis) (Michel, Reichhart et al. 2001) and *PGRP-LC*/TM6B (Gottar, Gobert et al. 2002) mutant strains were obtained from Julien Royet, IMBC du CNRS, Strasbourg, France.

***Bacterial infection & survival of* Drosophila.**

GFP-expressing Gram- *Escherichia coli* OP-50 were a gift from J.J. Ewbank (INSERM, Marseille-Luminy, France). Gram+ *Staphylococcus aureus* were kind gifts from Philippe Bulet (IBMC, Strasbourg, France).

Bacteria were cultured to OD_{600} = 0.7, pelleted, washed and resuspended in phosphate-buffered saline (PBS) to final concentrations (*E. coli* at OD_{600}=0.1, *S. aureus* at OD_{600}=0.01). 69 nl of the bacterial suspension or PBS for controls were injected into the *Drosophila* thorax with a nano-injector (Nanoject, Drummond) while the fruit flies were anesthetized with CO_2. Dead fruit flies were counted daily and removed over a period of at least 7 days. The results generally are representative of at least three independent experiments, each carried out with more than 50 fruit flies per batch.

3.3 DNA METHODS

cDNA production.

Total RNA was extracted from *A. gambiae* tissue with TRIzol reagent (Invitrogen, San Diego, CA) and treated with DNAseI. After Phenol-Chloroform extraction and photo-spectrometric concentration determination, first strand cDNA synthesis was performed using 1µl of 0.5 µg/µl oligo dT primer (Sigma) with ~5µg of total RNA preparations (though cDNA synthesis was successfully performed with as little as 0.2µg total RNA from single mosquitoes).

The mixture (10µl) was denatured at 68°C for 7 min and a cDNA synthesis reaction mix was added (1µl Reverse Transcriptase (SuperScript™II Rnase H- Reverse Transcriptase (Invitrogen) or MLV Reverse Transcriptase (Gibco BRL)); 1.5µl dNTPs (10mM) (peqlab); 1.5µl DTT (Gibco BRL); 4µl 5x "First Strand" Buffer (Gibco BRL); 1µl RNAse Inhibitor (Roche)) and incubated at 42°C. After 50 min the Reverse Transcriptase was inactivated by heating the mixture for 15 min at 70°C.

The cDNA was then treated with 100µg RNAseA per for 30' at 37°C, a Phenol/Chloroform Extraction performed and resuspended in an appropriate volume of ddH_2O.

3.4 RNA METHODS

total RNA extraction.

9 volumes of TRIZOL are used for one volume of tissue (or 400µl TRIZOL for 1-20 mosquitoes) according to manufacturer's instructions. Total RNA is resuspended in ~70µl ddH_2O.

Semiquantitative RT-PCR.

cDNA was produced as described above. Several dilutions of cDNA and numbers of PCR cycles were usually tested. The PCR reactions (50µl) were performed using 10pmol of gene-specific primers, 0.1 U of Amplitaq (Roche), 1µl 10mM (10nmol) dNTPs and 5µl of 10x PCR Buffer including 1.5mM MgCl2 (Roche).

After electrophoresis, the agarose gel was stained with the sensitive SYBR green dye (Molecular Probes) for 30-45min and analyzed with a fluorimager (Fuji FLA-2000).

Quantitative real-time RT-PCR.

The SYBR Green PCR Master Mix is a convenient premix of all the components, except primers, template and water necessary to perform real-time PCR using SYBR® Green I Dye. Direct detection of polymerase chain reaction (PCR) product is monitored by measuring the increase in fluorescence caused by the binding of SYBR Green dye to double-stranded (ds) DNA.

cDNA template was diluted 1:10 or 1:20 and triplets of the same reaction were analyzed. PCR reactions were performed according to the ABI Bulletin #2, whereby 25µl

reactions were analyzed according to the "Relative Standard Curve Method" with the ribosomal protein S7 as internal calibrator. In this method, target quantity for experimental samples is determined from the standard curve and divided by the target quantity of the calibrator S7. Thus, the calibrator becomes the 1x sample, and all other quantities are expressed as x-fold difference relative to the calibrator.

For each experimental sample, the amount of target and calibrator were determined from the corresponding standard curve. Then, the target amount was divided by the calibrator amount to obtain a normalized target amount.

An Applied Biosystems 7500 Real-Time PCR System was used according to manufacturer's instructions. One PCR reaction (25µl) consisted of 12.5µl 2xSYBR® Green PCR Master Mix, pretested concentration (mostly 900 nM) of forward and reverse primers, and cDNA template dilutions.

Thermal cycling parameters were

1. 95°C, 10min

2. 40 cycles of: 95°C, 15sec;

60°C, 1min

dsRNA production for RNAi.

dsRNA is either produced by 1. using the pLL10 vector or 2. PCR with T7 primers.

1. Target gene fragments are cloned into the vector pLL10 (derived from the pLL17 vector with the pLL7 polylinker (Levashina, Moita et al. 2001)) and thus flanked by T7 polymerase promoters. The construct is linearized on either end of the insert with the

appropriate restriction enzymes in two separate reactions. The linearized constructs serve as template for the RNA transcription using the MEGAscript T7 Kit (Ambion).

After overnight incubation of the T7 reaction mixture at 37°C, the reactions are treated with DNaseI and cleaned up by Phenol:Chloroform extraction or with the RNeasy kit (QIAGEN). The ssRNAs is resuspended in 20µl RNase-free water. The concentration of the ssRNAs is determined by spectrometric concentration analysis (OD_{260}).

Minus- and plus- RNA strands are diluted to the concentration of 3µg/µl, mixed and denatured in a big beaker of boiling water for 5min. The mixture is then allowed to cool down to RT and anneal (several hours, preferably ON). The resulting dsRNA is analyzed on 1.5% agarose gel supplemented with 0,5µg/ml Ethidium bromide.

2. PCR amplicons obtained from PCRs with primers tailed with the T7 promoter sequences (GAATTAATACGACTCACTATAGGG) at their 5' end are cleaned up with the QIAquick Gel extraction kit or QIAquick PCR Purification kit (QIAGEN) and resuspended in nuclease free water. The purified PCR amplicons are then used to synthesize dsRNA with the MEGAscript T7 Kit (Ambion).

Just like in method 1, the reaction is treated with DNaseI and cleaned up by Phenol:Chloroform extraction or with the RNeasy kit (QIAGEN). The dsRNA is resuspended in 20µl RNase-free water, but <u>not</u> denatured and annealed by a boiling and cooling cycle.

The concentration of the dsRNA is determined by photo-spectrometric concentration analysis (OD_{260}) and adjusted to 3µg/µl by concentration with a vacuum centrifuge. The dsRNA is also analyzed on a >1.5% agarose gel supplemented with 0,5µg/ml Ethidium bromide.

3.5 MICROARRAYS

EST and oligonucleotide microarray analysis.

Mosquito EST microarray construction, hybridization and analysis were performed as described (Dimopoulos, Christophides et al. 2002). Developmental profiling was performed using embryonic, 4th larval instars, pupal and newly emerged female adult stages of *A. gambiae*, Suakoko strain. A pool of total RNA prepared from all stages was used as reference sample. For immune challenges, 2 to 3-day old female mosquitoes of the *Plasmodium*-susceptible strain 4a r/r (Collins, Sakai et al. 1986) were pricked with either a sterile needle or dipped in thick suspension of *E. coli* or *S. aureus*. Total RNA was collected 12 hrs after challenge.

For malaria infection experiments, 4a r/r mosquitoes were fed on control Balb/c mice or on mice infected with *P. berghei*, and mosquito RNA samples collected at 24 hrs, 28 hrs, 6 days, 11 days and 16 days post-infection.

Cell line 4a3B (Muller, Dimopoulos et al. 1999) was challenged with paraformaldehyde-fixed *E. coli* and *S. aureus* (OD 0,05), PGN (10 μg/ml) and H_2O_2 (2 □ M). Duplicated RNA samples were collected 12 hrs after challenge and hybridized to arrays as described (Dimopoulos, Christophides et al. 2002). RNA prepared from naïve cells was used as reference.

Microarray construction/printing.

4k Microarray slides were printed as described (Dimopoulos, Christophides et al. 2002). Oligonucleotide primers were designed to amplify individual genes from a cDNA library or adult genomic DNA (average probe-length 500bp). PCR products were purified with ion exchange columns (Macherey-Nagel GmbH & Co. KG, Dueren, Germany) and spotted at 500ng/µl in 3X SSC. Spotting was performed on aminosilane coated glass slides using the Omnigrid arrayer (GeneMachines, San Carlos, CA) and Telechem Stealth Pins (Telechem International, Sunnyvale, CA).

Sample amplification.

First Strand Synthesis: 100pmol T7-d(T)24 primer and 5µg total RNA are used to produce cDNA with 5' T7 overhangs. The reaction mix consists of 4 µl 5x first strand synthesis buffer (Gibco BRL); 2 µl DTT 0.1 M (Gibco BRL); 1ul dNTP mix 10 mM (Gibco BRL); 1 µl 200 U/µl SuperScript™II Rnase H- Reverse Transcriptase (Invitrogen) or MLV Reverse Transcriptase (Gibco BRL). Incubate at 37°C for 1 hour.

Second Strand Synthesis: To the reaction is added: 10 ul 5x Second Strand Synthesis Buffer

1 µl dNTP mix 10 mM

0.35 ul DNA Ligase 10 U/µl

1.3 µl DNA Polymerase I 10 U/µl

0.7 µl RNase H 10 U/µl

16.7 µl DEPC-H2O

Incubate sample at 16°C for 2h and add 100µl H_2O for a final volume of 150µl. Stop the reaction by adding 10µl 0.5 M EDTA and clean up the cDNA by with Phenol-Chloroform.

cDNA Amplification: Use the T7 RNA Polymerase (Ambion T7 Megascript kit) to amplify the cDNA. For that 6µl ds-cDNA sample are mixed with 1.5µl 10x Reaction Buffer, 1.5µl of each ribonucleotide (ATP, CTP, GTP, UTP) and 1.5µl T7 Polymerase enzyme mixture. The final volume is 15µl. After 4h at 37°C, 0.5µl DNaseI are added and the sample incubated at 37°C for 15min. Then chill on ice, clean up the amplified RNA with the QIAGEN RNeasy mini kit, determine cmRNA concentration and purity spectrophotometrically (the expected yield is 60 to 100µg) and dilute it to a final concentration of 1 µg/µl.

Sample labeling.

The following mix is prepared on ice and being placed at 70°C for 5 min:

5µl poly A+ RNA 1 µg/µl

1µl random hexamers 3 µg/µl

1µl oligo dT 1 µg/µl

3µl H_2O (final reaction volume10 µl)

Then add:

4µl Superscript II First Strand Synthesis Buffer

2µl 0.1 M DTT

2µl Cy-3 or Cy-5

0.4µl dT-NTP mix (25mM dA, dC, dG; 10 mM dT)

1.5µl Superscript II

and incubate at 42°C for 2 h. Stop the reaction with 1µl 1M NaOH / 20 mM EDTA and incubate at 70°C for 5 min. 7. Add 60µl H_2O to Cy-3 reaction, combine with Cy-5 reaction and add 10µl 3M NaAcetate, pH ~5. Then purify the labeled cDNA through a Qia-quick PCR purification column (Quiagen) (elution volume: 2x30µl) and add poly(A) DNA (herring sperm) to a 0.4 µg/µl final concentration in the hybridization solution. 9. Then dry the sample in speed-vac and resuspend into the desired volume of hybridization solution (50µl) just before applying onto the array or store at 4°C in the dark for later user.

Array hybridization & data acquisition.

Slides were pre-hybridized in 5xSSC, 0.1% SDS, and 1% BSA at 42°C for 45 min; washed at RT with deionized water, dipped in 100% isopropanol; and dried. Equal volumes of Cy-3 and Cy-5 labeled samples were combined in hybridization buffer, supplemented with Poly(A)-DNA and hybridized to microarray slides at 42°C over night. Hybridized arrays were washed, air-dried and scanned using the ScanArray 3000 (GSI Lumonics, Billerica, MA).

Data analysis.

Two independent experiments with dye swaps were performed. Expression ratios were calculated and analyzed using the TM4 microarray software package (www.tigr.org/software/tm4/). Following spot quality filtering, expression data were normalized using the Lowess (locfit) algorithm, and clustered with *k*-Mean and hierarchical clustering with the TMEV software (TIGR) (Hegde, Qi et al. 2000). Elements with at least 500 signal intensities were considered.

3.6 PROTEIN METHODS

***Protein expression in* E. coli.**

Escherichia coli bacteria are transformed by making them electro-competent. 1 µl of a 1:100 plasmid dilution is added to the freshly thawed bacteria in a ice-cold electro-transformation vial. This is exposed to electricity for 4.5 seconds (or until sparks are produced. Bacteria are then incubated in 1ml of SOC bacterial growth medium for 1 hour at 37°C and consequently plated out on LB agar plates with the appropriate antibiotic (Kanamycin in the case of pEMT-11, pEMT-13 or pEMT-60 vectors (refer to chapter 10.4)).

From the LB agar plate a single colony is picked and grown overnight in different growth media (with appropriate antibiotic) at 37°C and constant shaking to an OD_{600} of 0.5 to 0.7. Expression of the vector insert is then induced by addition of IPTG to a final concentration of 0.01 to 0.1 mM and a 1 to 3 hours incubation at 37°C with constant shaking.

The growth medium is then replaced with 1x PBS, 0.1% Triton and a Protease Inhibitor Cocktail ("Complete EDTA-free" Tablets from Roche). Sonicate the bacteria on ice and spin down at 4°C at maximal speed. Run a Tris-glycine SDS-polyacrylamide protein gel (Sambrook and T. 1989) with a lane for the non-induced bacteria, the induced bacteria, the induced supernatant and the inclusion bodies of the induced bacteria.

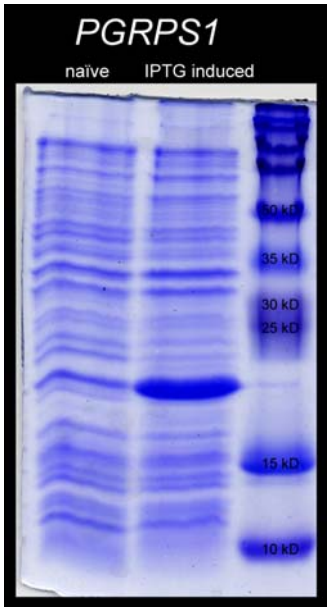


Fig. 3.6.1 IPTG induced expression of PGRPS1

Ideally there will be a fat protein band of the expected size visible in the lanes representing the induced bacteria and the induced supernatant (compare to Fig 3.6.1). This will mean that your protein is expressed and soluble. In all other cases the conditions (bacterial strains, growth media, IPTG concentration and incubation time) need to be varied.

Protein gel electrophoresis.

Clean equipment and glass plates with isopropanol. Prepare the required amount of resolving gel in small Becher according to Maniatis 2004 Table 10.2A.1.

Pour resolving gel and overlay it with Isopropanol to eliminate bubbles and air-contact and let it polymerize for 20-30 minutes at RT (you can then leave the gel ON at 4˚C if desired, but overlay the gel with ddH_2O and pack in saran wrap to prevent it from drying).

Decant the isopropanol and wash with ddH_2O to remove unpolymerized acrylamide - drain fluids thoroughly with a paper towel.

Prepare the required amount of stacking gel according to Maniatis 2004 Table 10.2A.1, pour and slowly gel combs. Let it polymerize for 20-30 min at RT. While stacking gel is polymerizing, boil samples in loading buffer for 3 minutes to denature proteins. After polymerization, remove combs under running tab water, set up, load and run gel (0.03 A per mini-gel; 100 V initially and 135 V as soon as proteins reached resolving gel).

After run, remove gel from glass with razorblade and mark upper left corner by removing it and either use for Western blotting or stain gel in Coomassie from 1h to 24h at RT. Rinse gel in water and destain in Destaining solution with tissue to take up the blue color – gel can be kept in Destaining solution for weeks.

Laemmli loading buffer (3X stock):

1M Tris-Cl pH 6.8	2.4 ml
20% SDS	3 ml
Glycerol (100%)	3 ml
B-mercaptoethanol	1.6 ml
Bromophenol blue	0.006g
DTT	0.45g

10 ml (store 4°C)

Coomassie-Destaining Solution:

30%	Methanol	300 ml
10%	Acetic acid	100 ml
60%	H_2O	fill up to 1 liter

Western blotting.

Blotting:

Separate the protein samples using protein gel electrophoresis as described above. In the meantime, prepare a sheet of Hybond-P® protein transfer membrane. Cut the membrane and 2 Whatman 3MM paper to the size of the resolving gel and pre-wet the membrane in ~5ml 100% methanol for ~10 sec. Wash the membrane in ~50 ml of distilled

water for five minutes and equilibrate the membrane in 50ml protein transfer buffer for at least 10 minutes. After the protein gel is done, remove the stacking gel and orient the resolving gel by clipping a corner.

Assemble the electro-blotting cassette and place the blotting setup (top to bottom: whatman-gel-membrane-whatman) between the electrodes in the blotting unit (according to the manufacturer's instructions) wetting everything with the transfer buffer. Transfer with 220mA for 30-45 min.

Following the transfer, remove the membrane from the blotting cassette, mark the orientation of the gel on the membrane and the molecular marker bands with a pencil and rinse the membrane briefly in PBS. Trim the membrane/blot.

(Optional: check transfer with PonceauS solution. Stain the blot for 5-10 minutes at RT – the bands come up in red. (PonceauS solution can be reused multiple times – store at RT in the dark) Wash off with H_2O after staining. Mark molecular weight marker bands with pencil).

Blots may be used immediately or stored for later use (membrane should be air dried before storage. Blots may be stored between sheets of Whatman 3MM paper wrapped in Saran Wrap at 2-8° C for up to 3 months. Once dry, the membrane will require pre-wetting with Methanol/H_2O).

Immunodetection:

Pre-wet the membrane in 100% methanol if previously dried, then wash 5 minutes in distilled water. Block non-specific binding sites on the membrane in by incubating for one hour at RT in ~50ml PBS-T with 5% (w/v) dried skimmed milk (if done over night, at 4°C).

Rinse membrane twice briefly in PBS-T (as large a volume of wash buffer as possible: >=1-2 ml/cm^2) and wash the blot for 5 minutes with an excess volume of PBS-T.

Incubate the blot with the primary antibody at the optimized dilution in PBS-T for 1 h (good first approximation for antibody dilution: 1:500 to 1:20000) – the primary antibody can also be incubated ON at 4˚C.

Briefly rinse the blot in PBS-T and wash twice for 10 minutes with excess PBS-T.

Incubate the blot with appropriate secondary antibody at the recommended dilution in PBS-T for 1 h. Do NOT incubate secondary antibody over night – even 45 min at RT is enough.

Briefly rinse in PBS-T and wash three times 10 min in excess PBS-T. Proceed with the appropriate detection system. (for HRP: Western Lightning Chemiluminescence Reagent (Perkin Elmer) and expose to Kodak Film (X-OMAT or Bioluminescence Film (more expensive)) for ~30sec and develop the film. If necessary repeat exposure with varied exposure time for optimal signal.

Protein sample loading buffer:
4ml distilled H_2O
1ml 0.5M Tris-HCl, pH 6.8
0.8ml Glycerol
1.6ml 10% (w/v) Sodium dodecyl sulphate (SDS)
0.4ml β-Mercaptoethanol
0.5ml 0.05% (w/v) Pyronin Y (e.g. Sigma code P-7017)
(store in dark at room temp for a maximum of 2 weeks)

Protein transfer buffer:
3.03g Trizma-base
14.4g glycine
dissolve in 900ml H_2O w/o Methanol add 1/10th volume of Methanol before use
Add approximately 650ml of distilled H_2O. Mix to dissolve. Make up to a final volume of 900ml. Store @ 2-8˚ C. Inclusion of methanol in the buffer minimizes swelling of the gel during blotting.

PBS-Tween (PBS-T):
Dilute required volume of Tween 20 in PBS to give a 0.1% (v/v) solution. store @ 2-8˚ C. Make a liter of it – you will need it.
The choice of PBS or TBS will depend on the detection system of choice. The use of PBS is not advised with alkaline phosphatase based detection systems as phosphate ions are powerful inhibitors of this enzyme.

3.7 CELL CULTURE

A. gambiae *cell lines maintenance and thawing/freezing*.

Maintenance

The cell lines were established from minced neonate larvae of three *A. gambiae* strains: Suakoko 2La, 4a r/r, and L3-5 (Muller, Dimopoulos et al. 1999).

The cells are grown in Schneider medium (used: GibcoBRL-liquid, or Sigma-powder) supplemented with 10% FCS and 100 U/ml Pen.+100µg/ml Strep. The cell lines can be grown without antibiotics if special care is devoted to sterile conditions.

CO_2 gassing is not necessary. Cells are generally grown at 27°C. However, they can be kept at 25°C. For reduced growth, keep them at 18°C, with more medium than usual per flask and decreased FCS concentration (5%). Heat-inactivate FCS for one hour at 60°C prior to use. Generally: the more FCS is added to the Schneider Medium (up to 30% is possible), the faster the cells grow and attach less to the plastic (and the easier they are to resuspend).

For splitting, shake the bottle vigorously. This releases the majority of cells and split diluting the cells 1:2-1:3 (up to 1:10 is possible) when cells are approaching confluence.

Strongly adherent cells like Sua1B can be trypsinized using standard cell culture protocols, though the cells should be washed with PBS several times before trypsinizing them.

Some debris, especially in old cultures, may be observed – it will be phagocytosed by the cells!

Freezing/Thawing

Use a densely grown big flask (250ml) for 2-3 aliquots.

Remove all medium and add 2-3ml Schneider medium with 10% FCS and 10% DMSO. Scrape the cells to release them from the plastic surface and aliquot 1ml in cryo tubes. Immediately place the tubes on ice for 30-45 minutes and then put them in a polystyrene box overnight at -80°C. Maximal 48 hours later, transfer the tubes into liquid nitrogen.

For thawing, warm up the tube in your hand and liberally flame the top (all the cells are settled on the ground of the tube). Resuspend the tube contents 1ml of fresh medium and add to the medium in a 15ml Falcon tube. Repeat until all content of the cryo tube is in the tube. Centrifuge 5min at 1500 rpm, take away the medium and resuspend the cells in new medium in a flask.

You may also thaw by just adding the cryo tube contents in the flask with medium and letting the cells settle for a few hours. Then take off the DMSO containing medium and replace it with fresh medium. You may increase FCS to 20% at the beginning to let the cells recover.

A. gambiae *cell line challenge.*

Cells were challenged commercial bacterial cell wall component extracts (LPS, PGN), commercial fungi cell wall component extract (Zymosan) or heat killed (60˚C for 1h) bacteria (*S. aureus, E. coli, M. luteus*) and fungal spores (*B. bassiana*).

LPS and PGN were dissolved in water (10 mg/ml). 2.5µl per 5ml cell well were added for a 5µg/ml final concentration. Zymosan was dissolved in water for a 10 mg/ml concentration and 5µl of that was added per well for a 10 µg/ml final concentration.

M. luteus, *S. aureus*, *E. coli* (DH5α) bacteria and *B. bassiana* spores (American Type Culture and Collection, Manassas, VA) were heat-inactivated by incubation at 60˚C for 1h in phosphate buffered saline (PBS), diluted to a OD_{600}=0.2 and 175µl were added per well to result in a final OD_{600}=0.007.

Hybridoma cell maintenance and thawing/freezing.

Maintenance

Care must be taken at all times to work under sterile conditions. Cells usually need to be split 1:5 on Monday and Wednesday and 1:10 on Friday

Thawing

Thaw a vial of cells quickly at 37˚C, centrifuge in 8-10 ml of HM20-10% HCF. Discard the medium and resuspend the pellet in 1 ml of HM20-10% HCF. Transfer the cells into a well of a 24 well plate (let' s call it first well). Take 0.5mL of cells from the first well and add to a second well containing 1 ml of HM20-10% HCF.

Take 0.5ml of cells from the second well and add to a third well containing 1 ml of HM20-10% HCF. Then take 0.5ml of cells from the third well and add to a fourth well containing 1 ml of HM20-10% HCF. Then take 0.5ml of cells from the fourth well and put it back into the first well. Culture the cells at 37˚C in a humidified incubator with 5% CO_2.

When splitting cells, at each passage the HCF must be gradually decreased in concentration from 10% to 5%, 2.5% and no HCF. Check every time whether cells are in good shape and healthy. This way you will have resuspended cells at four different dilutions.

Once they have recovered you can either do the subcloning immediately or transfer into 6 well plate/small flask by gradually decreasing the percentage of HCF. You should end up with a culture with no HCF.

Ideally cells should be subcloned upon defrosting due to chromosomal instability inherent in hybridomas, which can be especially exaggerated by freeze/thawing.

Freezing

Aliquots of cells should be frozen as backup as soon as possible after defrosting. For that the cells are pelleted by centrifugation at 100g for 5 minutes. Resuspend approximately 107 cells/ml of 90% FCS /10% DMSO. Leave at -80˚C for at least 2 hours (not longer than 48 hours) and transfer to liquid nitrogen.

Solutions and chemicals used:

HM20 (Hybridoma medium 20% FCS)

FCS (Hyclone # SH30070.03)		100ml
L-Glutamine 200 mM (Gibco # 25030-024)	5ml	
Gentamicin 50mg/mL (Gibco #15750-037)	1ml	

DMEM

without Na pyruvate, with 4500mg/ml glucose,
with Piridoxine HCl (Gibco 41965-039) make up to 500ml

Hybridoma cloning factor (HCF):

Hybrid-MAX (Euroclone: Catlogue Number ECO 1021N)…use at 10%

3.8 SOFTWARE

Software	Task
ABI PRISM 7000 SDS	qRT-PCR analysis
Adobe Photoshop CS	Image processing
Adobe Illustrator CS	Image processing
AnoBase	EST retrieval (http://www.anobase.org/)
Artemis 7.0	Sequence analysis (http://www.sanger.ac.uk/Software/Artemis/)
BLAST	Basic Local Alignment Search Tool (http://www.ncbi.nlm.nih.gov/blast/)
ClustalW	Multiple sequence alignment (http://www.ebi.ac.uk/clustalw/)
ConSurf	Surface-mapping of phylogenetic information (http://consurf.tau.ac.il/)
Ensembl genome browser	Sequence retrieval (http://www.ensembl.org/index.html)
FlyBase	*Drosophila* sequence retrieval (http://www.flybase.org/)
GeneSpring	Software Platform for statistical analysis of microarray data
GenePix Pro	Image analysis software for microarrays
GenomeRNAi	RNAi primer design (http://www.dkfz.de/signaling/ernai/ernai_prime.html)
Genscan	Gene prediction (http://genes.mit.edu/GENSCAN.html)
OpenWetWare	Lab resource wiki (http://openwetware.org/wiki/Main_Page)
Primer3	Primer design (http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi)
Protein Explorer	Protein structure viewer (http://www.umass.edu/microbio/chime/pe/protexpl/frntdoor.htm)
PubMed	Literature Mining (www.pubmed.gov)
PyMOL	Protein structure viewer (http://pymol.sourceforge.net/)
Spotfire	Visual data analysis application
Tcoffee	Multiple sequence alignment (http://igs-server.cnrs-mrs.fr/Tcoffee/tcoffee_cgi/index.cgi)
TreeView	Phylogenetic analysis (http://taxonomy.zoology.gla.ac.uk/rod/treeview.html)

RESULTS

4.1 *A. GAMBIAE* ORTHOLOGS OF THE *TOLL* AND *IMD* PATHWAY

We have identified 242 genes (Christophides, Zdobnov et al. 2002) in the *A. gambiae* genome (Holt, Subramanian et al. 2002) from 18 gene families implicated in innate immunity in *D. melanogaster*. The *Toll* and *Imd* pathways are central to the *D. melanogaster* innate immune signaling and we thus decided to concentrate on them. The *A. gambiae* orthologs of the *Toll* and *Imd* pathway signaling components are presented in Fig. 4.1.1. However, a few key players are missing from the picture. Orthologous groups of genes, but not clear 1:1 orthologs could be assigned to the *Toll* receptor, *GNBP1* and *GNBP3*, while *Dif*, *PGRP-LE* and *PGRP-SD* are missing.

Seven PGRPs have been implicated in the *Imd* and *Toll* signaling pathways: *PGRP-LC* (Choe, Werner et al. 2002; Gottar, Gobert et al. 2002), *PGRP-SA* (Michel, Reichhart et al. 2001), *PGRP-SD* (Bischoff, Vignal et al. 2004), *PGRP-LE* (Takehana, Katsuyama et al. 2002; Takehana, Yano et al. 2004), PGRP-LB (Zaidman-Remy, Herve et al. 2006), PGRP-SC1 and PGRP-SC2 (Bischoff, Vignal et al. 2006) (Fig. 1.6.1).

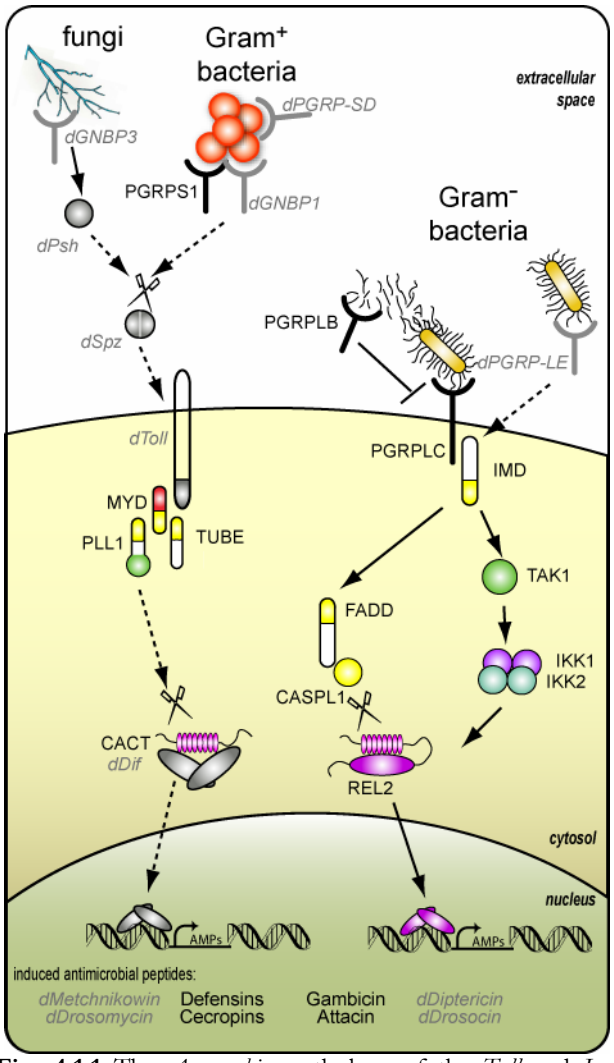


Fig. 4.1.1 The *A. gambiae* orthologs of the *Toll* and *Imd* immune signaling pathways in *D. melanogaster* (modified from (Meister, Koutsos et al. 2004)). Genes without a clear cut ortholog or missing entirely in the *Anopheles* genome are in grey and have a 'd' for prefix.

4.2 THE *PGRP* GENE FAMILY IN *A. GAMBIAE*

Multiple alignment & phylogenetic tree.

In *Drosophila* there are 13 genes containing 16 PGRP domains (Werner, Liu et al. 2000) as compared to only 7 genes with 10 PGRP domains in *Anopheles* (Christophides, Zdobnov et al. 2002). The *Anopheles* genes are named to correspond to their *Drosophila* orthologs. In general, *PGRP* genes can be classified into two classes. Among the 7 *Anopheles* genes there are three genes of the short, extracellular and secreted S-class: *PGRPS1* (ENSANGG00000014831), *PGRPS2* (ENSANGG00000022240), *PGRPS3* (ENSANGG00000010490), and four genes of the long, intracellular/transmembrane L-class: *PGRPLA* (ENSANGG00000007952), *PGRPLB* (ENSANGG00000011459), *PGRPLC* (ENSANGG00000007834), *PGRPLD* (ENSANGG00000024042).

All *D. melanogaster* and *A. gambiae* PGRP domain sequences were aligned with each other (Fig. 4.2.1) with Tcoffee (http://igs-server.cnrs-mrs.fr/Tcoffee/tcoffee_cgi/index.cgi) (Notredame, Higgins et al. 2000) and the alignment was then edited by hand. It served as the basis for the phylogenetic tree in Fig. 4.2.2b and to illustrate the extent of conservation across the PGRP domains of these two species. The secondary structure of the PGRP domain has been mapped above the multiple protein sequence alignment (Fig. 4.2.1).

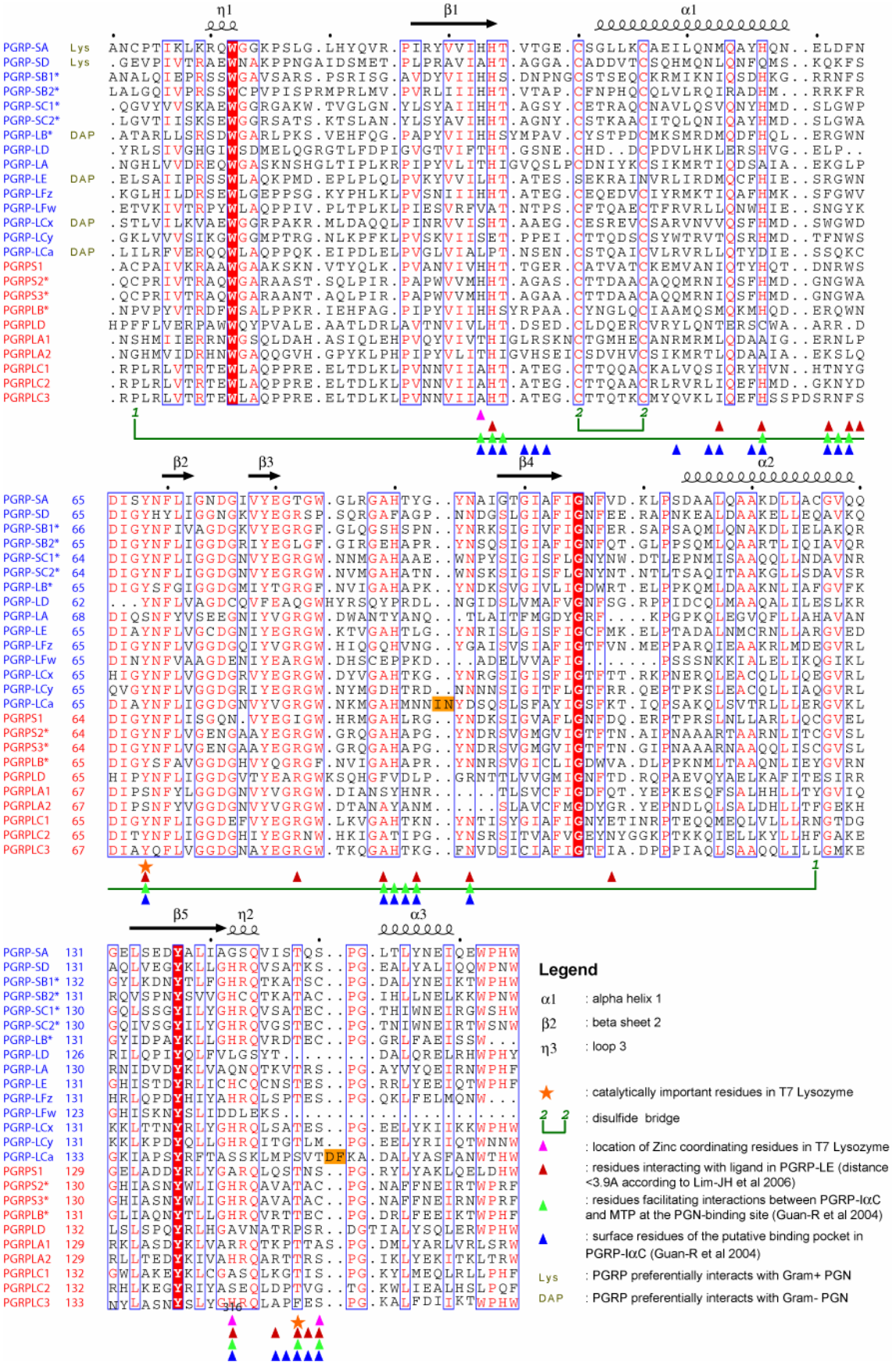

η1
β1
α1
β2
β3
β4
α2
β5
η2
α3
Legend
α1 : alpha helix 1
β2 : beta sheet 2
η3 : loop 3
: catalytically important residues in T7 Lysozyme
: disulfide bridge
: location of Zinc coordinating residues in T7 Lysozyme
: residues interacting with ligand in PGRP-LE (distance <3.9A according to Lim-JH et al 2006)
: residues facilitating interactions between PGRP-IαC and MTP at the PGN-binding site (Guan-R et al 2004)
: surface residues of the putative binding pocket in PGRP-IαC (Guan-R et al 2004)
Lys : PGRP preferentially interacts with Gram+ PGN
DAP : PGRP preferentially interacts with Gram- PGN

Fig. 4.2.1 (previous page) Multiple alignment of all *A. gambiae* (red) and *D. melanogaster* (blue) PGRP domains. The secondary structure (α: alpha helix; β: beta strand; η: loop) of *Drosophila* PGRP-SA is shown above the sequences. An asterisk next to the PGRP name indicates potential or experimentally proven Amidase activity. Pink triangles below the sequences indicate residues that correspond to the location of the zinc coordination residues in T7 Lysozyme. Orange stars mark the location of catalytically important residues in T7 Lysozyme. Red triangles indicate residues that interact with Ligands in PGRP-LE (Lim, Kim et al. 2006). Green triangles mark residues that interact with the ligand in human PGRP-Iα, and blue triangles represent residues on the surface of the putative PGN binding pocket (Guan, Roychowdhury et al. 2004) Green numbers connected with green lines highlight Cysteine positions that form a disulfide bridge when present in the sequences.

The multiple alignment was restricted to the conserved PGRP domains; the varying lengths of the full proteins and their extremely low conservation outside of the PGRP domains would have only served to dilute the signal for the alignment algorithm. The presented alignment is thus extremely compact and has very few gaps.

It can be clearly seen from this linear alignment, that the amino acids responsible for ligand interactions (highlighted with colored triangles at the bottom of the alignment) are distributed over the whole length of the PGRP domain. This is because the 3D structure of the domain is organized in such a way that the required amino acids come together – as discussed in chapter 4.3 in the case of *PGRPLC*. It is worth pointing out that the disulfide bridge 2 is conserved in all PGRPs except PGRP-LE and that it is disrupted by the single mutation in the loss-of-function *semmelweiss* mutant (Michel, Reichhart et al. 2001).

The phylogenetic relationships between the *Drosophila* and *Anopheles* PGRPs are shown in Fig. 4.2.2. This tree highlights that *PGRPLB* and *PGRPLD* are true (1:1) orthologs of their *Drosophila* counterparts while PGRPLA1, PGRPLA2 and PGRP-LA form an orthologous group. For *PGRPLC* and *PGRPS1* the phylogenetic relationships that can be derived from the tree are not as clear. In the case of *PGRPLC* the situation is confounded by the *Drosophila PGRP-LF* that seems to share a common ancestor with *PGRP-LC*, which was

duplicated within the *Drosophila* lineage to produce the two genes. The orthologous PGRPLC genes have a very similar gene organization that clearly distinguishes them from *PGRP-LF* (see chapter 4.3 'The PGRPLC gene cluster').

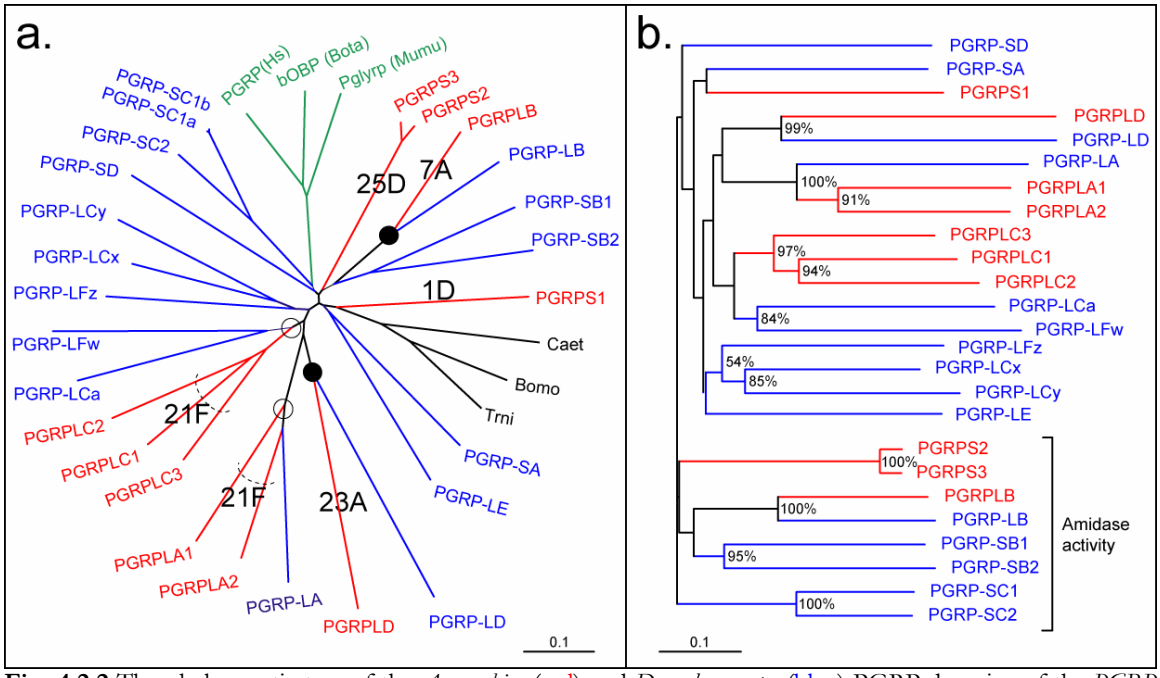


Fig. 4.2.2 The phylogenetic tree of the *A. gambiae* (red) and *D. melanogaster* (blue) PGRP domains of the *PGRP* gene family. **a.** Unrooted, radial tree with short mammalian (Hs, *Homo sapiens* (Kang, Liu et al. 1998; Liu, Xu et al. 2001); Bota, *Bos taurus* (Tydell, Yount et al. 2002); Mumu, *Mus musculus* (Kang, Liu et al. 1998)) and short insect (Bomo, *Bombyx mori* (Yoshida, Kinoshita et al. 1996; Ochiai and Ashida 1999); Caet, *Calpodes ethlius (Marcu and Locke 1998)*; Trni, *Trichoplusia ni* (Kang, Liu et al. 1998)) PGRP domains (green) used as out-groups. Labels on the branches denote the *Anopheles* genome chromosomal regions (cytological data from polytene chromosomes). Solid circles on the nodes indicate clear cut orthologs and hollow circles indicate orthologous groups. (adapted from: (Christophides, Zdobnov et al. 2002)) **b.** Phylogram version of the PGRP domain phylogenetic tree. PGRPs with potential or proven Amidase activity cluster together and the respective section of the tree is indicated. Percentages on nodes indicate branches with bootstrap support >50%.

PGRPS1.

The relation between *PGRPS1* and *PGRP-SA* is too distant to firmly establish their orthology from the phylogenetic tree alone. However, *PGRPS1* is *PGRP-SA*'s best reciprocal

blast hit, and both genes are located on the X chromosome; for that reason they can be regarded as orthologs (Zdobnov, von Mering et al. 2002).

PGRPS2 & -S3.

PGRPS2 and *PGRPS3* do not have clear orthologs in *D. melanogaster*. They are clustered in the *Anopheles* genome and are intronless. Their open reading frames are only 3712 nucleotides apart on opposite strands, and are 95% identical at the level of amino acids (9 aa different out of a total of 188 aa) and nucleotides (25 nt different out of a total of 564 nt).

We examined whether or not the two genes are haplotypes of the same gene that was artificially duplicated during the genome assembly process. The lack of similarity of the surrounding genomic regions suggested that these are two separate genes. We designed primers for *PGRPS2* and *PGRPS3* targeting regions, which specifically recognize only one of the genes (Exp25.1-F; Exp25.2-R; Exp25.3-R) and used them to perform PCR on the genomic DNA of eight individual G3 and L3-5 strain females and three females of the 69 and the Yaoundé strain each (Fig. 4.2.3). The results showed with a high degree of confidence, that *PGRPS2* and *PGRPS3* are independent genes, as the PCR on individual mosquitoes always produced 2 bands, one for each gene. This was consistent for all four strains (Fig. 4.2.3).

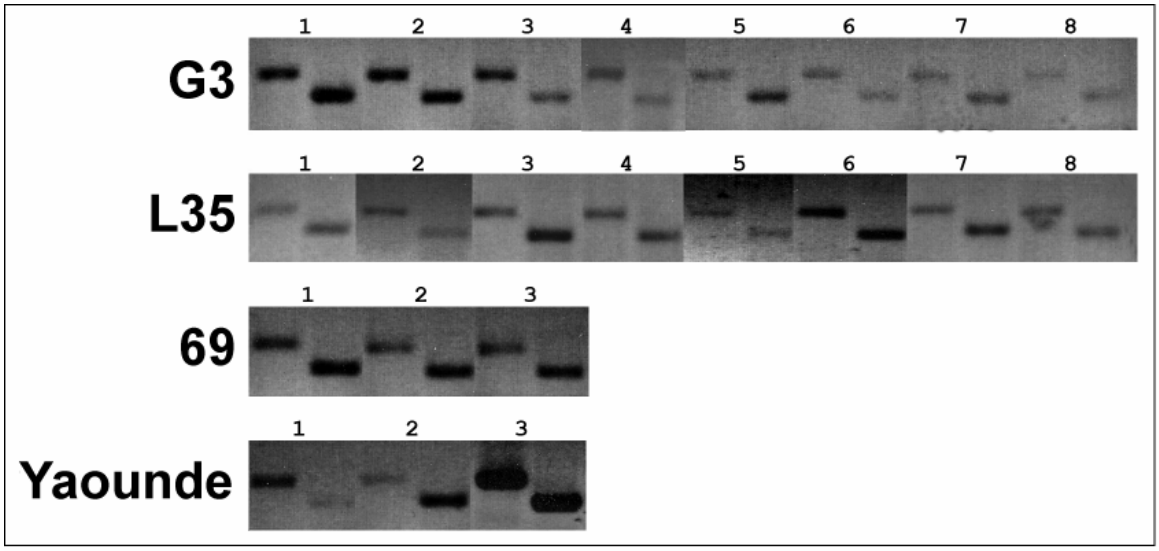


Fig. 4.2.3 Agarose gel electrophoresis of PCR performed using primers specific to either *PGRPS2* (higher bands) or *PGRPS3* (lower bands) on genomic DNA of individual G3 (1-8), L3-5 (1-8), 69 (1-3) and Yaoundé (1-3) strain females.

PGRPLA.

The mosquito PGRP-LA gene has two PGRP domains, similar to the fly and mosquito LCs and the fly LF, although the Ensembl *PGRPLA* (ENSANGG00000007952) annotation is missing the second PGRP domain. *PGRPLA* is in a gene cluster with *PGRPLC*, similar to its *Drosophila* ortholog. However, the *Drosophila PGRP-LA* only has a single PGRP domain (Fig. 4.3.1). The splice sites within the PGRP domains of all *PGRPLA*/*PGRP-LA* is conserved in both *Drosophila* and *Anopheles* (Fig. 4.3.3). The gene architecture suggests that *PGRPLA2* can produce functional mRNAs, which, however, could not be detected by PCR in a cDNA library. This suggests that the second domain is tightly regulated and not expressed under unchallenged conditions. A complete annotation of *PGRPLA* can be found in chapter 10.2 (Seq. 9.2.4 & 9.2.5) and 10.3 (Seq. 9.3.17 & Seq. 9.3.18) in the Appendix of this thesis.

note: as of June 2006, *PGRPLA* was not annotated correctly in the Ensembl genome browser (http://www.ensembl.org). It was annotated as a gene with a single instead of two PGRP domain.

PGRPLB.

PGRPLB is the clear-cut ortholog of *Drosophila PGRP-LB*: it clusters with PGRP-LB in the phylogenetic tree (Fig. 4.2.2) and is also its best reciprocal blast hit (Zdobnov, von Mering et al. 2002). The gene produces two transcripts, one of which has a transmembrane domain. Interestingly, transcription of the *Anopheles PGRPLB* gene is strongly upregulated in cells and adult mosquitoes challenged with various immune elicitors (Dimopoulos, Christophides et al. 2002). One of these challenges is malaria parasite infection.

PGRPLD.

PGRPLD is the ortholog of the *Drosophila PGRP-LD*. The official Ensembl *PGRPLD* annotation (ENSANGG00000024042) only encompasses the last exon of the ORF. A better annotation can be found in chapter 10.2 (Seq. 9.2.13) and 10.3 (Seq. 9.3.3) in the Appendix of this thesis.

PGRP KD and survival after bacterial infection.

All mosquito PGRP genes and isoforms (based on the different domains) were knocked down in adult female mosquitoes by RNAi (Blandin, Moita et al. 2002) by injection of dsRNA as described in Chapter 3 "Materials & Methods". Because of the high similarity of *PGRPS2* and *PGRPS3* on the sequence level, we could not knock down these two genes separately. Four days after dsRNA injection, the KD mosquitoes were injected with bacterial suspensions in Phosphate Buffered Saline (PBS) and their survival recorded. A significant reduction of mosquito survival relative to the control GFP dsRNA injected mosquitoes could only be observed after KD of PGRPLC, for both Gram- *E. coli* (Fig. 4.2.4) and

Gram+ *S. aureus* (Fig. 4.2.5) bacterial infections. The results of the different domains of *PGRPLC* are presented separately.

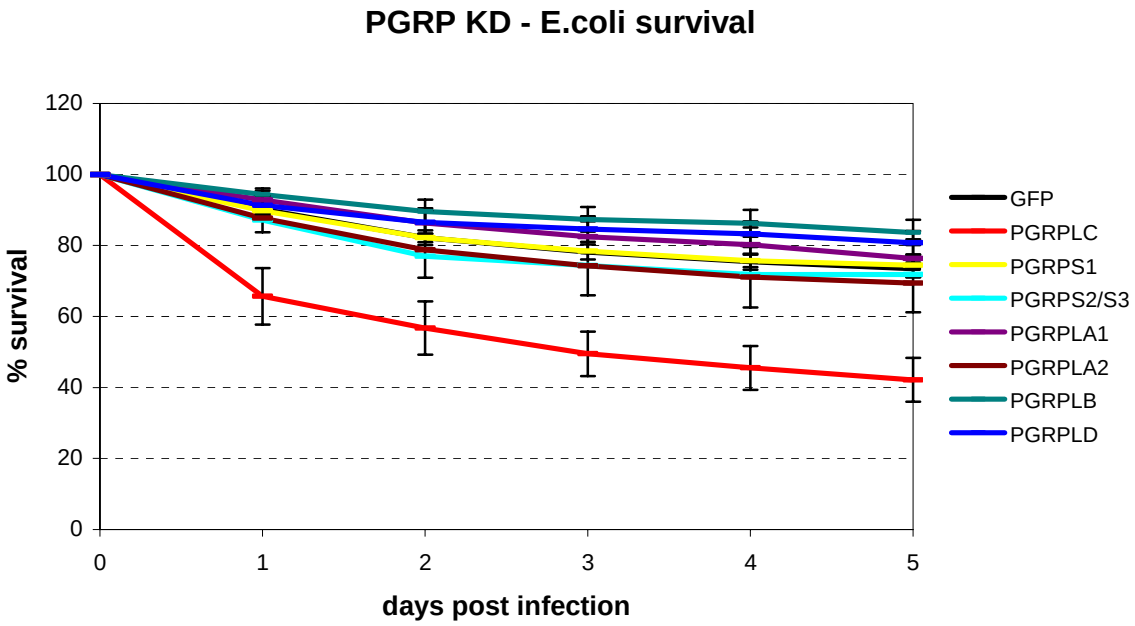


Fig. 4.2.4 Survival of adult female mosquitoes after gene KD by injection with dsRNA (PGRPLC, *PGRPS1*, *PGRPS2/3*, *PGRPLA1*, *PGRPLA2*, *PGRPLB* or *PGRPLD* or control GFP) and infection of *E. coli*. Error bars represent the Standard Error.

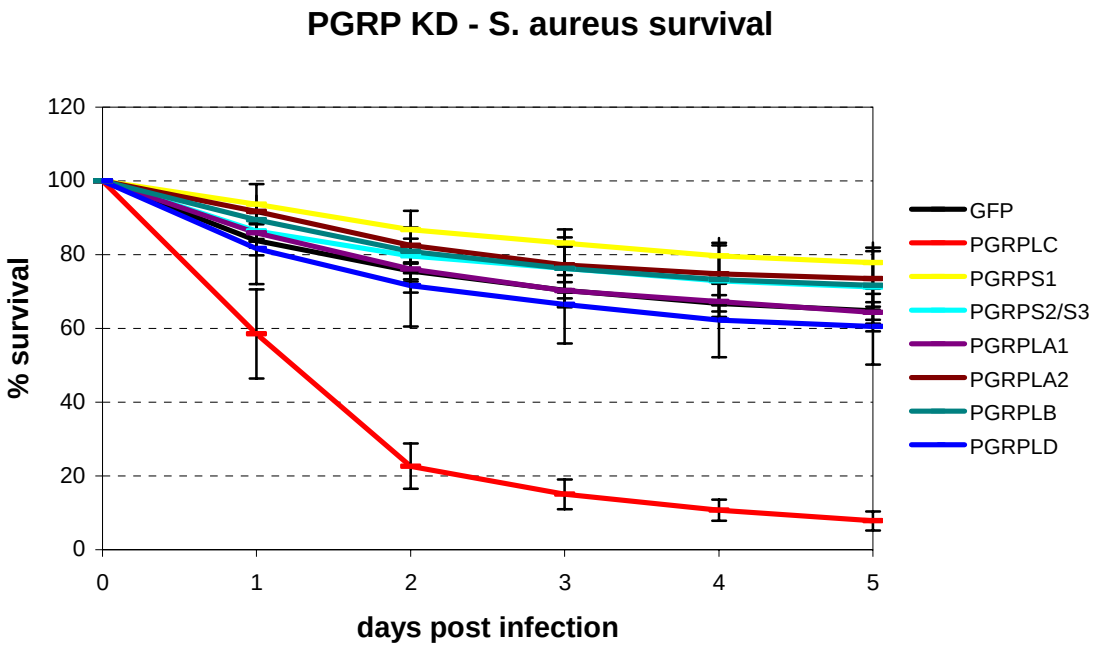


Fig. 4.2.5 Survival of adult female mosquitoes after gene KD by injection with dsRNA (PGRPLC, *PGRPS1*, *PGRPS2/3*, *PGRPLA1*, *PGRPLA2*, *PGRPLB* or *PGRPLD* or control GFP) and infection of *S. aureus*. Error bars represent the Standard Error.

Survival of Drosophila PGRP-LC *and* PGRP-SA^{seml} *mutants after bacterial infection.*

The results of the bacterial survival assay after gene KD by RNAi appeared to contradict the published results obtained in *D. melanogaster*. However, the experimental setups are different. Forward genetics was used to obtain the results in the fruit fly and the readout was mainly AMP reporter genes. In *Anopheles*, however, survival assays were used. For that reason we followed the experimental setup of *Anopheles* in a series of *Drosophila* experiments, to be able to compare the two species.

Mutant lines for the *PGRP-LC* (Gottar, Gobert et al. 2002) and the *PGRP-SA^{seml}* (*semmelweiss*) (Michel, Reichhart et al. 2001) were kindly provided by Dr. Julien Royet, IMBC du CNRS, Strasbourg (present address: Université de la Méditerranée, Marseille, France).

The white$^-$ mutant line that was used as control was obtained from Blades Biological Ltd (http://www.blades-bio.co.uk/drosophila_and_equipment.htm).

Opportunistic pathogens might conceivably overwhelm an immune compromised fruit fly when the cuticular barrier is breached by injection and the fly is thus further weakened. However, such complications were not observed: the differences between the 3 *Drosophila* strains were only minor in preliminary tests with saline solution injections (1x PBS) (Fig. 4.2.7).

As described in chapter 3 (Materials & Methods), two to three day old flies of the various strains were injected with 69 nl of *E. coli* (at OD_{600}=0.1) or *S. aureus* (at OD_{600}=0.01) and their survival monitored over five days (Figs 4.2.8 and 4.2.9, respectively). After *E. coli* injection, the *PGRP-LC* mutant flies were the only ones to show a significant decrease in survival. The white$^-$ and *PGRP-SA^{seml}* mutant flies were not significantly affected by *E. coli* infection.

However, both the *PGRP-SA^{seml}* and the *PGRP-LC* mutant flies showed a significant decrease in survival after *S. aureus* infection (>90% dead after 5 days). Though survival of the white$^-$ flies was also reduced, they were not as affected as the *PGRP-SA^{seml}* and the *PGRP-LC* mutant flies (Fig. 4.2.9). These results suggest that, similar to its *Anopheles* counterpart, the *Drosophila PGRP-LC* is essential for resistance to both bacterial infections.

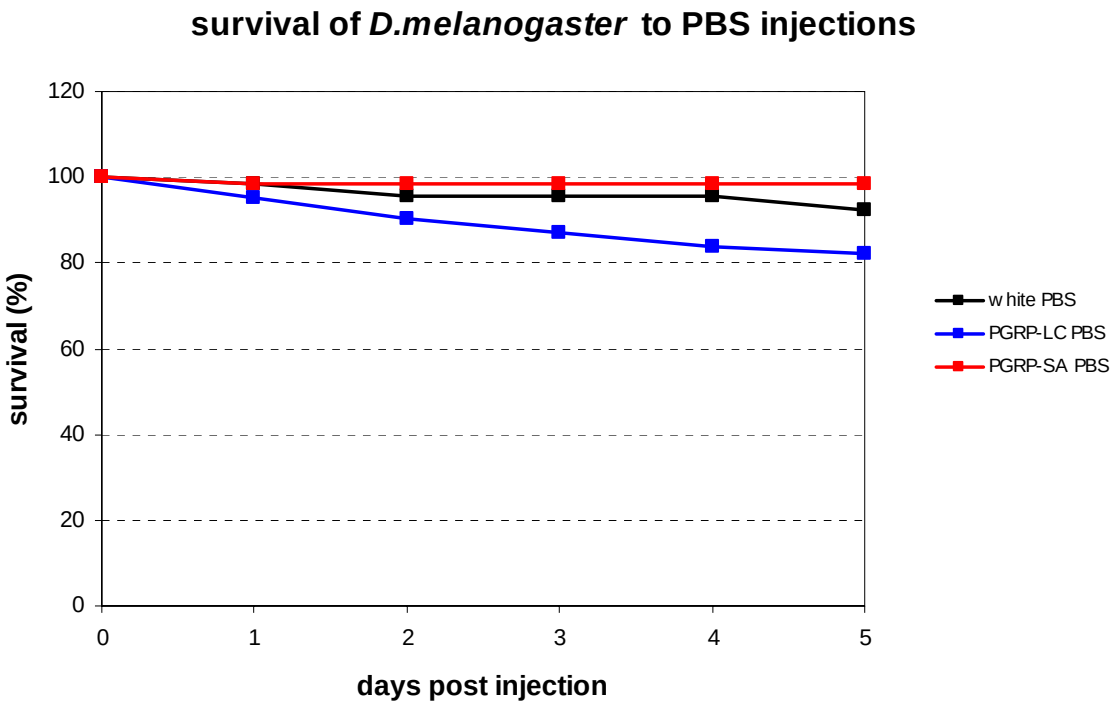


Fig 4.2.7 Survival over time of the *D. melanogaster* mutant strains white^{-}, *PGRP-LC*$^{-}$ and *PGRP-SA*^{seml} after PBS injections.

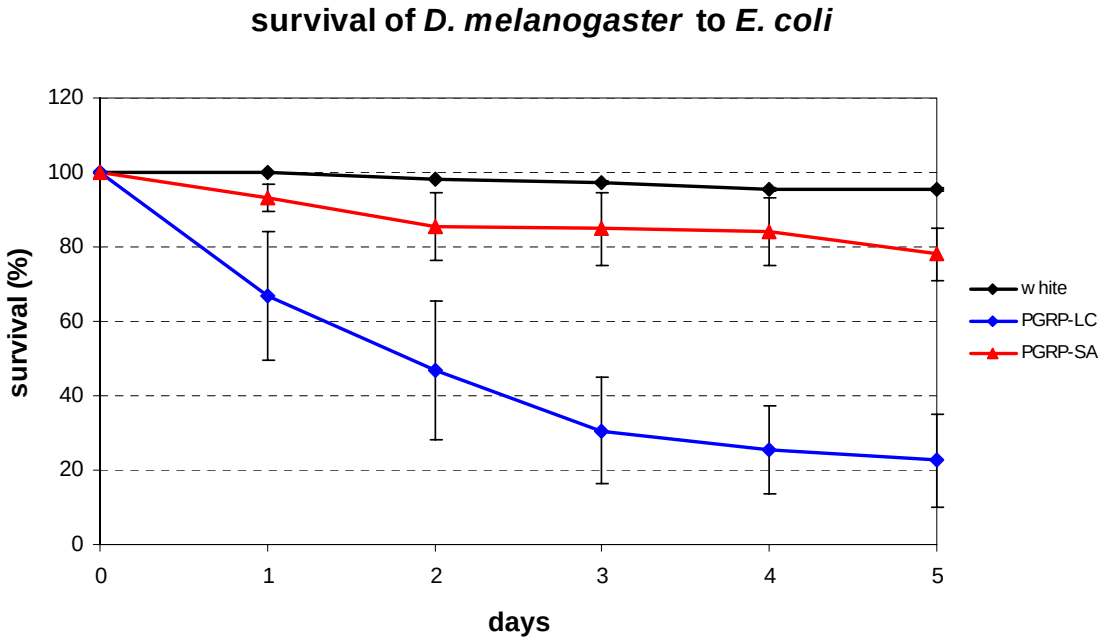


Fig. 4.2.8 Survival after Gram- *E. coli* infections of the *D. melanogaster* mutant strains white^{-}, *PGRP-LC*$^{-}$ and *PGRP-SA*^{seml}. Error bars represent the Standard error.

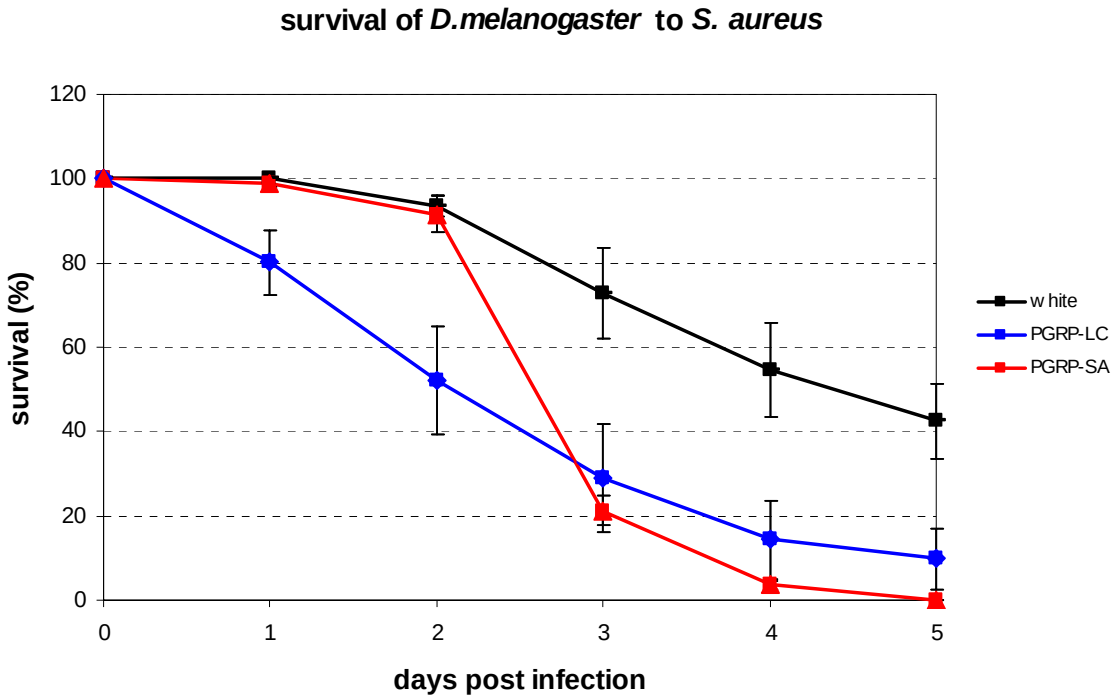


Fig. 4.2.9 Survival after Gram+ *S. aureus* infections of the *D. melanogaster* mutant strains white^{-}, *PGRP-LC*$^{-}$ and *PGRP-SA*^{seml}. Error bars represent the Standard error.

Is the Drosophila semmelweis *mutation rescued by* PGRPS1?

The predicted full length ORF of *PGRPS1* was amplified with the primers #9.1.54/55 (Table 9.1.1) and cloned into the pUAST *Drosophila* transformation vector using the polylinker restriction enzyme sites EcoRI and XbaI. The construct was sent to our collaborators at the CNRS, IBMC in Strasbourg, France. The plan was to transform the *D. melanogaster semmelweis* mutant of *PGRP-SA* and examine whether the mosquito ortholog *PGRPS1* can rescue the mutant phenotype. The *semmelweis* mutant is a single amino acid change of cysteine 80 into a tyrosine. This cysteine forms the structurally important disulfide bridge in Figure 4.2.1 (disulfide bridge 'number 2' indicated with green connector lines). At the time of writing this thesis, this work had not yielded conclusive results.

PGRP KD and **Plasmodium** *infection.*

As a first approach, the relevance of some *PGRPs* to malaria infection of the mosquito were tested by RNAi-based gene knockdown and subsequent infection of mosquitoes with *P. berghei* as described in chapter 3 (Materials & Methods). We used two transgenic *P. berghei* strains expressing GFP at the ookinete and early oocyst stages (Vlachou, Zimmermann et al. 2004) and all developmental stages (Franke-Fayard, Trueman et al. 2004). The mosquitoes used were from the susceptible G3 strain. Five to seven days post malaria infection, the mosquitoes were dissected and the number of oocysts and melanized ookinetes per midgut was counted. The results are presented as a percentage relative to the *dsGFP* injected control group of mosquitoes in Figure 4.2.6.

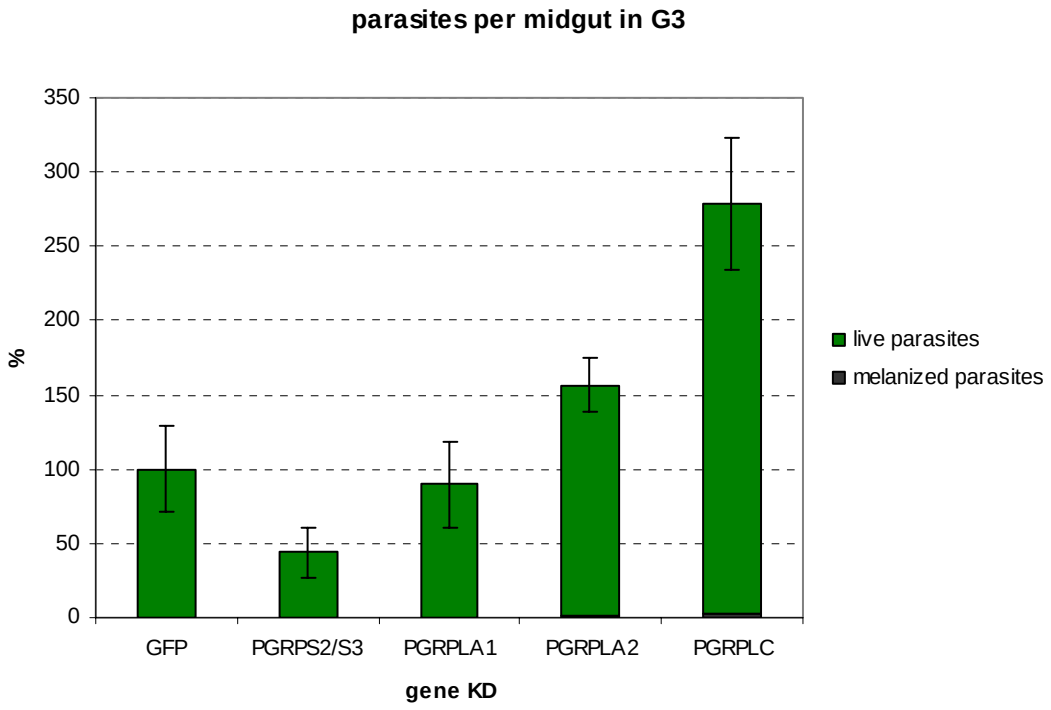


Fig. 4.2.6 Differences in G3 parasite numbers per midgut after gene KD. Green colored fractions represent live oocysts and black colored fractions melanized ookinetes. PGRPS2/S3, 22 midguts (P(KS-Test)=0.014, P(t-test)=0.096); PGRPLA1, 17 midguts (P(KS-Test)=0.19, P(t-test)=0.); PGRPLA2, 71 midguts (P(KS-Test)= 0.01, P(t-test)= 0.022); PGRPLC, 32 midguts (P(KS-Test)= 0.003, P(t-test)= 0.0005). Error bars are standard errors.

The KD of *PGRPLA2* and *PGRPLC* caused a statistically significant increased the parasite load per G3 midgut by ~56.6% and ~178%, respectively. In contrast, the co-silencing of *PGRPS2* and *PGRPS3* resulted in a reduction of parasite load by 55.9%. KD of *PGRPLA1* had no effect on the parasite development.

4.3 THE *PGRPLC* GENE CLUSTER

The previous sections indicated that, among the PGRPs in *A. gambiae*, only PGRPLC has an effect on the mosquitoes defense to bacterial infection, and only PGRPLC and PGRPLA seemed relevant to malaria infection of the mosquito. Both genes are clustered together in the genome. We have analyzed the *PGRPLA/LC* locus in detail, focusing more on *PGRPLC*. In *D. melanogaster PGRP-LC* is a pattern recognition receptor implicated in initiation of the *Imd* pathway (Choe, Werner et al. 2002; Gottar, Gobert et al. 2002; Ramet, Manfruelli et al. 2002; Werner, Borge-Renberg et al. 2003).

PGRPLC *gene architecture.*

The *Anopheles PGRPLC* gene encodes three alternative PGRP domains (Fig. 4.3.1a & Fig. 4.3.4) that can be alternatively used for the production of three protein isoforms. Furthermore, the PGRP domains have a very modular build: the fourth exon of the gene is common to all three PGRP domains (Fig. 4.3.1a pink exon), and the introns 5, 7 and 9 are at the exact same relative location in the sequence of all three PGRP domains (Fig. 4.3.1a & Fig. 4.3.3 grey solid triangle). This modular build suggests, that production of additional PGRP domains by recombination of the existing exons through splicing may be possible (Fig. 4.3.6 dashed splicing lines). This is an advantage that the *Anopheles PGRPLC* gene build has over that of the *Drosophila PGRP-LC*. In *Drosophila* the splicing has to be more rigid as the second splice site within the PGRP domain is lacking (Fig. 4.3.1b).

It is also worth mentioning that although both the fly and mosquito *PGRPLC* genes have three PGRP domains, these domains are not homologous: they are more similar within

than across species (Fig 4.2.2). This suggests that the common ancestor of *A. gambiae* and *D. melanogaster* probably had a single-PGRP-domain *LC* gene that duplicated independently in each species. This hypothesis is also supported by the exon-intron structure of the gene (Fig. 4.2.2, 4.2.1, 4.3.3), and the analogous introns (Fig. 4.3.2). In each case, the similarities between domains within a species is higher than across species.

-

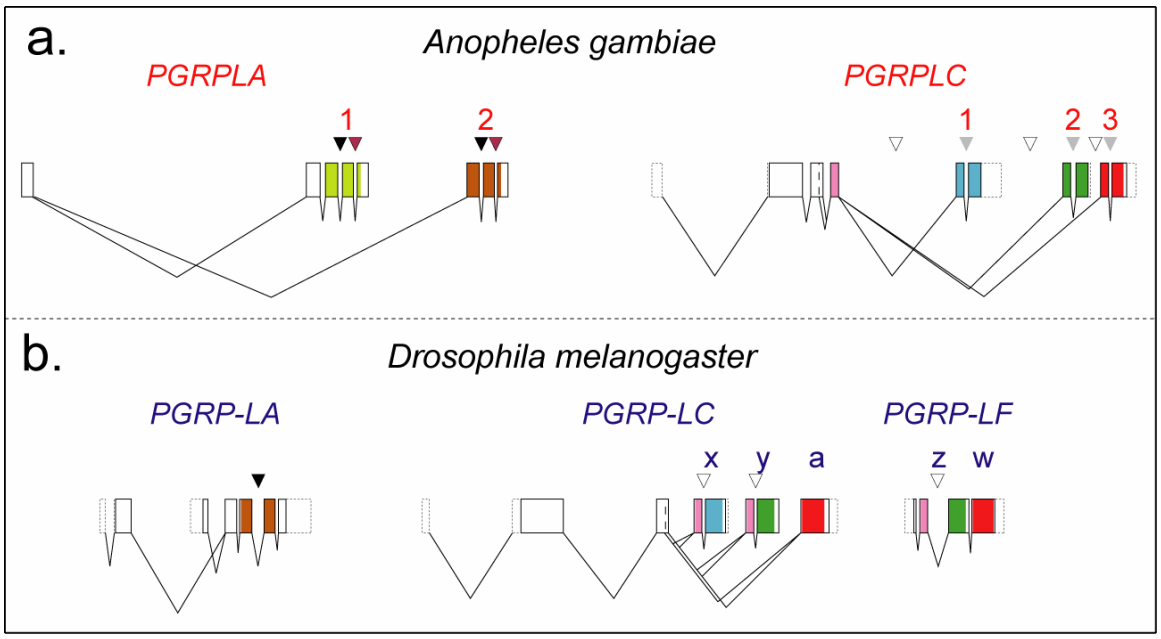


Fig. 4.3.1 The *PGRPLC* gene cluster architecture in **a.** *A. gambiae* (21.1 kb) (red) and **b.** its orthologous cluster in *D. melanogaster* (17.1 kb) (blue). Drawn to scale. Each box represents an exon (dash-framed boxes indicate untranslated exons) and connecting lines represent spliced out introns. Colored parts of exons contribute to the PGRP domains. Identical triangles above the introns indicate analogous locations of splice sites (compare to Fig. 4.3.3). (modified from (Christophides, Zdobnov et al. 2002))

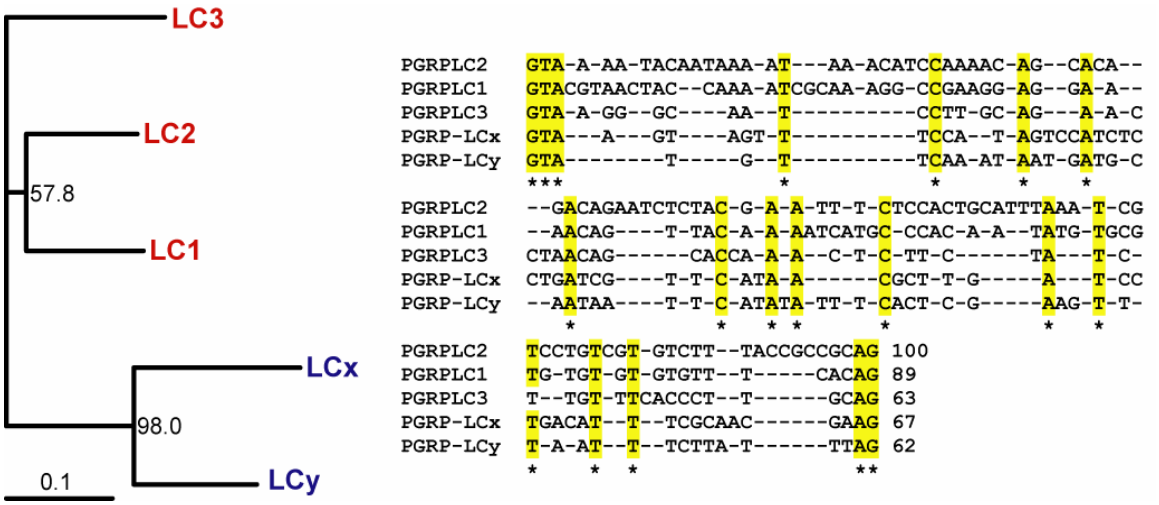


Fig. 4.3.2 Alignment and phylogenetic tree of the second introns of the *Anopheles PGRPLC* domains and the first introns of the *Drosophila PGRP-LC* domains. Numbers at the nodes are bootstrap percentages.

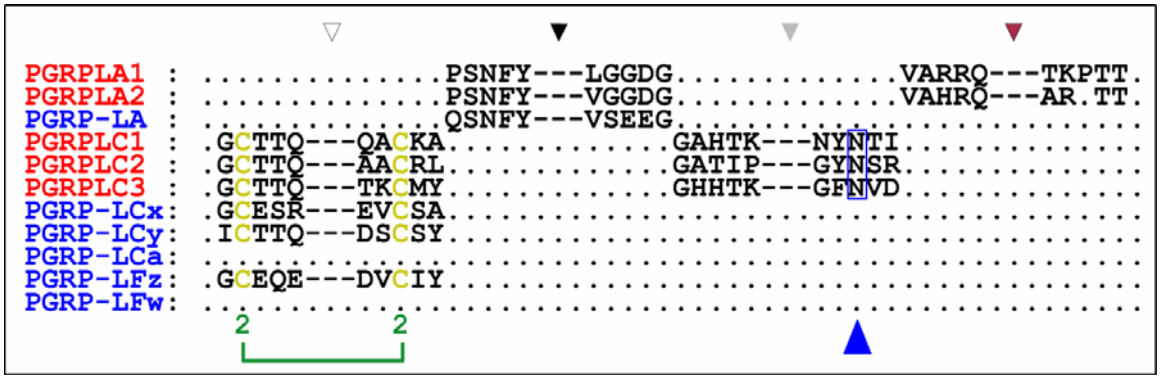


Fig. 4.3.3 Conservation across species of the *PGRPLC* and *-LA* gene splice sites of the PGRP domain in *A. gambiae* (red) and *D. melanogaster* (blue). Triangles above the introns indicate analogous locations of splice sites (triangles correspond to the ones in Fig. 4.3.1). The numbers connected with a green line highlight the Cysteine residues that form a disulfide bridge (compare to Fig. 4.2.1). The blue triangle indicates the location of a catalytically important residue in T7 lysozyme (modified from (Christophides, Zdobnov et al. 2002))

PGRPLC *splicing*.

We performed Reverse Transcriptase PCR (RT-PCR) on a cDNA pool prepared from bacterial challenged and unchallenged adult mosquitoes using a variety of primers distributed along the *PGRPLC* gene to examine the hypothesis that novel PGRP domains might be created by recombining the existing exons through alternative splicing. The design of PCR

primers was based on a gene structure that was derived from alignment of t he genomic sequence with all available EST sequences (http://web.bioinformatics.ic.ac.uk:8080/AnoEST/anoest.php) as shown in Fig. 4.3.4.

Most PCRs resulted in multiple products, depending on how the primers were combined (Fig. 4.3.5). The PCR products were separated by agarose gel electrophoresis, and purified with the QIAquick Gel Extraction Kit (QIAGEN Ltd.) and sequenced.

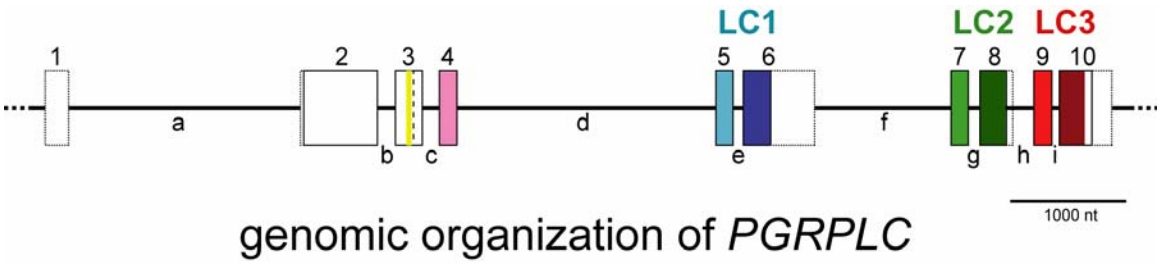


Fig. 4.3.4 Genomic organization of the *Anopheles PGRPLC* gene, drawn to scale. Each box represents an exon (dash-framed boxes indicate untranslated exons). Exons are numbered and introns are assigned letters. The yellow line indicates the position of the transmembrane domain. The PGRP domains LC1 (exons 4,5,6), LC2 (exons 4,7,8) and LC3 (4,9,10) are encoded by colored exons.

In none of the performed PCRs could the proposed hybrid *PGRPLC* PGRP domains be detected. In fact, the only time that the linear order of the PGRP-domain exons was disrupted by splicing, the resulting sequences encode for early stop codons and cannot produce functional multidomain proteins (Fig. 4.3.5 j & g).

However, the alternative use of a 75 nucleotide (nt) cassette (Seq. 9.2.11 & Seq. 9.2.12) in exon 3 was detected (Fig. 4.3.5 d, e, h). It appears that transcripts for the LC1 and LC2 PGRP domain are always spliced together with the short version of exon 3 (Seq. 9.2.12) whereas the LC3 PGRP domain can be combined with either version of exon 3. The significance of this is unclear, however, it is worth noting that the transmembrane domain is situated just before this cassette, and that the *D. melanogaster PGRP-LC* gene has a

homologous alternative cassette at the exact same position (Fig. 4.3.1, Seq. 9.2.14, Seq. 9.2.15).

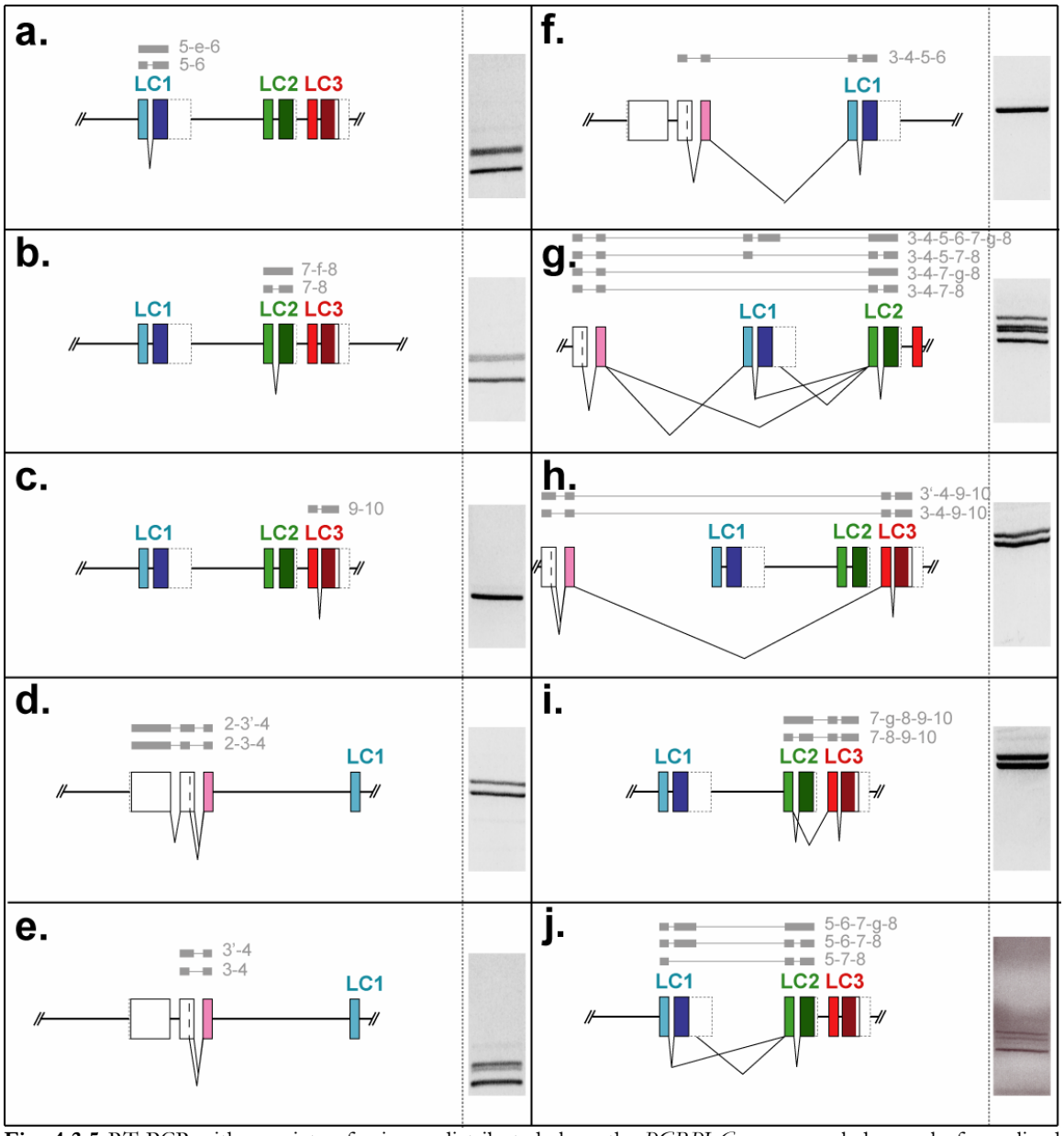


Fig. 4.3.5 RT-PCR with a variety of primers distributed along the *PGRPLC* gene revealed a pool of unspliced transcripts. Every subpanel (a.-j.) features a different PCR primer combination and the corresponding section of the *PGRPLC* gene model in Fig. 4.3.4. The grey bars above these sections represent the sequenced RT-PCR products shown on the right of each subpanel. The numbers and letters next to the gray bars spell out the intron & exon combination for clarity. These numbers and letters are derived from Fig. 4.3.4. A summary of this figure can be found in Fig. 4.3.6.

RT-PCR on the *LC1* and *LC2* PGRP domains revealed a significant amount of unspliced transcripts (Fig. 4.3.5 a, b). This was not due to contamination with genomic DNA as the samples of total RNA were treated with DNAse before reverse transcription. In addition, the control PCR reaction with primers targeting the ribosomal *S7* gene did not produce a double band even though the primers (S7-A & S7-B) span an intron. These data suggest the presence of a pool of nuclear unspliced or partly spliced transcripts that might be used for the production of protein coding transcripts upon a need for specific PGRPLC proteins (see below). Only for the LC3 PGRP domain could no unspliced transcript be detected (Fig. 4.3.5 c).

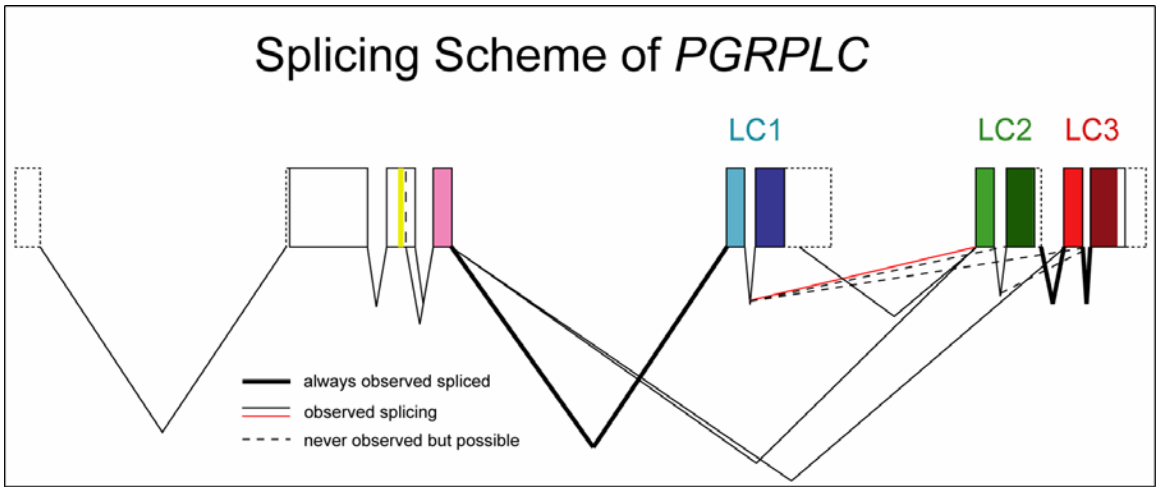


Fig. 4.3.6 Splicing of the *Anopheles PGRPLC* gene. Drawn to scale. Each box represents an exon (dotted boxes indicate untranslated exons) and connecting lines represent spliced out introns. The yellow line indicates the position of the transmembrane domain and the other colored parts of exons contribute to the PGRP domains.

PGRPLC *domain microarray*.

To determine if the *Anopheles PGRPLC* ortholog is transcriptionally activated during immune challenge, we constructed DNA microarrays containing probes representing the different domains of PGRPLC (PGRP domains of other mosquito PGRP genes were also included). This design allowed us to examine whether the various *PGRPLC* exons can respond to different challenges, leading to production of different protein isoforms via alternative splicing (Dimopoulos, Christophides et al. 2002). For the hybridization we challenged a hemocyte-like cell line, 4a3B (Muller, Dimopoulos et al. 1999) with paraformaldehyde-fixed *E. coli* and *S. aureus* (OD 0,05), PGN (10 μg/ml) and H_2O_2 (2 μM), the latter serving as control. Duplicate RNA samples were collected 12 hrs after challenge and hybridized to arrays as described (Dimopoulos, Christophides et al. 2002). RNA prepared from naïve cells was used as reference.

The results indicated that PGRP genes are differentially regulated in response to the different challenges (Christophides, Zdobnov et al. 2002) (Fig. 4.3.7). Interestingly, the three PGRP domains of PGRPLC appeared to respond differently to the diverse challenges, suggesting a role for alternative splicing in the response to infection. PGRPLC2 was also regulated transcriptionally when exposed to oxidative stress, but PGRPLC1 and LC3 were not. Dissecting the role of the alternatively spliced PGRP gene products may elucidate a novel step of innate immunity regulation at the level of RNA splicing.

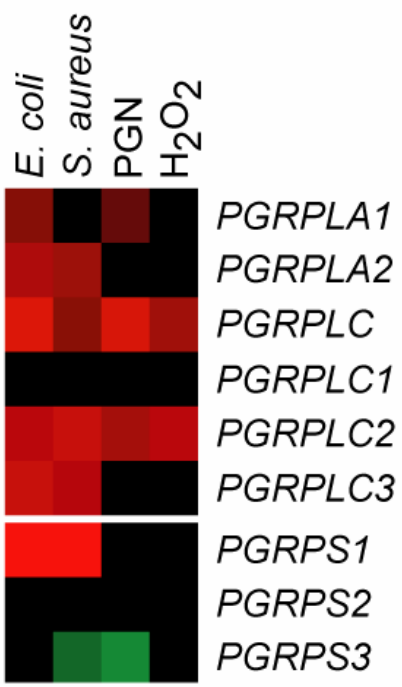


Fig. 4.3.7 Expression profiles of PGRP isoforms of female adult mosquitoes challenged with *E. coli, S. aureus*, peptidoglycan (PGN), and H_2O_2. Color intensities indicate fold regulation relative to reference (naïve) cells (see Chapter 3 "Materials & Methods"). Regulation values below 1.5-fold are masked. (modified from (Christophides, Zdobnov et al. 2002))

PGRPLC surface modeling.

The following results were obtained with the help of Dr. Bogos Agianian in the context of our collaboration on the *PGRP* structural models.

The sequence of the PGRP domains of *Anopheles* and *Drosophila* were modeled onto the 3D structure of the *Drosophila PGRP-LB* domain (Kim, Byun et al. 2003) (Fig. 4.3.8). PGRP sequences were firstly aligned with Tcoffee (http://igs-server.cnrs-mrs.fr/Tcoffee/tcoffee_cgi/index.cgi) (Notredame, Higgins et al. 2000) to allow for determination of a conservation score with ConSurf (http://consurf.tau.ac.il/) and based

on these scores, to then project each sequence onto the 3D model structure of PGRP-LB. The .pdb models of the PGRP domains obtained in this manner were then viewed with protein structure viewers like PyMOL (http://pymol.sourceforge.net/) or the Protein Explorer (http://www.umass.edu/microbio/chime/pe/protexpl/frntdoor.htm).

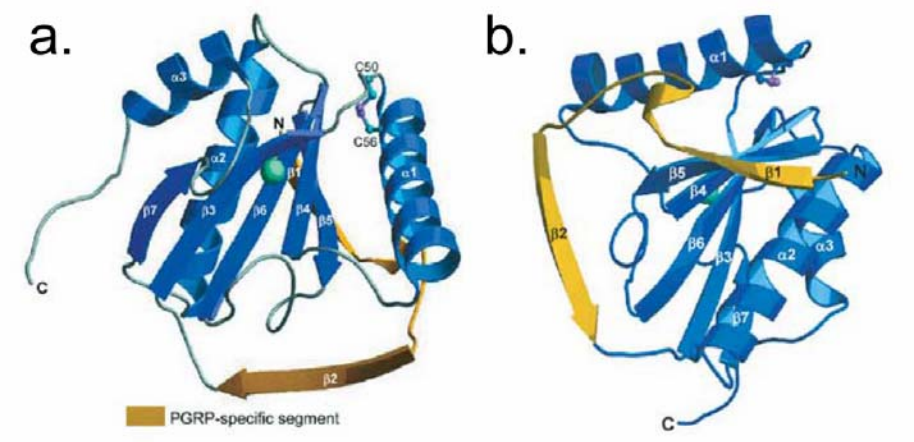


Figure 4.3.8 Ribbon diagram of structure of PGRP-LB. The PGRP-specific segment at the N terminus, colored in gold, is absent in T7 lysozyme. The single disulfide bond is shown as a 'ball-and-stick'. The bound zinc ion is shown as sphere. **a.** 'front' face of the PGRP domain. The isolated strand β2 adopts the typical backbone conformation of a β-strand. **b.** 'back' face of the PGRP domain. The PGRP-specific segment (in gold) is absent from T7 lysozyme. The segment adopts mainly extended β-strand and loop structures. (Figure modified from (Kim, Byun et al. 2003)).

As can be seen in Fig. 4.3.9 a, all three exons encoding the PGRP domains of PGRPLC isoforms contribute to the PGN binding site. The first exon is common for all three PGRP domains. This makes the prospect of diversification of PGRPs by potential exon-recombination, as suggested by the *Anopheles* gene architecture (Fig. 4.3.4), even more intriguing. In *Drosophila* (Fig. 4.3.9 b) each PGRP domain has their own first (red) exon and thus constitute three entirely distinct PGRP domains.

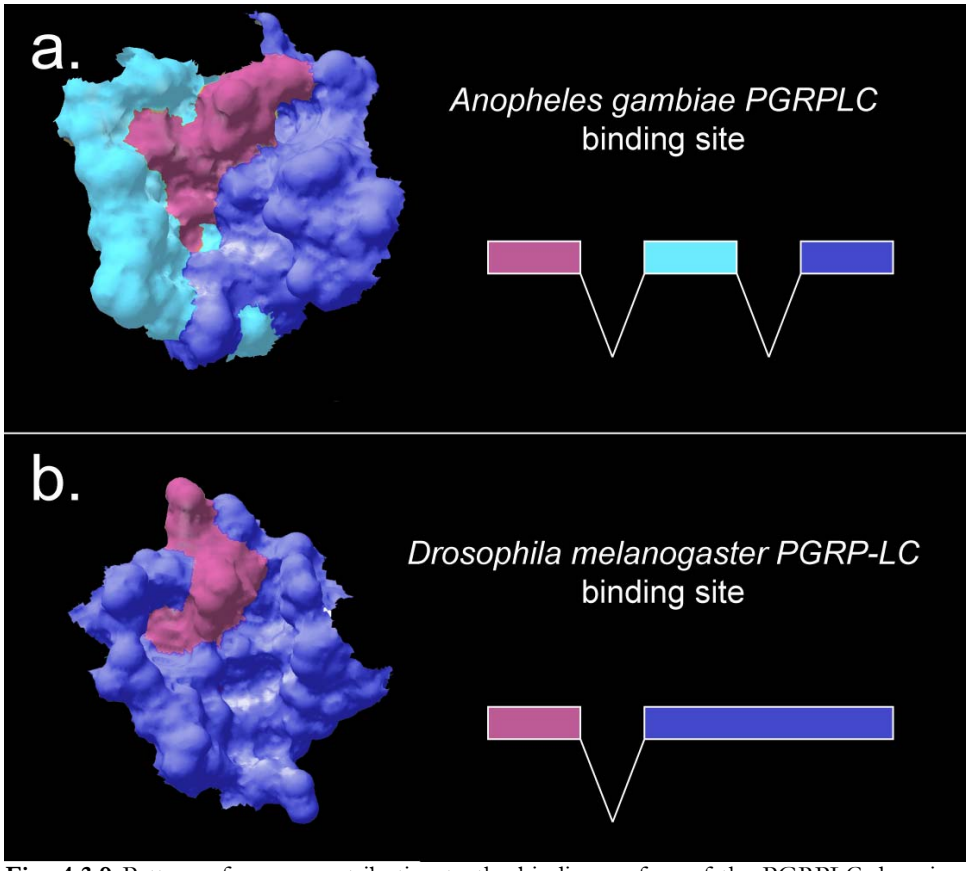


Fig. 4.3.9 Pattern of exons contributing to the binding surface of the PGRPLC domains. The colors represent the exons coding for the respective region of the surface. **a.** Surface of the *A. gambiae* PGRPLC1 domains. The first, common exon is colored pink, the second and third exons are colored light and dark blue, respectively (analogous to the colors in Fig. 4.3.4). **b.** Surface of the *D. melanogaster* PGRP-LC1 domains. Pink represents the first exon and the second exon is colored dark blue.

As mentioned previously, because of separate exon duplication events in the two lineages, the PGRP domains of the *Anopheles* and *Drosophila* PGRPLC protein isoforms are more closely related with other domains encoded by the same gene than with the analogous domains of the orthologous gene in the other species (Fig. 4.2.2). We attempted to visualize the consequences of this phenomenon by color coding sequence conservation and projecting it onto the surface model of the first PGRPLC domain of each species, PGRPLC1 and -LCx (Fig. 4.3.10).

Since the 'golden' PGRP-specific β sheets (β1 & β2) in Fig. 4.3.8 are contributed by the first, common 'pink' exon (Fig. 4.3.4), they are colored dark blue (highly conserved) on the backside of *A. gambiae PGRPLC*, indicating 100% conservation (Fig. 4.3.10). The groove formed by these β sheets is highly hydrophobic and as such ideal for potential protein-protein interactions.

The protruding ridges of the PGN binding sites are highly variant among the three PGRPLC domains of either species. This might reflect the need for varying PGN binding properties for PGRPLC to accommodate PGN from a variety of microbial sources in the binding site.

The ability of a PGRP domain to bind PGN specific to Gram- bacteria (DAP-type PGN) is determined by specific interactions in the pocket encoded by the first exon of *PGRPLC*. This region is invariable in *Anopheles*, and thus all three PGRP domains should be able to bind DAP-type PGN. However, it is likely that one or more of the *A. gambiae PGRPLC* domains are also able to bind Lys-type (gram+) PGN, as has been shown for *Drosophila (Werner, Borge-Renberg et al. 2003; Kaneko, Golenbock et al. 2005).*

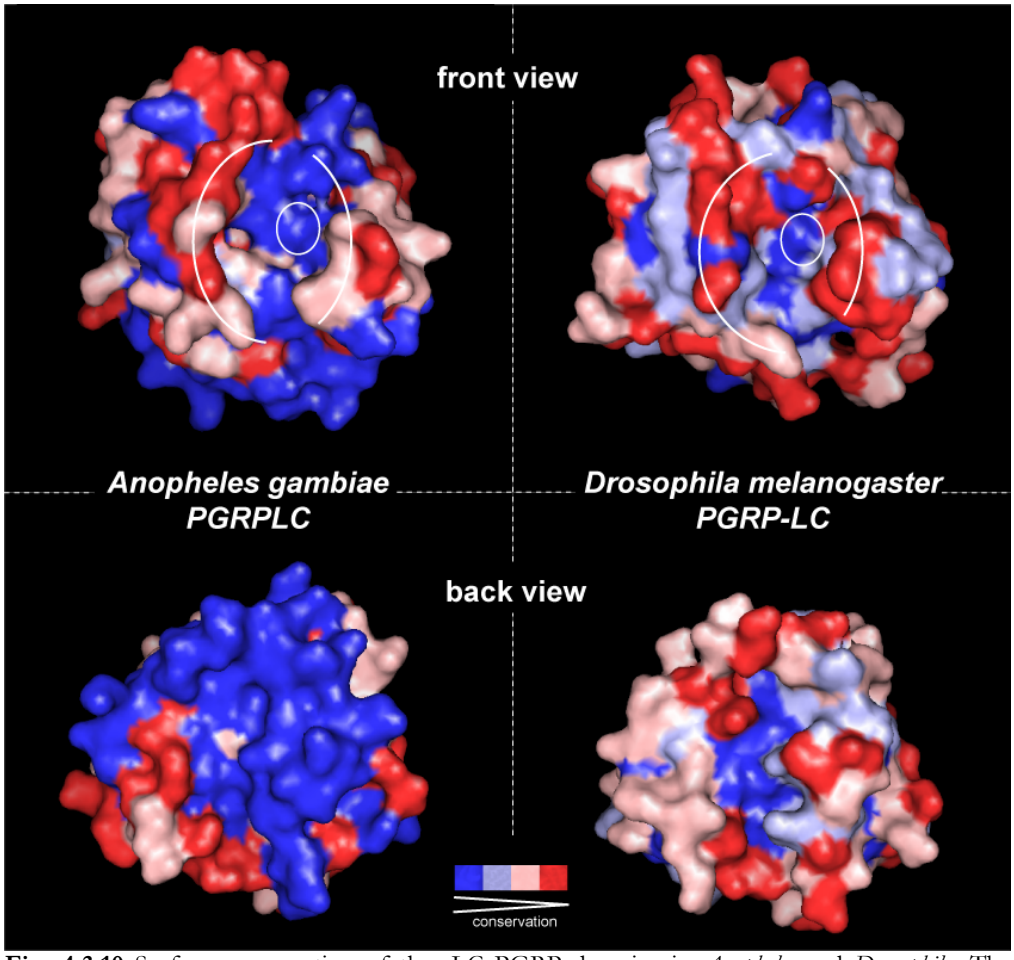


Fig. 4.3.10 Surface conservation of the -LC PGRP domains in *Anopheles* and *Drosophila*. The outer ellipse marks the ridge of the PGN binding site and the inner ellipse marks the catalytic Tyrosine68. Blue signifies greater and red lesser sequence conservation. The structure models of the first domains (LC1 and LCx) of each species have been used as scaffolds. 'front' refers to the PGN binding site, and 'back' refers to the opposite side; the relative orientation in space ('up' and 'down') of the models is the same as in Fig. 4.3.9 and Fig. 4.3.8.

PGRPLC1, 2 or 3 *specific KD in adult mosquitoes.*

The RNAi procedure was performed as described in chapter 3 (Materials & Methods) on 10 one to two day old female G3 mosquitoes. Their RNA was extracted four days later and used to assess the KD efficiency of specific PGRPLC domain transcripts by RT-PCR (Fig. 4.3.11) and quantitative realtime-PCR (qRT-PCR) (Fig. 4.3.12).

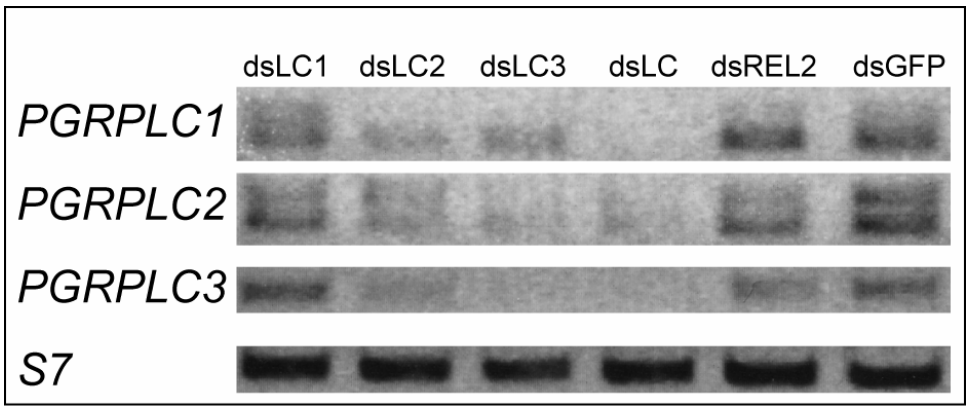


Fig. 4.3.11 RT-PCR on whole female G3 mosquitoes to determine *PGRPLC* mRNA levels after domain specific KD of *PGRPLC*. Columns: dsRNA used to inject adult G3 female mosquitoes for gene silencing. Rows: primers specific to the LC1, LC2, LC2 domains and the ribosomal protein S7 used for monitoring expression levels.

The targeted KD of *PGRPLC1* or *PGRPLC2* transcripts (Fig. 4.3.11 first and second columns) does not appear to be very effective according to the RT-PCR expression assays. However, KD of *PGRPLC3* and the entire *PGRPLC* gene (columns three and four) are effective. Intriguingly, there always seems to be a cross effect on the other domains, when one of them is targeted.

Apparent lack of effectiveness might be due in part to differences in the tissues of isoform origin in conjunction with differences in accessibility of injected dsRNAs to different tissues. To obtain more definite results, we used qRT-PCR as a better method of transcript quantification and also distinguished between expression levels in a specific organ, the midgut, the carcass (the remaining tissue after removal of head, wings, legs and gut), and in whole mosquitoes. The results (Fig. 4.3.12) confirmed the effectiveness and partial specificity of domain specific KD of *PGRPLC* by RNAi. The reduction in LC1 transcript levels after dsRNA-mediated silencing was in the range of 34% - 54% for KD based on the

dsLC1 RNA. LC2 and LC3 transcripts showed a specific apparent reduction following treatment by corresponding dsRNAs, to a range of 71% - 75% and 69% - 76%, respectively.

Overall, RNAi seems to function well in both midgut and carcasses, indicating that there, the circulating hemolymph is capable of delivering the injected dsRNA and possibly, the small interfering RNAs (siRNA) to the mosquito tissues.

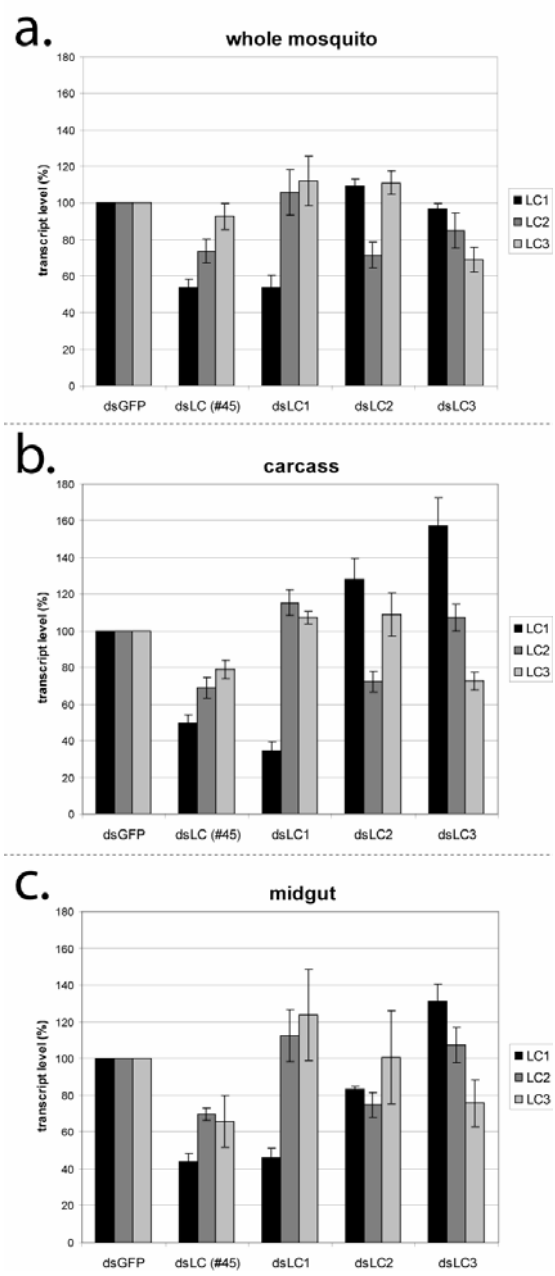


Fig. 4.3.12 PGRPLC domain transcript levels relative to ds*GFP* injected control levels as determined by qRT-PCR. **a.** qRT-PCR on cDNA from whole female mosquitoes. **b.** qRT-PCR on cDNA from female carcasses (no head, legs, wings and gut). **c.** qRT-PCR on cDNA from female midguts.

PGRPLC *KD and survival after bacterial infection.*

PGRPLC KD in one to two day old female G3 mosquitoes was performed as described in chapter 3 (materials & methods). Ampicillin resistant *E. coli* bacteria were grown to an OD of 0.7 to 1.0 (at λ=600nm), diluted to an OD of 0.01 in PBS of which 69 nl were then injected into the mosquito body cavity on day 4.

Survival of the mosquitoes was recorded over the next week (Fig. 4.3.13). Relative to the dsGFP injected control mosquitoes, *PGRPLC* KD mosquitoes showed a reduction in survival, with the most severe effect noted in silencing the *PGRPLC3* domain KD as well as the entire *PGRPLC* gene.

Survival to *S. aureus* infection was examined in the same way as survival to *E. coli*, except that a higher concentration of OD_{600}=0.4 was injected into the female G3 mosquitoes (Fig. 4.3.14). While the KD of the *PGRPLC2* domain did not seem to have an effect on survival of the mosquitoes, KD of *PGRPLC3* as well as the entire *PGRPLC* gene did. The effect of *PGRPLC1* KD was less prominent (Fig. 4.3.14).

The possibility of the *PGRPLC* KD by itself accounting for increased mortality in the mosquito was explored in the experiment presented in Fig. 4.3.15. and could be refuted.

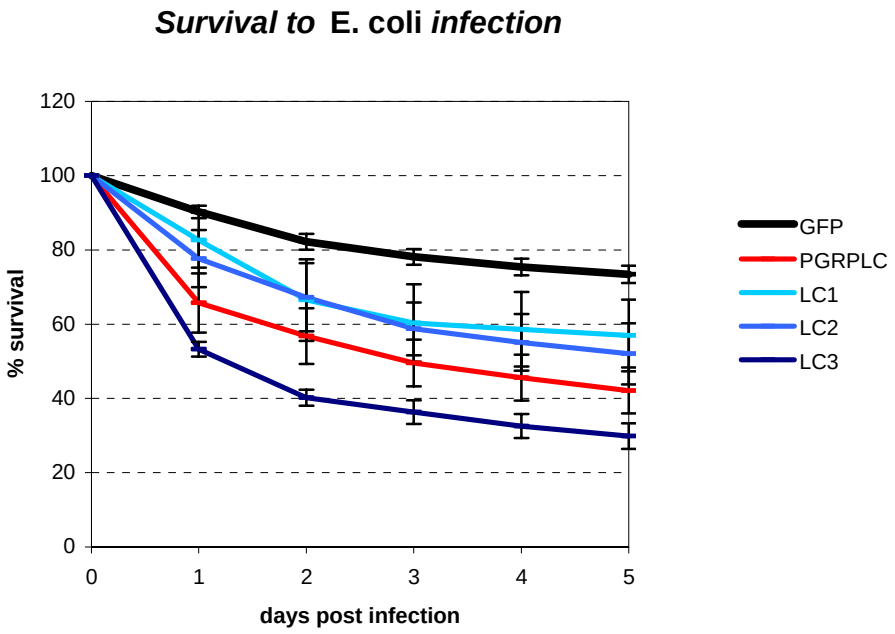


Fig. 4.3.13 Survival after *PGRPLC* KD and challenge with *E. coli* (69 nl of OD_{600}=0.01 in PBS). Number of repetitions per KD: *PGRPLC* 7x; *PGRPLC1* 4x; *PGRPLC2* 3x; *PGRPLC3* 4x.

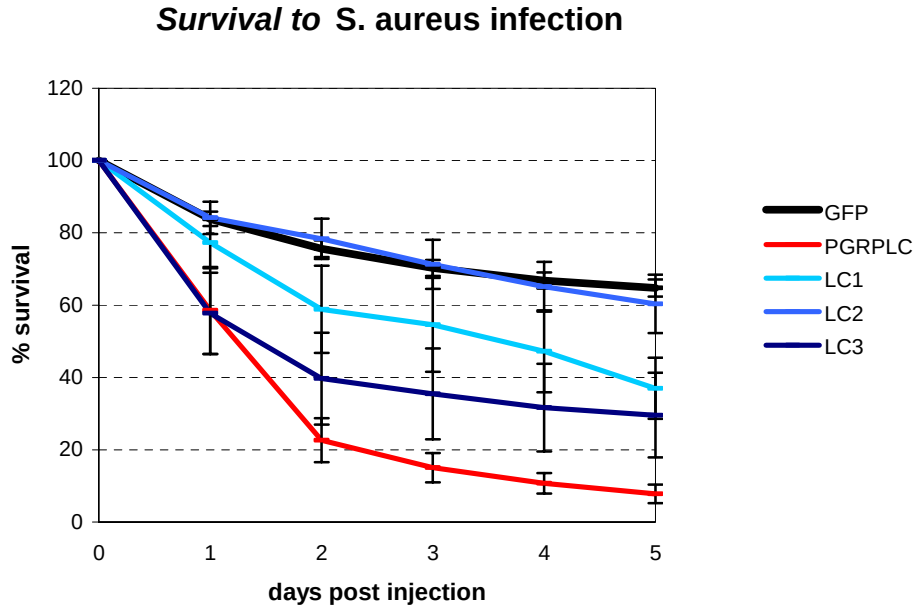


Fig. 4.3.14 Survival after *PGRPLC* KD and injection with *S. aureus* (69 nl of OD_{600}=0.4 in PBS). Number of repetitions per KD: *PGRPLC* 6x; *PGRPLC1* 3x; *PGRPLC2* 4x; *PGRPLC3* 5x.

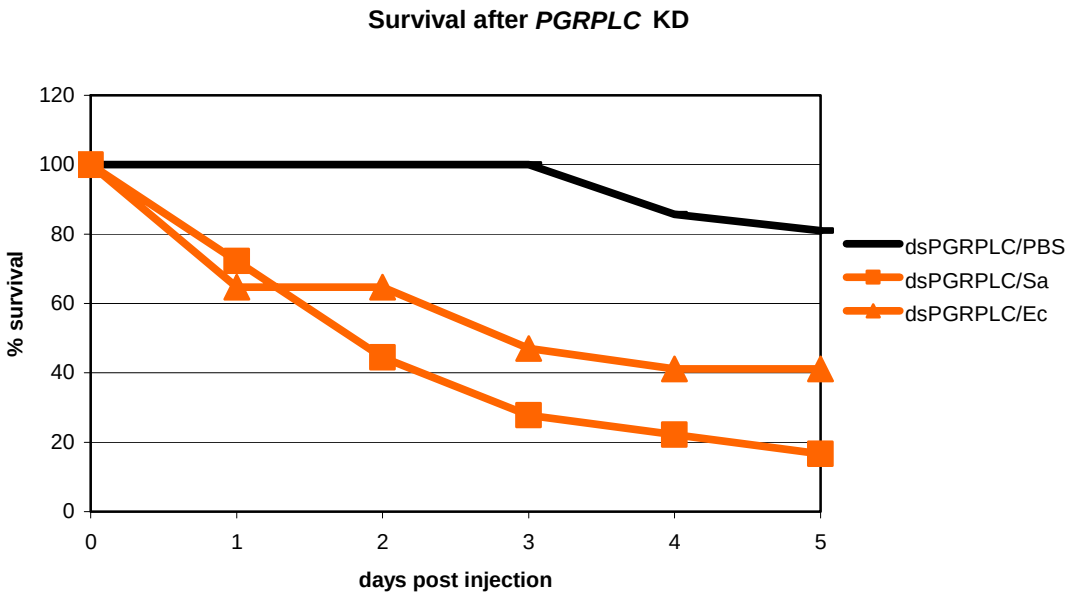


Fig. 4.3.15 Comparison of mosquito survival after *PGRPLC* KD and injection with either PBS, S. *aureus* or E. *coli* four days later.

PGRPLC *KD and* Plasmodium *infection.*

The *PGRPLC* KD and *P. berghei* infection were performed as described in chapter 3 (Materials & Methods). Two mosquito strains were used: the susceptible G3 strain and the refractory L3-5 strain (Collins, Sakai et al. 1986). Five to seven days post malaria infection, the mosquitoes where dissected and the number of oocyst and melanized ookinetes per midgut was counted. The results are presented as a percentage relative to the *dsGFP* injected control group of mosquitoes in Figure 4.3.16.

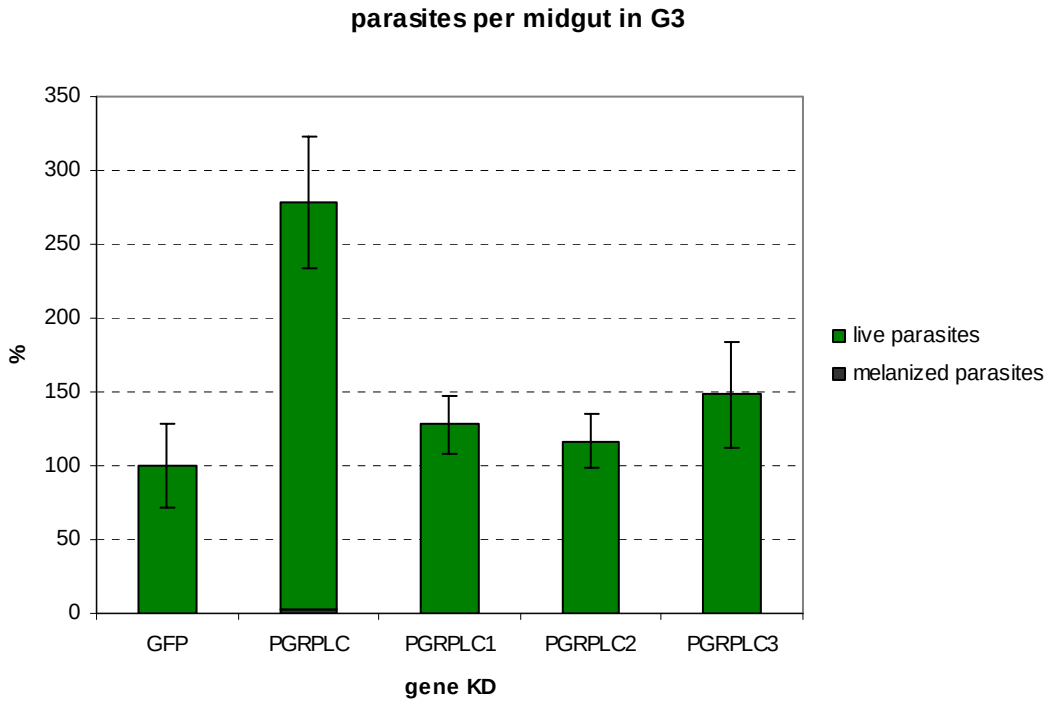


Fig. 4.3.16 *PGRPLC* KD and *P. berghei* infection in susceptible G3 mosquitoes. The number of dissected midguts are 32 for PGRLC, 49 for PGRPLC1, 45 for PGRPLC2 and 24 for PGRPLC3.

Whereas the single domain KDs did not appear to have a significant effect, the whole gene KD resulted in a 2.5-fold increase of the parasite load per midgut. According to the Kolmogorov-Smirnov (KS-)Test and the Student's t-Test, only the increase in parasite numbers due to the whole *PGRPLC* gene KD is significantly distinct from that of the *GFP* control (p values of 0.003 and 0.0005, respectively).

The KD of the entire *PGRPLC* gene also caused a small but significant increase of 2.53% in melanized ookinetes relative to the dsGFP injected control.

The results obtained in the refractory mosquito strain L3-5 (Collins, Sakai et al. 1986) showed that the KD of *PGRPLC* does not alter the capacity of these mosquitoes to kill

parasites by melanizing the ookinetes and preventing them from developing into oocysts, although this phenotype gets partially reversed by the KD of the whole *PGRPLC* gene (3.35% increase relative to the *dsGFP* injected control) (Fig. 4.3.17).

However, we observed a drastic increase in parasite numbers which was statistically significant only in the KD of the whole *PGRPLC* gene (181.07% increase relative to the control, with a Student's t-Test p-value of 0.031) and in the *PGRPLC3* KD (211.42% increase relative to control, with a Student's t-Test p-value of 0.046). The slight decrease in parasite numbers after *PGRPLC1* KD is not statistically significant.

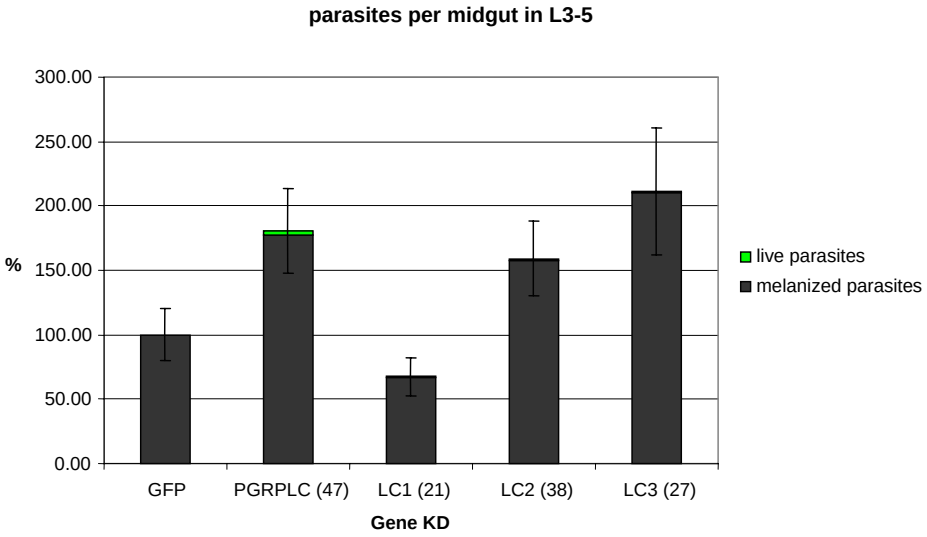


Fig. 4.3.17 PGRPLC KD and *Plasmodium* infection in refractory L3-5 mosquitoes.
The numbers of dissected midguts are given in brackets.

***PGRPLC and S1* antibodies.**

Transcripts encoding the PGRPLC1, 2 and 3 and PGRPS1 proteins were amplified from an abdominal cDNA library (Dimopoulos, Casavant et al. 2000) with the primers #9.1.15 to #9.1.32 (for primer sequences refer to Table 9.1.1) and directly cloned into the pGEM-T easy vector (for characteristics of the vectors used refer to Chapter 10.4 'vector sequences & maps') and verified by sequencing. The inserts were then excised with NcoI and EcoRI to be cloned into the expression vectors pETM-11 and pETM-60, and their sequences were verified.

The constructs for PGRPLC1 (pEMT-11), PGRPLC2 (pEMT-60), PGRPLC3 (pEMT-11) and PGRPS1 (pEMT-11) were expressed in *E. coli* bacteria with support of Dr. Ario de Marco of the EMBL Protein Expression and Purification Core Facility (Fig. 4.3.18 a.).

Antibodies were raised against the expressed PGRP domain of *PGRPS1* and unique sequences of the PGRPLC isoforms (see Fig. 4.3.6) at the EMBL Monoclonal Antibody Core Facility (MACF) in Monterotondo, Italy and the EMBL Animal House Unit for polyclonal antibody production in rabbits. Only PGRPLC1 and PGRPLC2 domains yielded monoclonal antibodies; the PGRPLC3 and PGRPS1 produced proteins that proved to be insoluble in 1M Urea, which is the highest possible concentration for antigen injection.

Western blots of whole protein extracts of three day old adult female mosquitoes hybridized with the *PGRPLC* antibodies, revealed several bands at a variety of molecular weights (Fig. 4.3.18 d.). Interestingly, these bands were mostly of the same molecular weight between monoclonal and polyclonal PGRPLC1 and PGRPLC2 antibodies.

Whole PGRPLC domain proteins were produced to examine whether these antibodies were able to distinguish between PGRPLC isoforms and not binding to other proteins in an unspecific manner. The chosen sequences were amplified from the abdominal cDNA library with the primers #9.1.37 - #9.1.40, cloned into the pGEM-T easy vector, excised with the NcoI and EcoRI restriction enzymes and subcloned into the pEMT-13 expression vector. The pEMT-13 vector has the advantage that the expressed proteins lack His-Tags, fusion proteins and protease cleavage sites that might interfere with their folding. The molecular weights for all three expressed PGRPLC protein fragments were ~ 17 kD, as expected (Fig. 4.3.18 b.).

When these whole PGRPLC domain proteins were used to test the antibodies, it was revealed that the monoclonal mouse αPGRPLC1 antibody was highly specific, but the monoclonal mouse αPGRPLC2 did not recognize the PGRPLC2 domain fragment – or any of the other PGRPLC (Fig. 4.3.18 c.). The rabbit polyclonal versions obtained with the same antigens produced a high background noise that quenched any signal that might have been detectable. Fig. 4.3.18 e. is representative for all three polyclonal αPGRPLC domain antibodies. In addition, the monoclonal mouse αPGRPLC2 hybridoma lines were tested with a dot blot to quickly determine the specificity and sensitivity of the antibodies produced by different lines. Table 4.3.1 shows that while 5 out of the 7 lines showed strong reactivity to the antigen, none of them was specific only for PGRPLC2.

The characterization of the PGRPLC proteins using the produced antibodies is still ongoing.

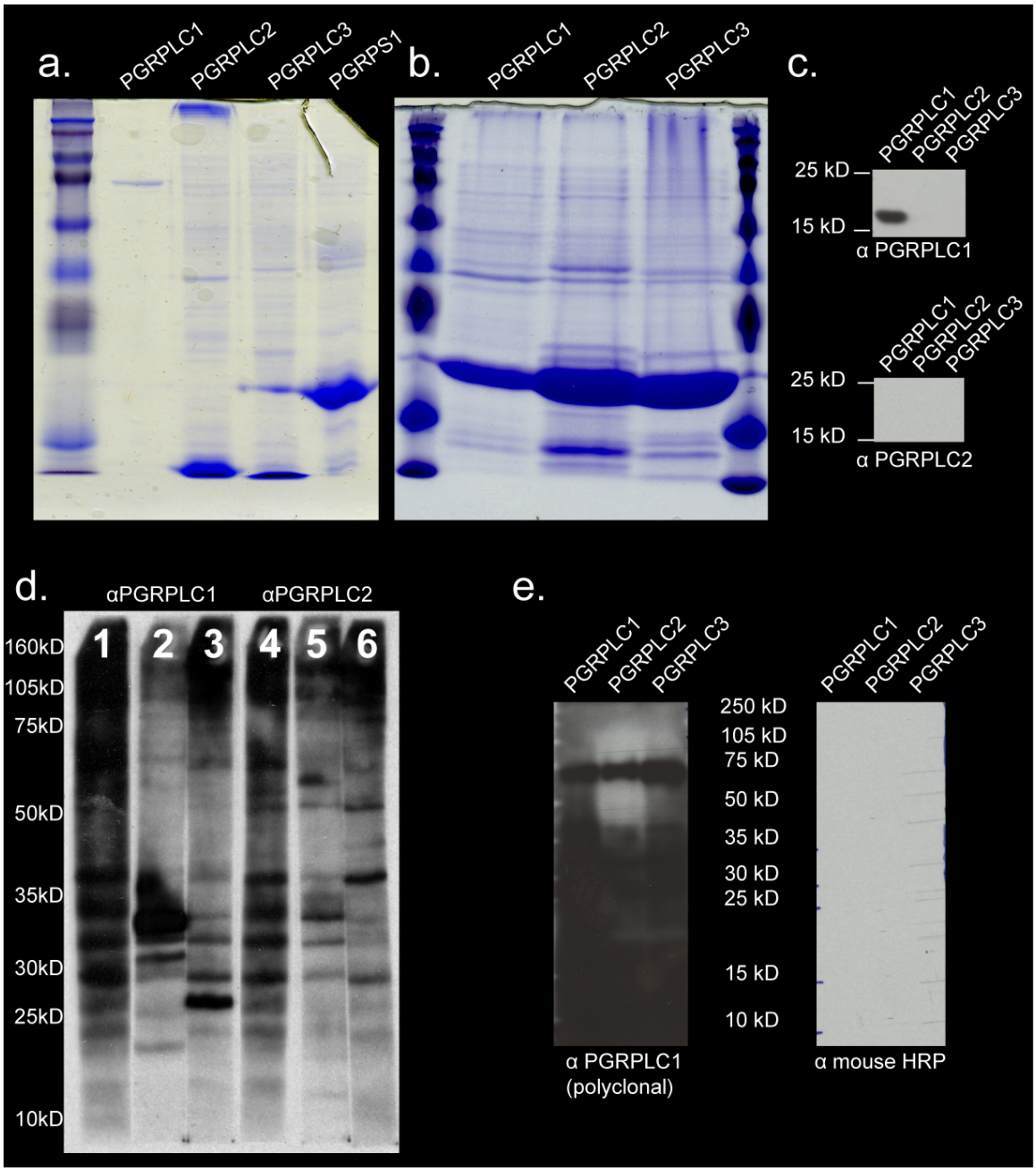
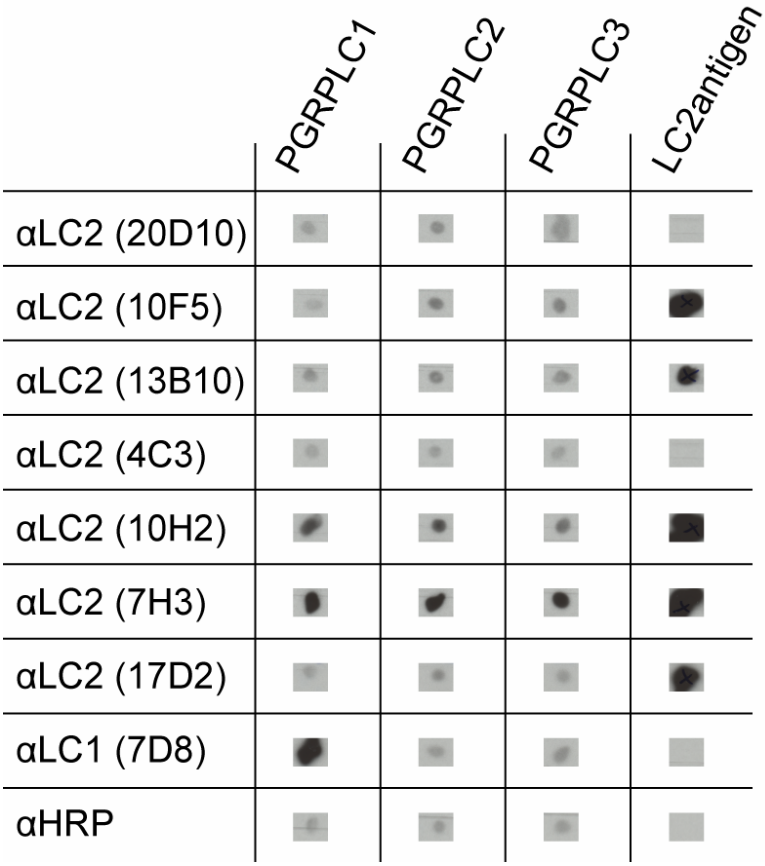


Fig. 4.3.18 a. Expressed antigens of *PGRPLC1*, *PGRPLC2*, *PGRPLC3* and *PGRPS1*. 1 µl per lane. The molecular weight rainbow marker was run for reference in the left lane. **b.** Expressed whole PGRPLC isoforms at 17 kDa. **c.** Western blot with the whole domain PGRPLC proteins probed with the monoclonal mouse αPGRPLC1 and αPGRPLC2 antibodies. **d.** Western blot with protein extract of adult female mosquitoes treated with monoclonal mouse (lanes 1,4) and polyclonal rabbit antibodies (lanes 2,3,5,6). lanes 1-3: α*PGRPLC1;* lanes 4-6: α*PGRPLC2*. **e.** Western blot with the whole domain PGRPLC proteins probed with the polyclonal rabbit αPGRPLC1 AB (left blot) and the secondary αmouse HRP only (right blot). αPGRPLC1: 1:10,000 dilution; αmouse HRP: 1:30,000

Table 4.3.1: Dotblot of monoclonal αPGRPLC2 AB probing full length PGRPLC domain fragments. The monoclonal αPGRPLC1 and secondary αHRP AB served as control.



4.4 THE INTRACELLULAR *RELISH* PATHWAY*

REL2 *gene architecture and phylogenetic analysis.*

The genomic organization of *REL2* was determined by RT-PCR using RNA from adult female *A. gambiae*. The gene consists of 11 exons and encompasses 10.9 kb of genomic DNA (Fig. 4.4.1b). It encodes putative proteins with a Glutamine-/Histidine-rich region (Q/H) that is potentially implicated in protein–protein interactions, followed by a Relish homology domain (RHD), an IPT/TIG domain (DNA binding), a nuclear localization signal (NLS), Ankyrin-repeats (ANK) and a death domain (DD) (Fig. 4.4.1c).

We detected two *REL2* transcripts: the full-length, *REL2*-F, and a shorter form, *REL2*-S. The latter encompasses a unique exon at its 3' end. For mRNA splicing, this exon uses a splice acceptor that is 4 bp upstream to that used in *REL2*-F and results in a downstream early termination codon. Thus, *REL2*-F encodes a protein with all the domains encoded by the gene, whereas *REL2*-S is missing the ANK and DD. Additional alternative splicing at the 5' end of *REL2* may also produce two transcript variants for each form (Fig. 4.4.1c).

REL2-F shows significant similarity to the *A. aegypti* (55%) and *D. melanogaster* (29%) Relish, as well as the mouse NF-κB factors p100 (19%) and p105 (22%) (Table 4.4.1). Phylogenetic analysis shows that REL2-F clusters with the related Relish proteins in *A. aegypti* and *D. melanogaster* and not with the other NF-κB proteins like the *Drosophila* Dif and Dorsal (Fig. 4.4.2).

* Some parts of this work were conducted by people in the laboratory of Dr. Liangbiao Zheng at the Yale University School of Medicine, New Haven, CT, USA in the context of collaboration and is presented here for completeness. These parts are indicated with an asterisk.

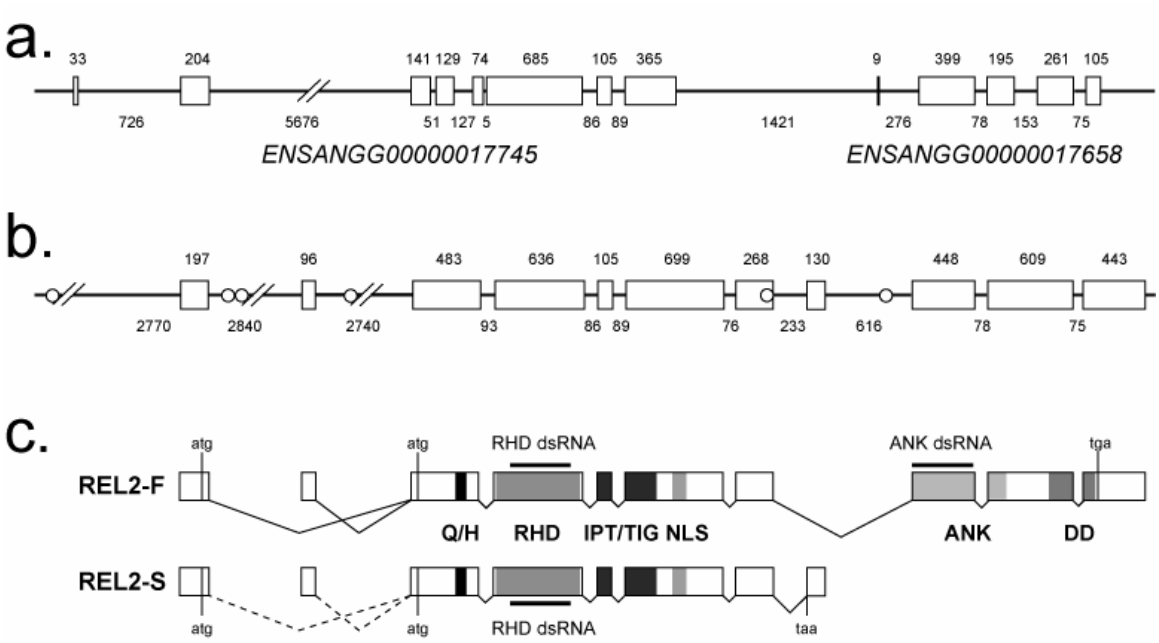


Fig. 4.4.1 Genomic organization and alternative splicing of the *Anopheles REL2* gene. Introns (lines) and exons (open boxes) are based on **a.** Ensembl gene prediction and **b.** cDNA sequencing data. Numbers below and above the introns and exons show the length (bp). Circles indicate putative NF-*κ*B elements. **c.** *REL2* has two alternative 5' exons. The 5' end of *REL2-S* has not been confirmed and is indicated by dashed intron lines. Alternative splicing occurs at the 3' end with a *REL2-S*-specific exon. Protein domains are depicted by shades of gray inside the exons. Q/H, glutamine and histidine-rich region; RHD, Relish homology domain; IPT/TIG, DNA-binding domain; NLS, nuclear localization signal; ANK, Ankyrin-repeats domain; DD, death domain. Putative start- and stop-codons and target regions of dsRNA constructs (RHD dsRNA and ANK dsRNA) are indicated. (from (Meister, Kanzok et al. 2005))

Table 4.4.1 Protein sequence similarities for some selected NF-*κ*B factors. Aa *A. aegypti*; Ag *A. gambiae*; Dm *D. melanogaster* (fruit fly); Mm *Mus musculus* (mouse). (sequences used: Seq. 9.3.15; 9.3.12; 9.3.9; 9.3.7; 9.3.8).

	*Ag*REL-2-F				
*Aa*REL2	55%	*Aa*REL2			
*Dm*Relish	29%	28%	*Dm*Relish		
*Mm*p105	22%	22%	19%	*Mm*p105	
*Mm*p100	19%	22%	18%	39%	*Mm*p100

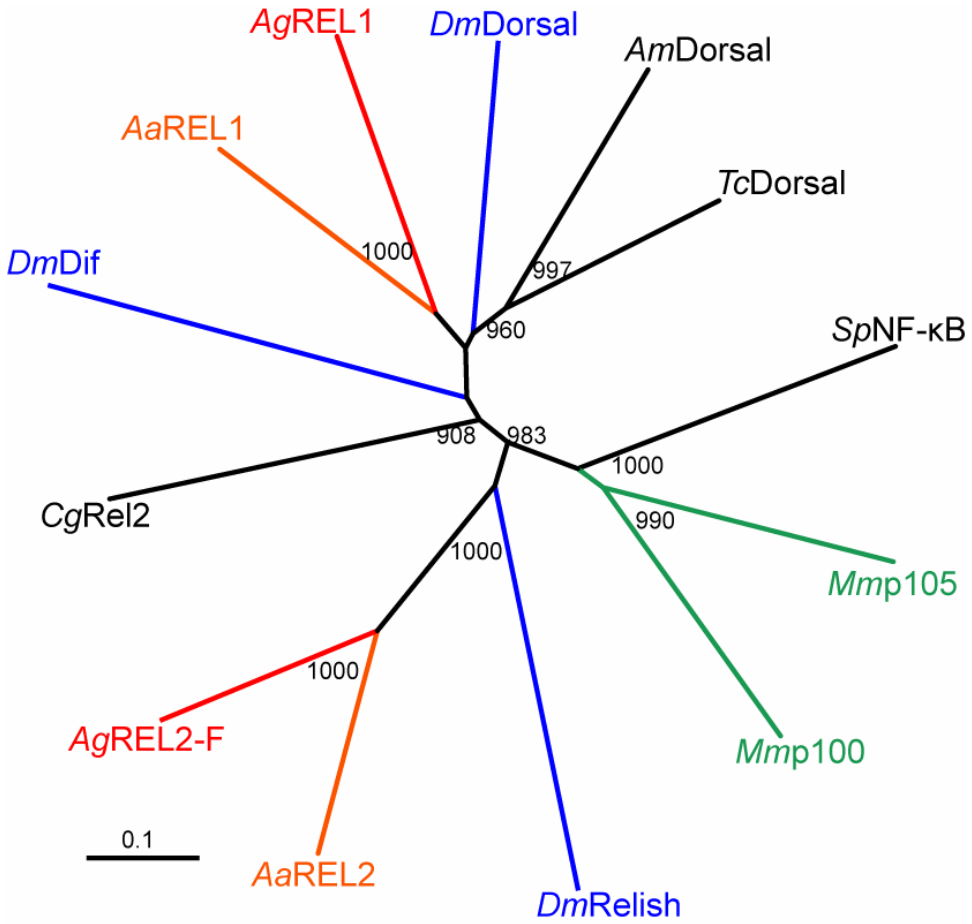


Fig. 4.4.2 Phylogenetic analysis of NF-κB proteins. The N-J tree was calculated from a ClustalW (program version 1.82) protein alignment with default settings. Bootstrap values (1000 iterations) at nodes of the tree provide an estimate of the statistical significance of particular branching points. Aa *A. aegypti*; Am *Apis melifera* (honey bee); Ag *A. gambiae*; Cg *Crassostrea gigas* (portuguese oyster); Dm *D. melanogaster* (fruit fly); Mm *Mus musculus* (mouse); Sp *Strongylocentrotus purpuratus* (purple sea urchin); Tc *Tribolium castaneum* (red flour beetle). (sequences used: Seq. 9.3.4 through 16) (modified from (Meister, Kanzok et al. 2005)).

REL2 *KD expression profile.**

REL2-F and -S transcripts are expressed constitutively throughout *A. gambiae* development, albeit at different levels: *REL2*-F is expressed more strongly (Fig. 4.4.3a). Both

transcripts are also expressed in cultured cell lines, Sua1B (Fig. 4.4.3b) and 4a3A (data not shown).

Cells challenged with heat-killed *E. coli* (Gram-) or purified *E. coli* lipopolysaccharides (LPS) up-regulate *REL2*-F significantly and *REL2*-S to a lesser degree; however, we cannot exclude the possibility that induction after LPS challenge was due to peptidoglycan contaminants that are frequently found in commercial LPS preparations (Leulier, Parquet et al. 2003). Challenge with the fungus *B. bassiana* or the Gram+ bacterium *Micrococcus luteus* did not result in a significant transcriptional up-regulation of *REL2* (Fig 4.4.3b).

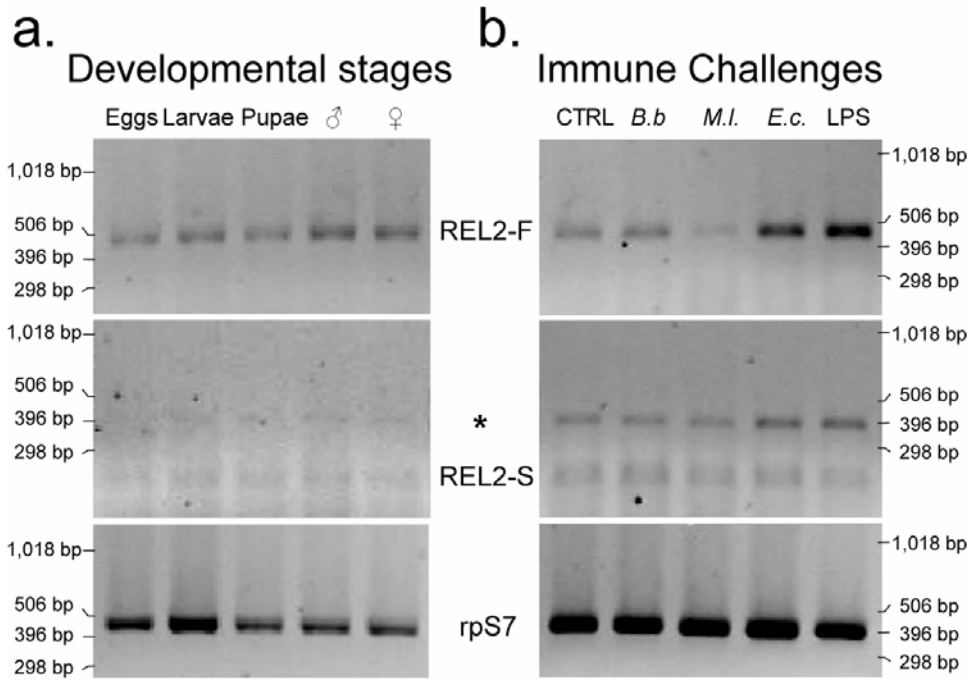


Fig 4.4.3 Expression profile of *REL2* in mosquito cultured cells and developmental stages. a. Relative abundance of *REL2-F* and *REL2-S* transcripts at different stages of mosquito development. **b.** Transcriptional regulation of *REL2* transcripts in immune challenged cultured Sua1B cells. *REL2-F* was detected using primer pairs targeting the ANK domain (#9.1.84/85). *REL2-S* was detected with primers #9.1.86/87. The asterisk indicates a band which includes the intron sequence of *REL2-S*. It is unclear whether this transcript is part of an additional splice variant of *REL2*. The identity of *REL2-S* bands was confirmed by sequencing. Control RT-PCR with *rpS7* showed no genomic DNA contamination. ♂, males; ♀, females; CTRL, Control; *B.b., B. bassiana; M.l., Micrococcus luteus; E.c., Escherichia coli.*

REL2 *KD efficiency in adult mosquitoes.*

RNAi was performed as described in chapter 3 (Materials & Methods). In brief, dsRNA was injected into one to two day old adult female mosquitoes that were allowed to recover for four days before the KD efficiency was tested by real-time quantitative RT-PCR (qRT-PCR). For qRT-PCR we used oligonucleotide primers spanning the alternative introns after the 7th REL2 exon (Fig. 4.4.1). This way both *REL2*-S (primers #9.1.80/81) and *REL2*-F (primers #9.1.82/83) could be targeted. Because *A. gambiae* shows no transitive RNAi effects, i.e., dsRNAs are effective only against mRNAs bearing the corresponding sequences (Hoa, Keene et al. 2003), and thus targeting the RHD should silence both REL2 forms, whereas ANK-domain KD should affect only *REL2*-F. The results confirmed that RHD dsRNA injection silences both *REL2*-S and *REL2*-F transcripts equally (~50%), whereas ANK dsRNA decreases the levels of only *REL2*-F (Fig. 4.4.4). Interestingly, *REL2*-S transcript levels were up-regulated by 1.5-fold in dsANK-injected mosquitoes.

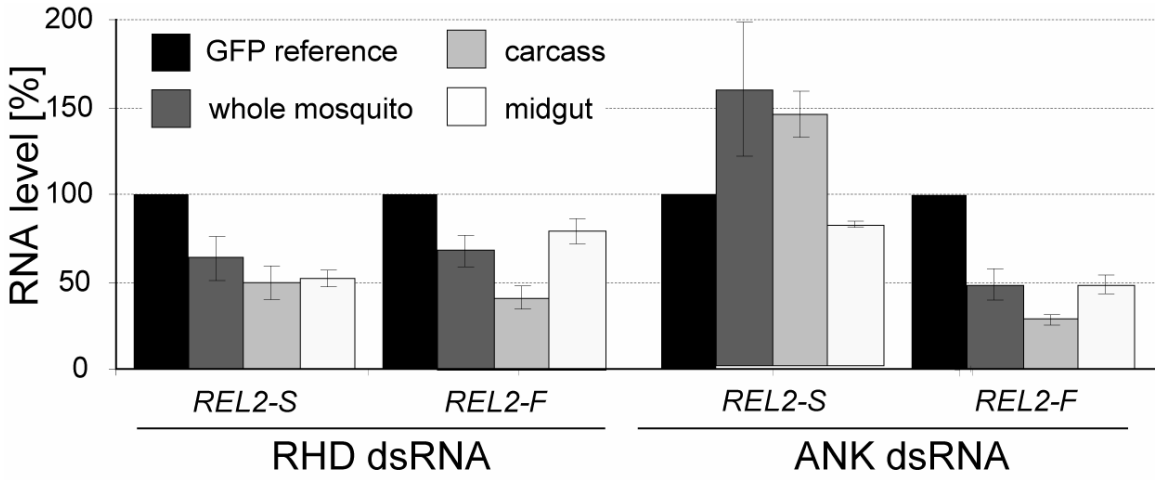


Fig. 4.4.4 Efficacy of RNAi-mediated *REL2* gene silencing in adult mosquitoes after injections with either RHD or ANK dsRNA. Relative *REL2*-S and *REL2*-F transcript levels in KD compared with control GFP dsRNA-treated mosquitoes were determined in the carcass, midgut, and whole mosquitoes 4 days after injections by using qRT-PCR (modified from (Meister, Kanzok et al. 2005)).

REL2/IMD *immunity pathway and defense to bacterial infection.*

We investigated, by RNAi analysis, the possible role of *REL2* and other components of the *Imd* pathway in the defense against bacterial infections in adult mosquitoes. As described above and in chapter 3 (Materials & Methods), dsRNA was injected into one to two days old adult female mosquitoes that were allowed to recover for four days and then challenged with bacteria representative of the Gram+ (*S. aureus*) or the Gram- (*E. coli*) types.

Mosquito survival was assessed daily for 7 days after bacterial infection (Fig. 4.4.5). The results revealed that simultaneous KD of *REL2*-F and *REL2*-S transcripts (RHD dsRNA) significantly compromised mosquito defense against both *S. aureus* and *E. coli*. In contrast, KD of *REL2*-F alone (ANK dsRNA) reduced the survival of *S. aureus* – but not of *E. coli* – infected mosquitoes.

KD of *CASPL1* (the ortholog of the *Drosophila Dredd*) did not have a significant effect on the mosquito anti-bacterial and anti-parasitic immunity and thus it is not discussed further or shown in the figures. For the KD we used the primers #9.1.50/51 (Table 9.1.1).

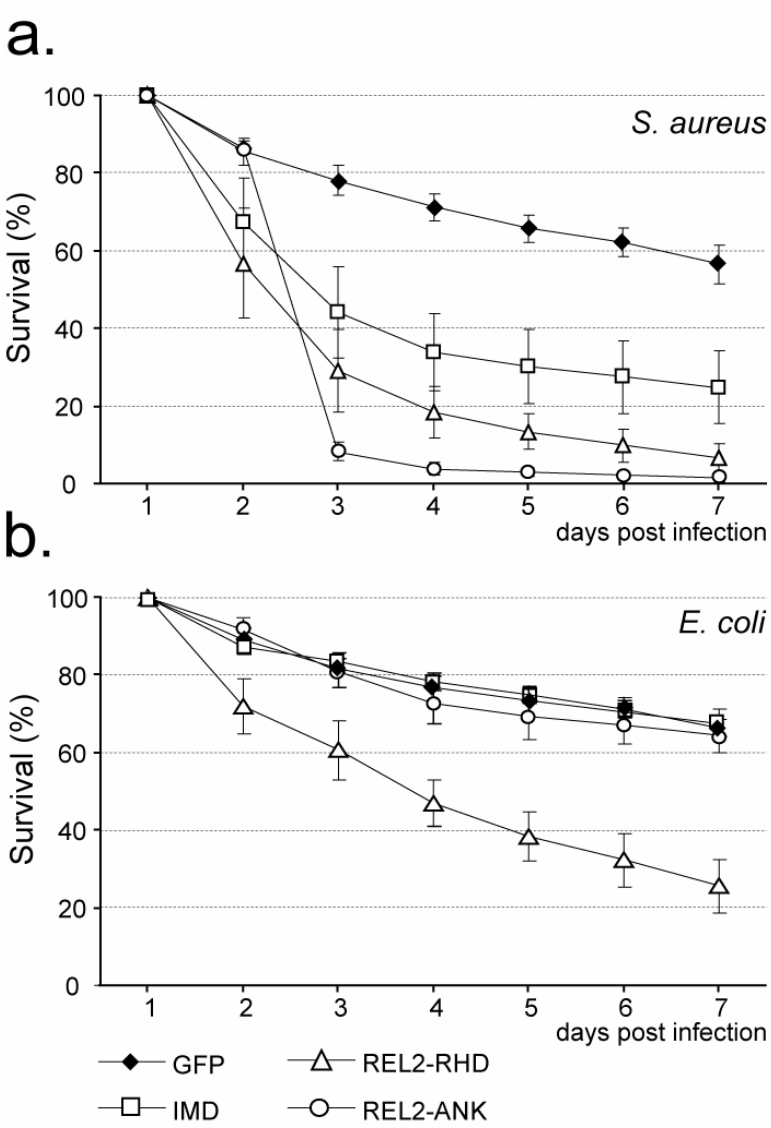


Fig. 4.4.5 Mosquito survival after gene KD and bacterial infection. Survival of adult mosquitoes to **a.** Gram-positive *S. aureus* and **b.** Gram-negative *E. coli* infections after KD of *IMD*, *REL2*-RHD, and *REL2*-ANK. *GFP* dsRNA-treated mosquitoes were used as control. The survival assay was repeated at least 3 and up to 10 times per gene. Error bars represent the standard error (from (Meister, Kanzok et al. 2005)).

REL2/IMD *KD and* Plasmodium *infection*.

It has been reported previously that *CEC1* and other AMP genes are up-regulated when *Plasmodium* crosses the *Anopheles* midgut (Dimopoulos, Richman et al. 1997; Vizioli, Bulet et al. 2000; Christophides, Zdobnov et al. 2002). Furthermore, microarray results (Meister, Kanzok et al. 2005) indicated that *REL2* regulates expression of the major *Plasmodium* antagonist LRIM1 (Osta, Christophides et al. 2004). We tested the effect of *REL2* and *IMD* KD on *P. berghei* survival at the oocyst stage in the mosquito midgut. We also tested the effect of the *REL2/IMD* pathway silencing on *P. berghei* development in the refractory L3-5 mosquitoes. These mosquitoes melanize all the parasites at the ookinete stage during midgut invasion. In the previous chapter, we showed that KD of *PGRPLC*, the putative receptor of the *IMD/REL2* pathway, leads to a drastic increase of parasite numbers, which however are all killed by melanization. In this next set of experiments, *LRIM1* was used as a control. KD of *LRIM1* leads to an increase in parasite numbers and prevents parasite melanization. The *REL2* and *IMD* KD revealed the same phenotype as that of the *PGRPLC* KD: an increase in number of melanized parasites (especially in REL2 KD mosquitoes; Fig. 4.4.14). This suggests that PGRPLC might be indeed the receptor of the *IMD* pathway, and that, similar to the susceptible G3 mosquitoes, L3-5 mosquitoes kill and lyse a large number of *P. berghei* parasites before melanization of the remaining ones. It also indicates that although the *LRIM1* gene is one of the targets of the *IMD/REL2* pathway, the effect of the pathway on P. *berghei* survival is much more complex.

Silencing of *REL2* (RHD dsRNA) or *IMD* in G3 mosquitoes resulted in a statistically significant twofold increase of oocyst numbers, suggesting that the pathway is, indeed, implicated in parasite killing before or during midgut invasion (Fig. 4.4.6, Table 4.4.2).

Quantitatively similar results were obtained when silencing *REL2*-F alone (ANK dsRNA), indicating that *REL2*-F, rather than *REL2*-S, is implicated in this reaction. KDs of both *REL2*-RHD and *REL2*-ANK, but not of *IMD*, led to melanization of some parasites (~5% of the ookinetes). This finding would suggest that *REL2*-F is involved in the control of the melanization reaction, in a manner independent of the *IMD*.

Table 4.4.2 Effect of *REL2* pathway gene silencing on parasite development.

Gene KD	midguts	par/midgut	Prev (%)	Mel (%)	*P* (KS)	*P* (t)
CONTROL (5)	79	86	87.34	0.07		
IMD (5)	79	153	96.20	0.13	0.044	0.017
CONTROL (4)	49	51	87.76	0.63		
REL2-RHD (4)	49	104	91.84	4.37	0.147	0.035
CONTROL (3)	42	82	97.62	0.00		
REL2-ANK (3)	42	166	95.24	4.87	0.093	0.008

Numbers of oocysts and melanized parasites are reported as three experimental datasets for knocked down genes or *REL2* domains and control dsRNA-treated mosquitoes (GFP). The number of replicates is indicated in parenthesis in the first column. Each replicate used different batches of mosquitoes, which were fed on the same infected mouse. Identical numbers of midguts from control and KD mosquitoes were assessed for each replicate (see also Supporting Table 5). The probabilities (*P*) indicate whether the distributions of parasites in KD and control midguts in each dataset are similar, and were determined by KS and Student's t test. Prev, prevalence; mel, melanized parasites (ookinetes); par, parasite number (oocysts and melanized ookinetes).

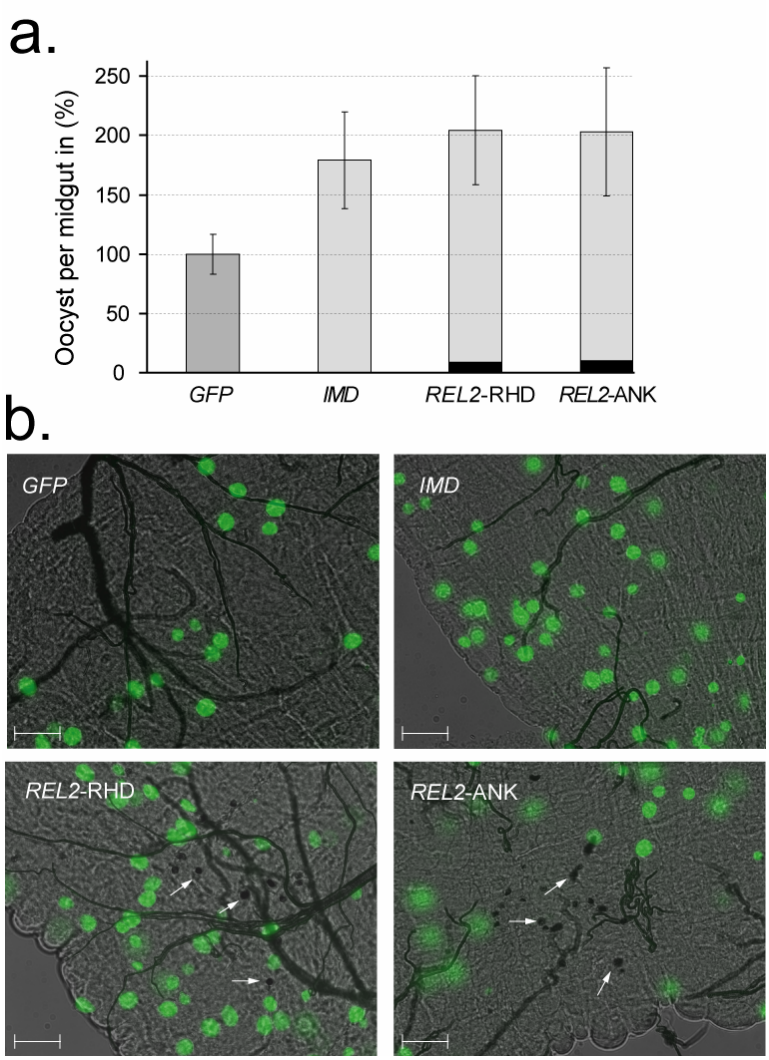


Fig. 4.4.6 *REL2*- and *IMD*-mediated Plasmodium killing and melanization in the mosquito vector. **a.** Averaged *P. berghei* oocyst load per midgut after KD of *IMD*, *REL2*-ANK, and *REL2*-RHD. The oocyst load after individual gene KD was normalized to the oocyst load of parallel dsGFP treatment and presented as a percentage. Every gene KD was repeated three times or more. Black sections of columns represent melanized ookinete fraction. Error bars represent the standard error. **b.** Representative sections of infected *A. gambiae* midguts infected with GFP-tagged *P. berghei*. Arrows point to melanized ookinetes (from (Meister, Kanzok et al. 2005)).

CEC1 *is under* REL2 *transcriptional control.**

A putative AMP target gene of *Anopheles* immune signaling pathways is *CEC1* (previously designated as *Cecropin A*) (Vizioli, Bulet et al. 2000; Zheng and Zheng 2002). Compared with *Drosophila* Cecropins, which are mostly active against Gram-negative bacteria, *CEC1* has a wide spectrum of antibacterial activities and is induced by *Plasmodium* and both Gram types of bacteria (Vizioli, Bulet et al. 2000). Sua1B and 4a3A cells have high levels of constitutive *CEC1* expression that are further enhanced by treatment with heat-killed *E. coli* or *M. luteus* (data not shown). We examined the effect of REL proteins on *CEC1* gene expression in Sua1B cells after transfection with a *CEC1*-promoter luciferase-reporter construct (Zheng and Zheng 2002), silencing of *REL1* or *REL2* genes by RNAi, and subsequent monitoring of luciferase activity. Targeting either the RHD or ANK domains of *REL2* resulted in substantial reduction of luciferase activity, typically 5- to 10-fold. In contrast, *REL1* knockdown had no significant effect (Fig. 4.4.7).

We observed that KD using the ANK dsRNA results in a major (~91%) reduction in *CEC1*-promoter expression, essentially indistinguishable from that observed with RHD dsRNA. Thus, it appears that REL2-F largely controls the constitutive *CEC1*-promoter expression that is detected in the Sua1B cells, whereas REL2-S has no major effect (Fig. 4.4.7). Similar results were obtained for 4a3A cells (data not shown). Proteolytic cleavage is necessary to remove the inhibitory ANK domain of *Drosophila* Relish, and this processing would also be expected for REL2-F activation. By extension, these results suggest that REL2-F is constitutively activated in these mosquito cells.

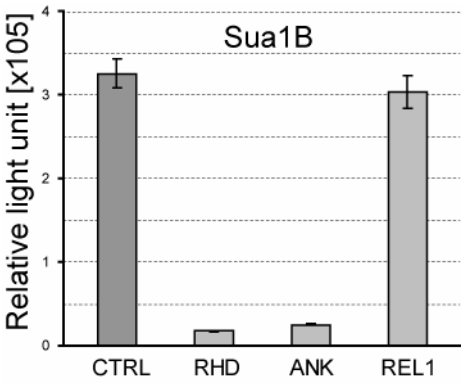


Fig. 4.4.7 *CEC1* promoter activity in cultured *Anopheles* cells. *REL2*, but not *REL1*, is required for expression of *CEC1* promoter in Sua1B cells, which is equally affected by dsRNAs encompassing the RHD and the ANK-repeats domain. CTRL, control (modified from (Meister, Kanzok et al. 2005)).

Other *Anopheles* orthologs of *Drosophila* Imd-pathway components, the *Imd* adaptor (*IMD*) and caspase *Dredd* (*CASPL1*), were previously identified (Christophides, Zdobnov et al. 2002); orthologs of *IKKβ* (*IKK1* or *ENSANGG00000014263*) and *IKKγ* (*IKK2* or *ENSANGG00000018475*) were identified herein: (Meister, Kanzok et al. 2005). To examine the degree of functional *Imd*-pathway conservation between *Anopheles* and *Drosophila*, we silenced these genes in Sua1B and 4a3A cells by using RNAi and observed significant down-regulation (Fig. 4.4.8). In both cell lines, *IMD* KD reduced *CEC1*-promoter expression (~60% and 68%, respectively), although not as efficiently as *REL2* KD. This is consistent with evidence from *Drosophila imd* mutants, in which *Cecropin A* expression is reduced but not abolished (Onfelt Tingvall, Roos et al. 2001).

Surprisingly but consistently, silencing of *CASPL1* and *IKK1* or *IKK2* had opposite effects in the two cell lines. In Sua1B cells, KD of *CASPL1* but not of *IKK1* or *IKK2*, substantially reduced *CEC1* promoter expression, whereas, in 4a3A cells, KD of *IKK1* or *IKK2*, but not of *CASPL1*, reduced *CEC1* promoter expression.

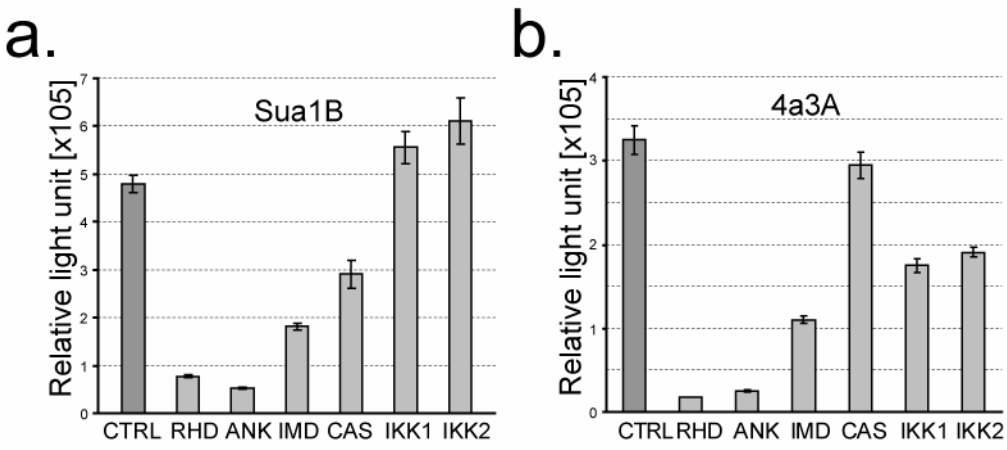


Fig. 4.4.8 *CEC1* promoter activity in cultured *Anopheles* cells. **a.** *REL2*, *IMD*, and *CASPL1*, but not *IKK1* or *IKK2*, are required for *CEC1* expression in Sua1B cells. **b.** *REL2*, *IMD*, *IKK1*, and *IKK2*, but not *CASPL1*, are required for *CEC1* expression in 4a3A cells. CTRL, control; CAS, *CASPL1* (modified from (Meister, Kanzok et al. 2005)).

We determined the level of gene silencing for *IKK1*, *IKK2* and *CASPL1* (Fig. 4.4.9). Importantly, KD of *IKK1* and *IKK2* in Sua1B was highly effective, suggesting that these genes are not implicated in the regulation of *CEC1* in Sua1B cells. KD of *CASPL1* in the same cells was less effective (~40%), and yet associated with substantial reduction in *CEC1* activity (Fig. 4.4.8a). A plausible explanation that requires further investigation is that IKK1/IKK2 and CASPL1 are implicated in alternative branches of the *CEC1*-activation pathway that are differentially active in these two cell types.

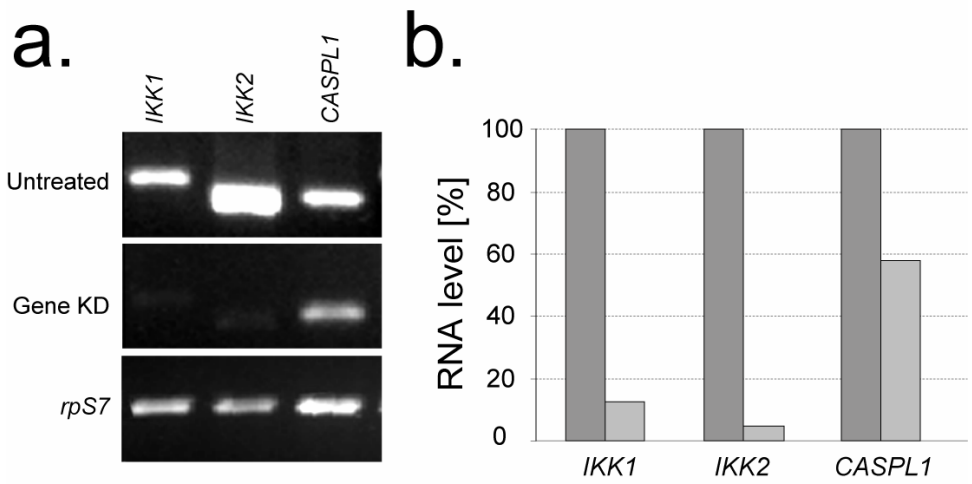


Fig. 4.4.9 Efficacy of RNAi-mediated gene silencing of *IKK1*, *IKK2*, and *CASPL1* genes in Sua1B cultured cells. RNAi effectiveness was determined **a.** qualitatively by RT-PCR and **b.** quantitatively by qRT-PCR. Ribosomal protein *S7 rpS7* served as internal reference and untreated cells served as calibrator (modified from (Meister, Kanzok et al. 2005)).

Other genes under **REL2** *transcriptional control*[*].

A DNA microarray analysis (Dimopoulos, Christophides et al. 2002) was performed to identify genes regulated by REL2 in *Anopheles* 4a3A cells by comparing the expression profiles of cells treated with either *REL2* (RHD) or control *lacZ* dsRNAs. Cells were challenged 24 h after dsRNA addition, by a 12-h exposure to heat-killed *E. coli*, and expression profiles were assessed at 0, 2, 6, and 12 h. From the 3,840 EST clones present on the microarray, nine were consistently down-regulated in *REL2* KD cells by at least 1.7-fold and one by 1.2-fold (Table 4.4.4). Six of these genes have been categorized previously as defense/immunity-related (Dimopoulos, Casavant et al. 2000). They encode the AMPs CEC3 and GAM1 (previously CecB and Gambicin, respectively), the serine protease CLIPB14, another two CLIP-domain serine proteases, a fibrinogen-domain lectin, a Brix-domain protein implicated in ribosome biogenesis, a protein disulfide isomerase, KIN1, a kininogen-domain protein (see below), and LRIM1, a leucine-rich repeat immune protein

with strong anti-parasitic activity (Osta, Christophides et al. 2004). The *CEC3* gene is tightly clustered with *CEC1* and *CEC2* and its promoter region includes two NF-κB binding sites (Zheng and Zheng 2002), suggesting regulatory similarities to *CEC1*. At least 1, and as many as 10, κB elements were detected in adjacent genomic regions of all regulated genes (Table 4.4.4).

Table 4.4.4 *REL2* regulated genes as determined by microarray (Meister, Kanzok et al. 2005).

Ensembl gene ID	Folds	# κB	Gene description
ENSANGG00000010552	-2.85	2	*LRIM1*
ENSANGG00000013355	-2.6	4	CLIP domain SP
ENSANGG00000009506	-2.3	2	*CEC3*
ENSANGG00000015633	-2.28	8-10	**CLIPB14 *(SP)***
ENSANGG00000018623	-2.2	1	Brix domain
ENSANGG00000013381	-2.19	1	CLIP domain SP
ENSANGG00000015896	-2.1	2	Disulfide isomerase
ENSANGG00000015703	-2.01	4	Fibrinogen lectin
ENSANGG00000010776	-1.7	2	*GAM1*
ENSANGG00000011642	-1.2	2	*KIN1*

The number (#) of putative NF-κB sites in the vicinity of each gene is listed; SP, serine protease.

KIN1 encodes a short mature cationic polypeptide with a kininogen domain, which is enriched in histidine (26 residues, 36%), plus an amino-terminal signal peptide. *KIN1* is strongly up-regulated by immune challenges (Dimopoulos, Christophides et al. 2002) and *Plasmodium* infection (Vlachou, Schlegelmilch et al. 2005). Examination of a *KIN1*-promoter

luciferase construct in Sua1B cells showed that REL2 is, indeed, required for *KIN1* expression. RHD and ANK dsRNA resulted in more than 50% reduction of KIN1 promoter activity (Fig. 4.4.10).

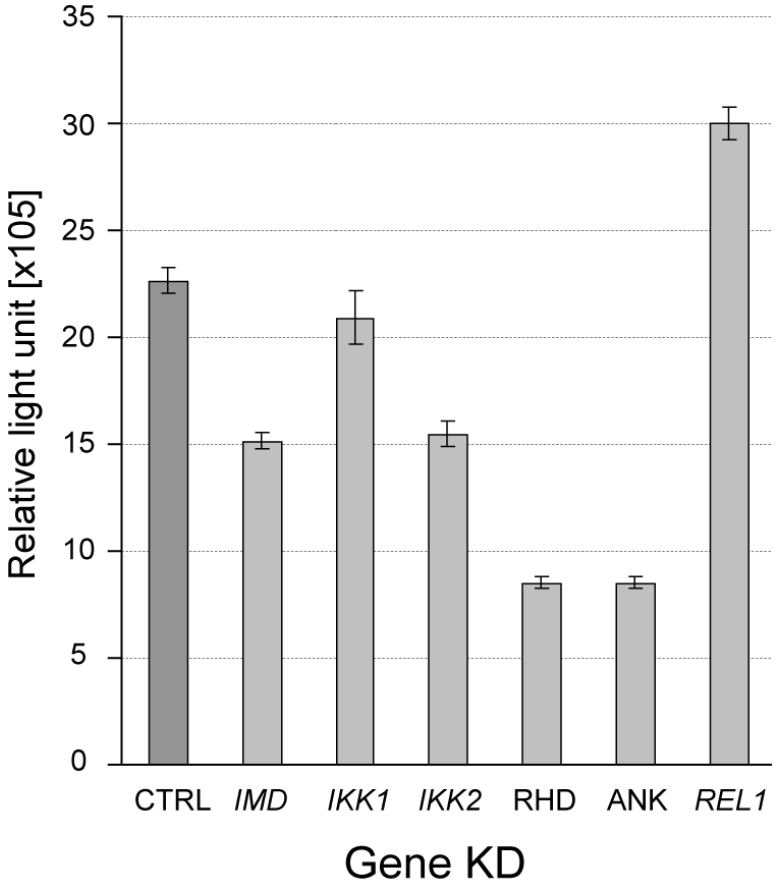


Fig. 4.4.10 *KIN1* gene is transcriptionally regulated by *REL2*. The *KIN1* promoter was amplified from genomic DNA with *KIN1*-promoter primers (#9.1.88/89) and cloned into pGL3 (Promega, Madison WI). *KIN1* KD and the luciferase assay in Sua1B cell lines was carried out as described in chapter 3 (Materials & Methods). CTRL, control (modified from (Meister, Kanzok et al. 2005)).

The Anopheles TOLL/REL1 *pathway in anti-bacterial defense.*

The importance of the *IMD/REL2* pathway in resistance to infection with both Gram+ and Gram- bacteria prompted us to expand our analysis to the other NF-κB factor in the *Anopheles* genome, *REL1* (the ortholog of *Dorsal*), that we had used as control thus far (Fig. 4.4.7 & 4.4.10). *REL1* lacks an IκB inhibitory domain, similar to its fly ortholog. Thus, *Dorsal* (and *Dif*) in *Drosophila* is inhibited from constitutive translocation to the nucleus by the IκB-domain protein Cactus, which also has an *Anopheles* clear-cut ortholog, CACT. We used dsRNA injection in adult mosquitoes, as described in the Materials & Methods (Chapter 3), to KD REL1 (primers #9.1.68/69), CACT (primers #9.1.66/67)or both genes simultaneously. The mosquitoes which were then infected with *E. coli* and *S. aureus* bacteria and their resistance to infection, expressed by their survival rates, was monitored.

However, none of the KDs led to reduced mosquito survival after infection with *E. coli* (Fig 4.4.11) or *S. aureus* (Fig 4.4.12) bacteria. The fact that KD of *REL1* had no measurable effect on mosquito survival after bacterial challenge, suggests that *REL2* is sufficient to confer resistance to bacterial infections.

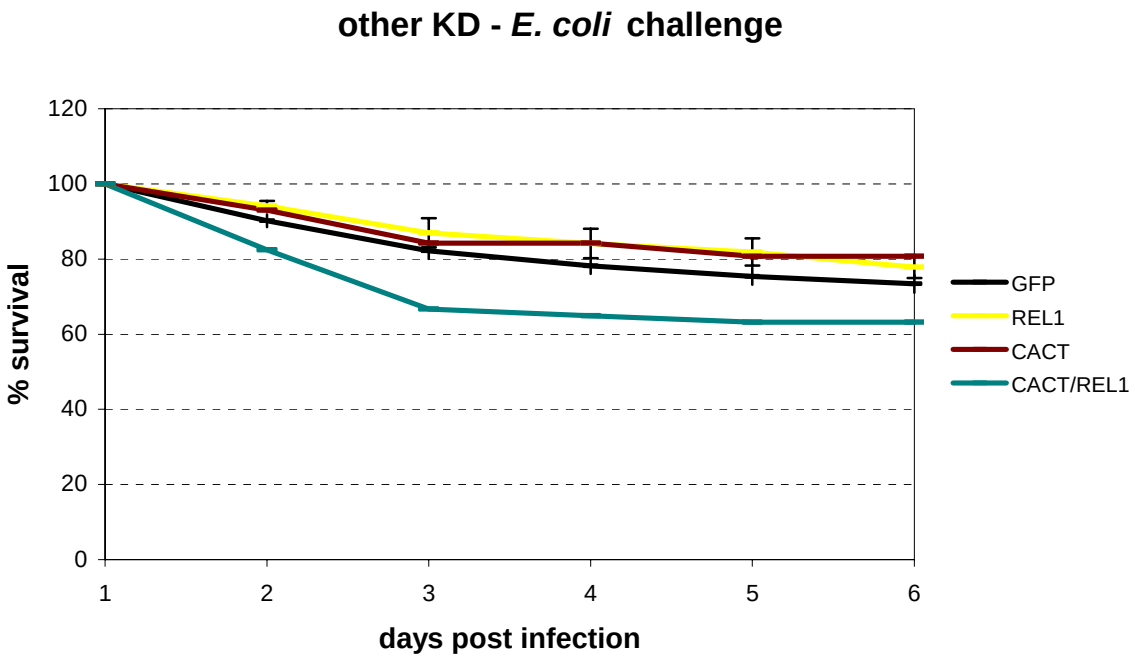


Fig. 4.4.11 Survival of adult female mosquitoes after gene KD by injection with dsRNA (*IMD (5 repetitions), REL1 (3 repetitions), CASPL1 (3 repetitions), CACT (1 repetition), CACT/REL1-doubleKD (1 repetition)* or control *GFP*) and injection of *E. coli.*

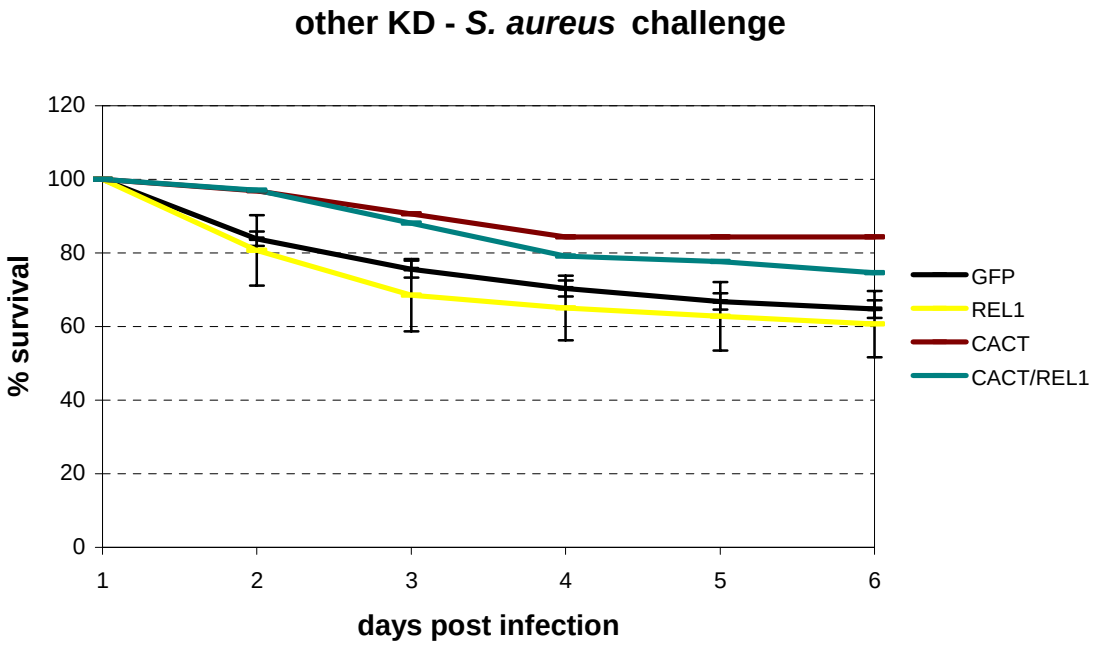


Fig. 4.4.12 Survival of adult female mosquitoes after gene KD by injection with dsRNA; *REL1 (3 repetitions), CASPL1 (3 repetitions), CACT (1 repetition), CACT/REL1-doubleKD (1 repetition)* or control *GFP*) and injection of *S. aureus.*

The role of the **Toll/REL1** *pathway in* **Plasmodium** *infection.*

We assessed the role of *REL1* in mosquito infection with the rodent malaria parasite *P. berghei* as described previously. KD of *REL1* did not have any effect on the number of oocysts developing in the mosquito midgut, suggesting that only *REL2*, and its associated *IMD* pathway is involved in the *P. berghei* killing that is documented before or during midgut invasion (Fig 4.4.13). However, the evidence from a massive screen that was conducted in the laboratory suggested that KD of the IκB protein CACT leads to extensive parasite killing and melanization. Indeed, our results, presented in Fig 4.4.13 showed that *CACT* KD results in extensive parasite killing and elimination, possibly by lysis, that reaches 80% of the parasites compared to the control ds*GFP* treated mosquitoes, while the remaining 20% of the parasites are melanized.

To further explore the role of *CACT* in inhibition of the mosquito NF-κB pathways, we simultaneously silenced *CACT* and either of the two NF-κB factors, *REL1* and *REL2*.

As REL2-S is lacking an inhibitory IκB domain (compared to REL2-F), this experiment was critical to examine whether its nuclear translocation is inhibited by CACT. However, the results showed that only silencing of *REL1* was able to rescue the wt phenotype of *CACT* KD mosquitoes. Double KD of *CACT* and *REL2* resulted in a phenotype similar to that of *CACT*, namely extensive parasite lysis and some ookinete melanization.

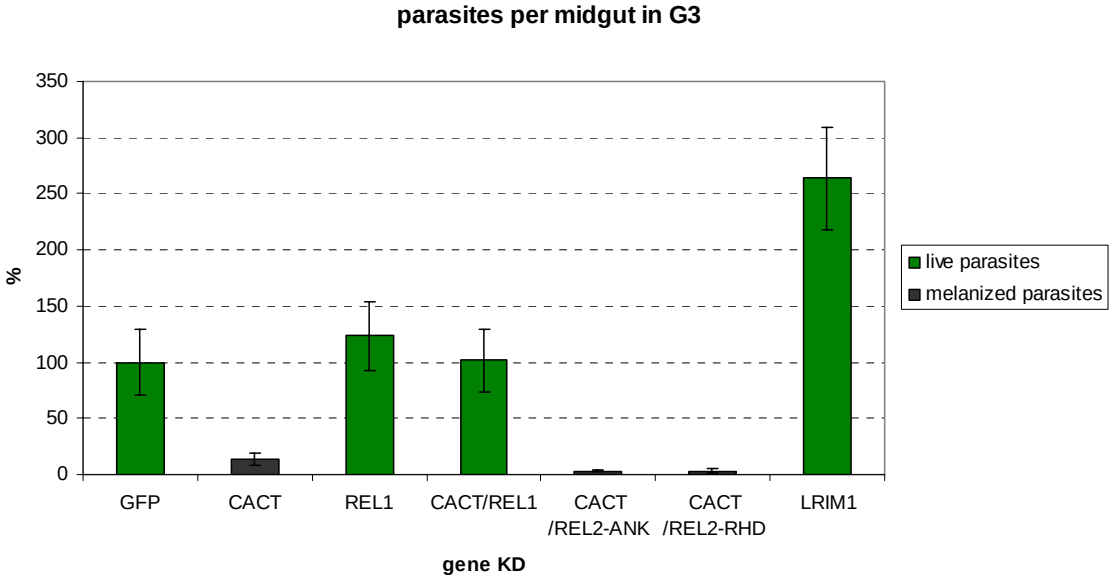


Fig 4.4.13 Differences in G3 parasite numbers per midgut after gene KD. KD of the *P. berghei* antagonist *LRIM1* was used as control. Green colored fractions represent live oocysts and grey colored fractions melanized ookinetes. Sample sizes: CACT, 43 midguts; REL1, 52 midguts; CACT/REL1, 30 midguts; CACT/ANK, 21 midguts; CACT/RHD, 12 midguts; LRIM1, 67 midguts. Error bars represent standard errors.

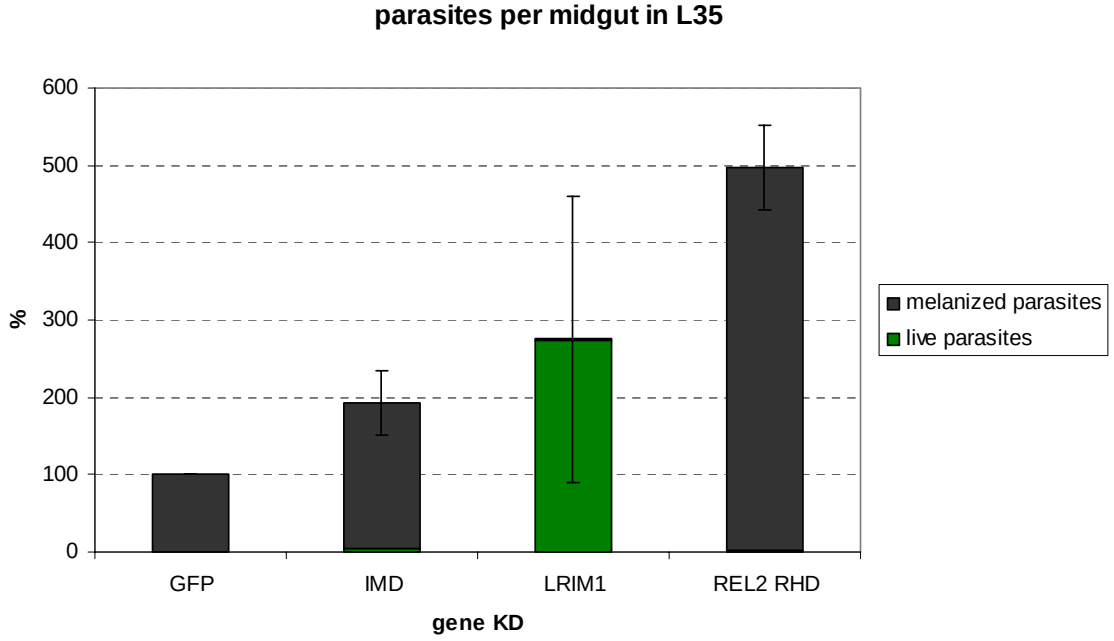


Fig 4.4.14 Differences in L3-5 parasite numbers per midgut after gene KD. Green colored bars represent live oocyst fractions and grey colored bars melanized ookinetes. Sample sizes: *IMD*, 79 midguts; *LRIM1*, 7 midguts; *REL2*, 31 midguts. Error bars represent standard errors.

A. gambiae *immunity to fungal infection.*

Innate immune defense to fungi in *D. melanogaster* is mediated by the *Toll* pathway and the NF-κB transcription factor *Dif* (Lemaitre, Nicolas et al. 1996). Similarly, the A. *aegypti REL1* transcription factor, which is orthologous to the A. *gambiae REL1* and believed to be a component of the *Toll* pathway, mediates immunity to the filamentous entomopathogenic fungus *B. bassiana* (Bian, Shin et al. 2005; Shin, Kokoza et al. 2005).

We examined the role of the A. *gambiae REL1* and *REL2* in defense to *B. bassiana*. The mosquitoes were exposed to fungi spores using the protocols described in the Materials & Methods (chapter 3). As shown in Fig. 4.7.1, the effect of a mosquito infection with spores of *B. bassiana* proved to be dramatic and the fungi overwhelm the mosquito innate immune system and all mosquitoes die within the following four days (Fig. 4.7.2).

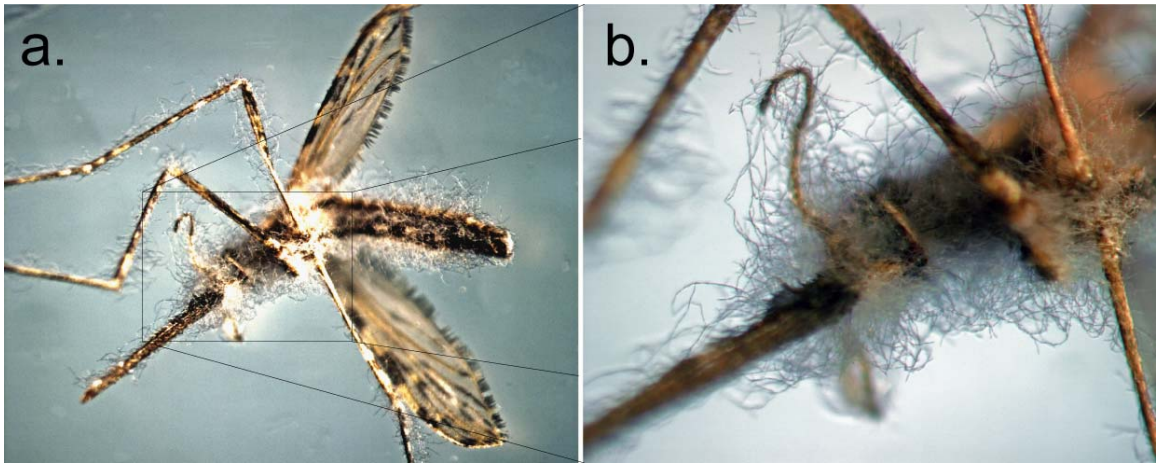


Fig 4.7.1 a. *A. gambiae* overcome by *B. bassiana* 6 days after infection. **b.** Close-up of head region

The effect was so drastic, that any additional effect from a gene KD was masked by the rate of natural death from the infection (Fig. 4.7.3). We also examined whether KD of *CACT* can lead to an increase in mosquito survival. However, the rate of infection again did not allow to observe any possible effect (data not shown).

These results indicated that the *A. gambiae* infection with *B. bassiana* needed further optimization in order for it to be used as a diagnostic tool to examine the role of immune pathways during fungal infections.

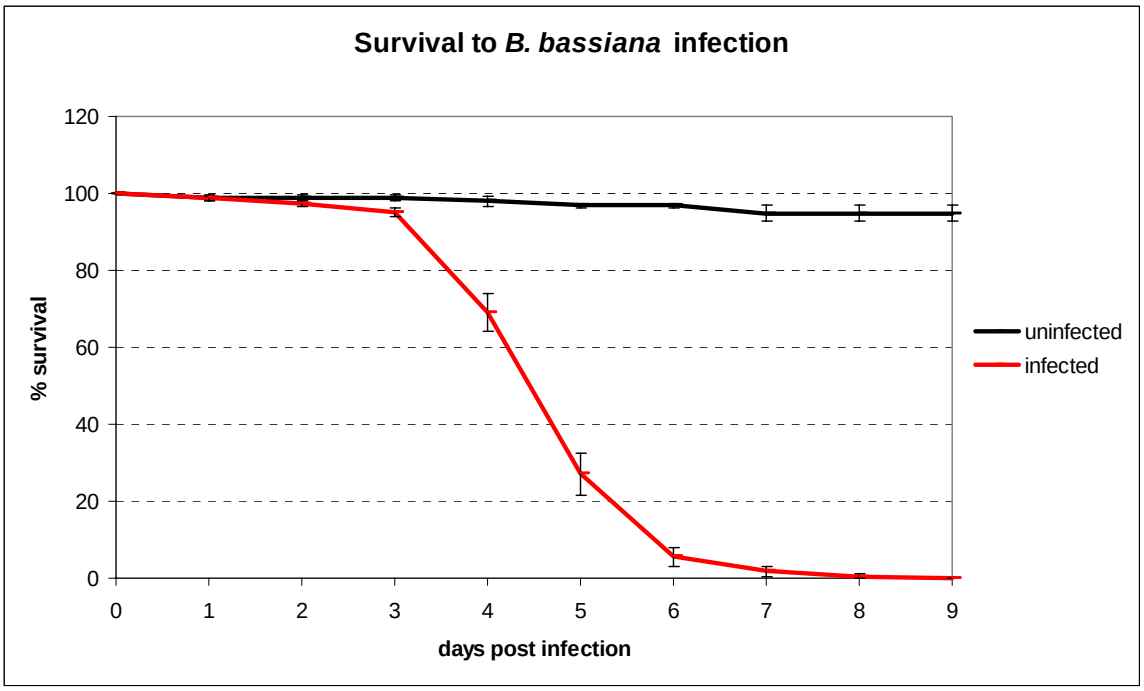


Fig 4.7.2 *A. gambiae* survival to infection with *B. bassiana*. The infection was repeated 11 times – the uninfected mosquitoes (kept in the same location) were sampled twice.

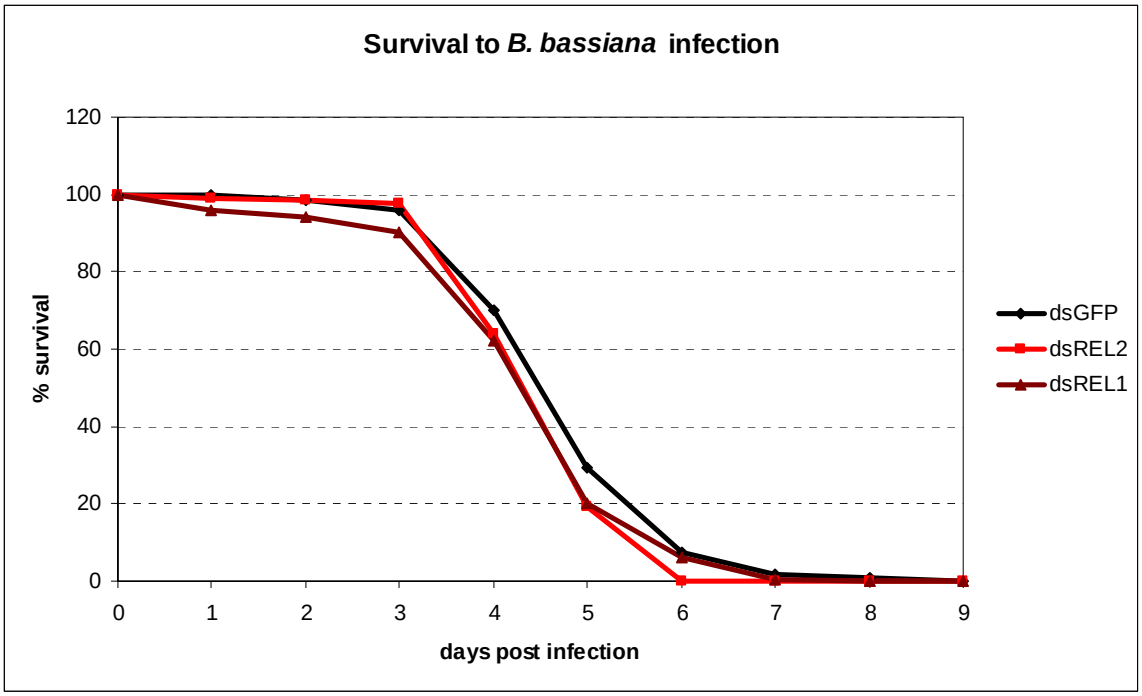


Fig 4.7.3 *A. gambiae* survival to infection with *B. bassiana* after KD of *REL1, REL2* or treatment with ds*GFP*.

DISCUSSION

5.1 THE *A. GAMBIAE* PGRP GENE FAMILY

The first step on the way to activation of the immune signaling pathways is the recognition of PAMPs (pathogen associated molecular patterns) by specific receptor proteins (Janeway and Medzhitov 2002). In insects, the PGRP and GNBP protein families of receptors seem to have the main responsibility of pathogen recognition and pathway activation (Royet, Reichhart et al. 2005).

The strong upregulation of the *PGRPLB* gene in cells challenged with immune elicitors and in the mosquito upon malaria infection (Dimopoulos, Christophides et al. 2002) drew our attention to the PGRP family, even though the gene KD by RNAi proved not to have a significant effect on mosquito survival after subsequent bacterial challenge.

Members of the family of PGRPs have one or more PGRP domains and are important components of insect immune reactions. The first PGRP was characterized in *Bombyx mori* as a hemolymph protein that binds PGN and activates the PPO cascade (Yoshida, Kinoshita et al. 1996). In *Drosophila*, secreted PGRP-SA is essential for activating the *Toll* signaling pathway mediated response to Gram+ bacteria, but not to fungi (Michel, Reichhart et al. 2001), while PGRP-LC activates an alternative immune signaling pathway (*imd*) that responds to Gram- bacteria to induce certain AMPs (Choe, Werner et al. 2002; Gottar, Gobert et al. 2002).

Of the 13 *Drosophila* genes that encode PGRP domains (Werner, Liu et al. 2000), seven are classified as short (S) and encode secreted proteins, while six genes of the long (L) subfamily encode transmembrane or intracellular products. The *Anopheles* genome includes three members of the short subfamily (*PGRPS1*, *PGRPS2* and *PGRPS3*) and four of the long

subfamily (*PGRPLA*, *PGRPLB*, *PGRPLC* and *PGRPLD*). The latter are clear orthologs of correspondingly named *D. melanogaster* genes.

Among the short PGRPs, *PGRPS1* was identified as the ortholog of the *Drosophila PGRP-SA* gene, but our results, using reverse genetics and bacterial challenge, could not implicate it in bacterial defense in *Anopheles*. It remains to be seen if *PGRPS1* is able to rescue the *Drosophila PGRP-SA* mutant *semmelweis*. These experiments are currently ongoing.

The sequence of the other two short PGRPs (PGRPS2 and S3) differ only in 9 amino acids which reveals an amazing subtlety in specificity and prompts us to speculate that these genes may have an important role in the mosquito defenses. The recognition of their putative ligands has to be critical enough, so as to justify the maintenance of these two almost identical genes in a gene cluster. An alternative explanation would a gene regulatory mechanism involving RNAi, as these genes are on opposing DNA strands. Preliminary evidence obtained from the first version of the *A. aegypti* genome suggests that there, too, exist two very closely related, clustered short PGRPs.

Relatively little is known about *PGRPLA* and *PGRPLD*. They could not be implicated in bacterial defense, and the *PGRPLA1* isoform does not appear to have a role in defense against the malaria parasite. However, KD of *PGRPLA2* increases the parasite load per infected midgut in the G3 mosquito strain by 56.6%. Interestingly, we couldn't detect this gene by PCR in freshly prepared adult mosquito cDNA or various cDNA libraries (larval, thoracical, abdominal). This might indicate that the gene is only expressed at very low levels or that its expression is restricted to specific tissue or cell types.

5 PGRPs in *Drosophila* (PGRP-SB1, PGRP-SB2, PGRP-SC1, PGRP-SC2, PGRP-LB) and 3 PGRPs in *Anopheles* (PGRPS2, PGRPS3, PGRPLB) have all the required attributes

and features (disulfide bridges and catalytically important residues) of a catalytically active Amidase. Importantly, the PGRPs with putative Amidase activity cluster together on the phylogenetic tree which leads to the intriguing conclusion, that the PGRPs not phylogenetically clustering with these catalytically active PGRPs are likely to be involved in signaling instead.

PGRPS2, *PGRPS3*, *PGRPLD*, *PGRPLB*, *PGRPLA1* and *PGRPLA2* genes all have either no assigned ortholog or their ortholog has not been implicated in signaling. A recent report suggested that human PGRPs kill bacteria by interacting with the PGN of their cell wall (Lu, Wang et al. 2005) – unlike AMPs which permeabilize their membranes (Zasloff 2002; Ganz 2003; Brogden 2005). The authors postulated, that PGRPs represent a new class of bactericidal and bacteriostatic proteins with a structure and mechanism of action different from that of classical AMPs (Lu, Wang et al. 2005). It is possible that some of the aforementioned *Anopheles* PGRPs fulfill a similar bactericidal role in insects.

The PGRPLC *gene.*

PGRPLC was the only gene having an effect on mosquito survival after bacterial challenge when knocked down. It also had an effect on parasite development in the mosquito. This prompted us to investigate the *PGRPLC* gene cluster in *A. gambiae*, which is 21.1 kb long and encompasses the *PGRPLA* and *PGRPLC* genes. *PGRPLA* has two PGRP domains and *PGRPLC* has three domains. Compared to the analogous gene cluster in *D. melanogaster*, it is slightly bigger by 4 kb because of longer intronic and intergenic sequences, and has lost the *PGRP-LF* gene and gained a PGRP domain in *PGRPLA*. It is very likely that the common ancestor of flies and mosquitoes had just one single domain *PGRPLC*

gene, because a phylogenetic analysis showed that the PGRP domains are more similar within each species than across species and because the second splice site of all the *Anopheles* PGRPLC domains are unique to *A. gambiae*.

The PGRP domains of both *PGRPLC* and *PGRPLA* are encoded by three exons per domain which are of exactly the same length between the domains, and within a gene and can thus in theory be rearranged by alternative splicing and still yield functional PGRP domains. In an attempt to detect potential hybrid *PGRPLC* domains, RT-PCR with a variety of primers distributed along the length of the gene was performed. However, only one novel, though not functional (frameshift) PGRP domain could be detected. We could also detect a pool of unspliced transcripts for the *PGRPLC1* and *PGRPLC2* domains. Only *PGRPLC3* appeared to always be spliced. These transcripts were not due to genomic DNA contamination and although the presence of nuclear unspliced mRNA cannot be ruled out, it is an insufficient explanation as the ribosomal control gene S7 appeared to be completely spliced. We hypothesize that this unspliced pool of transcripts may act as a immature transcript reservoir to decrease reaction timing in case of infection and to direct the splicing to the *PGRPLC* isoform that is best suited to deal with the specific invader. Indeed we were able to detect by microarray analysis that the *PGRPLC* domains are differentially regulated in response to various pathogens.

The apparent efficiency and specificity of KDs by RNAi was limited and may be complicated both by questions of PGRP turnover rates and by what appears to be a reserve of unspliced transcripts. Knocked down transcripts could be replenished from this pool and other mRNA isoform transcript levels could be affected. Even more worryingly, the other

PGRPLC domains appeared to compensate for the reduction of the targeted domain by increasing their transcript levels.

Nevertheless, both by RT-PCR and by qRT-PCR we observed significant reduction of transcript levels. The clearest results were obtained by qRT-PCR in dissected midguts and carcass, where the apparent reduction in transcript levels after dsRNA-mediated silencing was in the range of 34 to 76 percent.

Structural modeling of the PGRP domains.

The rational of splice isoforms was best investigated by exploring the 3D structure of the PGRPLC domain. Also, we aimed to detect structural difference between the three domains of *PGRPLC* that might reflect different recognition specificities. As PGRP domains display an exceptional level of conservation, we modeled the domains in question onto solved 3D structures of *Drosophila* PGRP domains. To this date, crystal structures of *Drosophila* PGRP-LB and PGRP-SA and human PGRP-IαC and -S have been reported (Kim, Byun et al. 2003; Chang, Pili-Floury et al. 2004; Guan, Malchiodi et al. 2004; Guan, Wang et al. 2005). Importantly, all these structures adopt a conserved surface groove that has been demonstrated to be the PGN-docking groove by mutational and structural analysis (Chang, Pili-Floury et al. 2004; Guan, Roychowdhury et al. 2004). We finally used the structural model of the *Drosophila PGRPLB* (Kim, Byun et al. 2003) as a template to model the individual domains of *PGRPLC* onto.

The bottom of the PGN binding site is conserved between the three domains as it is encoded mainly by the first common exon. However, the ridges of the pocket encoded by the unique exons of each domain are rather divergent and are thought to confer different

binding specificities. When the PGN recognition sites between isoforms within each insect species were compared, diversity at the edges of the binding pocket was intensified by small loop insertions or deletions (Fig. 4.3.10), suggesting isoform-specific topological changes. Most probably, these changes of isoforms are "tailored" to be specific to PGN ligands. When mutating amino acids on both edges of this groove region of PGRP-SA in *Drosophila*, almost every position tested led to impaired PGN binding and Toll signaling activity (Chang, Pili-Floury et al. 2004). Thus, further structural studies of the *Anopheles* PGRPs could provide the basic understanding of ligand recognition.

The postulated rearrangement of exons in the mRNA transcript would thus form new hybrid domains that are indeed likely to produce new recognition surfaces. As all three exons contribute equally to the binding site, we could speculate that a possible hybrid domain may increase the binding spectrum and specificity of *PGRPLC* pattern recognition capacity. Indeed, *Drosophila* PGRP-LC domain homo- and heterodimers have since been shown to distinguish between both polymeric and monomeric forms of Gram- DAP-type PGN (Kaneko, Goldman et al. 2004; Stenbak, Ryu et al. 2004) and even to mount a response to Lys-type PGN (Kaneko, Golenbock et al. 2005; Kaneko and Silverman 2005).

Just like the PGN docking pocket, the 'back face' groove displays a conserved hydrophobic surface with extremely variable edges. This was also observed when the *Drosophila* PGRPs were compared (Fig. 4.3.10). This hints at the fact that the 'back face' of PGRP domains might serve an equally important role to pattern recognition as the front, as growing evidence suggests that they might interact with immunity effector proteins (Pili-Floury, Leulier et al. 2004), or other PAMPs such as LPS (Werner, Borge-Renberg et al. 2003; Tydell, Yuan et al. 2006) or 1,3-β-Glucan (Lee, Osaki et al. 2004).

The whole of the PGRP domain is extremely rigid and no conformational changes can be invoked to explain a signal cascade initiation upon ligand binding (Kim, Byun et al. 2003). There has since been evidence accumulating that PGRPs dimerize when brought in physical proximity (Werner, Borge-Renberg et al. 2003; Mellroth, Karlsson et al. 2005) through binding to the same surface and thus initiate signaling with their hydrophobic 'back face'.

In the case of the *Anopheles PGRPLC*, the two beta strands representing the putative protein interaction region on the 'back face' of each domain are encoded by the common exon, and are thus likely to interact with the same downstream signaling partners. This means that the signals originating form the various *PGRPLC* domains most likely are integrated to amplify their combined signal. This is in agreement with the fact that we showed *PGRPLC* to be important to the immune response to both *E. coli* and *S. aureus* bacteria.

The mosquito **PGRPLC** *is implicated in defense to bacteria.*

We confirmed the relevance of the *PGRPLC* gene to the *A. gambiae* innate immune system by gene knockdown (KD) and subsequent infection with Gram+ (*S. aureus*) and Gram- (*E. coli*) bacteria. In both cases, KD of the gene severely impairs mosquito resistance the bacteria. As far as individual PGRPLC domains in these assays are concerned, PGRPLC3 appears to be the most important one as its KD reduces survival after bacterial challenge to a similar extent as the whole gene KD does. This is interesting in the context that the *PGRPLC3* transcript is the one that is always readily spliced. PGRPLC2, on the other hand, does not seem to play an important role in defense to the Gram+ *S. aureus*.

Thus we speculate that PGRPLC could act as a PRR for both Gram types of bacteria unlike its *Drosophila* counterpart which is believed to be only necessary to defend against Gram- bacteria. However, this result prompted us to reinvestigate *PGRP-LC*'s role in *Drosophila* by subjecting a *PGRP-LC* mutant fly strain (Gottar, Gobert et al. 2002) to the same assays with the same bacterial strains that we had used on the mosquito. We also tested the *PGRP-SA*^{seml} (*semmelweiss*) mutant strain (Michel, Reichhart et al. 2001) for further comparison. The results showed that, as expected, the *semmelweis* mutant had increased mortality after Gram+ (*S. aureus*) but not after Gram- (*E. coli*) challenge. However, the mutant *PGRP-LC* fly stain displayed a reduction in survival after challenge with either type of bacteria. This latter result was surprising, but exactly analogous to the results obtained in the mosquito.

From this data, we conclude that there has not necessarily been a switch in pathways in *Anopheles* as we initially believed after the mosquito survival experiments. An alternative explanation for the reduced survival after Gram+ bacteria infection in the *PGRP-LC* mutant could be the fact that *PGRP-LC* has multiple modes of action. It can activate AMP expression (Choe, Werner et al. 2002; Gottar, Gobert et al. 2002) and/or initiated phagocytosis (Ramet, Manfruelli et al. 2002) and thus different functions of the gene might result in reduction in survival for different bacteria, i.e. failure of AMP expression in the case of *E. coli*, but lack of phagocytosis in the case of *S. aureus*. However, RNAi screens in *Drosophila* and *Anopheles* cultured cells point to *PGRP-LC* as a key player for phagocytosis of only Gram- (*E. coli*) but not Gram+ (*S. aureus*) bacteria (Ramet, Manfruelli et al. 2002; Moita, Wang-Sattler et al. 2005).

Another hypothesis was that the constitutive activation of *PGRP-LC* keeps commensal bacteria in check and thus its KD would leave the immune system exposed to such bacteria. The insect would thus be very fragile as a result and the subsequent experimental infection is not the cause of the observed phenotype. However, we tested this hypothesis by comparing the survival rates of PGRP KD mosquitoes that were infected with bacteria or injected with just saline solution. The results falsified this hypothesis as no increase of the death rate was observed in the non-infected mosquitoes.

A role for **PGRPLC** *in malaria parasite killing.*

As the immune system is known to be activated by the malaria parasite, we investigated the involvement of *PGRPLC* in the defense against *P. berghei*. KD of the whole *PGRPLC* gene in susceptible G3 mosquitoes results in a 2.5-fold increase of parasites in infected mosquito midguts. In addition, a residual number (2.5%) of parasites gets melanized. However, the specific KDs of individual *PGRPLC* domains do not seem to have an effect on the malaria parasite, suggesting that it is the concerted action of the different isoforms that mediates parasite killing.

The numbers of malaria parasites per midgut obtained when infecting the refractory mosquito strain L3-5 after *PGRPLC* KD (Fig. 4.3.17), are slightly different from those obtained in G3. The increase in parasites in L3-5 is not as high after *PGRPLC* gene KD - though the melanization reaction at large is unperturbed (3.35% percent of the parasites are now alive as compared to none in the GFP control). However, the domain specific KD of *PGRPLC3* alone increases parasite numbers in the L3-5 strain, unlike in the G3 strain. In

fact, *PGRPLC3* can probably be held accountable for the increase in parasite numbers we see in the whole gene KD.

This slight discrepancy between the two strains is not without precedent – the two strains have displayed this kind of inconsistent behavior after KD of the same gene in the past ((Volz, Osta et al. 2005) and (Fig. 4.4.14)). Nevertheless, the *PGRPLC* KD results are in good accordance with the results obtained in G3 *Plasmodium* infected midguts after KD of *IMD* and *REL2* (Fig. 4.4.6). Parasite numbers roughly double after KD of this pathway and the balance of melanization gets disturbed which results in a low but consistent number of melanized parasites per midgut.

Three other PGRP genes also show an effect on parasite numbers in susceptible mosquitoes. *PGRPLA2* shows a similar phenotype to that of *PGRPLC* although the numbers are not as much increased. Notably, *PGRPLA2* is not detectably expressed in the various mosquito tissues tested. It is most likely expressed at very low levels in specific tissues or cells (e.g. hemocytes). In contrast, the simultaneous KD of the – very similar – *PGRPS2* and *S3* genes leads to lower parasite numbers (approximately 44%). Both genes encode proteins with a predicted amidase activity and thus, based on evidence from *Drosophila*, may function as clearing and inactivating PGN. If this is true, the proposed model of action of *PGRPLC* during *Plasmodium* infection is the following: commensal bacteria that reside in the mosquito gut (and/or propagate during a blood meal) constitutively activate the PGRPLC-triggered pathway that consequently keeps their numbers under control. A side effect of this pathway activation is killing of a large number of parasites that attempt to invade the mosquito gut. This may be achieved through parasite agonists such as LRIM1 (Osta, Christophides et al. 2004) and CLIPB14/B15 (Volz, Osta et al. 2005), all of which are

targets of the PGRPLC-activated *IMD/REL2* pathway (Meister, Kanzok et al. 2005). At the same time, catalytic PGRPs, such as PGRPS2 and S3, negatively regulate the pathway via scavenging of the bacterially shed PGN. Silencing of these PGRPs constitutively activates the pathway, resulting in increased killing of *Plasmodium* parasites.

5.2 THE *A. GAMBIAE IMD* PATHWAY

We showed that *PGRPLC*, and thus its downstream immune signaling pathway, the *Imd* pathway (Choe, Werner et al. 2002; Gottar, Gobert et al. 2002), is very important for the mosquito defense against bacteria as well as the malaria parasite. Two PGN recognition proteins have thus far been shown in *Drosophila* to mediate activation of the *Imd* pathway in *Drosophila*: the transmembrane PGRP-LC (Choe, Werner et al. 2002; Gottar, Gobert et al. 2002) and the extracellular PGRP-LE (Takehana, Katsuyama et al. 2002). After an initial recognition event, the signal is passed on directly to Imd (Choe, Lee et al. 2005), an adaptor protein carrying a Death domain, which is commonly associated with proteins controlling apoptosis (Lemaitre, Kromer-Metzger et al. 1995). Interestingly, the second *Imd*-pathway-associated receptor, *PGRP-LE*, has no *Anopheles* orthologue (preliminary data suggests that it might be present in *A. aegypti)*. In contrast, intracellular components of the *Imd* pathway are highly conserved in *Anopheles* (Christophides, Zdobnov et al. 2002).

Downstream of *Imd*, the pathway appears to fork into two parallel branches (Fig. 1.6.1). One path involves the *Drosophila* homolog of the MAPKK (MAP-Kinase-Kinase) TAK1 and the mammalian I-κB kinase complex (IKK) (Silverman, Zhou et al. 2000). The complex phosphorylates the inhibitory carboxyterminal I-κB domain of Relish, the third member of the NF-κB family of transcription factors, which acts similarly to the *Toll* pathway inhibitor Cactus, preventing the nuclear translocation of the transcription factor domain when the pathway is inactive (Dushay, Asling et al. 1996). This I-κB domain is thus marked for cleavage and subsequently degraded by the proteasome.

The other path signals through the death domain carrying FADD to the caspase Dredd which proteolytically cleaves the I-κB domain of Relish, freeing the transcription factor to enters the nucleus. Whether phosphorylation of the I-κB domain precedes cleavage by Dredd is unclear (Stoven, Ando et al. 2000). Once inside the nucleus, Relish induces transcription of downstream effector genes such as AMPs.

Orthologs in *Anopheles* have been identified in the genome for all the aforementioned genes (*PGRPLC, IMD, IKK1, IKK2, REL2, CASPL1, TAK1, FADD*) except for *PGRP-LE*, (Christophides, Zdobnov et al. 2002; Christophides, Vlachou et al. 2004) (Fig. 4.1.1).

The intracellular components of the Anopheles IMD *pathway.*

We examined the intracellular part of the *IMD* pathway in *A. gambiae* by RNAi. Silencing the receptor adaptor protein IMD resulted in reduction of survival after *S. aureus* (Gram+) infection, similar to the putative receptor PGRPLC. However, *IMD* KD appeared to have no effect when the mosquitoes were exposed to *E. coli* infection. These results could indicate that although PGRPLC may recognize both *S. aureus* and *E. coli* (or both Gram+ and Gram- bacteria), it only passes the signal through IMD in the case of Gram+ *S. aureus*, and that in the case of *E. coli* it either triggers a different signaling cascade or promotes bacterial phagocytosis which is independent of IMD. The fact that resistance to *E. coli* is mediated by a form of Relish (REL2-S; see below) that lacks the inhibitory I-κB domain supports the former hypothesis, although the latter hypothesis cannot be excluded.

CASPL1, *IKK1* and *IKK2* are likely to also be part of the *IMD* pathway in *Anopheles*, as their KD resulted in reduced activation of a *CEC1* reporter construct in cell lines, similar to that of *IMD* and *REL2*-F, although silencing of *CASPL1* and *IKK1* or *IKK2* surprisingly but

consistently had opposite effects in the two cell lines. In Sua1B cells, KD of *CASPL1* but not of *IKK1* or *IKK2* substantially reduced *CEC1* promoter expression. In 4a3A cells, however, KD of *IKK1* or *IKK2* but not of *CASPL1* reduced *CEC1* promoter expression (Fig. 4.4.8).

We used semiquantitative and quantitative RT-PCR to confirm silencing of the genes. Importantly, KD of *IKK1* and *IKK2* in Sua1B was highly effective, suggesting that these genes are not implicated in the regulation of *CEC1* in Sua1B cells. KD of *CASPL1* in the same cells was less effective (~40%) and, yet, associated with substantial reduction in CEC1 activity. A plausible explanation that requires further investigation is that the pathway branches of *IKK1*/*IKK2* and *CASPL1* are alternative branches of the *CEC1*-activation pathway that are differentially active in these two cell types. Recent, preliminary results show that *CEC1* activation following *S. aureus* infection is indeed mediated by PGRPLC.

Thus, one would expect *CASPL1* and *IKK1*/*IKK2* to be situated downstream of *IMD* in the pathway, and only responsive to *S. aureus*. However, KD of *CASPL1* in adult mosquitoes did not show any effect following *S. aureus* (or *E. coli*) infection. This result could suggest that either both branches of the pathway are needed or that *CASPL1* is not important for resistance to *S. aureus* (*IKK1* and *IKK2* have not been tested).

REL2 mediated defense against bacteria and malaria parasites.

The transcription factor of the *Drosophila Imd* pathway is the NF-κB factor *Relish* (Dushay, Asling et al. 1996), and its ortholog in *Anopheles* is *REL2* (Christophides, Zdobnov et al. 2002). However, there is a major difference between the two genes. *REL2* produces two isoforms, a full length form, *REL2*-F, and a short form, *REL2*-S, which lacks the I-κB

inhibitory domain. The REL2-F situation is a similar setup as with two (p52/p100; p50/p105) of the five mammalian NF-κB factors (Rel-A (p65), NF-kappaB$_1$ (p50/p105), NF-kappaB$_2$ (p52/p100), c-Rel and Rel-B (Ghosh, May et al. 1998)) that also produce proteins that carry their own Ankyrin (I-κB) domains. Furthermore, *REL2*-F and the mammalian p100/p52 and p105/p50, but not the *Drosophila* Relish, contain a death domain at their carboxy terminal, after the I-κB domain. This could imply a drastic difference in the transcription factor activation and thus the signaling events between the mosquito and the fly NF-κB pathways. It is noteworthy that the ortholog of *REL2* in the mosquito *A. aegypti* has similar structural characteristics and is also able to produce two isoforms of the transcriptional activator, one with the I-κB and death domains and one without (Shin, Kokoza et al. 2002).

Indeed, our results reveal that the two forms of REL2 are differentially involved in defense against Gram+ and Gram- bacteria: REL2-F confers resistance to *S. aureus* infections and REL2-S to *E. coli*. In this way, the single mosquito gene *REL2* mediates alternative immune responses, which, in the fruit fly require two genes: *Relish* and *Dif*. Interestingly, Dif is not present in the mosquitoes (Meister, Kanzok et al. 2005; Shin, Kokoza et al. 2005). Thus, in addition to the use of *REL1*, posttranscriptional processing of *REL2* could be another strategy applied by *Anopheles* (and possibly *Aedes*) to compensate for the absence of *Dif*.

In addition to compensating for the absence of *Dif*, this data could imply a major difference in the evolution of the NF-κB signaling pathways during the ~250 million years from the last common ancestor of *Drosophila* and *Anopheles* (Yeates and Wiegmann 1999; Gaunt and Miles 2002). On the one hand, the *Anopheles PGRPLC*, *IMD* and *REL2*-F are

used for defense against the Gram+ bacterium *S. aureus*, whereas the orthologous and structurally similar cascade in the fruit fly is required for defense against Gram- bacteria. On the other hand, *PGRPLC* and *REL2*-S are responsible for defense against the Gram- bacterium *E. coli*, possibly through a different signaling cascade, as *IMD* is not involved in this reaction. *REL2*-S is structurally analogous to *Dif* and *Dorsal* which deal with fruit fly responses against Gram+ bacteria as none of them has an inhibitory I-κB domain. Together with our data about the mosquito *REL1*, the ortholog of *Dorsal*, which is discussed further below, these results may suggest that the function that the *Toll* pathway has in the *Drosophila* anti bacterial defense is likely to have been assumed by a different pathway of the mosquito, which derived from the original *IMD* pathway and involves *PGRPLC* and *REL2*-S, but not *IMD*.

As discussed above, *REL2-F*, unlike its fly counterpart *Relish* but similar to orthologous genes in other higher eukaryotes, encodes a death domain located at the carboxyl terminus of the deduced protein. This domain might be used for oligomerization with the respective domain of IMD or other death domain-encoding proteins (such as *CASPL1*). Pertinent information is available from studies on the *Relish* gene of *A. aegypti*, the vector of yellow and dengue fever, which belongs to a different mosquito subfamily (*Culicinae* rather than *Anophelinae*). In addition to the two isoforms that are orthologous to the *Anopheles* REL2-F and REL2-S, the *Aedes REL2* produces a third transcript that encodes a protein encompassing only the ANK and Death domain (Shin, Kokoza et al. 2002). It is pertinent to speculate that a similar REL2 protein form having is expressed in *Anopheles*, which we have missed in our analysis. In such a case, this protein could function as the inhibitor of the REL2-S nuclear translocation. As discussed also in the next chapter, we have

shown that this role is not served by CACT, as one would expect since Cactus is the inhibitor of Dif and Dorsal in *Drosophila*.

In *Aedes*, transgenic overexpression of a truncated version of the *REL2*-S homologue (C8), which lacks the putative glutamine- and histidine-rich transactivator domain, results in susceptibility to Gram- but not to Gram+ bacteria (Shin, Kokoza et al. 2003), similar to the results we obtained in *Anopheles*. However, additional studies are required to determine whether C8 indeed acts as a dominant negative allele of *REL2-S*, as suggested. In this case, our postulated change in pathway function between *Anopheles* and *Drosophila* would probably have occurred in the common ancestor of anopheline and culicine mosquitoes. In strong support of our hypothesis about the difference in immune pathways between mosquitoes and flies, more recent studies have shown that the *Aedes* orthologue of *REL1* is not involved in defense against bacteria but only against fungal infections (Shin, Kokoza et al. 2005). Likewise, our results that are discussed below, indicate that the *Anopheles REL2* may be the only NF-κB factor implicated in the antibacterial defense against *E. coli* and *S. aureus* and that *REL1* has no role in these reactions. Our preliminary data implicates *REL1* in anti fungal response. Thus, it is likely that the two mosquitoes use similar strategies to deal with infections.

An important finding of our work is that the PGRPLC/IMD/REL2-F signaling cascade is involved in limiting the number of *P. berghei* oocysts that develop in the mosquito midgut. As discussed previously, this reaction could be due to activation of the pathway by bacteria residing in the mosquito gut. However, it is believed that the malaria parasite itself is able to elicit immune responses in the mosquito, since challenge of mosquito cell lines with *P. berghei* ookinetes under otherwise septic conditions leads to transcriptional activation of

immunity genes (G.K. Christophides, personal communication), some of which are targets of the *IMD* pathway. Whether *P. berghei* is recognized by receptors of the *IMD* pathway (or PGRPLC itself) remains to be investigated.

Nevertheless, the observed parasite killing through the *PGRPLC/IMD/REL2*-F pathway is thought to be the result of the transcriptional activation of effector genes that are targets of this pathway.

A recent study has shown that ectopic overexpression of *CEC1* drastically reduces susceptibility of mosquitoes to *Plasmodium* (Kim, Koo et al. 2004). Other REL2-regulated proteins (Table 4.4.4) that might explain the observed phenotype are *Gambicin* (*GAM1*), another AMP with anti-parasitic activity *in vitro* (Vizioli, Bulet et al. 2001), *LRIM1*, a key antagonist of ookinete-to-oocyst development (Osta, Christophides et al. 2004), and *CLIPB14/B15*, of which the KD also leads to an increase in parasite numbers (Volz, Osta et al. 2005).

The silencing of *PGRPLC* and *REL2-F* also leads to melanization of some ookinetes during their passage through the *Anopheles* midgut epithelium. This melanization reaction is independent of IMD (melanization is not observed after *IMD* KD), pointing to an alternative signaling cascade. In *Drosophila*, the *Imd* pathway seems to regulate melanization upstream of Serpin27A (Takehana, Yano et al. 2004), which is an inhibitor of the final steps of prophenoloxidase activation and is controlled by the Toll pathway (Ligoxygakis, Pelte et al. 2002).

5.3 THE *A. GAMBIAE TOLL/ REL1* PATHWAY

Our results that far had shown that the *IMD/REL2* pathway, but not the *Toll/REL1*, is involved in the mosquito defense to bacterial infections. The same pathway is used to control the numbers of *P. berghei* parasites that develop to oocysts on the mosquito midgut. In contrast, the *REL1* pathway remains inactive during *P. berghei* infection and we had used *REL1* as control in our experiments in cell lines and when we tried to determine whether CACT was the I-κB factor for REL2-S. The results made it apparent that REL1 played an important role in defense against *Plasmodium* and that we needed to conduct further experiments to examine the role of *REL1* and the *Toll* pathway.

Upstream of the *Drosophila* Toll pathway, the soluble PGN recognition protein PGRP-SA (Michel, Reichhart et al. 2001) and GNBP1 (a β-1,3 glucan binding protein) (Pili-Floury, Leulier et al. 2004) act in concert to recognize Gram+ bacteria and mediate activation of the pathway (Gottar, Gobert et al. 2002). Furthermore, another member of the GNBP family, GNBP3, was implicated in the recognition of fungi which also active the Toll pathway (Leclerc and Reichhart 2004). Upon bacterial and fungal recognition, the signal is conveyed to a proteolytic cascade, ultimately leading to cleavage of the cytokine-like polypeptide Spaetzle (Spz), which in turn binds to and activates the receptor activity of Toll (Lemaitre, Nicolas et al. 1996). Subsequent conformational changes of Toll result in the intracellular recruitment of at least three cytoplasmic proteins, MyD88, Tube and Pelle. The latter is a serine-threonine kinase believed to play an indirect role in the phosphorylation and subsequent proteolytic degradation of Cactus, a member of the I-κB family of proteins, that

normally bind to and prevents nuclear translocation of the NF-κB transcription factors, Dorsal and Dif. Whereas Dorsal is essential for developmental processes, Dif is mostly implicated in the transcription of AMPs through specific binding to cis-acting elements (κB) found in the promoter sequences of these genes.

Orthologs of *PGRP-SA* (*PGRPS1*) and the intracellular signaling molecules (*MYD, TUBE, PLL1*) were found in *Anopheles*; however, no orthologs of *Dif* and *GNBP1* were detected, and *Toll* forms an orthologous group together with four mosquito genes (*TOLL1A*, *TOLL5A*, *TOLL1B*, *TOLL5B*) (Christophides, Zdobnov et al. 2002; Luna, Wang et al. 2002).

Our results suggest, that unlike in *Drosophila,* where the *Toll* pathway appears to be important for bacterial and fungal infections, the pathway might have a secondary role in defense against bacterial infections in *Anopheles*. Preliminary results suggest that the *Anopheles Toll* pathway may have a role in the antifungal defense, similar to the *Drosophila Toll*. Interestingly, *Drosomycin*, the prime anti-fungal peptide in *Drosophila* has no ortholog in *Anopheles*.

The mosquito *Toll* pathway also appears to play an important role during infection with the malaria parasite *Plasmodium*. Our results suggest that the pathway either fails to be activated or is actively evaded by the parasite. The consequences of *Toll* pathway activation are lethal for the parasite as REL1 mediates effective killing of all *P. berghei* parasites by both lysis and melanization. This can be achieved by elimination of CACT, the suppressor of REL1 nuclear migration (see below). The implication of the *Toll* pathway in parasite melanization is not surprising, given that there is a direct link between the *Toll* pathway and melanization in *Drosophila* (Ligoxygakis, Pelte et al. 2002).

We cannot comment at this point on the conservation of the intracellular signaling events between the *Anopheles* and the *Drosophila Toll* pathways, however, the transcriptional NF-κB activator appears to be conserved as *REL1* is the ortholog of *Dorsal*.

CACT *is the I-κB inhibitor of* REL1.

The absence of an ortholog of the NF-κB-like transcription factor *Dif* in *Anopheles* may imply that in a putative *Toll* pathway, one of the other two *Anopheles* NF-κB-like proteins, *REL1* (the *Dorsal* orthologue) or *REL2* (the *Relish* orthologue), may substitute for the role of *Dif*. The *Drosophila* Dorsal has been mainly implicated in developmental processes (reviewed in (Belvin and Anderson 1996)), but recent reports suggested additional roles for Dorsal in immunity (Bettencourt, Asha et al. 2004). A third NF-κB gene with an RHD (Rel homology domain) and IPT/TIG transcriptional activation domain is present in the *Anopheles* genome (ENSANGG00000014870) and would certainly be another potential candidate. However, this gene has an ortholog (CG11172) in *D. melanogaster* that not much is known about other than that its structural characteristics.

The inhibitor of nuclear migration of Dorsal and Dif is the I-κB protein Cactus which has a clear-cut ortholog in *Anopheles*, CACT. Thus it was expected that CACT would be the I-κB inhibitor of REL1. In addition, the fact that REL2-S is lacking the inhibitory Ankyrin repeats domain and serves some for the roles of the missing Dif makes it a likely candidate to also be inhibited by CACT.

To detect the genetic relationship between CACT and REL1 or REL2-S, we used a powerful genetic tool that was revealed by silencing of CACT and subsequent infection with *P. berghei*. *CACT* KD mosquitoes are fully refractory to *P. berghei* infections: more than 80%

of ookinete staged parasites are lysed, and the remaining 20% of parasites are melanized (Fig 4.4.13). The effect of the *CACT* KD on *P. berghei* ookinetes suggests that a previously repressed antagonistic module of the malaria parasite was activated by the KD. It was thus expected that the double KD of *CACT* and this antagonist should rescue the wild-type (wt) phenotype, susceptibility to *P. berghei* parasite development.

However, the double KD of *CACT* and *REL2* produced the same refractory phenotype as the *CACT* KD did. In fact, the phenotype appeared aggravated as the parasite numbers were even more reduced (Fig 4.4.13). This suggested that nuclear translocation of REL2-S is not inhibited by CACT, but by a different I-κB protein, e.g. a third form of REL2, that has only the ANK domain as discussed in a previous section. In contrast, the double KD of *CACT* and *REL1* was able to completely rescue the wt phenotype. Also, KD of *REL1* alone had no significant effect on the malaria infection load of the midgut (Fig 4.4.13). It thus appears likely that *REL1* is the partner of *CACT* in *A. gambiae*; however, this signaling pathway remains silent during *Plasmodium* infection and REL1 is retained in the cytosol. As mentioned above, this latter circumstance may have its origin either in failure of pathway activation (absence or evasion of recognition by the parasite) or in active pathway suppression by the malaria parasite.

Regulation of the melanization reaction is thought to be achieved through serpins (serine protease inhibitors), which act as suicide substrates of the PPO activating serine proteases (PPAEs) (De Gregorio, Han et al. 2002; Ligoxygakis, Pelte et al. 2002; Ligoxygakis, Pelte et al. 2002; Zhu, Wang et al. 2003). A proposed model in *Drosophila* is founded on a preactivation balance between the inhibitory serpin and the PPAE, which changes in favor of the PPAE upon Toll pathway activation (Ligoxygakis, Pelte et al. 2002; Ligoxygakis, Pelte

et al. 2002). This model would explain why some parasites are melanized upon activation of the *Toll* pathway by KD of *CACT* and subsequent nuclear translocation of REL1.

5.4 THE CURRENT WORKING MODEL OF *ANOPHELES* INNATE IMMUNE SIGNALLING.

***An updated view of the innate immune pathways in* A. gambiae.**

The conclusions about signaling pathways from the work reported in this thesis are summarized in Figure 5.3.1. *PGRPLC* is the only PRR we know of at this time that signals to the innate immune pathways in *A. gambiae*. It mediates defense to both Gram- and Gram+ bacteria as the KD of *PGRPLC* results in reduced survival after bacterial challenge with either Gram-type.

The intracellular signaling after infection with Gram+ (*S. aureus*) bacteria is likely to proceed through IMD. Essentially, the *Imd* pathway of *Drosophila,* which mediates defense against Gram- bacteria, seems to be intact and to have changed its recognition properties in *Anopheles*.. CASPL1, IKK1 and IKK2 appear to also be situated downstream of IMD in the mosquito, just like in the fruit fly. The signaling cascade going through IMD results in the activation of the NF-κB factor REL2-F whose KD leaves the mosquito defenseless to Gram+ (*S. aureus*) bacterial infections (Fig. 4.4.5a).

At this time, it is unclear what the intracellular signaling events are in the case of a Gram- (*E. coli*) bacterial infection; however, the signal seems to activate the short form of REL2 (REL2-S). The I-κB factor that retains REL2-S in the cytosol also remains unknown unclear at this point (CACT is not involved in this reaction). It could be speculated that another isoform of REL2, carrying only the ANK domain may function as the inhibitor of REL2-S or that REL2-S in the cytosol exists as a dimer with REL2-F. Preliminary data

reveals that the *Toll/REL1* pathway is also involved in defense to fungal infections with *B. bassiana*.

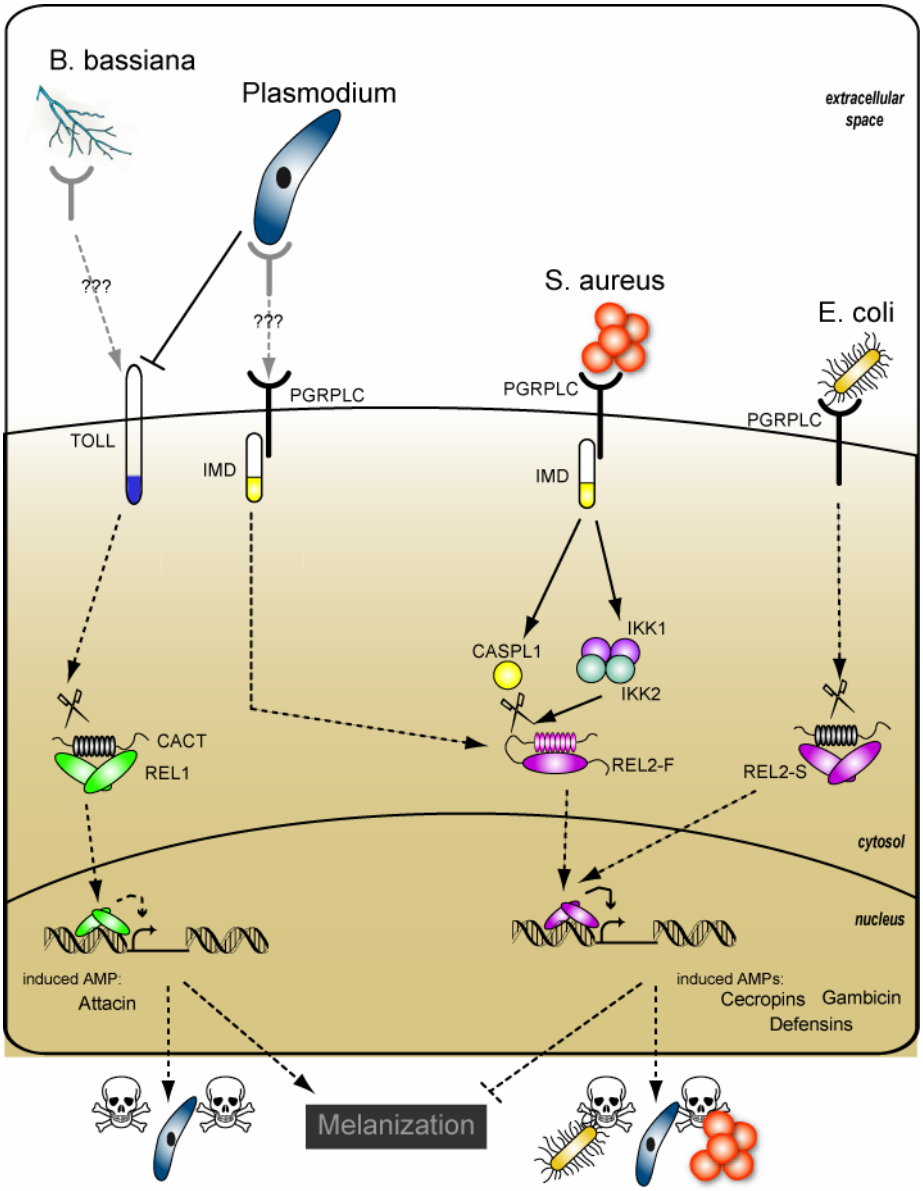


Fig 5.3.1 An updated view of the *A. gambiae REL1* and *REL2* innate immune pathways, based on the results reported in this thesis. Solid arrows signify direct interactions; dashed arrows represent unknown intermediate steps.

The pathway of the *PGRPLC/IMD/REL2*-F genes appears to also be capable of conferring some resistance to *Plasmodium* infections as their KD can double the parasite load in a mosquito midgut. *PGRPLC* and REL2 also influence the expression of genes inhibiting the ookinete melanization reaction. Whether the signaling in response to *Plasmodium* involves CASPL1, IKK1 and IKK2 like in the case of Gram+ bacteria, was not tested, but seems likely.

A working model for the defense of Anopheles *against* Plasmodium.

Infection of *A. gambiae* with the rodent malaria parasite *P. berghei* can be responded to by both the *Toll/REL1* and the *IMD/REL2* pathway, which are capable of driving lysis of parasites before or during invasion of the mosquito midgut. Melanization of the *Plasmodium* ookinetes appears to be also positively controlled by the *TOLL/REL1* pathway (Fig. 4.4.13), whereas the *IMD/REL2* seems to negatively affect ookinete melanization. Although the *REL2* pathway is a factor affecting the final number of parasites developing in the mosquito, the *REL1* pathway either fails to be activated during infection is actively inhibited by the parasite. When integrating these results, we can construct a working model of the *Anopheles* defense against *Plasmodium* that is shown in Fig. 5.3.2.

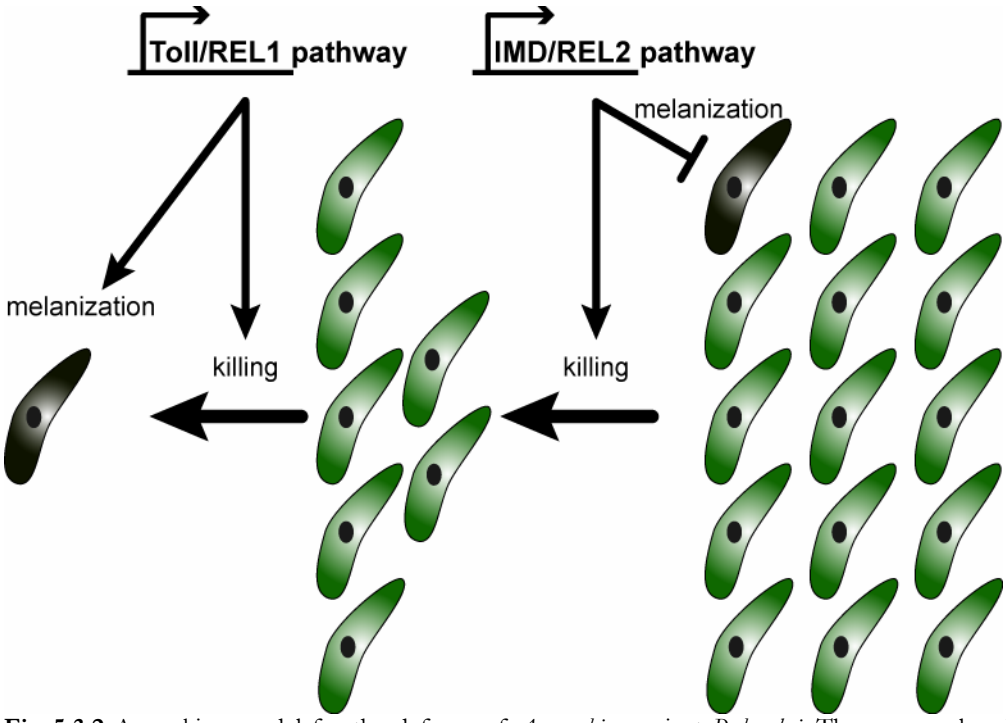


Fig 5.3.2 A working model for the defense of *A. gambiae* against *P. berghei*. The green colored parasites represent live parasites, black parasites symbolized melanized parasites. The 7 green parasites in the center are the ones that will develop to the oocyst stage in naïve infected mosquitoes. The 15 parasites on the right are what we see after KD of *PGRPLC/IMD/REL2* and the one on the left is the picture presented to us after *CACT* KD.

According to this working model, the *PGRPLC/IMD/REL2* pathway kills and lyses about half of the initial parasite population and through an unknown mechanism inhibits parasite melanization. As discussed in previous sections, this lysis reaction could be a side-effect of the pathway activation in response to the bacterial proliferation in the mosquito midgut after the blood meal. We do not know if the mosquito immune system is capable of directly recognizing the *Plasmodium* parasite, and if so, what receptor would mediate this recognition.

The next step in Fig. 5.3.2 is generally suppressed in the mosquito, possibly by means of immune evasion by *Plasmodium*. This would not be unprecedented, e.g. another apicomplexan parasitic protozoa, *Toxoplasma*, inhibits the phosphorylation and thus the

activation of NF-κB factors (Shapira, Harb et al. 2005). However, when the *TOLL/REL1* pathway is activated by KD of *CACT* all parasites are killed, about 80% of parasites are lysed and the remaining 20% are melanized. It may not be possible or desirable for the parasite to shut down/evade the *IMD/REL2* pathway as well, as it may be needed to contain the strong bacterial growth in the mosquito midgut following a blood meal. Some of the AMPs secreted in response to these bacteria have however indeed been shown to harm the parasite (*GAM1* (Vizioli, Bulet et al. 2001); *CEC* (Kim, Koo et al. 2004)).

Microarray expression analysis of *REL2* KD in cell lines revealed several target genes of this pathway. Among these, was the major parasite antagonist *LRIM1* and the AMP *GAM1*, which has been shown to have in vitro activity against *Plasmodium* (Vizioli, Bulet et al. 2001; Osta, Christophides et al. 2004).

We have also done a whole mosquito *REL2* KD microarray hybridization (Data not shown) and preliminary results indicate that among the genes down regulated after *REL2* KD are: *GAM1* (Vizioli, Bulet et al. 2001), *KIN1*, *CEC1*, *CEC3*, *CLIPB14* (Volz, Osta et al. 2005), *CLIPB15* (Volz, Osta et al. 2005), *DEF1*, *LRIM1* (Osta, Christophides et al. 2004) – all of which have either been shown or are likely to reduce the ability of the parasite to survive in the mosquito.

A similar preliminary microarray analysis of the KD of *CACT* in combination with KD of *REL1* in the whole mosquito revealed numerous targets of the REL1 pathway: *DEF1*, *CEC1*, *GAM1*, Serine Proteases, *KIN1*, *CLIPB15*, *TEP15*, a *GNBP*, *SOD*, *CLIPA3* and a *Lysozyme* (Data not shown). These results, albeit preliminary, clearly show that there is extensive overlap in the regulated target genes of *REL1* and *REL2*. The only genes specific

to *REL1* appear to be *SOD* and *CLIPA3*, while the only gene truly specific to *REL2* transcriptional activation seems to be *LRIM1*.

Among the experiments to be done in the future would thus be the double KD of *CACT* and *LRIM1*. It will be interesting to see if *LRIM1* is able to reverse the *CACT* KD phenotype, suggesting a fine interplay between the two pathways. Another would be the double KD of *CTL4* and *CACT* or *REL2*, as it would be interesting to know if *CTL4* is activated by the *REL1* pathway and if the parasite killing following these KDs is cumulative and the mechanisms by which it is achieved is the same or distinct.

Finally, another interesting experimental setup would be to co-infect malaria and a fungus. Assuming that the fungus is activating the *TOLL/REL1* pathway and that the malaria parasite is evading or modulating the pathway activation, we would predict that mosquitoes infected with fungi became partially or completely refractory to malaria.

In conclusion, our data suggests significant divergence of immune signaling between the *A. gambiae* mosquito and the fruit fly *D. melanogaster*. The documented differences most likely reflect their different lifestyles and, consequently, different infectious agents that the two insects encounter during their lifetimes. In mosquitoes, one of these agents is the malaria parasite *Plasmodium*.

Perspectives.

Our knowledge about the innate immune pathways to date almost exclusively derives from studies in *D. melanogaster* (reviewed in: (Hoffmann 2003)). The genetic tools available for *Drosophila* are by far superior to that of most other model organisms. However, this

means that pathway information was for the most part obtained by creation of or screening for mutant flies and using a few genes as pathway readout: *Diptericin* for the *Imd* pathway and the antifungal gene *Drosomycin* for the *Toll* pathway. This relatively simple reductionist approach has revealed the hierarchy of genes in the pathways and is the origin of the convenient assignment of Gram- bacteria to the *Imd* pathway and Gram+ bacteria to the *Toll* pathway. Only slowly are studies published suggesting that these assignments might be an oversimplification and the situation is in fact more complicated. Rather than the Gram type of bacteria, the PGN type (Dap or Lys) seems to be the distinguishing factor for the two immune pathways and the upregulation of AMPs seems to be far less selective than was initially thought (Hedengren-Olcott, Olcott et al. 2004). It is noteworthy, that while the Gram+ genera *Bacillus* and *Clostridium* have the DAP type PGN (found in Gram- bacteria), the Lys type PGN is exclusively found in the rest of Gram+ bacteria.

In order not to loose sight of the big picture, survival assays, like the ones used in this work, could prove to be very useful. Rather than sampling a single reporter gene, survival probability of the whole organism is recorded. We are aware that this approach might well lead to results different to the ones obtained in *Drosophila*; especially genes further downstream in the pathway might not prove to be as relevant for survival as crosstalk between the pathways might be able to compensate for the knock down.

On the other hand – the effect of a knock down might be especially dramatic, the further up in the signaling chain it happens. This may explain why *PGRPLC* has a drastic effect on survival for both Gram types of bacteria, yet it was always only associated with the *Imd* pathway in *Drosophila*.

It has been proposed that the function of the *Toll* pathway is mainly to deal with fungal infections. In fact, in *Drosophila* only the anti-fungal gene *Drosomycin* is exclusively activated by the *Toll* pathway – all other AMPs appear to be activated independently of the Gram type of the infecting bacteria (Hedengren-Olcott, Olcott et al. 2004). We were not able to record any effect on survival for the NF-κB transcriptional activator *REL1* – the supposed functionally equivalent gene of *Dif* in *Anopheles*. This result reinforces the hypothesis that the *Toll* pathway's main function is to deal with fungal infections, a hypothesis that is supported by a study conducted in the mosquito *A. aegypti* (Bian, Shin et al. 2005).

A better and more sensitive indicator of the importance of a knocked down gene for innate immunity would probably have been bacterial survival and/or propagation in the mosquito instead of survival of the mosquito itself. Bacterial numbers in the mosquito could be recorded using bacterial strains with antibiotic resistances and smearing mosquitoes onto antibiotics treated agar plates after a certain time, or by performing RT-PCR on whole mosquitoes with primers targeting strain specific bacterial genes.

SUMMARY

6.1 SUMMARY.

Innate immunity is the first line of defense against invading microorganisms and provides clues to adaptive immunity for the development of memory for subsequent infections. Insects, similar to other invertebrates, do not have adaptive immunity and thus rely on their innate immune system to combat infections. We have analyzed the role of the peptidoglycan (PGN) receptor protein (PGRP) family and other components of innate immune signaling pathways in the immune defense of the mosquito *Anopheles gambiae*, the main vector of human malaria in sub-Saharan Africa. The *PGRP* gene family consists of seven genes with ten PGRP domains. We have shown that from all *PGRP* genes only *PGRPLC* has a role in the resistance to bacterial infections of both Gram types. In our experiments we have used the Gram-positive bacterium *Staphylococcus aureus* and the Gram-negative bacterium *Escherichia coli*. The *PGRPLC* gene encodes at least three isoforms that derive from infection-driven alternative splicing of a pool of immature transcripts. Each isoform contains a different PGRP domain, encoded by three exons that all contribute equally to the PGN binding pocket. Structural modeling of the PGRPLC isoforms revealed a potential for all isoforms to bind both types of PGN, the Lys-type, which is mostly found in Gram-positive bacteria and the DAP-type, mostly found in Gram-negative bacteria. The isoform PGRPLC3 seems to be the most important in the defense against both bacteria species, although PGRPLC1 also has a crucial role in the defense against *S. aureus*.

Bacterial defense is mediated by the NF-κB transcription factor REL2, which is the ortholog of the *Drosophila* Relish. The *REL2* gene also encodes two protein isoforms: REL2-S, which only has the NF-κB domain, and REL2-F, which carries an I-κB inhibitory domain

and a death domain in addition. REL2-F functions together with the receptor adaptor protein IMD to deal with *S. aureus* infections, whereas REL2-S has a role in the defense against *E. coli*. The *PGRPLC/IMD/REL2*-F pathway (*IMD*) is also partly responsible for the losses of *Plasmodium berghei*, which can be observed during the first stages of malaria infection of *A. gambiae*. *P. berghei* is a rodent malaria parasite, which has been used as a model in our studies. Whether the pathway is able to recognize the malaria parasite through PGRPLC or another associated receptor is still unclear. Another possibility is that the pathway is activated by the proliferation of commensal bacteria in the mosquito gut following a blood meal. However, we have shown that more than one of the three main isoforms of PGRPLC are required for the reaction to *P. berghei*. Other PGRP genes, which have been proven to play a role during infection with *P. berghei,* are *PGRPLA2,* which also mediates parasite killing, and the almost identical and thus hardly indistinguishable *PGRPS2* and *S3,* which appear to inhibit parasite killing. This is possibly achieved by negative regulation of the *IMD* pathway through sequestration of PGN, which derives from the commensal bacteria and constitutively activates the pathway.

We have shown that REL1, the second NF-κB transcription factor of *A. gambiae*, which is orthologous to the *Drosophila* Dorsal (Dif does not exist in *Anopheles*), is not involved in the mosquito resistance to bacterial infections. This fact provides additional evidence that the *REL2*-associated pathways are of utmost importance in the *A. gambiae* defense to bacteria. In addition, REL1 has no role in the documented *P. berghei* killing. However, silencing of the REL1 inhibitor CACT (the ortholog of the *Drosophila* Cactus) during a parasite infection leads to a very strong refractoriness phenotype: most of the midgut-invading ookinetes are eliminated (presumably by lysis) and the remaining of the

ookinetes are melanized. We thus assume that under wild-type infection conditions the parasite is either evading recognition by the REL1-associated receptors or actively modulating activation of REL1.

In conclusion, the data reported in this PhD thesis suggest significant divergence of immune signaling between the mosquito *A. gambiae* and the fruit fly *D. melanogaster*. The observed differences most likely reflect their different lifestyles and, consequently, different infectious agents, which the two insects encountered during their evolutionary lifetimes. In mosquitoes one of these agents is the malaria parasite.

6.2 ZUSAMMENFASSUNG.

Angeborene Immunität ist die primäre Verteidigungsstrategie gegen eindringende Mikroorganismen und liefert dem adaptiven Immunsystem Signale für die Entwicklung von Gedächniszellen für nachfolgende Infektionen. Ähnlich wie andere Invertebraten, verfügen Insekten nicht über eine adaptive Immunreaktion und verlassen sich deshalb voll auf ihr angeborenes Immunsystem, um Infektionen zu bekämpfen. Wir haben analysiert, welche Rolle die Familie der Peptidoglycan (PGN) Rezeptor Proteine (PGRP) und andere Komponenten der angeborenen Immunsignalkaskaden bei der Immunantwort des Moskitos *Anopheles gambiae* spielen. *A. gambiae* ist der Hauptüberträger der menschlichen Malaria im südlich der Sahara gelegenen Teil Afrikas. Die Genfamilie der PGRPs besteht aus sieben Genen mit zehn PGRP Domänen. Wir konnten zeigen, dass von allen PGRP Genen nur *PGRPLC* eine Rolle in der Verteidigung gegen bakterielle Infektionen, egal welchen Gramtyps, spielt. In unseren Experimenten haben wir das grampositive Bakterium *Staphylococcus aureus* und das gramnegative Bakterium *Escherichia coli* benutzt. Das *PGRPLC* Gen kodiert mindestens 3 Isoformen, die – je nach Infektion – aus einem Pool von unreifen Transkripten durch alternatives Splicing gebildet werden. Jede Isoform hat eine andere PGRP Domäne, welche jeweils von drei Exons kodiert wird, die alle gleich viel zur PGN Erkennungstasche beitragen. Strukturelle Modelle von PGRPLC zeigten, dass alle Isoformen dazu in der Lage sind, beide Arten von PGN zu binden, wobei die Lys-Form hauptsächlich in grampositiven Bakterien und die DAP-Form hauptsächlich in gramnegativen Bakterien vorkommt. Die PGRPLC3 Isoform scheint die wichtigste Rolle bei der Verteidigung gegen

die beiden bakteriellen Formen zu haben, obwohl PGRPLC1 auch eine wichtige Rolle bei der Verteidigung gegen S. *aureus* zu spielen scheint.

Die Verteidigung gegen Bakterien wird von dem NF-κB Transkriptionsfaktor REL2 bewerkstelligt, welcher das Ortholog des *Drosophila Relish* Gens ist. Das *REL2* Gen kodiert zwei Protein Isoformen: REL2-S, das lediglich die NF-κB Domäne hat, und REL2-F, das zusätzlich noch eine inhibierende I-κB und eine Death Domäne hat. REL2-F arbeitet mit dem Rezeptor-Adaptor Protein IMD zusammen, um *S. aureus* Infektionen zu bekämpfen, wogegen REL2-S eine Rolle in der Verteidigung gegen *E. coli* spielt. Die *PGRPLC/IMD/REL2*-F Signalkaskade (*IMD*) ist auch teilweise verantwortlich für die Verluste, die *Plasmodium berghei*, ein Nager Malariaparasit, während der ersten Stadien der Malaria Infektion in *A.gambiae* erleidet. Dieser Malariaparasit ist als Modelsystem für unsere Studien verwendet worden. Ob die *IMD* Signalkaskade auch fähig ist, den Malariaparasiten direkt durch PGRPLC oder einen anderen beteiligten Rezeptor zu erkennen, ist immer noch unklar. Nach einer Blutmahlzeit vermehren sich die residenten Bakterien im Moskitodarm. Es ist durchaus möglich, dass die Signalkaskade dadurch aktiviert wird. Unabhängig davon, konnten wir zeigen, dass mehr als eine der drei PGRPLC Isoformen benötigt werden, um eine Reaktion auf *P. berghei* hervorzurufen. Weitere PGRP Gene, die nachweislich eine Rolle während der Infektion mit *P. berghei* spielen, sind *PGRPLA2* und *PGRPS2* und *S3*. PGRPLA2 begünstigt den Kampf gegen den Parasiten, während PGRPS2 und S3, die nahezu identisch und deshalb kaum zu unterscheiden sind, den Kampf gegen den Parasiten zu behindern scheinen. Letzteres könnte durch Sequestration des PGN der residenten Bakterien erreicht werden, welches zu einer Inhibierung der *IMD* Signalkaskade führen würde.

Wir haben gezeigt, dass der zweite NF-κB Transkriptionsfaktor in *A. gambiae*, REL1, der ein Ortholog des *Drosophila* Dorsal ist (Dif existiert nicht in *Anopheles*), nicht in der Moskito Verteidigung gegen bakterielle Infektionen involviert ist. Dieser Umstand beweist noch einmal mehr, welche große Bedeutung den *REL2* Signalkaskaden in *A. gambiae* im Kampf gegen die Bakterien zukommt. Zudem scheint REL1 keine Rolle in der Verteidigung gegen *P. berghei* zu spielen. Setzt man jedoch den REL1 Inhibitor CACT (das Ortholog des *Drosophila* Cactus) während einer Malaria Infektion außer Kraft, so kommt es zu einem sehr stark refraktorischen Phenotypus: Die meisten der Ookineten, die versuchen in den Moskito Mitteldarm einzudringen, werden eliminiert (vermutlich lysiert) und die restlichen Ookineten werden melanisiert. Wir vermuten, dass unter den Bedingungen einer natürlich hervorgerufenen Infektion der Parasit entweder die Erkennung durch die Rezeptoren der REL1 Signalkaskade vermeidet oder aber gezielt die Aktivierung von REL1 verhindert.

Zusammenfassend lässt sich sagen, dass die Daten, die in dieser Doktorarbeit angeführt werden, eine signifikante Abweichung der Signalkaskaden des Moskito *A. gambiae* von denen der Fruchtfliege *D. melanogaster* beschreiben. Die beobachteten Unterschiede lassen sich wahrscheinlich auf die unterschiedlichen Lebensumstände und infektiösen Organismen zurückführen, denen diese zwei Insekten während ihrer Evolution ausgesetzt waren. Für den Moskito war einer dieser Organismen der Malaria Parasit.

(THIS PAGE WAS INTENTIONALLY LEFT BLANK)

BIBLIOGRAPHY

7.1 REFERENCES

Abraham, E. G., S. B. Pinto, et al. (**2005**). "An immune-responsive serpin, SRPN6, mediates mosquito defense against malaria parasites." *Proc Natl Acad Sci U S A* **102**(45): 16327-32.

Aderem, A. and D. M. Underhill (**1999**). "Mechanisms of phagocytosis in macrophages." *Annu Rev Immunol* **17**: 593-623.

Alavi, Y., M. Arai, et al. (**2003**). "The dynamics of interactions between Plasmodium and the mosquito: a study of the infectivity of Plasmodium berghei and Plasmodium gallinaceum, and their transmission by Anopheles stephensi, Anopheles gambiae and Aedes aegypti." *Int J Parasitol* **33**(9): 933-43.

Anderson, K. V., L. Bokla, et al. (**1985**). "Establishment of dorsal-ventral polarity in the Drosophila embryo: the induction of polarity by the Toll gene product." *Cell* **42**(3): 791-8.

Asling, B., M. S. Dushay, et al. (**1995**). "Identification of early genes in the Drosophila immune response by PCR-based differential display: the Attacin A gene and the evolution of attacin-like proteins." *Insect Biochem Mol Biol* **25**(4): 511-8.

Baton, L. A. and L. C. Ranford-Cartwright (**2004**). "Plasmodium falciparum ookinete invasion of the midgut epithelium of Anopheles stephensi is consistent with the Time Bomb model." *Parasitology* **129**(Pt 6): 663-76.

Baton, L. A. and L. C. Ranford-Cartwright (**2005**). "How do malaria ookinetes cross the mosquito midgut wall?" *Trends Parasitol* **21**(1): 22-8.

Belvin, M. P. and K. V. Anderson (**1996**). "A conserved signaling pathway: the Drosophila toll-dorsal pathway." *Annu Rev Cell Dev Biol* **12**: 393-416.

Bettencourt, R., H. Asha, et al. (**2004**). "Hemolymph-dependent and -independent responses in Drosophila immune tissue." *J Cell Biochem* **92**(4): 849-63.

Bian, G., S. W. Shin, et al. (**2005**). "Transgenic alteration of Toll immune pathway in the female mosquito Aedes aegypti." *Proc Natl Acad Sci U S A* **102**(38): 13568-73.

Billingsley, P. F., G. F. Medley, et al. (**1994**). "Relationship between prevalence and intensity of Plasmodium falciparum infection in natural populations of Anopheles mosquitoes." *Am J Trop Med Hyg* **51**(3): 260-70.

Billker, O., V. Lindo, et al. (**1998**). "Identification of xanthurenic acid as the putative inducer of malaria development in the mosquito." *Nature* **392**(6673): 289-92.

Bischoff, V., C. Vignal, et al. (**2004**). "Function of the drosophila pattern-recognition receptor PGRP-SD in the detection of Gram-positive bacteria." *Nat Immunol* **5**(11): 1175-80.

Bischoff, V., C. Vignal, et al. (**2006**). "Downregulation of the Drosophila Immune Response by Peptidoglycan-Recognition Proteins SC1 and SC2." *PLoS Pathog* **2**(2): e14.

Blandin, S., L. F. Moita, et al. (**2002**). "Reverse genetics in the mosquito Anopheles gambiae: targeted disruption of the Defensin gene." *EMBO Rep* **3**(9): 852-6.

Blandin, S., S. H. Shiao, et al. (**2004**). "Complement-Like Protein TEP1 Is a Determinant of Vectorial Capacity in the Malaria Vector Anopheles gambiae." *Cell* **116**(5): 661-70.

Brand, A. H. and N. Perrimon (**1993**). "Targeted gene expression as a means of altering cell fates and generating dominant phenotypes." *Development* **118**(2): 401-15.

Brehelin, M., D. Zachary, et al. (**1978**). "A comparative ultrastructural study of blood cells from nine insect orders." *Cell Tissue Res* **195**(1): 45-57.

Brogden, K. A. (**2005**). "Antimicrobial peptides: pore formers or metabolic inhibitors in bacteria?" *Nat Rev Microbiol* **3**(3): 238-50.

Bulet, P., J. L. Dimarcq, et al. (**1993**). "A novel inducible antibacterial peptide of Drosophila carries an O-glycosylated substitution." *J Biol Chem* **268**(20): 14893-7.

Bulet, P., C. Hetru, et al. (**1999**). "Antimicrobial peptides in insects; structure and function." *Dev Comp Immunol* **23**(4-5): 329-44.

Chang, C. I., S. Pili-Floury, et al. (**2004**). "A Drosophila pattern recognition receptor contains a peptidoglycan docking groove and unusual L,D-carboxypeptidase activity." *PLoS Biol* **2**(9): E277.

Cheng, X., X. Zhang, et al. (**1994**). "The structure of bacteriophage T7 lysozyme, a zinc amidase and an inhibitor of T7 RNA polymerase." *Proc Natl Acad Sci U S A* **91**(9): 4034-8.

Chipman, D. M. and N. Sharon (**1969**). "Mechanism of lysozyme action." *Science* **165**(892): 454-65.

Choe, K. M., H. Lee, et al. (**2005**). "Drosophila peptidoglycan recognition protein LC (PGRP-LC) acts as a signal-transducing innate immune receptor." *Proc Natl Acad Sci U S A* **102**(4): 1122-6.

Choe, K. M., T. Werner, et al. (**2002**). "Requirement for a peptidoglycan recognition protein (PGRP) in Relish activation and antibacterial immune responses in Drosophila." *Science* **296**(5566): 359-62.

Christophides, G. K., D. Vlachou, et al. (**2004**). "Comparative and functional genomics of the innate immune system in the malaria vector Anopheles gambiae." *Immunol Rev* **198**(1): 127-148.

Christophides, G. K., E. Zdobnov, et al. (**2002**). "Immunity-related genes and gene families in Anopheles gambiae." *Science* **298**(5591): 159-65.

Collins, F. H., R. K. Sakai, et al. (**1986**). "Genetic selection of a Plasmodium-refractory strain of the malaria vector Anopheles gambiae." *Science* **234**(4776): 607-10.

Danielli, A., F. C. Kafatos, et al. (**2003**). "Cloning and characterization of four Anopheles gambiae serpin isoforms, differentially induced in the midgut by Plasmodium berghei invasion." *J Biol Chem* **278**(6): 4184-93.

De Gregorio, E., S. J. Han, et al. (**2002**). "An immune-responsive Serpin regulates the melanization cascade in Drosophila." *Dev Cell* **3**(4): 581-92.

De Pauw, P., C. Neyt, et al. (**1995**). "Characterization of human serum N-acetylmuramyl-L-alanine amidase purified by affinity chromatography." *Protein Expr Purif* **6**(3): 371-8.

Dimarcq, J. L., D. Hoffmann, et al. (**1994**). "Characterization and transcriptional profiles of a Drosophila gene encoding an insect defensin. A study in insect immunity." *Eur J Biochem* **221**(1): 201-9.

Dimopoulos, G., T. L. Casavant, et al. (**2000**). "Anopheles gambiae pilot gene discovery project: identification of mosquito innate immunity genes from expressed sequence tags generated from immune-competent cell lines." *Proc Natl Acad Sci U S A* **97**(12): 6619-24.

Dimopoulos, G., G. K. Christophides, et al. (**2002**). "Genome expression analysis of Anopheles gambiae: responses to injury, bacterial challenge, and malaria infection." *Proc Natl Acad Sci U S A* **99**(13): 8814-9.

Dimopoulos, G., A. Richman, et al. (**1997**). "Molecular immune responses of the mosquito Anopheles gambiae to bacteria and malaria parasites." *Proc Natl Acad Sci U S A* **94**(21): 11508-13.

Dimopoulos, G., D. Seeley, et al. (**1998**). "Malaria infection of the mosquito Anopheles gambiae activates immune-responsive genes during critical transition stages of the parasite life cycle." *Embo J* **17**(21): 6115-23.

Dushay, M. S., B. Asling, et al. (**1996**). "Origins of immunity: Relish, a compound Rel-like gene in the antibacterial defense of Drosophila." *Proc Natl Acad Sci U S A* **93**(19): 10343-7.

Dziarski, R. (**2004**). "Peptidoglycan recognition proteins (PGRPs)." *Mol Immunol* **40**(12): 877-86.

Evans, C. J., V. Hartenstein, et al. (**2003**). "Thicker than blood: conserved mechanisms in Drosophila and vertebrate hematopoiesis." *Dev Cell* **5**(5): 673-90.

Evans, J. D. (**2004**). "Transcriptional immune responses by honey bee larvae during invasion by the bacterial pathogen, Paenibacillus larvae." *J Invertebr Pathol* **85**(2): 105-11.

Fehlbaum, P., P. Bulet, et al. (**1994**). "Insect immunity. Septic injury of Drosophila induces the synthesis of a potent antifungal peptide with sequence homology to plant antifungal peptides." *J Biol Chem* **269**(52): 33159-63.

Ferrandon, D., J. L. Imler, et al. (**2004**). "Sensing infection in Drosophila: Toll and beyond." *Semin Immunol* **16**(1): 43-53.

Franke-Fayard, B., H. Trueman, et al. (**2004**). "A Plasmodium berghei reference line that constitutively expresses GFP at a high level throughout the complete life cycle." *Mol Biochem Parasitol* **137**(1): 23-33.

Ganz, T. (**2003**). "Defensins: antimicrobial peptides of innate immunity." *Nat Rev Immunol* **3**(9): 710-20.

Gaunt, M. W. and M. A. Miles (**2002**). "An insect molecular clock dates the origin of the insects and accords with palaeontological and biogeographic landmarks." *Mol Biol Evol* **19**(5): 748-61.

Gelius, E., C. Persson, et al. (**2003**). "A mammalian peptidoglycan recognition protein with N-acetylmuramoyl-L-alanine amidase activity." *Biochem Biophys Res Commun* **306**(4): 988-94.

Ghosh, S., M. J. May, et al. (**1998**). "NF-kappa B and Rel proteins: evolutionarily conserved mediators of immune responses." *Annu Rev Immunol* **16**: 225-60.

Gobert, V., M. Gottar, et al. (**2003**). "Dual activation of the Drosophila toll pathway by two pattern recognition receptors." *Science* **302**(5653): 2126-30.

Gottar, M., V. Gobert, et al. (**2002**). "The Drosophila immune response against Gram-negative bacteria is mediated by a peptidoglycan recognition protein." *Nature* **416**(6881): 640-4.

Guan, R., E. L. Malchiodi, et al. (**2004**). "Crystal structure of the C-terminal peptidoglycan-binding domain of human peptidoglycan recognition protein Ialpha." *J Biol Chem* **279**(30): 31873-82.

Guan, R., A. Roychowdhury, et al. (**2004**). "Structural basis for peptidoglycan binding by peptidoglycan recognition proteins." *Proc Natl Acad Sci U S A* **101**(49): 17168-73.

Guan, R., Q. Wang, et al. (**2005**). "Crystal structure of human peptidoglycan recognition protein S (PGRP-S) at 1.70 A resolution." *J Mol Biol* **347**(4): 683-91.

Han, Y. S., J. Thompson, et al. (**2000**). "Molecular interactions between Anopheles stephensi midgut cells and Plasmodium berghei: the time bomb theory of ookinete invasion of mosquitoes." *Embo J* **19**(22): 6030-40.

Hedengren-Olcott, M., M. C. Olcott, et al. (**2004**). "Differential activation of the NF-kappaB-like factors Relish and Dif in Drosophila melanogaster by fungi and Gram-positive bacteria." *J Biol Chem* **279**(20): 21121-7.

Hegde, P., R. Qi, et al. (**2000**). "A concise guide to cDNA microarray analysis." *Biotechniques* **29**(3): 548-50, 552-4, 556 passim.

Hemingway, J., L. Field, et al. (**2002**). "An overview of insecticide resistance." *Science* **298**(5591): 96-7.

Hillyer, J. F. and B. M. Christensen (**2002**). "Characterization of hemocytes from the yellow fever mosquito, Aedes aegypti." *Histochem Cell Biol* **117**(5): 431-40.

Hillyer, J. F., S. L. Schmidt, et al. (**2003**). "Rapid phagocytosis and melanization of bacteria and Plasmodium sporozoites by hemocytes of the mosquito Aedes aegypti." *J Parasitol* **89**(1): 62-9.

Hoa, N. T., K. M. Keene, et al. (**2003**). "Characterization of RNA interference in an Anopheles gambiae cell line." *Insect Biochem Mol Biol* **33**(9): 949-57.

Hoffmann, J. A. (**2003**). "The immune response of Drosophila." *Nature* **426**(6962): 33-8.

Hoijer, M. A., M. J. Melief, et al. (**1996**). "Purification and characterization of N-acetylmuramyl-L-alanine amidase from human plasma using monoclonal antibodies." *Biochim Biophys Acta* **1289**(1): 57-64.

Holt, R. A., G. M. Subramanian, et al. (**2002**). "The genome sequence of the malaria mosquito Anopheles gambiae." *Science* **298**(5591): 129-49.

Hultmark, D. (**2003**). "Drosophila immunity: paths and patterns." *Curr Opin Immunol* **15**(1): 12-9.

Imler, J. L. and P. Bulet (**2005**). "Antimicrobial peptides in Drosophila: structures, activities and gene regulation." *Chem Immunol Allergy* **86**: 1-21.

Janeway, C. A., Jr. (**1989**). "Approaching the asymptote? Evolution and revolution in immunology." *Cold Spring Harb Symp Quant Biol* **54 Pt 1**: 1-13.

Janeway, C. A., Jr. and R. Medzhitov (**2002**). "Innate immune recognition." *Annu Rev Immunol* **20**: 197-216.

Kaneko, T., W. E. Goldman, et al. (**2004**). "Monomeric and polymeric gram-negative peptidoglycan but not purified LPS stimulate the Drosophila IMD pathway." *Immunity* **20**(5): 637-49.

Kaneko, T., D. Golenbock, et al. (**2005**). "Peptidoglycan recognition by the Drosophila Imd pathway." *J Endotoxin Res* **11**(6): 383-9.

Kaneko, T. and N. Silverman (**2005**). "Bacterial recognition and signalling by the Drosophila IMD pathway." *Cell Microbiol* **7**(4): 461-9.

Kang, D., G. Liu, et al. (**1998**). "A peptidoglycan recognition protein in innate immunity conserved from insects to humans." *Proc Natl Acad Sci U S A* **95**(17): 10078-82.

Kim, M. S., M. Byun, et al. (**2003**). "Crystal structure of peptidoglycan recognition protein LB from Drosophila melanogaster." *Nat Immunol* **4**(8): 787-93.

Kim, W., H. Koo, et al. (**2004**). "Ectopic expression of a cecropin transgene in the human malaria vector mosquito Anopheles gambiae (Diptera: Culicidae): effects on susceptibility to Plasmodium." *J Med Entomol* **41**(3): 447-55.

Kiselev, S. L., O. S. Kustikova, et al. (**1998**). "Molecular cloning and characterization of the mouse tag7 gene encoding a novel cytokine." *J Biol Chem* **273**(29): 18633-9.

Kumar, S., G. K. Christophides, et al. (**2003**). "The role of reactive oxygen species on Plasmodium melanotic encapsulation in Anopheles gambiae." *Proc Natl Acad Sci U S A* **100**(24): 14139-44.

Kylsten, P., C. Samakovlis, et al. (**1990**). "The cecropin locus in Drosophila; a compact gene cluster involved in the response to infection." *Embo J* **9**(1): 217-24.

Leclerc, V. and J. M. Reichhart (**2004**). "The immune response of Drosophila melanogaster." *Immunol Rev* **198**: 59-71.

Lee, M. H., T. Osaki, et al. (**2004**). "Peptidoglycan recognition proteins involved in 1,3-beta-D-glucan-dependent prophenoloxidase activation system of insect." *J Biol Chem* **279**(5): 3218-27.

Lehane, M. J., D. Wu, et al. (**1997**). "Midgut-specific immune molecules are produced by the blood-sucking insect Stomoxys calcitrans." *Proc Natl Acad Sci U S A* **94**(21): 11502-7.

Lemaitre, B., E. Kromer-Metzger, et al. (**1995**). "A recessive mutation, immune deficiency (imd), defines two distinct control pathways in the Drosophila host defense." *Proc Natl Acad Sci U S A* **92**(21): 9465-9.

Lemaitre, B., E. Nicolas, et al. (**1996**). "The dorsoventral regulatory gene cassette spatzle/Toll/cactus controls the potent antifungal response in Drosophila adults." *Cell* **86**(6): 973-83.

Leulier, F., C. Parquet, et al. (**2003**). "The Drosophila immune system detects bacteria through specific peptidoglycan recognition." *Nat Immunol* **4**(5): 478-84.

Levashina, E. A., E. Langley, et al. (**1999**). "Constitutive activation of toll-mediated antifungal defense in serpin-deficient Drosophila." *Science* **285**(5435): 1917-9.

Levashina, E. A., L. F. Moita, et al. (**2001**). "Conserved role of a complement-like protein in phagocytosis revealed by dsRNA knockout in cultured cells of the mosquito, Anopheles gambiae." *Cell* **104**(5): 709-18.

Levashina, E. A., S. Ohresser, et al. (**1995**). "Metchnikowin, a novel immune-inducible proline-rich peptide from Drosophila with antibacterial and antifungal properties." *Eur J Biochem* **233**(2): 694-700.

Ligoxygakis, P., N. Pelte, et al. (**2002**). "Activation of Drosophila Toll during fungal infection by a blood serine protease." *Science* **297**(5578): 114-6.

Ligoxygakis, P., N. Pelte, et al. (**2002**). "A serpin mutant links Toll activation to melanization in the host defence of Drosophila." *Embo J* **21**(23): 6330-7.

Lim, J. H., M. S. Kim, et al. (**2006**). "Structural basis for preferential recognition of diaminopimelic acid-type peptidoglycan by a subset of peptidoglycan recognition proteins." *J Biol Chem.*

Liu, C., Z. Xu, et al. (**2001**). "Peptidoglycan recognition proteins: a novel family of four human innate immunity pattern recognition molecules." *J Biol Chem* **276**(37): 34686-94.

Lowenberger, C. (**2001**). "Innate immune response of Aedes aegypti." *Insect Biochem Mol Biol* **31**(3): 219-29.

Lu, X., M. Wang, et al. (**2005**). "Peptidoglycan recognition proteins are a new class of human bactericidal proteins." *J Biol Chem*.

Luna, C., X. Wang, et al. (**2002**). "Characterization of four Toll related genes during development and immune responses in Anopheles gambiae." *Insect Biochem Mol Biol* **32**(9): 1171-9.

Marcu, O. and M. Locke (**1998**). "A cuticular protein from the moulting stages of an insect." *Insect Biochem Mol Biol* **28**(9): 659-69.

Medzhitov, R. and C. Janeway, Jr. (**2000**). "Innate immunity." *N Engl J Med* **343**(5): 338-44.

Meister, M. and M. Lagueux (**2003**). "Drosophila blood cells." *Cell Microbiol* **5**(9): 573-80.

Meister, S., S. M. Kanzok, et al. (**2005**). "Immune signaling pathways regulating bacterial and malaria parasite infection of the mosquito Anopheles gambiae." *Proc Natl Acad Sci U S A*.

Meister, S., A. C. Koutsos, et al. (**2004**). "The Plasmodium parasite-a 'new' challenge for insect innate immunity." *Int J Parasitol* **34**(13-14): 1473-82.

Mellroth, P., J. Karlsson, et al. (**2005**). "Ligand-induced dimerization of Drosophila peptidoglycan recognition proteins in vitro." *Proc Natl Acad Sci U S A* **102**(18): 6455-60.

Mellroth, P., J. Karlsson, et al. (**2003**). "A scavenger function for a Drosophila peptidoglycan recognition protein." *J Biol Chem* **278**(9): 7059-64.

Michel, K., A. Budd, et al. (**2005**). "Anopheles gambiae SRPN2 facilitates midgut invasion by the malaria parasite Plasmodium berghei." *EMBO Rep* **6**(9): 891-7.

Michel, T., J. M. Reichhart, et al. (**2001**). "Drosophila Toll is activated by Gram-positive bacteria through a circulating peptidoglycan recognition protein." *Nature* **414**(6865): 756-9.

Moita, L. F., R. Wang-Sattler, et al. (**2005**). "In vivo identification of novel regulators and conserved pathways of phagocytosis in A. gambiae." *Immunity* **23**(1): 65-73.

Muller, H. M., G. Dimopoulos, et al. (**1999**). "A hemocyte-like cell line established from the malaria vector Anopheles gambiae expresses six prophenoloxidase genes." *J Biol Chem* **274**(17): 11727-35.

Nappi, A. J. and E. Vass (**1998**). "Hydrogen peroxide production in immune-reactive Drosophila melanogaster." *J Parasitol* **84**(6): 1150-7.

Niare, O., K. Markianos, et al. (**2002**). "Genetic loci affecting resistance to human malaria parasites in a West African mosquito vector population." *Science* **298**(5591): 213-6.

Notredame, C., D. G. Higgins, et al. (**2000**). "T-Coffee: A novel method for fast and accurate multiple sequence alignment." *J Mol Biol* **302**(1): 205-17.

Ochiai, M. and M. Ashida (**1999**). "A pattern recognition protein for peptidoglycan. Cloning the cDNA and the gene of the silkworm, Bombyx mori." *J Biol Chem* **274**(17): 11854-8.

Onfelt Tingvall, T., E. Roos, et al. (**2001**). "The imd gene is required for local Cecropin expression in Drosophila barrier epithelia." *EMBO Rep* **2**(3): 239-43.

Osta, M. A., G. K. Christophides, et al. (**2004**). "Effects of mosquito genes on Plasmodium development." *Science* **303**(5666): 2030-2.

Paskewitz, S. and M. A. Riehle (**1994**). "Response of Plasmodium refractory and susceptible strains of Anopheles gambiae to inoculated Sephadex beads." *Dev Comp Immunol* **18**(5): 369-75.

Paskewitz, S. M., M. R. Brown, et al. (**1988**). "Ultrastructure of the encapsulation of Plasmodium cynomolgi (B strain) on the midgut of a refractory strain of Anopheles gambiae." *J Parasitol* **74**(3): 432-9.

Pili-Floury, S., F. Leulier, et al. (**2004**). "In vivo RNA interference analysis reveals an unexpected role for GNBP1 in the defense against Gram-positive bacterial infection in Drosophila adults." *J Biol Chem* **279**(13): 12848-53.

Pringle, G. (**1966**). "A quantitative study of naturally-acquired malaria infections in Anopheles gambiae and Anopheles funestus in a highly malarious area of East Africa." *Trans R Soc Trop Med Hyg* **60**(5): 626-32.

Ramet, M., P. Manfruelli, et al. (**2002**). "Functional genomic analysis of phagocytosis and identification of a Drosophila receptor for E. coli." *Nature* **416**(6881): 644-8.

Reiser, J. B., L. Teyton, et al. (**2004**). "Crystal structure of the Drosophila peptidoglycan recognition protein (PGRP)-SA at 1.56 A resolution." *J Mol Biol* **340**(4): 909-17.

Richman, A. M., P. Bulet, et al. (**1996**). "Inducible immune factors of the vector mosquito Anopheles gambiae: biochemical purification of a defensin antibacterial peptide and molecular cloning of preprodefensin cDNA." *Insect Mol Biol* **5**(3): 203-10.

Richman, A. M., G. Dimopoulos, et al. (**1997**). "Plasmodium activates the innate immune response of Anopheles gambiae mosquitoes." *Embo J* **16**(20): 6114-9.

Royet, J., J. M. Reichhart, et al. (**2005**). "Sensing and signaling during infection in Drosophila." *Curr Opin Immunol* **17**(1): 11-7.

Sambrook, J. and M. T. (1989). Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press.

Schleifer, K. H. and O. Kandler (**1972**). "Peptidoglycan types of bacterial cell walls and their taxonomic implications." *Bacteriol Rev* **36**(4): 407-77.

Shai, Y. (**2002**). "Mode of action of membrane active antimicrobial peptides." *Biopolymers* **66**(4): 236-48.

Shapira, S., O. S. Harb, et al. (**2005**). "Initiation and termination of NF-kappaB signaling by the intracellular protozoan parasite Toxoplasma gondii." *J Cell Sci* **118**(Pt 15): 3501-8.

Shin, S. W., V. Kokoza, et al. (**2002**). "Characterization of three alternatively spliced isoforms of the Rel/NF-kappa B transcription factor Relish from the mosquito Aedes aegypti." *Proc Natl Acad Sci U S A* **99**(15): 9978-83.

Shin, S. W., V. Kokoza, et al. (**2005**). "REL1, a homologue of Drosophila dorsal, regulates toll antifungal immune pathway in the female mosquito Aedes aegypti." *J Biol Chem* **280**(16): 16499-507.

Shin, S. W., V. Kokoza, et al. (**2003**). "Relish-mediated immune deficiency in the transgenic mosquito Aedes aegypti." *Proc Natl Acad Sci U S A* **100**(5): 2616-21.

Silverman, N., R. Zhou, et al. (**2000**). "A Drosophila IkappaB kinase complex required for Relish cleavage and antibacterial immunity." *Genes Dev* **14**(19): 2461-71.

Sinden, R. E. (**1999**). "Plasmodium differentiation in the mosquito." *Parassitologia* **41**(1-3): 139-48.

Sinden, R. E. (**2002**). "Molecular interactions between Plasmodium and its insect vectors." *Cell Microbiol* **4**(11): 713-24.

Smalley, M. E. and R. E. Sinden (**1977**). "Plasmodium falciparum gametocytes: their longevity and infectivity." *Parasitology* **74**(1): 1-8.

Soderhall, K. and L. Cerenius (**1998**). "Role of the prophenoloxidase-activating system in invertebrate immunity." *Curr Opin Immunol* **10**(1): 23-8.

Stenbak, C. R., J. H. Ryu, et al. (**2004**). "Peptidoglycan molecular requirements allowing detection by the Drosophila immune deficiency pathway." *J Immunol* **173**(12): 7339-48.

Stewart-Tull, D. E. (**1980**). "The immunological activities of bacterial peptidoglycans." *Annu Rev Microbiol* **34**: 311-40.

Stoven, S., I. Ando, et al. (**2000**). "Activation of the Drosophila NF-kappaB factor Relish by rapid endoproteolytic cleavage." *EMBO Rep* **1**(4): 347-52.

Tahar, R., C. Boudin, et al. (**2002**). "Immune response of Anopheles gambiae to the early sporogonic stages of the human malaria parasite Plasmodium falciparum." *Embo J* **21**(24): 6673-80.

Takehana, A., T. Katsuyama, et al. (**2002**). "Overexpression of a pattern-recognition receptor, peptidoglycan-recognition protein-LE, activates imd/relish-mediated antibacterial defense and the prophenoloxidase cascade in Drosophila larvae." *Proc Natl Acad Sci U S A* **99**(21): 13705-10.

Takehana, A., T. Yano, et al. (**2004**). "Peptidoglycan recognition protein (PGRP)-LE and PGRP-LC act synergistically in Drosophila immunity." *Embo J* **23**(23): 4690-700.

Theopold, U., D. Li, et al. (**2002**). "The coagulation of insect hemolymph." *Cell Mol Life Sci* **59**(2): 363-72.

Thomasova, D., L. Q. Ton, et al. (**2002**). "Comparative genomic analysis in the region of a major Plasmodium-refractoriness locus of Anopheles gambiae." *Proc Natl Acad Sci U S A* **99**(12): 8179-84.

Tydell, C. C., N. Yount, et al. (**2002**). "Isolation, characterization, and antimicrobial properties of bovine oligosaccharide-binding protein. A microbicidal granule protein of eosinophils and neutrophils." *J Biol Chem* **277**(22): 19658-64.

Tydell, C. C., J. Yuan, et al. (**2006**). "Bovine peptidoglycan recognition protein-s: antimicrobial activity, localization, secretion, and binding properties." *J Immunol* **176**(2): 1154-62.

Tzou, P., S. Ohresser, et al. (**2000**). "Tissue-specific inducible expression of antimicrobial peptide genes in Drosophila surface epithelia." *Immunity* **13**(5): 737-48.

Valinger, Z., B. Ladesic, et al. (**1982**). "Partial purification and characterization of N-acetylmuramyl-L-alanine amidase from human and mouse serum." *Biochim Biophys Acta* **701**(1): 63-71.

Vernick, K. D., H. Fujioka, et al. (**1995**). "Plasmodium gallinaceum: a refractory mechanism of ookinete killing in the mosquito, Anopheles gambiae." *Exp Parasitol* **80**(4): 583-95.

Vizioli, J., P. Bulet, et al. (**2000**). "Cloning and analysis of a cecropin gene from the malaria vector mosquito, Anopheles gambiae." *Insect Mol Biol* **9**(1): 75-84.

Vizioli, J., P. Bulet, et al. (**2001**). "Gambicin: a novel immune responsive antimicrobial peptide from the malaria vector Anopheles gambiae." *Proc Natl Acad Sci U S A* **98**(22): 12630-5.

Vlachou, D. and F. C. Kafatos (**2005**). "The complex interplay between mosquito positive and negative regulators of Plasmodium development." *Curr Opin Microbiol* **8**(4): 415-21.

Vlachou, D., T. Schlegelmilch, et al. (**2005**). "Functional genomic analysis of midgut epithelial responses in Anopheles during Plasmodium invasion." *Curr Biol* **15**(13): 1185-95.

Vlachou, D., T. Zimmermann, et al. (**2004**). "Real-time, in vivo analysis of malaria ookinete locomotion and mosquito midgut invasion." *Cell Microbiol* **6**(7): 671-85.

Volz, J., M. A. Osta, et al. (**2005**). "The roles of two clip domain serine proteases in innate immune responses of the malaria vector Anopheles gambiae." *J Biol Chem* **280**(48): 40161-8.

Wang, Z. M., X. Li, et al. (**2003**). "Human peptidoglycan recognition protein-L is an N-acetylmuramoyl-L-alanine amidase." *J Biol Chem* **278**(49): 49044-52.

Weber, A. N., S. Tauszig-Delamasure, et al. (**2003**). "Binding of the Drosophila cytokine Spatzle to Toll is direct and establishes signaling." *Nat Immunol* **4**(8): 794-800.

Werner, T., K. Borge-Renberg, et al. (**2003**). "Functional diversity of the Drosophila PGRP-LC gene cluster in the response to lipopolysaccharide and peptidoglycan." *J Biol Chem* **278**(29): 26319-22.

Werner, T., G. Liu, et al. (**2000**). "A family of peptidoglycan recognition proteins in the fruit fly Drosophila melanogaster." *Proc Natl Acad Sci U S A* **97**(25): 13772-7.

Wicker, C., J. M. Reichhart, et al. (**1990**). "Insect immunity. Characterization of a Drosophila cDNA encoding a novel member of the diptericin family of immune peptides." *J Biol Chem* **265**(36): 22493-8.

Wiegmann, B. M., D. K. Yeates, et al. (**2003**). "Time flies, a new molecular time-scale for brachyceran fly evolution without a clock." *Syst Biol* **52**(6): 745-56.

Wirth, D. F. (**2002**). "Biological revelations." *Nature* **419**(6906): 495-6.

Yeates, D. K. and B. M. Wiegmann (**1999**). "Congruence and controversy: toward a higher-level phylogeny of Diptera." *Annu Rev Entomol* **44**: 397-428.

Yoshida, H., K. Kinoshita, et al. (**1996**). "Purification of a peptidoglycan recognition protein from hemolymph of the silkworm, Bombyx mori." *J Biol Chem* **271**(23): 13854-60.

Zaidman-Remy, A., M. Herve, et al. (**2006**). "The Drosophila Amidase PGRP-LB Modulates the Immune Response to Bacterial Infection." *Immunity* **24**(4): 463-73.

Zasloff, M. (**2002**). "Antimicrobial peptides of multicellular organisms." *Nature* **415**(6870): 389-95.

Zdobnov, E. M., C. von Mering, et al. (**2002**). "Comparative genome and proteome analysis of Anopheles gambiae and Drosophila melanogaster." *Science* **298**(5591): 149-59.

Zheng, L., A. J. Cornel, et al. (**1997**). "Quantitative trait loci for refractoriness of Anopheles gambiae to Plasmodium cynomolgi B." *Science* **276**(5311): 425-8.

Zheng, L., S. Wang, et al. (**2003**). "Quantitative trait loci in Anopheles gambiae controlling the encapsulation response against Plasmodium cynomolgi Ceylon." *BMC Genet* **4**(1): 16.

Zheng, X. L. and A. L. Zheng (**2002**). "Genomic organization and regulation of three cecropin genes in Anopheles gambiae." *Insect Mol Biol* **11**(6): 517-25.

Zhu, Y., Y. Wang, et al. (**2003**). "Manduca sexta serpin-3 regulates prophenoloxidase activation in response to infection by inhibiting prophenoloxidase-activating proteinases." *J Biol Chem*.

7.2 PUBLICATIONS & MANUSCRIPTS IN PREPARATION

1. **Meister S**, Agianian B, Kafatos FC, Christophides GK. running title:"The *Anopheles gambiae PGRPLC* gene cluster."; in preparation

2. Koutsos AC, Blass C, **Meister S**, Schmidt S, Soares MB, Collins FH, Benes V, Zdobnov E, Kafatos FC, Christophides GK. "Lifecycle Transcriptomics of the Malaria Mosquito *Anopheles gambiae* and Comparison with the Fruitfly *Drosophila melanogaster*."; submitted

3. **Meister S**, Kanzok S, Zheng XL, Luna C, Li TR, Hoa NT, Clayton JR, White KP, Kafatos FC, Zheng L, Christophides GK. "Immune signaling pathways regulating bacterial and malaria parasite infection of the mosquito Anopheles gambiae." *Proc Natl Acad Sci U S A.* **2005** Aug 9;102(32):11420-5

4. **Meister S**, Koutsos AC and Christophides GK. "The Plasmodium parasite - a 'new' challenge for insect innate immunity." *Int J Parasitol.* **2004** Dec;34(13-14):1473-82. Review.

5. Kumar S, Christophides GK, Cantera R, Charles B, Han YS, **Meister S**, Dimopoulos G, Kafatos FC, Barillas-Mury C. "The role of reactive oxygen species on Plasmodium melanotic encapsulation in Anopheles gambiae." *Proc Natl Acad Sci U S A.* **2003** Nov 25; 100(24): 14139-44

6. Christophides GK, Zdobnov E, Barillas-Mury C, Birney E, Blandin S, Blass C, Brey PT, Collins FH, Danielli A, Dimopoulos G, Hetru C, Hoa NT, Hoffmann JA, Kanzok SM, Letunic I, Levashina EA, Loukeris TG, Lycett G, **Meister S**, Michel K, Moita LF, Muller HM, Osta MA, Paskewitz SM, Reichhart JM, Rzhetsky A, Troxler L, Vernick KD, Vlachou D, Volz J, von Mering C, Xu J, Zheng L, Bork P, Kafatos FC. "Immunity-related genes and gene families in Anopheles gambiae: A comparative genomic analysis." *Science* **2002** Oct 4;298(5591):159-65

7. Holt RA, Subramanian GM, Halpern A, Sutton GG, Charlab R, Nusskern DR, Wincker P, Clark AG, Ribeiro JM, Wides R, Salzberg SL, Loftus B, Yandell M, Majoros WH, Rusch DB, Lai Z, Kraft CL, Abril JF, Anthouard V, Arensburger P, Atkinson PW, Baden H, de Berardinis V, Baldwin D, Benes V, Biedler J, Blass C, Bolanos R, Boscus D, Barnstead M, Cai S, Center A, Chatuverdi K, Christophides GK, Chrystal MA, Clamp M, Cravchik A, Curwen V, Dana A, Delcher A, Dew I, Evans CA, Flanigan M, Grundschober-Freimoser A, Friedli L, Gu Z, Guan P, Guigo R, Hillenmeyer ME, Hladun SL, Hogan JR, Hong YS, Hoover J, Jaillon O, Ke Z, Kodira C, Kokoza E, Koutsos A, Letunic I, Levitsky A, Liang Y, Lin JJ, Lobo NF, Lopez JR, Malek JA, McIntosh TC, **Meister S**, Miller J, Mobarry C, Mongin E, Murphy SD, O'Brochta DA, Pfannkoch C, Qi R, Regier MA, Remington K, Shao H, Sharakhova MV, Sitter CD, Shetty J, Smith TJ, Strong R, Sun J, Thomasova D, Ton LQ, Topalis P, Tu Z, Unger MF, Walenz B, Wang A, Wang J, Wang M, Wang X, Woodford KJ, Wortman JR, Wu M, Yao A, Zdobnov EM, Zhang H, Zhao Q, Zhao S, Zhu SC, Zhimulev I, Coluzzi M, della Torre A, Roth CW, Louis C, Kalush F, Mural RJ, Myers EW, Adams MD, Smith HO, Broder S, Gardner MJ, Fraser CM, Birney E, Bork P, Brey PT, Venter JC, Weissenbach J, Kafatos FC, Collins FH, Hoffman SL. "The genome sequence of the malaria mosquito Anopheles gambiae." *Science* **2002** Oct 4;298(5591):129-49

8. Dimopoulos G, Christophides GK, **Meister S**, Schultz J, White KP, Barillas-Mury C, Kafatos FC. "Genome Expression Analysis of Anopheles gambiae: Responses to Injury, Bacterial Challenge and Malaria Infection." *Proc Natl Acad Sci U S A.* **2002** Jun 25;99(13):8814-9.

ABBREVIATIONS

4a r/r	susceptible *Anopheles gambiae* strain
aa	amino acid
AB	antibody
ANK	Ankyrin-repeat domain
bp	base pair
BSA	bovine serum albumin
CACT	*Anopheles gambiae* Cactus ortholog
CASPL1	*Anopheles gambiae Caspase8/Dredd* ortholog
CDS	coding sequence
CLIPA	clip domain serine protease A
CNRS	Centre National de la Recherche Scientifique
CTL	C-type Lectin
CTLMA	Mannose binding CTL
Dap	diaminopimelic acid
DD	Death Domain
DNA	deoxyribonucleic acid
Dnr1	*Drosophila melanogaster* inhibitor of *Dredd*
Dredd	*Drosophila melanogaster* gene
dsRNA	double stranded ribonucleic acid
EMBL	European Molecular Biology Laboratory
FADD	Fas-associated death domain protein
FCS	fetal calf serum
G3	wild type *Anopheles gambiae* strain
GFP	green fluorescent protein
GNBP	Gram-negative binding protein
Gram+	Gram-positive bacteria
Gram-	Gram-negative bacteria
IMBC	l'Institut de Biologie Moléculaire et Cellulaire
Imd	immune deficiency gene
L3-5	refractory *Anopheles gambiae* strain
LB medium	Lysogeny Broth (not: Luria broth, Lennox broth, or Luria-Bertani medium) for bacterial growth
LPS	Lipopolysaccharide
LRIM1	*Anopheles gambiae* gene leucine rich immune molecule 1
Lys	Lysine
MACF	EMBL Monoclonal Antibody Core Facility
MDP	Muramyl Dipeptide
ml	milliliter
µl	microliter
NAG	β-1,4-linked N-Acetylglucosamine
NAM	N-Acetyl Muramic Acid
NAMLAA	N-acetylmuramoyl-L-alanine amidase
Nec	blood serpin Necrotic
nl	nanoliter
NLS	nuclear localization signal
nt	nucleotide

ORF	open reading frame
ON	over night
PAMP	pathogen associated molecular pattern
PBS	Phosphate buffered saline
PCR	polymerase chain reaction
PGN	Peptidoglycan
PGRP	Peptidoglycan Recognition Protein
PO	Phenoloxidase
PPO	Prophenoloxidase
PPAE	PPO activating enzyme
PRR	pattern recognition receptor
Psh	Persephone protease
REL1	*Anopheles gambiae dorsal* ortholog
REL2	*Anopheles gambiae relish* ortholog
RHD	Relish homology domain
RNA	ribonucleic acid
RNAi	RNA interference
RT	room temperature
seml	*semmelweis* mutation of *PGRP-SA*
serpin	serine protease inhibitor
siRNA	small interfering RNA
SOB	Super Optimal Broth (bacterial growth medium)
SOC	Super Optimal Catobolite Repression (SOB medium with added glucose)
SOD	Superoxide Dismutase
Spz	Spaetzle gene
TAK1	*Anopheles gambiae TAK1* ortholog
TEP	Thioester-containing protein
wt	wild-type

(THIS PAGE WAS INTENTIONALLY LEFT BLANK)

APPENDIX

9.1 PRIMER SEQUENCES

Table 9.1.1. Sequences of primers used.

#	Gene	Primer Name	Primer Sequence (5'-3')	Amplicon Size	Remarks
9.1.1	*S7*	AgS7 qRT-PCR F	GTGCGCGAGTTGGAGAAGA	78 bp	300nM for qRT-PCR
9.1.2		AgS7 qRT-PCR R	ATCGGTTTGGGCAGAATGC		900nM for qRT-PCR
9.1.3	N/A	T7-d(T)24	GGCCAGTGAATTGTAATACGACTCACTATAGGGAGGCGG(T)24		
9.1.4	*S7*	S7-A	GGCGATCATCATCTACGTGC	460 bp	semiquantitative RT-PCR primers
9.1.5		S7-B	GTAGCTGCTGCAAACTTCGG		
9.1.6	*PGRPS2&3*	Exp25.1-F	*GAATTAATACGACTCACTATAGGG* AGA ACAACTTCCTGGTCGGTGAG		w/ T7 overhang
9.1.7	*PGRPS2*	Exp25.2-R	*GAATTAATACGACTCACTATAGGG* AGA TCACCTGTCACAATGGTCGT	533 bp	w/ T7 overhang
9.1.8	*PGRPS3*	Exp25.3-R	*GAATTAATACGACTCACTATAGGG* AGA CCCCACATTAAGCTACGTTTC	423 bp	w/ T7 overhang
9.1.9	*PGRPLC1*	PGRPLC1-Forw	AAACCACACCAACTACGGTGA	323 bp (+89 bp intron)	semiquantitative RT-PCR primers
9.1.10		PGRPLC1-Rev	AATGTGGCAAAAGCCTCAG		
9.1.11	*PGRPLC2*	PGRPLC2-Forw	ATGGATGGCAAAAACTACGAC	309 bp (+100 bp intron)	semiquantitative RT-PCR primers
9.1.12		PGRPLC2-Rev	CAGTGCTTCGATTAGCCACTT		
9.1.13	*PGRPLC3*	PGRPLC3-Forw	TTTAGCGACATTGCGTATCA	431 bp(+63 bp intron)	semiquantitative RT-PCR primers
9.1.14		PGRPLC3-Rev	GGAAGTCATCAAGGACACTTGG		
9.1.15	*PGRPLC1*	PGRP-LC1-AB-F	AAAGTTGGAGCCCACACCAAA	234 bp, 78 aa	for AB production
9.1.16		PGRP-LC1-AB-R	CAAAAGCCTCAGCTGTTC		
9.1.17	*PGRPLC2*	PGRP-LC2-AB-F	ACCATTCCCGGTTACAATTC	222 bp, 74 aa	for AB production
9.1.18		PGRP-LC2-AB-R	CAAGCTGTGCAGTGCTTCGAT		
9.1.19	*PGRPLC3*	PGRP-LC3-AB-F	AAAGGGTTCAACGTGGACAGC	282 bp, 94 aa	for AB production
9.1.20		PGRP-LC3-AB-R	TCGTCGGTTTGTGGTGTCGTT		
9.1.21	*PGRPS1*	PGRP-S1-AB-F	GCCCAGGACGAGCCGGCCCAG	510 bp, 170aa	for AB production
9.1.22		PGRP-S1-AB-R	GTCAAGCTCTTGCAGCTTCG		
9.1.23	*PGRPLA1*	PGRP-LA1-AB-F	CTCGGCAACAGCCACATGATC	438 bp, 146aa	for AB production
9.1.24		PGRP-LA1-AB-R	TCTAACCTGTCTTCGTGCCACAAGC		
9.1.25	*PGRPLA2*	PGRP-LA2-AB-F	CTCGGCAACGGGCACATGGTC	429 bp, 143aa	for AB production
9.1.26		PGRP-LA2-AB-R	TCGATGTGCGACGATTTTGTAG		
9.1.27	*PGRPLB*	PGRP-LB-AB-F	CCGGTGCCCTACGTGACGCGA	477 bp, 159aa	for AB production
9.1.28		PGRP-LB-AB-R	TTTGATCTCCTCGAACAGCCG		
9.1.29	*PGRPLD*	PGRP-LD-AB-F	CCGTTCTTTTTGGTCGAGCG	487 bp, 162aa	for AB production
9.1.30		PGRP-LD-AB-R	ATCGCTCCAGCTGACTGTAC		
9.1.31	*PGRPS2 & PGRPS3*	PGRP-2-AB forward	CGCATCGTGACCCGCGCCCAG	480 bp, 160 aa	for AB production of both S2 & S3
9.1.32		PGRP-2-AB reverse	GCGCGGCCAGGTGCGGATCTC		
9.1.33	*PGRPLC;* qRT-PCR primers	PGRPLC-RTPCR-Forward	CCGGTTAATAACGTCATCATTGC		900nM for qRT-PCR
9.1.34		Exp48.50	TCACCGTAGTTGGTGTGGTT	119 bp	900nM for qRT-PCR (LC1 Reverse)
9.1.35		Exp48.51	CGTCGTAGTTTTTGCCATCC	118 bp	900nM for qRT-PCR (LC2 Reverse)
9.1.36		Exp48.52	TCGGGACTCGAATGAAACTC	107 bp	900nM for qRT-PCR (LC3 Reverse)

9.1.37	*PGRPLC;* whole PGRP domains	PGRPLC-dom-F	*CCCATGGGT* CCAGATCCGAGACCGTTACGG		whole PGRP domain, with 5'NcoI overhang
9.1.38		PGRPLC1-dom-R	*GGAATTCCC* AAAATGTGGCAAAAGCCTCA	507bp, 169aa	5' EcoRI overhang
9.1.39		PGRPLC2-dom-R	*GGAATTCCC* AAACTGGGGCAAGCTGTGCAG	508bp, 169aa	5' EcoRI overhang
9.1.40		PGRPLC3-dom-R	*GGAATTCCC* CCAGTGGGGCCAAGTTTTGATG	510bp, 170aa	5' EcoRI overhang
9.1.41	N/A	M13/pUC-Forward	CGCCAGGGTTTTCCCAGTCACGAC	N/A	plasmid sequencing (pLL10 etc)
9.1.42		M13/pUC-Reverse	AGCGGATAACAATTTCACACAGGA		
9.1.43		SP6 Promoter	CATACGATTTAGGTGACACTATAG		
9.1.44		T7 Promoter 20mer	TAATACGACTCACTATAGGG		
9.1.45		T3 Promoter 23mer	GCAATTAACCCTCACTAAAGGGA		
9.1.46	*GFP*	GFP-T7-Forward	*TAATACGACTCACTATAGGG* CAAGACACGTGCTGAAGTCAA	~450 bp	w/ T7 overhang, for RNAi
9.1.47		GFP-T7-Reverse	*TAATACGACTCACTATAGGG* GCCTGAATTTAACCAGGAACC		
9.1.48	*LacZ*	LacZ-T7-Forward	*TAATACGACTCACTATAGGG* AGAATCCGACGGGTTGTTACT	~550 bp	w/ T7 overhang, for RNAi
9.1.49		LacZ-T7-Reverse	*TAATACGACTCACTATAGGG* CACCACGCTCATCGATAATTT		
9.1.50	*CASPL1*	CASPL1-T7-F	*TAATACGACTCACTATAGGG* TATGCGCTGCAAATTCTCAC	~400 bp	w/ T7 overhang, for RNAi
9.1.51		CASPL1-T7-R	*TAATACGACTCACTATAGGG* CGGACTGTTTCAAACCCAAC		
9.1.52	*DEF1*	Exp9-F	*TAATACGACTCACTATAGGG* TACCCTTCTGGACGAACTGC	240 bp	w/ T7 overhang, for RNAi
9.1.53		Exp9-R	*TAATACGACTCACTATAGGG* CTTCCCAGGATGCTAAGCTG		
9.1.54	*PGRPS1*	Exp1-F	*AACCGGAATTC* AAGCAAACGTCATTCTGAGAGTC	689 bp	5' EcoRI overhang
9.1.55		Exp1-R	*ACTAGTCTAGA* GCCTTTTGCAAGTAGAGAGCA		5' XbaI overhang
9.1.56	*REL2*	Exp13-F	*GAATTAATACGACTCACTATAGGG* AGA AATCCGACGCAACGATACG	427 bp	REL2-ANK domain, w/ T7 overhang for RNAi
9.1.57		Exp13-R	*GAATTAATACGACTCACTATAGGG* AGA GACCGCAATGTGAAGGATG		
9.1.58	*REL2*-B-Form	Exp 18-F	*GAATTAATACGACTCACTATAGGG* AGA CGAACCTCAGCAATGGAGTAG	90 bp	targeting exon of REL2-B form, w/ T7 overhang for RNAi
9.1.59		Exp 18-R	*GAATTAATACGACTCACTATAGGG* AGA CTCTGCAAGTGTTAAAAACAGTGA		
9.1.60	*PGRPLC*	PGRPLC RTPCR F	CACGCACCTGGCAATCTAGTCT	96 bp	qRT-PCR primers
9.1.61		PGRPLC RTPCR R	TGGCACACAGGACAATCATCA		
9.1.62	*Dnr1*, Dredd inhibitor	Exp20-For	*GAATTAATACGACTCACTATAGGG* TTTACCAGCTGATTGCTCAGG	314 bp	w/ T7 overhang, for RNAi; ENSANGG00000015649
9.1.63		Exp20-Rev	*GAATTAATACGACTCACTATAGGG* CTTCAGCTCCACGTACGTCTC		
9.1.64	*IMD*	2nd IMD RT-PCR F	TCGCAAATGATGCAGAGCC	65 bp	900nM for qRT-PCR
9.1.65		2nd IMD RT-PCR R	AAATGGCGCGACACCGTAT		900nM for qRT-PCR
9.1.66	*CACT*	Exp31-CACT-F	*GAATTAATACGACTCACTATAGGG* AGA GTCCGCTCTACACATCAGCA	318 bp	w/ T7 overhang, for RNAi
9.1.67		Exp31-CACT-R	*GAATTAATACGACTCACTATAGGG* AGA CCGTTCGGGTTAATGATGAC		
9.1.68	*REL1*	Exp31-REL1-F	*GAATTAATACGACTCACTATAGGG* AGA ATCAACAGCACGACGATGAG	386 bp	w/ T7 overhang, for RNAi
9.1.69		Exp31-REL1-R	*GAATTAATACGACTCACTATAGGG* AGA TCGAAAAAGCGCACCTTAAT		

9.1.70	*PGRPS1*	Exp46-PGRPS1-F	*GAATTAATACGACTCACTATAGGG* AGA GGACGAGCCGGCCCAGGAATC	493 bp	w/ T7 overhang, for RNAi ; same as MMC2 primers
9.1.71		Exp46-PGRPS1-R	*GAATTAATACGACTCACTATAGGG* AGA CAGCTTCGCGTACAGATAC		
9.1.72	*PGRPLD*	Exp45-PGRPLD-F	*GAATTAATACGACTCACTATAGGG* AGA GACTCGGAGGATTGTCTGGA	281 bp	w/ T7 overhang, for RNAi
9.1.73		Exp45-PGRPLD-R	*GAATTAATACGACTCACTATAGGG* AGA CGGATCGACTCGGTGATAAA		
9.1.74	*PGRPLB*	Exp45-PGRPLB-F	*GAATTAATACGACTCACTATAGGG* AGA CCATCCCGTACGTCATCATA	274 bp	w/ T7 overhang, for RNAi
9.1.75		Exp45-PGRPLB-R	*GAATTAATACGACTCACTATAGGG* AGA ATGTTCTTCGGTGGCAAATC		
9.1.76	*PGRPLA1*	Exp45-PGRPLA1-F	*GAATTAATACGACTCACTATAGGG* AGA CCGACATTCCAAGCAACTTT	280 bp	w/ T7 overhang, for RNAi
9.1.77		Exp45-PGRPLA1-R	*GAATTAATACGACTCACTATAGGG* AGA ACCAGCCTAGCGTACAGCAT		
9.1.78	*PGRPLA2*	Exp45-PGRPLA2-F	*GAATTAATACGACTCACTATAGGG* AGA TGCTGATAACGCACATAGGC	124 bp	w/ T7 overhang, for RNAi
9.1.79		Exp45-PGRPLA2-R	*GAATTAATACGACTCACTATAGGG* AGA TTGCTCGGTATGTCTTGCAG		
9.1.80	*REL2*-S	Exp34-RT-REL2-For	ACCGATACGGAAAGTGTGCT	228 bp	300nM for qRT-PCR
9.1.81		Exp34-REL2-S-Rev	CGGTGCTCCTCGTAATGACT		300nM for qRT-PCR
9.1.82	*REL2*-F	Exp34-RT-REL2-For	as above	215 bp	900nM for qRT-PCR
9.1.83		Exp34-REL2-F-Rev	GTATCGTTGCGTCGGATTG		300nM for qRT-PCR
9.1.84	*REL2*-F	Dev-REL2-F-For	AATCCGACGCAACGATACG		primers for developmental profile(Meister, Kanzok et al. 2005)
9.1.85		Dev-REL2-F-Rev	CATCCTTCACATTGCGG		
9.1.86	*REL2*-S	Dev-REL2-S-For	ACCGATACGGAAAGTGTGCTGGGACGGGC		
9.1.87		Dev-REL2-S-Rev	CAATGGAGTAGTCATTACGAGGAG		
9.1.88	*KIN1*	KIN1-Promoter-F	*CGAGCTCG* GCAAAGGAATTATCCGGTGA		5' SacI overhang
9.1.89		KIN1-Promoter-R	*GAAGATCTTC* GTTCAGCTCTGGATCGCACT		5' BglII overhang
9.1.90	*IMD*	IMD-KD-F	*GAATTAATACGACTCACTATAGGG* AGA GGGAATTTCCCAAATGGTGTG	229 bp	w/ T7 overhang, for RNAi; (reorder)
9.1.91		IMD-KD-R	*GAATTAATACGACTCACTATAGGG* AGA TGTGTAGATTGCTCGCGTTC		
9.1.92	*REL2*	ENSANGT00000020234R6920	*AGCCCCAACAATATG* AGCGAGCCGCATCTGG		MMC2-Primer, Plate 4, Row 1, Column 3
9.1.93		ENSANGT00000020234S6920	*CTACTCAGTCAACGG* TACGGCCGCCTCCTTCT		
9.1.94	*LRIM1*	ENSANGT00000013041R6542	*AGCCCCAACAATATG* AATATCTATCTCGCGAACAATAA		MMC2-Primer, Plate 2, Row 4, Column 11
9.1.95		ENSANGT00000013041S6542	*TCAGATTAGCTCAGT* TGGCACGGTACACTCTTCC		

9.2 NUCLEOTIDE SEQUENCES

>Seq. 9.2.1 **PGRPS1**
ATGATGCTGTTCGGAGGCATTTTACCGTATCTGCTTGTGCTATGGCAGCTACTGGCTTCGGTCGATGATTT
GTTGCATATTCTAGTCGCCCGGGCCCAGGACGAGCCGGCCCAGGAATCCGCCTGCCCAGCCATCGTGAAGC
GGGCGGCCTGGGGTGCAGCCAAGTCGAAGAACGTCACCTACCAGCTGAAACCGGTCGCGAACGTGATCGTG
CACCATACGACCGGGGAGCGGTGCGCGACAGTGGCCACCTGCAAGGAGATGGTGGCAAACATACAGACCTA
CCACCAGACCGACAACCGGTGGAGTGACATCGGCTACAACTTCCTAATCAGCGGGCAGAATGTGTACGAGG
GCATCGGCTGGCACCGGATGGGGGCGCATCTGCGCGGCTACAACGACAAATCGATCGGCGTCGCCTTCCTG
GGCAACTTTGACCAAGAGCGGCCGACCCCCCGCAGCCTCAACCTGCTTGCCCGGCTACTCCAGTGCGGTGT
CGAGCTGGGCGAGCTAGCCGACGACTACCGGCTGTACGGTGCCCGCCAGCTCCAGAGCACCAACAGCCCGG
GCCGGTATCTGTACGCGAAGCTGCAAGAGCTTGACCACTGGCAGGCGCAA

>Seq. 9.2.2 **PGRPS2** (ENSANGP00000012978/agCP5898)
ATGAACAAGTTCGCGACCGTTCTGGTGCTGGCCAGCATCTGTCTGGCGGGCGTTTCGGCCCAGTGCCCGCG
CATCGTGACCCGCGCCCAGTGGGGAGCCCGTGCCGCCAGCACCTCGCAGCTGCCGATCCGTCCGGCCCCGT
GGGTCGTGATGCACCATACGGCCGGTGCGTCCTGCACGACCGATGCGGCCTGCGCCCAGCAGATGCGCAAC
ATCCAGTCCTTCCACATGGACGGCAACGGCTGGGCTGACATTGGCTACAACTTCCTGGTCGGTGAGAACGG
TGCGGCCTACGAGGGACGCGGCTGGGGCCGCCAGGGCGCTCACGCTCCCGGCTACAACGATCGCTCGGTCG
GCATGGGCGTGATCGGCACGTTCACCAACGCCATCCCGAATGCGGCCGCCCGTACCGCAGCCCGTAACCTG
ATCACCTGCGGTGTGTCGCTCGGCCACATTGCCTCGAACTACTGGCTGATTGGACATCGTCAGGCCGTCGC
CACCGCCTGCCCGGGCAATGCGTTCTTCAACGAGATCCGCACCTGGCCGCGCTTCAACCCGAACGTG

>Seq. 9.2.3 **PGRPS3** (ENSANGP00000012979/agCP5906)
ATGAACAAGTTCGCGACCGTTCTGGTGCTGGCCAGCATCTGCCTGGCGGGCGTTTCGGCCCAGTGCCCGCG
CATCGTGACCCGCGCCCAGTGGGGAGCCCGTGCCGCCAACACCGCCCAGCTGCCGATCCGTCCGGCCCCGT
GGGTCGTGATGCACCATACGGCCGGTGCGGCCTGCACGACCGATGCGGCCTGCGCCCAGCAGATGCGCAAC
ATCCAGTCCTTCCACATGGACGGCAACGGCTGGGCTGACATTGGCTACAACTTCCTGGTCGGTGAGAACGG
TGCGGCGTACGAGGGACGCGGCTGGGGCCGCCAGGGCGCTCACGCTCCCGGCTACAACGATCGCTCGGTCG
GCATGGGCGTGATTGGCACGTTCACCAACGGCATCCCGAATGCGGCCGCCCGCAACGCTGCCCAGCAGCTG
ATCAGCTGTGGTGTGTCGCTCGGCCACATTGCCTCGAACTACTGGCTGATCGGACATCGTCAGGCCGTTGC
CACTGCCTGCCCCGGAAATGCCTTCTTCAACGAGATCCGAAACTGGCCGCGCTTCAACCCGAACGTG

>Seq. 9.2.4 **PGRPLA1** partial CDS
ATGGCCACGAACCATCAGAATGGGCTCAGCACTGGCAATGGCGGAAATACGACCGTGCAACCGGCGGCCAC
AAGCGTCATCAATCTGTCGAACTCCAGCGACGTCGTCATTGGCCCGATGACACAGTACCAAGGTTCCGTTA
CTATCTACCAGTACATGGATGCGACCGTCGAAGCTAGCCGAATTGCAAGTAAGTTTCCGCTTCCTATGCGT
AGCAGGAGCTTGACTGAATTTTATGACCTCGTCTTCCCATCGCACCTTTCCTTTCCTACAGCCGGCCGCAA
CAATCAGAACCGCGGACTACCTTCAGCGGAACCATCGACCCTGCGTCAGGAGCGCTACTTCATTTACGGAG
CGCTGATATTTTTCGCGATCGTCGGATTCTCAACCGCCATCTACTTCATCGTGAACCAGGTGCGCGGTCCA
GCTGACATTGACCGGGAGATACTGTTCGACACCAACTACCACCGGCGGACGATGCCCAATCTCGGCAACAG
CCACATGATCATCGAACGGCGTAACTGGGGTTCGCAGCTAGATGCGCACGCATCGATCCAATTGGAGCATC
CGGTTCAGTACGTGATCGTTACGCACATCGGCCTGCGCTCCAAAAATTGCACCGGCATGCACGAATGTGCG
AACAGGATGCGCATGCTACAGGATGCAGCCATCGGTGAGCGGAATCTGCCCGACATTCCAAGCAACTTTTA
TCTTGGTGGAGATGGCAACGTGTACGTGGGCCGTGGTTGGGATATTGCAAATTCGTATCACAATCGAACAC
TGTCCGTCTGCTTCATAGGAGACTTCCAAACCTACGAACCGAAAGAATCTCAATTCTCCGCCTTGCACCAT
CTCCTGACGTACGGTGTGATACAGCGAAAGCTTGCCAGCGACTATAAGCTTGTGGCACGAAGACAGACCAA
ACCTACCACAGCCAGCCCTGGCGACATGCTGTACGCTAGGCTGGTTAGACTATCCCGGTGGAATCCTTGTG
GAACGCAAGCATACGCACACTGTGGGGCCGAATTGGGCTTTCCGAGCGTTTGGGACGATGAGCACGATTTC
AATGAGCGGAACGGAATTCATAACCCATTGTTAATTCATCGGAAAAGCACAGAAGAAGAA

>Seq. 9.2.5 **PGRPLA2** (ENSANGG00000007842/agCP15114)
ATGGCCACGAACCATCAGAATGGGCTCAGCACTGGCAATGGCGGAAATACGACCGTGCAACCGGCGGCCAC
AAGCGTCATCAATCTGTCGAACTCCAGCGACGTCGTCATTGGCCCGATGACACAGTACCAAGGTTCCGTTA
CTATCTACCAGTACATGGATGCGACCGTCGAAGCTAGCCGAATTGCAATTCCCAATCTCGGCAACGGGCAC
ATGGTCATCGATCGCCACAACTGGGGTGCACAGCAGGGTGTACATGGACCGTACAAGCTGCCACATCCCAT
TCCGTACGTGCTGATAACGCACATAGGCGTCCACTCGGAAATCTGCTCCGATGTGCACGTTTGCTCCATCA
AGATGCGCACGCTGCAAGATGCAGCAATTGCCGAGAAAAGTCTGCAAGACATACCGAGCAATTTTTATGTC
GGTGGGGACGGAAACGTTTACGTTGGTCGTGGATGGGATACGGCCAATGCGTACGCTAATATGTCACTGGC
CGTGTGCTTCATGGGCGATTACGGGCGGTATGAACCGAACGATCTTCAGCTTTCAGCACTGGACCATCTAT
TGACGTTCGGCGAGAAACACCGCCTTTTAACGGAAGACTACAAAATCGTCGCACATCGACAGGCACGAACT
ACACGCAGTCCTGGGATAAAACTGTACGAGAAAATAACGAAACTCACGCGATGGTACCCGTGCGGACTACC
AGGTTACGCAAAGTGTGGCGTCGAAATAGGCCTACCAACAGTCTGGGATCAAGAATACCCAAAATCGATCC
CTACAATAGTTCCGAATGCTAGTGCTAATAGTACTCAAACCAAT

>Seq. 9.2.6 **PGRPLB** transcript 1 (ENSANGT00000013948)
CGTCCGCCACAGTATCATTTCAACCTGCGCCGGAGCCATCCGCGATAATAAGTGCGCGTACTTGAGTGCGT
ACCACCGACCAGTTGTGCGCCATCCCGTGCCGAACAAAAACAATACGTTTCGCGTTTCCACGCTATCAACA
ATCGCGTCGAGACCACGCGCGACACTGCTCTCTGTACGCAAATAATTCGATCGAGTGGGGCTGCTCTTTGG
GTTGCGTTGTCAAGCTCACGCGCAGGCTACTAGATACCGCACCTGCAGACAGACATGAAAGTGCAATCGGT
AACCTTCCCGTGGTTCTTATCCCGACAGTAGTGACAGTGATTTATTTGCTGAACGCACCAACGCACACAAC
ACCGTACAACAAATTCCTCACGAAAATACTGCTTATAAAAAAACTCCAGTGGCGGTGAGGAGAAAGATTGT
CCAAGAAGCGATTTCGGAACGGTTTAACGCAGAAGTTAGTGCACTAAATTATTGCAACAAAATGGACTTTG
TGAAAGACTTTTGTTACTACTTCGGCGTCATAGCATTTACCCTATTTTACGTTGCGGTTGATAGTAAAGCA
TGTGACCCGGTGCCCTACGTGACGCGAGACTTTTGGAGTGCTCTGCCACCGAAACGGATCGAACACTTCGC
CGGTCCCATCCCGTACGTCATCATACATCATTCGTACCGGCCGGCGGCATGCTACAACGGGCTGCAGTGCA
TCGCGGCAATGCAATCGATGCAGAAGATGCACCAGGACGAGCGCCAGTGGAACGACATCGGGTACAGCTTT
GCGGTCGGTGGCGACGGGCACGTGTACCAGGGCCGGGGATTCAATGTGATTGGCGCTCATGCGCCCCGATA
CAACAACCGCAGCGTTGGGATCTGTCTGATCGGTGATTGGGTTGCGGATTTGCCACCGAAGAACATGCTCA
CGGCCGCCCAAAACCTCATCGAGTATGGTGTCCGGAATGGGCTGATTGCACAGAACTACACCTTGCTGGGG
CATCGGCAGGTACGCACGACCGAATGTCCGGGCGATCGGCTGTTCGAGGAGATCAAAACCTGGCCCCATTT
CGATCCGATGACGGACATTGTCGATCAGAACAGTGTTTGACAGTGGCGAACGGAACCGGCCCCTACATGCT
ACAAGTATTGAAGAAACTCTTAAATAGCTTATTAACGTGCGCTACCTCTCCTTTGCATACGGTAACCATCA
TGTTCCTAAAGTGTGTGTTTTTTTTTTTTGGAATAAAACCATTCAAAACA

>Seq. 9.2.7 **PGRPLB** transcript 2 (ENSANGT00000024009)
ATGAATTTCTTATCACTGAAATACCATTTACGTTGGCAAACAGCTGACCAAACATTAGATTACAAATTGGC
AGTACGTTGTTGTGTTTACTTCCAGCTACTGCCCCCGGTAGTTTACTCAAATGCCCGCATATTAAGAGCCT
ACTGCCAGTTGGACAGTTTAATCATGATCGTGTACGTGATCGTGATCATTGCTTCCGTCATTCAGCTGCAT
GCTGCAGTCATTCGTGATGCGGTAATGGAACTTTTCCCCTTTTCCGACGATTCCGATACTACTACTGCACC
AACAATGACGTACGGTGCAAACCCGGTGCCCTACGTGACGCGAGACTTTTGGAGTGCTCTGCCACCGAAAC
GGATCGAACACTTCGCCGGTCCCATCCCGTACGTCATCATACATCATTCGTACCGGCCGGCGGCATGCTAC
AACGGGCTGCAGTGCATCGCGGCAATGCAATCGATGCAGAAGATGCACCAGGACGAGCGCCAGTGGAACGA
CATCGGGTACAGCTTTGCGGTCGGTGGCGACGGGCACGTGTACCAGGGCCGGGGATTCAATGTGATTGGCG
CTCATGCGCCCCGATACAACAACCGCAGCGTTGGGATCTGTCTGATCGGTGATTGGGTTGCGGATTTGCCA
CCGAAGAACATGCTCACGGCCGCCCAAAACCTCATCGAGTATGGTGTCCGGAATGGGCTGATTGCACAGAA
CTACACCTTGCTGGGGCATCGGCAGGTACGCACGACCGAATGTCCGGGCGATCGGCTGTTCGAGGAGATCA
AAACCTGGCCCCATTTCGATCCGATGACGGACATTGTCGATCAGAACAGTGTTT

>Seq. 9.2.8 **PGRPLC1** partial CDS
GTGAACGCATACGCAACGATTCGGTCATCGCCTCACCGTCATCGGGCCCATCCAGTACGCGCCCATCAAAA
CGGGAAGTCCGTGTTCAGGAATTATACAAGAAAAAAACTTCTTTTCTTGCTAGACAAAGCGTGTTACCACA
GCAACACACTTCCAAGAACGAGAATCTGTAATCTGTGTATTCCTGTTGGTGGAGGAAAAGCAGAGACTGGA
GCTACTGCGAGGCTAAGATGTCTCTCAACACTTCCGTGAGATAAAGAATAGCCACACTACAAACACAAACA
AACGGAAGCTTTCGGAGATACTTGCAAGTCGTGTGACTGTGACAGTGCCCGTTGCAGCAGTAAACGAAACC

CGAACGATGGCACTGGCAAGGAACCTACAGCTAGAAAATGATACCTCAGTCAACACCGGATACCACACTAA
GGCACAGTCGATGTCGGCGAGCGGCGGTGATACGATCGCTGGCCTAAACAACTCGGGCACCGGGACCGGCA
AGTCAGATAGCGCTAAGGGCAGCAGTCTTTCCAATATCTCGGCGTCATCCATCAAACGAACAGACAACGGT
GGCGATGATGACGACTCGAGTGTCTTCGATAGCAGTAGTAGCGATACCGAGTGTGACGATGATTCTATCAA
GCGAGCGATTGATCGTATCCCGGGTACACTGGCACCGGGGGAGGCACGCGTGCTCCCGAACGCGAACGTTA
AGGTTGCCGTGGAAGCCATCGCTCCCGGGGTACGGCCAAGCCCGGCTGCATCCACGATCGGGGCGATCGCG
GTCCACAACTCGTCGGACATCACGTTCGGCAACAAAACGTACATCAAGGGGCAGGTTGTGATAAAGAACAT
TTACCAGGATCTACCATCCGCAACCAAACCTCCCGAGCCACGCACCTGGCAATCTAGTCTTAAGACAATCA
TAAAGGATAAACCCTTGATCAGCTTTATCGTAATGGTATCGCTGATGATTGTCCTGTGTGCCATTGTAGCG
GTCATATCGATCCTCACCGCGTCAGAGGATGATCTATTCCCAGATCCGAGACCGTTACGGTTAGTGACGAG
AACGGAATGGCTGGCGCAACCGCCCAGAGAGGAATTGACCGATCTGAAACTGCCGGTTAATAACGTCATCA
TTGCTCACACTGCCACCGAAGGTTGCACTACTCAGCAAGCATGCAAAGCACTAGTGCAATCAATTCAAAGG
TATCACGTAAACCACACCAACTACGGTGATATCGGTTACAACTTTTTGATCGGTGGTGACGAATTTGTGTA
CGAGGGCCGCGGATGGCTTAAAGTTGGAGCCCACACCAAAAATTACAACACGATCAGCTACGGGATAGCAT
TCATCGGTAACTACGAGACTATCAATCGACCAACGGAACAGCAGATGGAACAGTTGGTACTGCTGCTTCGC
AACGGTACCGACGGCGGATGGCTTGCCAAGGAGTACAAGCTGTGCGGTGCCAGCCAGCTGAAGGGCACCAT
CAGCCCGGGCAAATATCTAATGGAACAGCTGAGGCTTTTGCCACATTTTAGTGAACGTTTGTAATTATGAT
ACGCATAGGTGATTGAACGTACGTAGTGTAGGTGCATGTGCCATTGAAGCGAAAGTGATGAATGCAGAATC
GAGTAAGACTGAGTGGCACCCTCTAAGATGAAGAACAACTCGTTAACCACGTTAACCCTGTCACTGGCAAC
CAAAAAAAACATCATATACCAGGCAACCGGGCTGCGCGGTTGGAAGCTACGTTCGAGGCTTGTAACTTTTG
AATACGTAATCCAATATCAATGCAATACTGTGATGAAAATTGAATAAACTAATGTAGTTTCTTTAACTGTG
CATAGATTGGTTCAACTCGCTACCGTTTTTATGTTATAGCTTTTTAAAGTTGACAGGATTTTATTGAAAAA
AAAAAACAAAATGAAA

>Seq. 9.2.9 **PGRPLC2** partial CDS
GTGAACGCATACGCAACGATTCGGTCATCGCCTCACCGTCATCGGGCCCATCCAGTACGCGCCCATCAAAA
CGGGAAGTCCGTGTTCAGGAATTATACAAGAAAAAAACTTCTTTTCTTGCTAGACAAAGCGTGTTACCACA
GCAACACACTTCCAAGAACGAGAATCTGTAATCTGTGTATTCCTGTTGGTGGAGGAAAAGCAGAGACTGGA
GCTACTGCGAGGCTAAGATGTCTCTCAACACTTCCGTGAGATAAAGAATAGCCACACTACAAACACAAACA
AACGGAAGCTTTCGGAGATACTTGCAAGTCGTGTGACTGTGACAGTGCCCGTTGCAGCAGTAAACGAAACC
CGAACGATGGCACTGGCAAGGAACCTACAGCTAGAAAATGATACCTCAGTCAACACCGGATACCACACTAA
GGCACAGTCGATGTCGGCGAGCGGCGGTGATACGATCGCTGGCCTAAACAACTCGGGCACCGGGACCGGCA
AGTCAGATAGCGCTAAGGGCAGCAGTCTTTCCAATATCTCGGCGTCATCCATCAAACGAACAGACAACGGT
GGCGATGATGACGACTCGAGTGTCTTCGATAGCAGTAGTAGCGATACCGAGTGTGACGATGATTCTATCAA
GCGAGCGATTGATCGTATCCCGGGTACACTGGCACCGGGGGAGGCACGCGTGCTCCCGAACGCGAACGTTA
AGGTTGCCGTGGAAGCCATCGCTCCCGGGGTACGGCCAAGCCCGGCTGCATCCACGATCGGGGCGATCGCG
GTCCACAACTCGTCGGACATCACGTTCGGCAACAAAACGTACATCAAGGGGCAGGTTGTGATAAAGAACAT
TTACCAGGATCTACCATCCGCAACCAAACCTCCCGAGCCACGCACCTGGCAATCTAGTCTTAAGACAATCA
TAAAGGATAAACCCTTGATCAGCTTTATCGTAATGGTATCGCTGATGATTGTCCTGTGTGCCATTGTAGCG
GTCATATCGATCCTCACCGCGTCAGAGGATGATCTATTCCCAGATCCGAGACCGTTACGGTTAGTGACGAG
AACGGAATGGCTGGCGCAACCGCCCAGAGAGGAATTGACCGATCTGAAACTGCCGGTTAATAACGTCATCA
TTGCTCACACTGCCACCGAAGGTTGCACTACTCAGGCCGCATGCCGTTTGCGGGTGCGCTTAATTCAGGAG
TTTCATATGGATGGCAAAAACTACGACGATATAACGTACAACTTTCTAATCGGTGGAGATGGACATATCTA
CGAGGGTCGAAATTGGCACAAAATTGGTGCCACCATTCCCGGTTACAATTCCCGCAGCATAACGGTGGCCT
TTGTGGGAGAGTACAATTACGGTGGCAAACCTACTAAAAAGCAAATCGAGCTGTTGAAATATCTACTGCAC
TTTGGTGCAAAGGAGCGGCACCTAAAGGAAGGTTACCGAATATATGCCTCGGAGCAGCTGGATCCAACGGT
TGGGACGGGCAAGTGGCTAATCGAAGCACTGCACAGCTTGCCCCAGTTTGTCGATAAGGAGCAAAATAAGG
ACGAGCGCCCCGATCACGAG

>Seq. 9.2.10 **PGRPLC3** partial CDS
GTGAACGCATACGCAACGATTCGGTCATCGCCTCACCGTCATCGGGCCCATCCAGTACGCGCCCATCAAAA
CGGGAAGTCCGTGTTCAGGAATTATACAAGAAAAAAACTTCTTTTCTTGCTAGACAAAGCGTGTTACCACA
GCAACACACTTCCAAGAACGAGAATCTGTAATCTGTGTATTCCTGTTGGTGGAGGAAAAGCAGAGACTGGA
GCTACTGCGAGGCTAAGATGTCTCTCAACACTTCCGTGAGATAAAGAATAGCCACACTACAAACACAAACA

AACGGAAGCTTTCGGAGATACTTGCAAGTCGTGTGACTGTGACAGTGCCCGTTGCAGCAGTAAACGAAACC
CGAACGATGGCACTGGCAAGGAACCTACAGCTAGAAAATGATACCTCAGTCAACACCGGATACCACACTAA
GGCACAGTCGATGTCGGCGAGCGGCGGTGATACGATCGCTGGCCTAAACAACTCGGGCACCGGGACCGGCA
AGTCAGATAGCGCTAAGGGCAGCAGTCTTTCCAATATCTCGGCGTCATCCATCAAACGAACAGACAACGGT
GGCGATGATGACGACTCGAGTGTCTTCGATAGCAGTAGTAGCGATACCGAGTGTGACGATGATTCTATCAA
GCGAGCGATTGATCGTATCCCGGGTACACTGGCACCGGGGGAGGCACGCGTGCTCCCGAACGCGAACGTTA
AGGTTGCCGTGGAAGCCATCGCTCCCGGGGTACGGCCAAGCCCGGCTGCATCCACGATCGGGGCGATCGCG
GTCCACAACTCGTCGGACATCACGTTCGGCAACAAAACGTACATCAAGGGGCAGGTTGTGATAAAGAACAT
TTACCAGGATCTACCATCCGCAACCAAACCTCCCGAGCCACGCACCTGGCAATCTAGTCTTAAGACAATCA
TAAAGGATAAACCCTTGATCAGCTTTATCGTAATGGTATCGCTGATGATTGTCCTGTGTGCCATTGTAGCG
GTCATATCGATCCTCACCGCGTCAGAGGATGATCTATTCCCAGATCCGAGACCGTTACGGTTAGTGACGAG
AACGGAATGGCTGGCGCAACCGCCCAGAGAGGAATTGACCGATCTGAAACTGCCGGTTAATAACGTCATCA
TTGCTCACACTGCCACCGAAGGTTGCACTACTCAGACGAAATGCATGTATCAGGTAAAGTTGATCCAGGAG
TTTCATTCGAGTCCCGACTCGAGGAACTTTAGCGACATTGCGTATCAATTTCTGGTGGGTGGTGATGGCAA
CGCGTACGAGGGTAGAGGATGGACAAAGCAAGGAGCGCACACGAAAGGGTTCAACGTGGACAGCATCTGCA
TCGCGTTCATTGGAACATTCATCGCGGATCCGCCACCGATCGCTCAGCTCAGTGCTGCGCAGCAGCTCATA
CTGCTAGGTATGAAGGAAAACTATCTTGCCTCCAACTACAGCCTGTACGGCCACCGACAGCTGGCACCGTT
CGAGAGTCCGGGCAAGGCACTGTTCGATATCATCAAAACTTGGCCCCACTGGTCGAACAAGCTGGGGTCCA
ATCACTGGGTCGAACCGGCGAACGACACCACAAACCGACGATGACGCAGGAAACGACTTGCCCCCATCACC
ATTCCTCCCCCCAAGTGTCCTTGATGACTTCCTCTCTCCTTCTCTCTCTCTTGAACCGGAAAAAGAGAACC
AAAAAGACTGTGATGATGAAAATAATCGAATAAACATTTTAAATATGGTGTAATTTAGCTCATACGATAAA
A

>Seq. 9.2.11 **PGRPLC 3rd exon** with alternative spliced extension cassette
TACCATCCGCAACCAAACCTCCCGAGCCACGCACCTGGCAATCTAGTCTTAAGACAATCATAAAGGATAAA
CCCTTGATCAGCTTTATCGTAATGGTATCGCTGATGATTGTCCTGTGTGCCATTGTAGCGGTCATATCGAT
CCTCACCGCGTCAGGTAAAGCTAGATTTAGGCTCCCCGTAGGCGACGGCGACGACAATCGTCCCAATATTC
CGCAGGACAAGGATATAG

>Seq. 9.2.12 **PGRPLC 3rd exon** without alternative spliced extension cassette
TACCATCCGCAACCAAACCTCCCGAGCCACGCACCTGGCAATCTAGTCTTAAGACAATCATAAAGGATAAA
CCCTTGATCAGCTTTATCGTAATGGTATCGCTGATGATTGTCCTGTGTGCCATTGTAGCGGTCATATCGAT
CCTCACCGCGTCAG

>Seq. 9.2.13 **PGRPLD** partial CDS
ACCAACGCAAGAGTGTCAATCCATCTCGAACCGAATCGACGCTGCTTTCCCGAATACAATCTTCAATTCCT
CCGCCTCCTCTATGGGCTAAGGGTTTGCGGGCCAGCGAAGGCACCGCTGCAGGCTCCCGACATCCTTATAT
TGTTGCCAACCTTCCATCCCTCCCAGCCAGCAATGTATCAGGCCGCGCGCACCGTTGCGGCGTCGAGCTCG
AGCCGGCTGCACTCCCCCGTTGCCTGGAGCCGCAAGTCCGTGCAGGAGGGCTACTACGATCTCGAGGCGGG
CGAAACGGAACGAACGCCCCTGCTGTACGTGCGCGACAATGGACGCTACCGGGCGAAGAGCCTGCACCGTG
CCAACGATCCGAACCGGGTGCATCCGGTCGCGCTGGGGCTGATGGTGGCGCTGCTCGTCTTCCTGCTGATC
GGTGTCGTGATCGGCGTCTATCTGCTGCTGCTGACGGTTCAACGGCCCTGGCCCGTATCGCATCCGTTCTT
TTTGGTCGAGCGACCCGCCTGGTGGCAGTATCCGGTGGCGCTCGAGGCGGCCACCCTCGACCGGCTCGCCG
TTACTAACGTGATCGTGCTGCACACCGACTCGGAGGATTGTCTGGATCAGGAGCGCTGTGTGCGCTACCTG
CAAAATACCGAGCGGAGCTGCTGGGCGGCACGGCGTGACCACATCCCGTACAACTTCTTGATCGGCGGGGA
CGGTGTGACGTACGAGGCGCGCGGCTGGAAGAGTCAGCACGGGTTCGTCGATCTGCCCGGCCGCAACACAA
CGCTCGTGGTGGGCATGATCGGCAACTTCACCGACCGTCAACCGGCGGAAGTACAGTACGCCGAGCTGAAG
GCATTTATCACCGAGTCGATCCGGCGGCTTAGCCTCTCGCCCCAGTACCGCCTGCACGGAGCGGTCAATGC
AACGCGTCCCTCGCGAGACGGTACAATCGCGCTGTACAGTCAGCTGGAGCGATGGCCACACTGGAAGGGGT
TCGTCGCTCGCGACCAACCC

>Seq. 9.2.14 *Drosophila* **PGRP-LC 3rd exon** with alternative spliced extension cassette
GCTCCAAACACGAAGATCCGGCGCAAACCCCACCAATTTGTCCTTTTCTGCCCAACACTGTCGGACGCAAG
GCCGTCACAGTTACAGTGGTTTTTGTAACTTTAACCTTCCTGCTGGGTATCGTACTGGCCACCACCACAAA
TCTCTTCGGAAAGACGTTGAACCAAAGTAAGATCAGAGATGACGACGATTACAGACAAAATATTCCAATCA
ATTCAACAATAG

>Seq. 9.2.15 *Drosophila* **PGRP-LC 3rd exon** without alternative spliced extension cassette
GCTCCAAACACGAAGATCCGGCGCAAACCCCACCAATTTGTCCTTTTCTGCCCAACACTGTCGGACGCAAG
GCCGTCACAGTTACAGTGGTTTTTGTAACTTTAACCTTCCTGCTGGGTATCGTACTGGCCACCACCACAAA
TCTCTTCGGAAAGACGTTGAACCAAA

9.3 PROTEIN SEQUENCES

>**Seq. 9.3.1 PGRPLB** transcript 1 (ENSANGP00000013948)
PSRGSYPDSSDSDLFAERTNAHNTVQQIPHENTAYKKTPVAVRRKIVQEAISERFNAEVSALNYCNKMDFV
KDFCYYFGVIAFTLFYVAVDSKACDPVPYVTRDFWSALPPKRIEHFAGPIPYVIIHHSYRPAACYNGLQCI
AAMQSMQKMHQDERQWNDIGYSFAVGGDGHVYQGRGFNVIGAHAPRYNNRSVGICLIGDWVADLPPKNMLT
AAQNLIEYGVRNGLIAQNYTLLGHRQVRTTECPGDRLFEEIKTWPHFDPMTDIVDQNSV

>**Seq. 9.3.2** PGRPLB transcript 2 (ENSANGP00000022756)
MNFLSLKYHLRWQTADQTLDYKLAVRCCVYFQLLPPVVYSNARILRAYCQLDSLIMIVYVIVIIASVIQLH
AAVIRDAVMELFPFSDDSDTTTAPTMTYGANPVPYVTRDFWSALPPKRIEHFAGPIPYVIIHHSYRPAACY
NGLQCIAAMQSMQKMHQDERQWNDIGYSFAVGGDGHVYQGRGFNVIGAHAPRYNNRSVGICLIGDWVADLP
PKNMLTAAQNLIEYGVRNGLIAQNYTLLGHRQVRTTECPGDRLFEEIKTWPHFDPMTDIVDQNSV

>**Seq. 9.3.3** PGRPLD partial CDS
TNARVSIHLEPNRRCFPEYNLQFLRLLYGLRVCGPAKAPLQAPDILILLPTFHPSQPAMYQAARTVAASSS
SRLHSPVAWSRKSVQEGYYDLEAGETERTPLLYVRDNGRYRAKSLHRANDPNRVHPVALGLMVALLVFLLI
GVVIGVYLLLLTVQRPWPVSHPFFLVERPAWWQYPVALEAATLDRLAVTNVIVLHTDSEDCLDQERCVRYL
QNTERSCWAARRDHIPYNFLIGGDGVTYEARGWKSQHGFVDLPGRNTTLVVGMIGNFTDRQPAEVQYAELK
AFITESIRRLSLSPQYRLHGAVNATRPSRDGTIALYSQLERWPHWKGFVARDQP

>**Seq. 9.3.4** AaREL1 *Aedes agypti* REL1
MGPNDNNVMEDDPSLPDLPNNLIVTDDILDLIIDIIDQEMPPTAQERPHVVITEQPQSKALRFRYECEGRS
AGSIPGVHSTPEQKTFPGIEIRGYKGRAVVVVSCVTKDPPYRPHPHNLVGKEGCKKGVCTVEINSSTMSYT
FSNLGIQCVKKKDIEEALRLREEIRVDPFKTGYGHARQPATIDLNAIRLCFQVFLEGQQRGRFTEPLQPVV
SDVIYDKKAMSDLVICKLSDVTAPVAGGREIILLCEKVAKEDIAVRFYEEQHGNIVWEEYGEFQHTNVHKQ
VAICFRTPRYRTLEVEHPVMVNIQLKRPSDGATSEPLPFELTPLDSAKRKKRKPNTWGDTPEHMPEAARHP
YYATLGNVPGEAIKTEPRDSPTHPFNFIGNYRQPNESPSPIERSPLYGNTTPVGQTSPYNAGRTSVTPTPG
GYQPNVMNPDFLLGGTQQFQNQYAFPTLPPTNGQPMYLQQPQQQQQQQQFQPPANMNNQQQFNVSNLIDPQ
DQLPQPGPSGVQAVPSFQPSGSIGEINISSPWLTMDSSELKSLMVHTLSDMKQEQQRIQQEAETLSDSFDR
LSTNLNNLNKP

>**Seq. 9.3.5** CgRel2 *Crassostrea gigas* Rel2 (Pacific oyster) (Acc#
AAK72691)
MAELNFGYDNPMGLGSNDLRLLMDGNHELLTNFDTADNDLVQLITQNQESFMTPQIPVTQPQPPKPQKQTT
HRGTPHKTQDGIYVEIVEQPKQRGLRFRYECEGRSAGSIPGEGSTSEKKTFPTIKIHNYTGTAVIVVSCVT
KDQPYEPHPHNLVGRDCKRGVCTLKVKDTNVISFPHLGIQCAKKKDVENNLKQRKEINVDPFQSGFKHLNK
INSIDLNVVRLCFQVFLPDENGKITRIVPPVVSHCIHDKKSLNELVICRVDRHSGKAKGGDEVFLLCEKIN
RDDIVVRFYEETECGECLWEDFADFSTNDIHRQYAIVFRTPPYKDTMITRPKEVKMQLKRNNEPETSDSIP
FIYMPEDPDPDRIMEKRKRKADQLKNWGFDMNVSGEDIKQRLKIKATKSRIKQEVPEVPPYTVENLLGGAH
QMPPVSNVTTTHTKDGTSLNVRTLDNTTMGNVTITPTESVRQQLANLPDDVQIKLLQNLAVRKLQEEATSR
INQPGQQLEKDHSTFLQSLLGGGMGMPQNSNPAQPLNQNLMMGGQAMGQCSGMDQENVDLMLQAYLGDQDQ
NMSFDSFGSNIGNMNFDPSVGANLNVTNDQGGLDETKTAVQSLGQS

>**Seq. 9.3.6** SpNF-kappaB *Strongylocentrotus purpuratus* (sea urchin) NF-
kappaB (Acc# T14892)
MEEKSDSVQSVEIPEDMLQQLIHQNGQVGLPMSGAENDFDSNDLMHLPVMKGLQEEVSKPHLKILEQPRQR
GFRFRYGCEGPSHGGLPGQNSQRGKRSFPSVEICNYKGSARIVVSLVTNEETPRPHAHSLVGKHCKDGLCT
VQVGPKDMTASFPNLGILHVTRKDVVPTLKTRILAQHRLYKDLINNSTPGESHWTEPSDAEIEKKAKEMAK
DMDLSVVRLCFQTYLPDISGHFTRPLDPVISVPVFDSKAPNATTLKICRMDKSAGCCTGGEEVYLLCDKVQ
KEDIQVKFFEISADGQMVWQSLAEFGPTDVHRQYAIVFKTPAYKDINIDKPVYVHVQLKRKSDNETSDPKP
FTFHPQVPDREGILRKRKKHLAHFNEYSSTYQQGGLGGSNGMGGGGGSTSFNFNNPNMGYGANTYNFSGNQ
MTSSTSQPQSVPQQIPSHQSGVVATGNSSQQMVYTTSTGAQNHQPLPSIATLARKPQEQLQQQQFQLQQQL

QQQQQQQQQQQQQQQRQQQQQQPMQMDFQMPPIQGSAMGGQAGGDRVKMESESKYCGNMDTAEGFQLDSNT
LRDCGPLPDLKELQDTGVLDLIPKGISRPKDMVEKNIQTDDITIDSDSMAWHVAQVTANALHDYAATGDIK
TILTVQRHLIAVEDDNGDTALHAAIINKKYDVTHALLSAVIKIPDQIIVNQTNHLKQTPLHLAVITNQSKM
VEVLLRCGANPNLCDHEGNTPLHLATMMGMTEGVNFLVRGPKAKAAIKPIKTDINPTNYEGLAPVHLAVIA
KNLDILKALVSSGADVNVADGKTGRTALHYAVEVESFPILGYLLIEAKVDINAVTFCGDSALHLASSLDLR
AVATLLIAAGADPKLENADLGDDSDSDVDGEDDDGEKDGEKKEEGVEEEEKEKHGKTPYQLATSAKMKEVL
SGHSTRRVEEEDFDYIETGSAISSVGSIQSEDHMKVNSWDQGYFSGKSSSSVSNLELKRTWPRVKSRMLGS
SLVTSKGDIGLLSETTTEKLGSMLDDNYPTTQSWFTLANRLGLSNMLNFLKLVPSPTAVILKQFEAMDGTI
KELRDVLSSMNHIEAVALLDQAVSWRKEKADEAHMYSSMHQSLGKMNLSDAHMHQHQFVY

>**Seq. 9.3.7** Mmp105 *Mus musculus* NF-kappa-B p105 subunit (Acc# P25799)
MADDDPYGTGQMFHLNTALTHSIFNAELYSPEIPLSTDGPYLQILEQPKQRGFRFRYVCEGPSHGGLPGAS
SEKNKKSYPQVKICNYVGPAKVIVQLVTNGKNIHLHAHSLVGKHCEDGVCTVTAGPKDMVVGFANLGILHV
TKKKVFETLEARMTEACIRGYNPGLLVHSDLAYLQAEGGGDRQLTDREKEIIRQAAVQQTKEMDLSVVRLM
FTAFLPDSTGSFTRRLEPVVSDAIYDSKAPNASNLKIVRMDRTAGCVTGGEEIYLLCDKVQKDDIQIRFYE
EEENGGVWEGFGDFSPTDVHRQFAIVFKTPKYKDVNITKPASVFVQLRRKSDLETSEPKPFLYYPEIKDKE
EVQRKRQKLMPNFSDSFGGGSGAGAGGGGMFGSGGGGGSTGSPGPGYGYSNYGFPPYGGITFHPGVTKSNA
GVTHGTINTKFKNGPKDCAKSDDEESLTLPEKETEGEGPSLPMACTKTEPIALASTMEDKEQDMGFQDNLF
LEKALQLARRHANALFDYAVTGDVKMLLAVQRHLTAVQDENGDSVLHLAIIHLHAQLVRDLLEVTSGLISD
DIINMRNDLYQTPLHLAVITKQEDVVEDLLRVGADLSLLDRWGNSVLHLAAKEGHDRILSILLKSRKAAPL
IDHPNGEGLNAIHIAVMSNSLPCLLLLVAAGAEVNAQEQKSGRTPLHLAVEYDNISLAGCLLLEGDAHVDS
TTYDGTTPLHIAAGRGSTRLAALLKAAGADPLVENFEPLYDLDDSWEKAGEDEGVVPGTTPLDMAANWQVF
DILNGKPYEPVFTSDDILPQGDMKQLTEDTRLQLCKLLEIPDPDKNWATLAQKLGLGILNNAFRLSPAPSK
TLMDNYEVSGGTIKELMEALQQMGYTEAIEVIQAAFRTPATTASSPVTTAQVHCLPLSSSSTRQHIDELRD
SDSVCDSGVETSFRKLSFTESLTGDSPLLSLNKMPHGYGQEGPIEGKI

>**Seq. 9.3.8** Mmp100 *Mus musculus* NF-kappa-B p100 subunit (Acc# Q9WTK5)
MDNCYDPGLDGIPEYDDFEFSPSIVEPKDPAPETADGPYLVIVEQPKQRGFRFRYGCEGPSHGGLPGASSE
KGRKTYPTVKICNYEGPAKIEVDLVTHSDPPRAHAHSLVGKQCSELGVCAVSVGPKDMTAQFNNLGVLHVT
KKNMMEIMIQKLQRQRLRSKPQGLTEAERRELEQEAKELKKVMDLSIVRLRFSAFLRASDGSFSLPLKPVI
SQPIHDSKSPGASNLKISRMDKTAGSVRGGDEVYLLCDKVQKDDIEVRFYEDDENGWQAFGDFSPTDVHKQ
YAIVFRTPPYHKMKIERPVTVFLQLKRKRGGDVSDSKQFTYYPLVEDKEEVQRKRRKALPTFSQPFGGGSH
MGGGSGGSAGGYGGAGGGGSLGFFSSSLAYNPYQSGAAPMGCYPGGGGGAQMAGSRRDTDAGEGAEEPRTP
PEAPQGEPQALDTLQRAREYNARLFGLAQRSARALLDYGVTADARALLAGQRHLLMAQDENGDTPLHLAII
HGQTGVIEQIAHVIYHAQYLGVINLTNHLHQTPLHLAVITGQTRVVSFLLQVGADPTLLDRHGDSALHLAL
RAGAAAPELLQALLRSGAHAVPQILHMPDFEGLYPVHLAVHARSPECLDLLVDCGAEVEAPERQGGRTALH
LATEMEELGLVTHLVTKLHANVNARTFAGNTPLHLAAGLGSPTLTRLLLKAGADIHAENEEPLCPLPSPST
SGSDSDSEGPERDTQRNFRGHTPLDLTCSTKVKTLLLNAAQNTTEPPLAPPSPAGPGLSLGDAALQNLEQL
LDGPEAQGSWAELAERLGLRSLVDTYRKTPSPSGSLLRSYKLAGGDLVGLLEALSDMGLHEGVRLLKGPET
RDKLPSTEVKEDSAYGSQSVEQEAEKLCPPPEPPGGLCHGHPQPQVH

>**Seq. 9.3.9** DmRel *Drosophila melanogaster* Relish Rel (CG11992-PA)
MNMNQYYDLDNGKNVMFMNDASSTSGYSSSTSPNSTNRSFSPAHSPKTMELQTDFANLNLPGGNSPHQPPM
ANSPYQNQLLNNGGICQLGATNLINSTGVSFGVANVTSFGNMYMDHQYFVPAPATVPPSQNFGYHQNGLAS
DGDIKHVPQLRIVEQPVEKFRFRYKSEMHGTHGSLNGANSKRTPKTFPEVTLCNYDGPAVIRCSLFQTNLD
SPHSHQLVVRKDDRDVCDPHDLHVSKERGYVAQFINMGIIHTAKKYIFEELCKKKQDRLVFQMNRRELSHK
QLQELHQETEREAKDMNLNQVRLCFEAFKIEDNGAWVPLAPPVYSNAINNRKSAQTGELRIVRLSKPTGGV
MGNDELILLVEKVSKKNIKVRFFEEDEDGETVWEAYAKFRESDVHHQYAIVCQTPPYKDKDVDREVNVYIE
LIRPSDDERSFPALPFRYKPRSVIVSRKRRRTGSSANSSSSGTESSNNSLDLPKTLGLAQPPNGLPNLSQH
DQTISEEFGREKHLNEFIASEDFRKLIEHNSSDLEKICQLDMGELQHDGHNRAEVPSHRNRTIKCLDDLFE
IYKQDRISPIKISHHKVEKWFIEHALNNYNRDTLLHEVISHKKDKLKLAIQTIQVMNYFNLKDVVNSTLNA
DGDSALHVACQQDRAHYIRPLLGMGCNPNLKNNAGNTPLHVAVKEEHLSCVESFLNGVPTVQLDLSLTNDD
GLTPLHMAIRQNKYDVAKKLISYDRTSISVANTMDGNNALHMAVLEQSVELLVLILDAQNENLTDILQAQN
AAGHTPLELAERKANDRVVQLLKNVYPEKGELAMTWIPCKVKEE

IDSSSDESSDAGQLEIKSEEMDIETKDEDSVELDLSSGPRRQKDESSRDTEMDNNKLQLLLKNKFIYDRLC
SLLNQPLGHGSDPQDRKWMQLARQTHLKQFAFIWLGAEDLLDHVKRKGASVEFSTFARALQAVDPQAYALL
VNPT

>**Seq. 9.3.10** DmDorsal *Drosophila melanogaster* Dorsal dl (CG6667-PA)
MFPNQNNGAAPGQGPAVDGQQSLNYNGLPAQQQQQLAQSTKNVRKKPYVKITEQPAGKALRFRYECEGRSA
GSIPGVNSTPENKTYPTIEIVGYKGRAVVVVSCVTKDTPYRPHPHNLVGKEGCKKGVCTLEINSETMRAVF
SNLGIQCVKKKDIEAALKAREEIRVDPFKTGFSHRFQPSSIDLNSVRLCFQVFMESEQKGRFTSPLPPVVS
EPIFDKKAMSDLVICRLCSCSATVFGNTQIILLCEKVAKEDISVRFFEEKNGQSVWEAFGDFQHTDVHKQT
AITFKTPRYHTLDITEPAKVFIQLRRPSDGVTSEALPFEYVPMDSDPAHLRRKRQKTGGDPMHLLLQQQQK
QQLQNDHQDGRQTNMNCWNTQNIPPIKTEPRDTSPQPFGLSYRAPPELTPSPQPLSPSSNYNHNSTPSPYN
MASAVTPTNGQQQLMSPNHPQQQQQQQQYGATDLGSNYNPFAQQVLAQQQQHQQQQQQHQHQHQQQHQQQQ
QQQQQQQQQSLQFHANPFGNPGGNSWESKFSAAAVAAAAATATGAAPANGNSNNLSNLNNPFTMHNLLTSG
GGPGNANNLQWNLTTNHLHNQHTLHQQQQLQQQQQQQYDNTAPTNNNANLNNNNNNNNTAGNQADNNGPTL
SNLLSFDSGQLVHINSEDQQILRLNSEDLQISNLSIST

>**Seq. 9.3.11** DmDif *Drosophila melanogaster* Dif (CG6794-PA)
MFEEAFGDIQEIINASMELNGGATGGGSVAGAVGGGGAAHHILSQSTSLPVMPSHIPLHLQNQNMNQNLPE
PSARSGPHLRIVEEPTSNIIRFRYKCEGRTAGSIPGMNSSSETGKTFPTIEVCNYDGPVIIVVSCVTSDEP
FRQHPHWLVSKEEADACKSGIYQKKLPPEERRLVLQKVGIQCAKKLEMRDSLVERERRNIDPFNAKFDHKD
QIDKINRYELRLCYQAFITVGNSKVPLDPIVSSPIYGKSSELTITRLCSCAATANGGDEIIMLCEKIAKDD
IEVRFYETDKDGRETWFANAEFQPTDVFKQMAIAFKTPRYRNTEITQSVNVELKLVRPSDGATSAPLPFEY
YPNPELLTKHNRRVAQKTVESLKRSLMSTNLHPSKQVKTSSQYTIFSKPQIATTTPQTQVSPGMPLMFPGG
SPNFVQDIKMENGFMDVDSQSSQCPSVERNFASPRSNCSTVDSIPPMQMGQNQTHLYLPDATNFTFNGNFA
SPSSNCSTVDSIPPFQIGQRNNHMYLPENSNFPVNGCSPTHFSGGSMTPINNNNNVLINNNNNDFLSQKMS
AISIPPQGNFGIKQVYQQTQQFLPQLQPESIPYLAQSHPEQSQYQQQQQPQEQQPPADEPTQSFSDLISSS
IGMAPIDTSELIQDIEAELNSLGIQPFK

>**Seq. 9.3.12** AaREL2 *Aedes aegypti* REL2 R6 isoform (Acc# AAM97895)
MSTLLNLDSYRHELFEHHQHQQQHHQLQQLSSSPTYSVLSMESVSPVSSSSASPAADTSKYYASNSPSSVS
NMSPKSTASDTSSFNMQNLNISNAYPYFAETHHQTMTTGEEGQLQTFTYNDPTQPRVELSVPQLTIVEQPV
DKFRFRYQSEMHGTHGSLMGVHTEKSKKTFPTVQLHGFQGEAKIRCSLFQVDPNRRAAHSHNLVIKSGEID
LNDPHDIDVNEECNYVAMFQGMGIIHTAKKNIAEELSKKMKKQRSVEMNRELTLREEYQLQKEAAEMAKTM
NLNQVCLCFQAFQVDASTGRWAQLCEPVYSNPINNMKSALTGELKICRLSATVGNVDGGEEVFMFVEKVCK
NNIKIRFYELDEYDQEIWQDWGSFSEADVHHQYAIAFKTPAYHNKDITEPVEVLMQLYRPRDKCQSEPVPF
KFKPRFNISRKRPRVSSGLLSNEIPTVVPNEPGSSRLPSFHHPFSMMPPMQAPIAEASGSSSSHTISKEYN
KSGIIQEILDSHIPTTIAGDISFSSNDFKEFVNCNSEELRKLITEIGEAQEESKLETDAVSVGHVDEAAAQ
FERTLEQYLEHNSQAKDVEILKKILAIIRLLFRKNYDQCRELISALWMSANKLKANCLHMAIERRNITIAC
KLVELLQDHHLLDMLDLFNERNETALHLAVSANLVEVVDVLLLTGSRISYCDSRGNSALHRAVYENALDSL
NVLLGHCKRNGFRLDSTNDDGFTALHLAVMCKNLKATKVLLDRGASYVLRDLKHGNNILHIAVESDSLDMV
NFILEGVDKTLADEPNNAGYTPLQLANARHLANANNKLIVRELLRYNPSGVLEKEHLTEEEDDDQEQEENL
PPSSDSQSETVLESMNVSCNRVEVIRLLEDYSPPNDEPKLTPLENVPYNSTLEDGSVYLFDEVCLSHLCGL
LNRKNLWREVGSLLEFNSFFMIWETSVNPAGMLLNYFEMQKVKLEHLIDVLQALDQKEAIHYIDEMICRQM
K

>**Seq. 9.3.13** AmDorsal *Apis mellifera* (honey bee) dorsal protein (Acc#
AAP23055)
MEQFHDMSDGNIHMSDVIEVIETDTKYNGREDQIPREMNTERLLPYVEIIEQPASKALRFRYECEGRSAGS
IPGVNSTSENKTFPTIKIVGYKGRALVVVSCVTKDQPYRPHPHNLVGKEACKQGVCTVEVSSENMTVTFAN
LGIQCVKKKDIEEALKIREEIRVDPFRTGFEHKRQPTSIDLNAVRLCFQVFLEGSQKRKFNVPLQPVVSDP
IFDKKAMSDLVICKLSHSNASVAGGMEMILLCEKVAKEDIQVRFFEEKDGQVLWEGFGDFQPVHVHKQTAI
AFRTPTYRMQQVEQPVQVYIQLKRPSDGATSEPFPFLMLPLGADDPDSLRRKRQKINNSQNALVLRHVQAE
AEKHAAMLYQYNFNIIISEPTERISPYGYPIGSGGSFPSLYPMATTSPQSQSTIQTLRPQVSPDRTSPMEY
RLYNPALIQSQPSPQYPSTSSHILQQPSIRTYTQQQFPYVHDTLQIQPQEQLTLSKVTSNYHEEFQSLNNA

VGEMEATNVTNILSMDNTQYNLDLSLPQLDSTELADLDISLSENLSSGLSISDSTKPETSKGINAEPSNIE
ESNNMTDSFTRIANNTIQELYTLNNMYKPTREVD

>**Seq. 9.3.14** TcDorsal *Tribolium castaneum* (red flour beetle) Dorsal (Acc# AAG22858)
MDENHLEQTLENAGMPDESINISDVIEVIETDPDFHESTMFNGGGEIPQNIIQQQPQQARRRAFVKIIEQP
ASKALRFRYECEGRSAGSIPGASSTPENKTFPSIQVMGYQGRAVVVVSCVTKDEPFRPHPHNLVGREGCKK
GVCTMEINSDTMCVTFSNLGIQCVKKKDIESALRLREEIKVDPFMTGFSHRNQPTSIDLNAVRLCFQVFLE
GERRGKFTVPLTPVVSEPIYDKKAMSDLMIVKLSHCNSYVDGGRNEIILLCEKVAKEDIQVRFFEEKNGKV
VWEGFGDFQPSQVHKQTAICFKAPRYHTLDITEPVKVFIQLRRPSDGATSEALPFELLPLDSEPGMLKRKR
QKYHDPSQLLRHVEESDKQRTHQPFLLNIPSEGIKVEPRDKAPSPYGGQMFEAPIPNVSPHYVTPTPPEAP
LFQNVLPPIGGMMNWNMPSTSVVGDMLNMNWNPSAQNDLSGVNLSETPSGISSLLNLDNQQLELKPIDLNS
GDLSMLETNNLSETFTQNLSLMDESEGRQEPNMTDSLTRLANSALDNICQASDMYKGAP

>**Seq. 9.3.15** AgREL2-F *Anopheles gambiae* REL2-F
MSTLLNLDSYRHELYEQQQLLGTSPIQYTVLSMDTPSPSSSSAAAAVVSVGEFTLGPGRTYASALSPSSSS
ASPSSPSSVASPNSRASNMSPESSASDQSAAYTLQNLNLSSSAGTMNYPGMGYQQQQQQQQQQQQHHQHQQ
LQQQQHHYYTPQLLNLDQEQLQTQTFTYVTSSNEAFAAPEPNYSEPHLVILEQPVDKFRFRYQSEMHGTHG
SLMGSRTEKSKKTFPTVELRGYGGEAKVRCSLYQVDPQRRAPHSHHLVIKSGELDLIDPHDLDVGGAAGAA
EPSEGVGGDGKYVATFQGMGIIHTAKKFIAEELYKKLRKHRLCELNREPTEREEQQMQKEAAVMARTMNLN
QVCLCFRAYRVEPGTGRWVPICEPVYSNPINNMKSALTGELKICRLSTTVSGVDGGEEVFMFVEKVCKNNI
KIRFYELDEYDQEVWQEMAIFSEADVHHQYAIAFKTPPYRHKDITEPVEVLMQLFRPRDRCQSEPVLFKYK
PRPGMMVVPGAGASGASRKRLRISSGNVSSEIPTVIQNDPNGPVGPAGVGGGGGGGGGGGGGVIGSGSTT
RLPPLHQPFPMLANHAGGGAIPEGQEPTSTTSLTTSAHHPDIMSGIGSTTTISKELTKASIIQEILNIPTT
IASDVAFDSSDFPCNSEEFNKLIQEIGNQQDLVKLETDAETAGGGATDTESVLGRAIADLVASGDDSRQGE
MLRKLLALIKLFAGDVNRSRQLLASHWTAANQQQLNCLHAAIRRNDTTIACKLIELLHEYQLAEELLDLPN
DRNETGLHLAVSCNSEPIVKALLGAGAKLHYCDYRGNTPLHRAVVENVPDMVRLLLLQGGLRLDCTNDDGL
TALQAAVYARNLKITRILLEAGASVREKDLKHGNNILHIAVDNDALDIVHYILEEVKEELGRERNNAGYTP
LQLADAKSHTGQGNNKLIVRELLRHYPDGLQKEVKKEVDAAEDDEEEEEEEQEEEEDEDEEGGEEHGQREA
SAPSSSVLDSMDLINGERASIARLLEEHEPEAEPQRKATKRSDSGPDRTEPDTLLDEQCLEELCRLLDAGS
GWRELGSLLDFHSFFTVWEQAPSPARMLLGYFEMQQLHLDRLIDMLRVLELRDPIRSIDEMICRRMK

>**Seq. 9.3.16** AgREL1 *Anopheles gambiae* REL1
MAGYVNGADAAFLPIEGILGESDLLDDIIHVIGKDIREEMPPVPNQRPYVEITEQPHPKALRFRYECEGRS
AGSIPGVNTTAEQKTFPSIQVHGYRGRAVVVVSCVTKEGPEHKPHPHNLVGKEGCKKGVCTVEINSTTMSY
TFNNLGIQCVKKKDVEEALRLRQEIRVDPFRTGFGHAKEPGSIDLNAVRLCFQVFLEGQQRGRFTEPLTPV
VSDIIYDKKAMSDLIICRLSDCTAPVSGGKEIILLCEKVVKEDIKVRFFEKKGNATVWENYAEFSHTDVHK
QVAISFRTPPYRTIDISDPVRVFVQLERPSDNTYSEARDFQFIPLDTVDLRRKRQKLTGSSNVFLPVVTAP
LPGPSPPSSLGMPGNIPNLSQLDATGGQSASTSGLPRGIYTYHNASAFQQMPKEEIKNEPGDSPSHNPSNQ
YQLQPMQPMFTAQSTSPGPDRSPATLTPSPGIGGPISPLDPGNVTPTPPAYTTLGGTGGTMGNLFGQFGTG
FSNAATQPNASVPMFETNLPGPSNGWVQPVPSHTQNGPSQPQPQQQNPFNLLNFGTLPLATGDPLMAMSTA
NTNNQPNVGPPSLGSDFYMNLDLANLDPVFNSSELRSVLGSLSTTDLNRLEQTANMQTSGGNYQQHSASN

>**Seq. 9.3.17** PGRP-LA1
MATNHQNGLSTGNGGNTTVQPAATSVINLSNSSDVVIGPMTQYQGSVTIYQYMDATVEASRIASKFPLPMR
SRSLTEFYDLVFPSHLSFPTAGRNNQNRGLPSAEPSTLRQERYFIYGALIFFAIVGFSTAIYFIVNQVRGP
ADIDREILFDTNYHRRTMPNLGNSHMIIERRNWGSQLDAHASIQLEHPVQYVIVTHIGLRSKNCTGMHECA
NRMRMLQDAAIGERNLPDIPSNFYLGGDGNVYVGRGWDIANSYHNRTLSVCFIGDFQTYEPKESQFSALHH
LLTYGVIQRKLASDYKLVARRQTKPTTASPGDMLYARLVRLSRWNPCGTQAYAHCGAELGFPSVWDDEHDF
NERNGIHNPLLIHRKSTEEE

>**Seq. 9.3.18** PGRP-LA2
MATNHQNGLSTGNGGNTTVQPAATSVINLSNSSDVVIGPMTQYQGSVTIYQYMDATVEASRIAIPNLGNGH
MVIDRHNWGAQQGVHGPYKLPHPIPYVLITHIGVHSEICSDVHVCSIKMRTLQDAAIAEKSLQDIPSNFYV

GGDGNVYVGRGWDTANAYANMSLAVCFMGDYGRYEPNDLQLSALDHLLTFGEKHRLLTEDYKIVAHRQART
TRSPGIKLYEKITKLTRWYPCGLPGYAKCGVEIGLPTVWDQEYPKSIPTIVPNASANSTQTN

>**Seq. 9.3.19** PGRPS1 (partial CDS)
MMLFGGILPYLLVLWQLLASVDDLLHILVARAQDEPAQESACPAIVKRAAWGAAKSKNVTYQLKPVANVIV
HHTTGERCATVATCKEMVANIQTYHQTDNRWSDIGYNFLISGQNVYEGIGWHRMGAHLRGYNDKSIGVAFL
GNFDQERPTPRSLNLLARLLQCGVELGELADDYRLYGARQLQSTNSPGRYLYAKLQELDHWQAQ

>**Seq. 9.3.20** PGRPS2 (partial CDS) (ENSANGP00000012978/agCP5898)
MNKFATVLVLASICLAGVSAQCPRIVTRAQWGARAASTSQLPIRPAPWVVMHHTAGASCTTDAACAQQMRN
IQSFHMDGNGWADIGYNFLVGENGAAYEGRGWGRQGAHAPGYNDRSVGMGVIGTFTNAIPNAAARTAARNL
ITCGVSLGHIASNYWLIGHRQAVATACPGNAFFNEIRTWPRFNPNV

>**Seq. 9.3.21** PGRPS3 (partial CDS)(ENSANGP00000012979/agCP5906)
MNKFATVLVLASICLAGVSAQCPRIVTRAQWGARAANTAQLPIRPAPWVVMHHTAGAACTTDAACAQQMRN
IQSFHMDGNGWADIGYNFLVGENGAAYEGRGWGRQGAHAPGYNDRSVGMGVIGTFTNGIPNAAARNAAQQL
ISCGVSLGHIASNYWLIGHRQAVATACPGNAFFNEIRNWPRFNPNV

>**Seq. 9.3.22** PGRPLC1 (partial CDS)
VNAYATIRSSPHRHRAHPVRAHQNGKSVFRNYTRKKLLFLLDKACYHSNTLPRTRICNLCIPVGGGKDVSQ
HFREIKNSHTTNTNKRKLSEILASRVTVTVPVAAVNETRTMALARNLQLENDTSVNTGYHTKAQSMSASGG
DTIAGLNNSGTGTGKSDSAKGSSLSNISASSIKRTDNGGDDDDSSVFDSSSSDTECDDDSIKRAIDRIPGT
LAPGEARVLPNANVKVAVEAIAPGVRPSPAASTIGAIAVHNSSDITFGNKTYIKGQVVIKNIYQDLPSATK
PPEPRTWQSSLKTIIKDKPLISFIVMVSLMIVLCAIVAVISILTASEDDLFPDPRPLRLVTRTEWLAQPPR
EELTDLKLPVNNVIIAHTATEGCTTQQACKALVQSIQRYHVNHTNYGDIGYNFLIGGDEFVYEGRGWLKVG
AHTKNYNTISYGIAFIGNYETINRPTEQQMEQLVLLLRNGTDGGWLAKEYKLCGASQLKGTISPGKYLMEQ
LRLLPHFSERL

>**Seq. 9.3.23** PGRPLC2 (partial CDS)
VNAYATIRSSPHRHRAHPVRAHQNGKSVFRNYTRKKLLFLLDKACYHSNTLPRTRICNLCIPVGGGKDVSQ
HFREIKNSHTTNTNKRKLSEILASRVTVTVPVAAVNETRTMALARNLQLENDTSVNTGYHTKAQSMSASGG
DTIAGLNNSGTGTGKSDSAKGSSLSNISASSIKRTDNGGDDDDSSVFDSSSSDTECDDDSIKRAIDRIPGT
LAPGEARVLPNANVKVAVEAIAPGVRPSPAASTIGAIAVHNSSDITFGNKTYIKGQVVIKNIYQDLPSATK
PPEPRTWQSSLKTIIKDKPLISFIVMVSLMIVLCAIVAVISILTASEDDLFPDPRPLRLVTRTEWLAQPPR
EELTDLKLPVNNVIIAHTATEGCTTQAACRLRVRLIQEFHMDGKNYDDITYNFLIGGDGHIYEGRNWHKIG
ATIPGYNSRSITVAFVGEYNYGGKPTKKQIELLKYLLHFGAKERHLKEGYRIYASEQLDPTVGTGKWLIEA
LHSLPQFVDKEQNKDERPDHE

>**Seq. 9.3.24** PGRPLC3 (partial CDS)
VNAYATIRSSPHRHRAHPVRAHQNGKSVFRNYTRKKLLFLLDKACYHSNTLPRTRICNLCIPVGGGKDVSQ
HFREIKNSHTTNTNKRKLSEILASRVTVTVPVAAVNETRTMALARNLQLENDTSVNTGYHTKAQSMSASGG
DTIAGLNNSGTGTGKSDSAKGSSLSNISASSIKRTDNGGDDDDSSVFDSSSSDTECDDDSIKRAIDRIPGT
LAPGEARVLPNANVKVAVEAIAPGVRPSPAASTIGAIAVHNSSDITFGNKTYIKGQVVIKNIYQDLPSATK
PPEPRTWQSSLKTIIKDKPLISFIVMVSLMIVLCAIVAVISILTASEDDLFPDPRPLRLVTRTEWLAQPPR
EELTDLKLPVNNVIIAHTATEGCTTQTKCMYQVKLIQEFHSSPDSRNFSDIAYQFLVGGDGNAYEGRGWTK
QGAHTKGFNVDSICIAFIGTFIADPPPIAQLSAAQQLILLGMKENYLASNYSLYGHRQLAPFESPGKALFD
IIKTWPHWSNKLGSNHWVEPANDTTNRR

9.4 VECTOR SEQUENCES & MAPS

pETM-11

Promoter:	T7/lac
Selection:	Kanamycin
Tag:	N-His
Protease cleavage site:	TEV
Origin:	pBR322
host:	BL21 (DE3)
Source:	G. Stier
note:	contains MAD insert in modified pET-24d

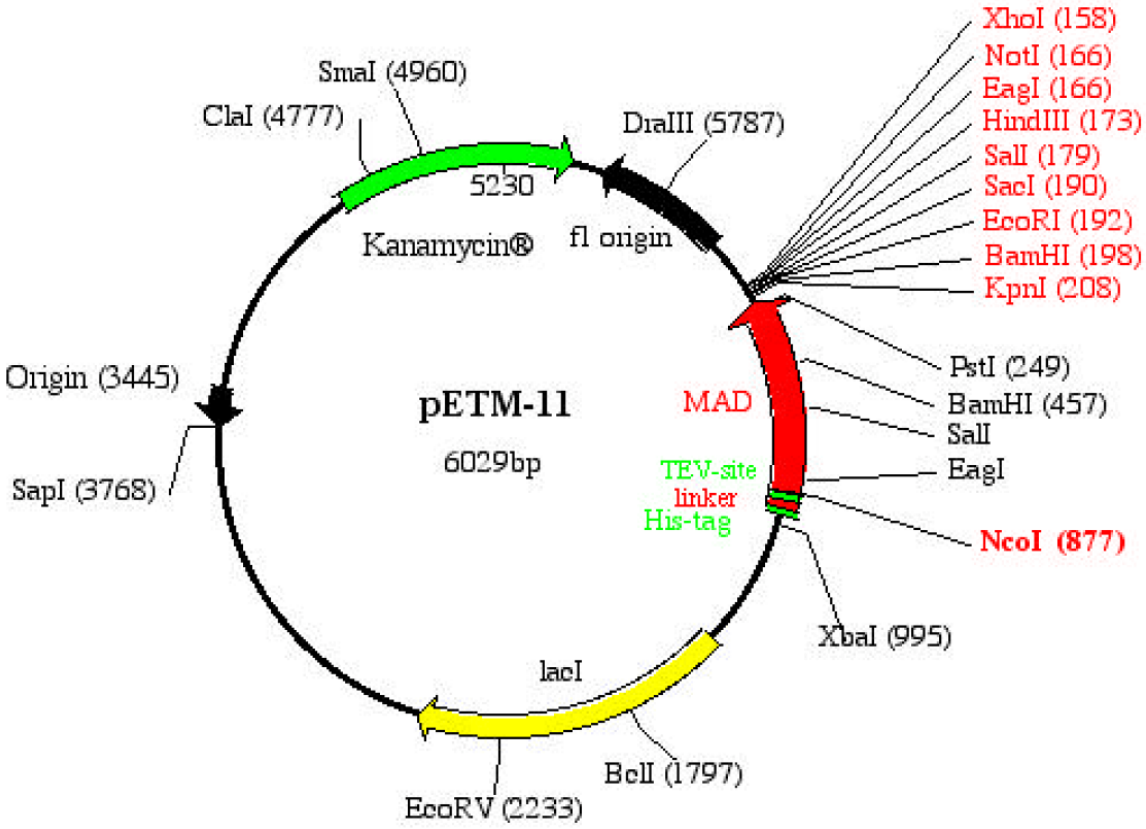


Fig. 9.4.1 The pETM-11 vector map.

```
    T7 promoter  -->          lac operator    XbaI
GAAATTAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCCCTCTAGAAAT
CTTTAATTATGCTGAGTGATATCCCCTTAACACTCGCCTATTGTTAAGGGGAGATCTTTA

                      rbs                          His.Tag
AATTTTGATTTAACTTTAAGAAGGAGATATACCATGAAACATCACCATCACCATCACCCC
TTAAAACTAAATTGAAATTCTTCCTCTATATGGTACTTTGTAGTGGTAGTGGTAGTGGGG
                                   METLysHisHisHisHisHisHisPro

                                 TEV site      NcoI
ATGAGCGATTACGACATCCCCACTACTGAGAATCTTTATTTTCAG GGCGCCATGGCGGCG
TACTCGCTAATGCTGTAGGGGTGATGACTCTTAGAAATAAAAGTC CCGCGGTACCGCCGC
MetSerAspTyrAspIleProThrThrGluAsnLeuTyrPheGln|GlyAlaMETAlaAla


GCGGTTCGGATGAAC..612bp..GACAGTCACAAGGCGTGTCTTGGTCTCTAACTAGTG
CGCCAAGCCTACTTG...MAD...CTGTCAGTGTTCCGCACAGAACCAGAGATTGATCAC
AlaValArgMetAsn..204aa..AspSerHisLysAlaCysLeuGlyLeu***

                                        NotI
      BamHI        SacI                 EagI
KpnI         EcoRI         SalI HindIII        XhoI   His-tag
GTACCGGATCCGAATTCGAGCTCCGTCGACAAGCTTGCGGCCGCACTCGAGCACCACCAC
CATGGCCTAGGCTTAAGCTCGAGGCAGCTGTTCGAACGCCGGCGTGAGCTCGTGGTGGTG
                                              HisHisHis
```

Single Cutters Listed by Site Order

80	Bpu1102I	204	Acc65I	1467	ApaBI	3768	SapI
80	EspI	204	Asp718I	1783	MluI	4057	BsiI
158	XhoI	208	KpnI	1797	BclI	4743	SpoI
160	SciI	210	SpeI	1964	BstEII	4743	NruI
166	NotI	212	Eco31I	1994	ApaI	4777	ClaI
173	HindIII	249	PstI	2194	BssHII	4958	XmaI
188	Ecl136II	877	NcoI	2233	EcoRV	4960	SmaI
188	EcoICRI	995	XbaI	2289	HpaI	5086	PvuI
190	SstI	1061	BglII	2865	MstI	5086	XorII
190	SacI	1102	SgrAI	3629	Tth111I	5787	DraIII
192	EcoRI						

Non Cutting Enzymes

AatII	AflII	AgeI	AhaIII	AscI	AsuII	AvrII
BalI	Bsp1407I	BspMI	Bsu36I	Csp45I	CspI	CvnI
DraI	Eam1105I	Eco72I	FseI	I-PpoI	MfeI	Mlu113I
MscI	MstII	NdeI	NheI	PacI	PinAI	PmaCI
PmeI	RleAI	SacII	SauI	ScaI	SfiI	SnaBI
SplI	SrfI	SstII	StuI	SunI	SwaI	

>**pETM-11;** Full length: 6029 bp

```
ATCCGGATATAGTTCCTCCTTTCAGCAAAAAACCCCTCAAGACCCGTTTAGAGGCCCCAAGGGGTTATGCT
AGTTATTGCTCAGCGGTGGCAGCAGCCAACTCAGCTTCCTTTCGGGCTTTGTTAGCAGCCGGATCTCAGTG
GTGGTGGTGGTGGTGCTCGAGTGCGGCCGCAAGCTTGTCGACGGAGCTCGAATTCGGATCCGGTACCACTA
```

GTTAGAGACCAAGACACGCCTTGTGACTGTCCTGCAGCTTTATTCTCTTGATGCTGGTGCTGGAATAGCCC
TCATCACTGCCGAGGCTCTGCATGCTGCCCCGCTCGTCAGAGTCGCTCACACTGCTGCTGCTCCAGTCCAG
ATCACCTGTGAGATAGTCCGTGCTCTCCACGTCAACGTCGATTTCTTCCCTGTCGGAGTCGGAGCGCTCCG
AGGAGACGGTGGAGCCGATGCTGTCCATCCGGATCCTCTCAATGCCCAGCTTCTCCAGCTGCCTCTTCAGG
TGTCGCTGCTCTCGCTGAAGCTGGTCGATTTGGTGAACGGCTTTTCTGTCACAATCTTCAAGTTTCTTTAT
GTGCAATTTGGCTTTTGTTAATAAACTCAACGTAGTGTGTCGACTTGATTCGGGTCCCAGTGGCACCAGCC
CCTTCAACTTCTCCAGGCACAAGCGAAGATGAGCCCGTCTATTCTTCTCCATTTCATTGTGAGTTGATCTG
CTACTGCTGTTATTCTTTTTGGATTTGTTCCTCCGTTTTAAGGCATCTCTGTCCTTGTTTTTGTATGGTAA
CATGGAGGCATAACCATGTTCAGCTTCTCTCTCCCGCCGCTCCAGATAGTCGGCCGCCTCCAGCAGCATCT
GGATGTTCATCCGAACCGCCGCCGCCATGGCGCCCTGAAAATAAAGATTCTCAGTAGTGGGGATGTCGTAA
TCGCTCATGGGGTGATGGTGATGGTGATGTTTCATGGTATATCTCCTTCTTAAAGTTAAATCAAAATTATT
TCTAGAGGGGAATTGTTATCCGCTCACAATTCCCCTATAGTGAGTCGTATTAATTTCGCGGGATCGAGATC
TCGATCCTCTACGCCGGACGCATCGTGGCCGGCATCACCGGCGCCACAGGTGCGGTTGCTGGCGCCTATAT
CGCCGACATCACCGATGGGGAAGATCGGGCTCGCCACTTCGGGCTCATGAGCGCTTGTTTCGGCGTGGGTA
TGGTGGCAGGCCCCGTGGCCGGGGGACTGTTGGGCGCCATCTCCTTGCATGCACCATTCCTTGCGGCGGCG
GTGCTCAACGGCCTCAACCTACTACTGGGCTGCTTCCTAATGCAGGAGTCGCATAAGGGAGAGCGTCGAGA
TCCCGGACACCATCGAATGGCGCAAAACCTTTCGCGGTATGGCATGATAGCGCCCGGAAGAGAGTCAATTC
AGGGTGGTGAATGTGAAACCAGTAACGTTATACGATGTCGCAGAGTATGCCGGTGTCTCTTATCAGACCGT
TTCCCGCGTGGTGAACCAGGCCAGCCACGTTTCTGCGAAAACGCGGGAAAAAGTGGAAGCGGCGATGGCGG
AGCTGAATTACATTCCCAACCGCGTGGCACAACAACTGGCGGGCAAACAGTCGTTGCTGATTGGCGTTGCC
ACCTCCAGTCTGGCCCTGCACGCGCCGTCGCAAATTGTCGCGGCGATTAAATCTCGCGCCGATCAACTGGG
TGCCAGCGTGGTGGTGTCGATGGTAGAACGAAGCGGCGTCGAAGCCTGTAAAGCGGCGGTGCACAATCTTC
TCGCGCAACGCGTCAGTGGGCTGATCATTAACTATCCGCTGGATGACCAGGATGCCATTGCTGTGGAAGCT
GCCTGCACTAATGTTCCGGCGTTATTTCTTGATGTCTCTGACCAGACACCCATCAACAGTATTATTTTCTC
CCATGAAGACGGTACGCGACTGGGCGTGGAGCATCTGGTCGCATTGGGTCACCAGCAAATCGCGCTGTTAG
CGGGCCCATTAAGTTCTGTCTCGGCGCGTCTGCGTCTGGCTGGCTGGCATAAATATCTCACTCGCAATCAA
ATTCAGCCGATAGCGGAACGGGAAGGCGACTGGAGTGCCATGTCCGGTTTTCAACAAACCATGCAAATGCT
GAATGAGGGCATCGTTCCCACTGCGATGCTGGTTGCCAACGATCAGATGGCGCTGGGCGCAATGCGCGCCA
TTACCGAGTCCGGGCTGCGCGTTGGTGCGGATATCTCGGTAGTGGGATACGACGATACCGAAGACAGCTCA
TGTTATATCCCGCCGTTAACCACCATCAAACAGGATTTTCGCCTGCTGGGGCAAACCAGCGTGGACCGCTT
GCTGCAACTCTCTCAGGGCCAGGCGGTGAAGGGCAATCAGCTGTTGCCCGTCTCACTGGTGAAAAGAAAAA
CCACCCTGGCGCCCAATACGCAAACCGCCTCTCCCCGCGCGTTGGCCGATTCATTAATGCAGCTGGCACGA
CAGGTTTCCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATTAATGTAAGTTAGCTCACTCATTAGGCAC
CGGGATCTCGACCGATGCCCTTGAGAGCCTTCAACCCAGTCAGCTCCTTCCGGTGGGCGCGGGGCATGACT
ATCGTCGCCGCACTTATGACTGTCTTCTTTATCATGCAACTCGTAGGACAGGTGCCGGCAGCGCTCTGGGT
CATTTTCGGCGAGGACCGCTTTCGCTGGAGCGCGACGATGATCGGCCTGTCGCTTGCGGTATTCGGAATCT
TGCACGCCCTCGCTCAAGCCTTCGTCACTGGTCCCGCCACCAAACGTTTCGGCGAGAAGCAGGCCATTATC
GCCGGCATGGCGGCCCCACGGGTGCGCATGATCGTGCTCCTGTCGTTGAGGACCCGGCTAGGCTGGCGGGG
TTGCCTTACTGGTTAGCAGAATGAATCACCGATACGCGAGCGAACGTGAAGCGACTGCTGCTGCAAAACGT
CTGCGACCTGAGCAACAACATGAATGGTCTTCGGTTTCCGTGTTTCGTAAAGTCTGGAAACGCGGAAGTCA
GCGCCCTGCACCATTATGTTCCGGATCTGCATCGCAGGATGCTGCTGGCTACCCTGTGGAACACCTACATC
TGTATTAACGAAGCGCTGGCATTGACCCTGAGTGATTTTTCTCTGGTCCCGCCGCATCCATACCGCCAGTT
GTTTACCCTCACAACGTTCCAGTAACCGGGCATGTTCATCATCAGTAACCCGTATCGTGAGCATCCTCTCT
CGTTTCATCGGTATCATTACCCCCATGAACAGAAATCCCCCTTACACGGAGGCATCAGTGACCAAACAGGA
AAAAACCGCCCTTAACATGGCCCGCTTTATCAGAAGCCAGACATTAACGCTTCTGGAGAAACTCAACGAGC
TGGACGCGGATGAACAGGCAGACATCTGTGAATCGCTTCACGACCACGCTGATGAGCTTTACCGCAGCTGC
CTCGCGCGTTTCGGTGATGACGGTGAAAACCTCTGACACATGCAGCTCCCGGAGACGGTCACAGCTTGTCT
GTAAGCGGATGCCGGGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTGTTGGCGGGTGTCGGGGCGCAG
CCATGACCCAGTCACGTAGCGATAGCGGAGTGTATACTGGCTTAACTATGCGGCATCAGAGCAGATTGTAC
TGAGAGTGCACCATATATGCGGTGTGAAATACCGCACAGATGCGTAAGGAGAAAATACCGCATCAGGCGCT
CTTCCGCTTCCTCGCTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCA
AAGGCGGTAATACGGTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCA
AAAGGCCAGGAACCGTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATC
ACAAAAATCGACGCTCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCT

GGAAGCTCCCTCGTGCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTC
GGGAAGCGTGGCGCTTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGC
TGGGCTGTGTGCACGAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCC
AACCCGGTAAGACACGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGT
AGGCGGTGCTACAGAGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCT
GCGCTCTGCTGAAGCCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCT
GGTAGCGGTGGTTTTTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTT
GATCTTTTCTACGGGGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAACAATA
AAACTGTCTGCTTACATAAACAGTAATACAAGGGGTGTTATGAGCCATATTCAACGGGAAACGTCTTGCTC
TAGGCCGCGATTAAATTCCAACATGGATGCTGATTTATATGGGTATAAATGGGCTCGCGATAATGTCGGGC
AATCAGGTGCGACAATCTATCGATTGTATGGGAAGCCCGATGCGCCAGAGTTGTTTCTGAAACATGGCAAA
GGTAGCGTTGCCAATGATGTTACAGATGAGATGGTCAGACTAAACTGGCTGACGGAATTTATGCCTCTTCC
GACCATCAAGCATTTTATCCGTACTCCTGATGATGCATGGTTACTCACCACTGCGATCCCCGGGAAAACAG
CATTCCAGGTATTAGAAGAATATCCTGATTCAGGTGAAAATATTGTTGATGCGCTGGCAGTGTTCCTGCGC
CGGTTGCATTCGATTCCTGTTTGTAATTGTCCTTTTAACAGCGATCGCGTATTTCGTCTCGCTCAGGCGCA
ATCACGAATGAATAACGGTTTGGTTGATGCGAGTGATTTTGATGACGAGCGTAATGGCTGGCCTGTTGAAC
AAGTCTGGAAAGAAATGCATAAACTTTTGCCATTCTCACCGGATTCAGTCGTCACTCATGGTGATTTCTCA
CTTGATAACCTTATTTTTGACGAGGGGAAATTAATAGGTTGTATTGATGTTGGACGAGTCGGAATCGCAGA
CCGATACCAGGATCTTGCCATCCTATGGAACTGCCTCGGTGAGTTTTCTCCTTCATTACAGAAACGGCTTT
TTCAAAAATATGGTATTGATAATCCTGATATGAATAAATTGCAGTTTCATTTGATGCTCGATGAGTTTTTC
TAAGAATTAATTCATGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCA
CATTTCCCCGAAAAGTGCCACCTGAAATTGTAAACGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGT
TAAATCAGCTCATTTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGA
GATAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTCCAACGTCAAAG
GGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACCCTAATCAAGTTTTTTGGGGTCG
AGGTGCCGTAAAGCACTAAATCGGAACCCTAAAGGGAGCCCCCGATTTAGAGCTTGACGGGGAAAGCCGGC
GAACGTGGCGAGAAAGGAAGGGAAGAAAGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCA
CGCTGCGCGTAACCACCACACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTCCCATTCGCCA

pETM-13

Promoter:	T7/lac
Selection:	Kanamycin
Tag:	none
Protease cleavage site:	none
Origin:	pBR322
host:	BL21 (DE3)
Source:	G. Stier
notes:	contains ABD insert in modified pET-24d

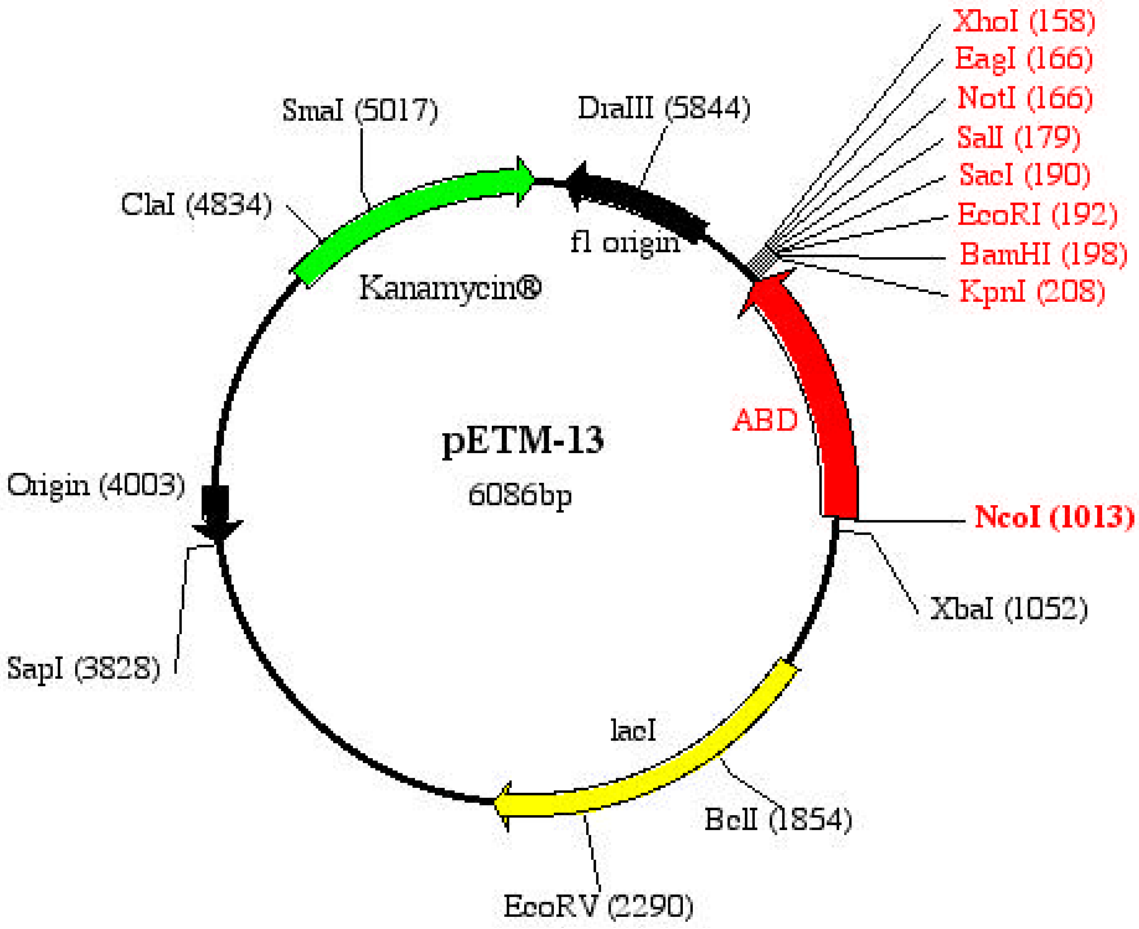


Fig. 9.4.2 The pETM-13 vector map.

```
     T7 promoter  -->  ____________ lac operator     XbaI
CGAAATTAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCCCTCTAGAAA
GCTTTAATTATGCTGAGTGATATCCCCTTAACACTCGCCTATTGTTAAGGGGAGATCTTT

                        ____rbs       NcoI
TAATTTTGTTTAACTTTAAGAAGGAGATATACCATGGGCCAGATAGAGCCCGGCGTGCAG
ATTAAAACAAATTGAAATTCTTCCTCTATATGGTACCCGGTCTATCTCGGGCCGCACGTC
                                    METGlyGlnIleGluProGlyValGln


                                                        KpnI   BamHI
TACAACTAC...744bp...CAGGCCGAGACAGCGGCTAACAGGTGAGGTACCGGATCC
ATGTTGATG....ABD....GTCCGGCTCTGTCGCCGATTGTCCACTCCATGGCCTAGG
TyrAsnTyr...248aa...GlnAlaGluThrAlaAlaAsnArg***


       SacI                 NotI
 EcoRI         SalI          EagI     XhoI _________C-His-tag
GAATTCGAGCTCCGTCGACAAGCTTGCGGCCGCACTCGAGCACCACCACCACCACCAC
CTTAAGCTCGAGGCAGCTGTTCGAACGCCGGCGTGAGCTCGTGGTGGTGGTGGTGGTG
                                          HisHisHisHisHisHis
```

Single Cutters Listed by Site Order

80	Bpu1102I	192	EcoRI	1052	XbaI	2922	MstI
80	EspI	198	BamHI	1118	BglII	3686	Tth111I
158	XhoI	204	Asp718I	1159	SgrAI	3825	SapI
160	SciI	204	Acc65I	1315	SphI	4114	BsiI
166	Eco52I	208	KpnI	1524	ApaBI	4357	AlwNI
166	XmaIII	272	SplI	1840	MluI	4800	NruI
166	EagI	272	SunI	1854	BclI	4800	SpoI
166	NotI	859	Bsu36I	2021	BstEII	4834	ClaI
179	SalI	859	SauI	2051	ApaI	5015	XmaI
188	EcoICRI	859	CvnI	2251	BssHII	5017	SmaI
188	Ecl136II	859	MstII	2290	EcoRV	5143	PvuI
190	SstI	917	BspMI	2346	HpaI	5143	XorII
190	SacI	1013	NcoI	2904	BglI	5844	DraIII

Non Cutting Enzymes

AatII	AflII	AgeI	AhaIII	AscI	AsuII	AvrII
BalI	Bsp1407I	Csp45I	CspI	DraI	Eam1105I	Eco31I
Eco72I	FseI	I-PpoI	MfeI	Mlu113I	MscI	NdeI
NheI	PacI	PinAI	PmaCI	PmeI	PstI	RleAI
SacII	ScaI	SfiI	SnaBI	SpeI	SrfI	SstII
StuI	SwaI					

>**pETM-13**; Full length: 6086 bp

```
ATCCGGATATAGTTCCTCCTTTCAGCAAAAAACCCCTCAAGACCCGTTTAGAGGCCCCAAGGGGTTATGCT
AGTTATTGCTCAGCGGTGGCAGCAGCCAACTCAGCTTCCTTTCGGGCTTTGTTAGCAGCCGGATCTCAGTG
```

GTGGTGGTGGTGGTGCTCGAGTGCGGCCGCAAGCTTGTCGACGGAGCTCGAATTCGGATCCGGTACCTCAC
CTGTTAGCCGCTGTCTCGGCCTGCTCCGCGCCCGCAAAAGCGTGGTAGAAGCAAGAGACGTACGTCATGAT
GGCTCTTTCATCGGGTTTAGGGGTGTTCACGATGTCTTCAGCATCCAACATTTTAGGAATATCCAGGTGCT
TCTCAGCGATTTCCATAGCCAGGTTAATATTTCCTATGGGGTCATCCTTGTTAAGCTTTGAGTAGTCAATG
AGGTCAGGCCGGTGTCGGTGGATGAGGGCACAGAGTCCAAGGCCATCTTTCCAGCTAGTATGGAAGTTCTG
AATGTTCACATTTCTATAAGGAGCAGTTTTCCTCTGACACCAAAGCAGCAGACCTTCTTTGGCAGATGTTT
CTTCAACCGAAATATCCTGAATAGCAAAGCGAAGGATGATGGTCCAGATCATACCCAGGGTCATTTTCACA
TTGCCATCAACAATTTCTTCAGCGCCGATGGACACCAGTTTCACCCCTTTGCTGGCTATGTAATCCAAAGC
TTTGTTGACATTAGCAATTTTGTGGAACCGCATTTTTCCCCGGTCAGGTTTGGGCAGCCTTTCCCCTGAGA
TGACTTCCAAAAGCAGCATGAGCTTNAGGCCATTCCTGAAGTCTTCCTCGATGTTCTCAATCTGGGTGCCG
GCTTTCCTTAGGTGGGAGTTACACCAGGCAGTGAAGGTCTTCCTCTGCTGCTTCTTCCAGGCTGGGTCCAG
GAGCAGGTCGCGGTCCCACTCCTCCTCCTGGATCATGTACTCATCCTCGTCGTACACGTAGTTGTACTGCA
CGCCGGGCTCTATCTGGCCCATGGTATATCTCCTTCTTAAAGTTAAACAAAATTATTTCTAGAGGGGAATT
GTTATCCGCTCACAATTCCCCTATAGTGAGTCGTATTAATTTCGCGGGATCGAGATCTCGATCCTCTACGC
CGGACGCATCGTGGCCGGCATCACCGGCGCCACAGGTGCGGTTGCTGGCGCCTATATCGCCGACATCACCG
ATGGGGAAGATCGGGCTCGCCACTTCGGGCTCATGAGCGCTTGTTTCGGCGTGGGTATGGTGGCAGGCCCC
GTGGCCGGGGGACTGTTGGGCGCCATCTCCTTGCATGCACCATTCCTTGCGGCGGCGGTGCTCAACGGCCT
CAACCTACTACTGGGCTGCTTCCTAATGCAGGAGTCGCATAAGGGAGAGCGTCGAGATCCCGGACACCATC
GAATGGCGCAAAACCTTTCGCGGTATGGCATGATAGCGCCCGGAAGAGAGTCAATTCAGGGTGGTGAATGT
GAAACCAGTAACGTTATACGATGTCGCAGAGTATGCCGGTGTCTCTTATCAGACCGTTTCCCGCGTGGTGA
ACCAGGCCAGCCACGTTTCTGCGAAAACGCGGGAAAAAGTGGAAGCGGCGATGGCGGAGCTGAATTACATT
CCCAACCGCGTGGCACAACAACTGGCGGGCAAACAGTCGTTGCTGATTGGCGTTGCCACCTCCAGTCTGGC
CCTGCACGCGCCGTCGCAAATTGTCGCGGCGATTAAATCTCGCGCCGATCAACTGGGTGCCAGCGTGGTGG
TGTCGATGGTAGAACGAAGCGGCGTCGAAGCCTGTAAAGCGGCGGTGCACAATCTTCTCGCGCAACGCGTC
AGTGGGCTGATCATTAACTATCCGCTGGATGACCAGGATGCCATTGCTGTGGAAGCTGCCTGCACTAATGT
TCCGGCGTTATTTCTTGATGTCTCTGACCAGACACCCATCAACAGTATTATTTTCTCCCATGAAGACGGTA
CGCGACTGGGCGTGGAGCATCTGGTCGCATTGGGTCACCAGCAAATCGCGCTGTTAGCGGGCCCATTAAGT
TCTGTCTCGGCGCGTCTGCGTCTGGCTGGCTGGCATAAATATCTCACTCGCAATCAAATTCAGCCGATAGC
GGAACGGGAAGGCGACTGGAGTGCCATGTCCGGTTTTCAACAAACCATGCAAATGCTGAATGAGGGCATCG
TTCCCACTGCGATGCTGGTTGCCAACGATCAGATGGCGCTGGGCGCAATGCGCGCCATTACCGAGTCCGGG
CTGCGCGTTGGTGCGGATATCTCGGTAGTGGGATACGACGATACCGAAGACAGCTCATGTTATATCCCGCC
GTTAACCACCATCAAACAGGATTTTCGCCTGCTGGGGCAAACCAGCGTGGACCGCTTGCTGCAACTCTCTC
AGGGCCAGGCGGTGAAGGGCAATCAGCTGTTGCCCGTCTCACTGGTGAAAAGAAAAACCACCCTGGCGCCC
AATACGCAAACCGCCTCTCCCCGCGCGTTGGCCGATTCATTAATGCAGCTGGCACGACAGGTTTCCCGACT
GGAAAGCGGGCAGTGAGCGCAACGCAATTAATGTAAGTTAGCTCACTCATTAGGCACCGGGATCTCGACCG
ATGCCCTTGAGAGCCTTCAACCCAGTCAGCTCCTTCCGGTGGGCGCGGGGCATGACTATCGTCGCCGCACT
TATGACTGTCTTCTTTATCATGCAACTCGTAGGACAGGTGCCGGCAGCGCTCTGGGTCATTTTCGGCGAGG
ACCGCTTTCGCTGGAGCGCGACGATGATCGGCCTGTCGCTTGCGGTATTCGGAATCTTGCACGCCCTCGCT
CAAGCCTTCGTCACTGGTCCCGCCACCAAACGTTTCGGCGAGAAGCAGGCCATTATCGCCGGCATGGCGGC
CCCACGGGTGCGCATGATCGTGCTCCTGTCGTTGAGGACCCGGCTAGGCTGGCGGGGTTGCCTTACTGGTT
AGCAGAATGAATCACCGATACGCGAGCGAACGTGAAGCGACTGCTGCTGCAAAACGTCTGCGACCTGAGCA
ACAACATGAATGGTCTTCGGTTTCCGTGTTTCGTAAAGTCTGGAAACGCGGAAGTCAGCGCCCTGCACCAT
TATGTTCCGGATCTGCATCGCAGGATGCTGCTGGCTACCCTGTGGAACACCTACATCTGTATTAACGAAGC
GCTGGCATTGACCCTGAGTGATTTTTCTCTGGTCCCGCCGCATCCATACCGCCAGTTGTTTACCCTCACAA
CGTTCCAGTAACCGGGCATGTTCATCATCAGTAACCCGTATCGTGAGCATCCTCTCTCGTTTCATCGGTAT
CATTACCCCCATGAACAGAAATCCCCCTTACACGGAGGCATCAGTGACCAAACAGGAAAAAACCGCCCTTA
ACATGGCCCGCTTTATCAGAAGCCAGACATTAACGCTTCTGGAGAAACTCAACGAGCTGGACGCGGATGAA
CAGGCAGACATCTGTGAATCGCTTCACGACCACGCTGATGAGCTTTACCGCAGCTGCCTCGCGCGTTTCGG
TGATGACGGTGAAAACCTCTGACACATGCAGCTCCCGGAGACGGTCACAGCTTGTCTGTAAGCGGATGCCG
GGAGCAGACAAGCCCGTCAGGGCGCGTCAGCGGGTGTTGGCGGGTGTCGGGGCGCAGCCATGACCCAGTCA
CGTAGCGATAGCGGAGTGTATACTGGCTTAACTATGCGGCATCAGAGCAGATTGTACTGAGAGTGCACCAT
ATATGCGGTGTGAAATACCGCACAGATGCGTAAGGAGAAAATACCGCATCAGGCGCTCTTCCGCTTCCTCG
CTCACTGACTCGCTGCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACG
GTTATCCACAGAATCAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACC

GTAAAAAGGCCGCGTTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGC
TCAAGTCAGAGGTGGCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGT
GCGCTCTCCTGTTCCGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGC
TTTCTCATAGCTCACGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCAC
GAACCCCCCGTTCAGCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACA
CGACTTATCGCCACTGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAG
AGTTCTTGAAGTGGTGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAG
CCAGTTACCTTCGGAAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTT
TTTTGTTTGCAAGCAGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGG
GGTCTGACGCTCAGTGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAACAATAAAACTGTCTGCTTA
CATAAACAGTAATACAAGGGGTGTTATGAGCCATATTCAACGGGAAACGTCTTGCTCTAGGCCGCGATTAA
ATTCCAACATGGATGCTGATTTATATGGGTATAAATGGGCTCGCGATAATGTCGGGCAATCAGGTGCGACA
ATCTATCGATTGTATGGGAAGCCCGATGCGCCAGAGTTGTTTCTGAAACATGGCAAAGGTAGCGTTGCCAA
TGATGTTACAGATGAGATGGTCAGACTAAACTGGCTGACGGAATTTATGCCTCTTCCGACCATCAAGCATT
TTATCCGTACTCCTGATGATGCATGGTTACTCACCACTGCGATCCCCGGGAAAACAGCATTCCAGGTATTA
GAAGAATATCCTGATTCAGGTGAAAATATTGTTGATGCGCTGGCAGTGTTCCTGCGCCGGTTGCATTCGAT
TCCTGTTTGTAATTGTCCTTTTAACAGCGATCGCGTATTTCGTCTCGCTCAGGCGCAATCACGAATGAATA
ACGGTTTGGTTGATGCGAGTGATTTTGATGACGAGCGTAATGGCTGGCCTGTTGAACAAGTCTGGAAAGAA
ATGCATAAACTTTTGCCATTCTCACCGGATTCAGTCGTCACTCATGGTGATTTCTCACTTGATAACCTTAT
TTTTGACGAGGGGAAATTAATAGGTTGTATTGATGTTGGACGAGTCGGAATCGCAGACCGATACCAGGATC
TTGCCATCCTATGGAACTGCCTCGGTGAGTTTTCTCCTTCATTACAGAAACGGCTTTTTCAAAAATATGGT
ATTGATAATCCTGATATGAATAAATTGCAGTTTCATTTGATGCTCGATGAGTTTTTCTAAGAATTAATTCA
TGAGCGGATACATATTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAA
GTGCCACCTGAAATTGTAAACGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTCATT
TTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGAGATAGGGTTGAGTG
TTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGGACTCCAACGTCAAAGGGCGAAAAACCGTC
TATCAGGGCGATGGCCCACTACGTGAACCATCACCCTAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGC
ACTAAATCGGAACCCTAAAGGGAGCCCCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAA
AGGAAGGGAAGAAAGCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTAACC
ACCACACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTCCCATTCGCCA

pETM-60

Promoter:	T7/lac
Selection:	Kanamycin
Tag:	N-NusA, N-His
Protease cleavage site:	TEV
Origin:	pBR322
host:	BL21 (DE3)
Source:	G. Stier
notes:	contains ABD insert in modified pET-24d

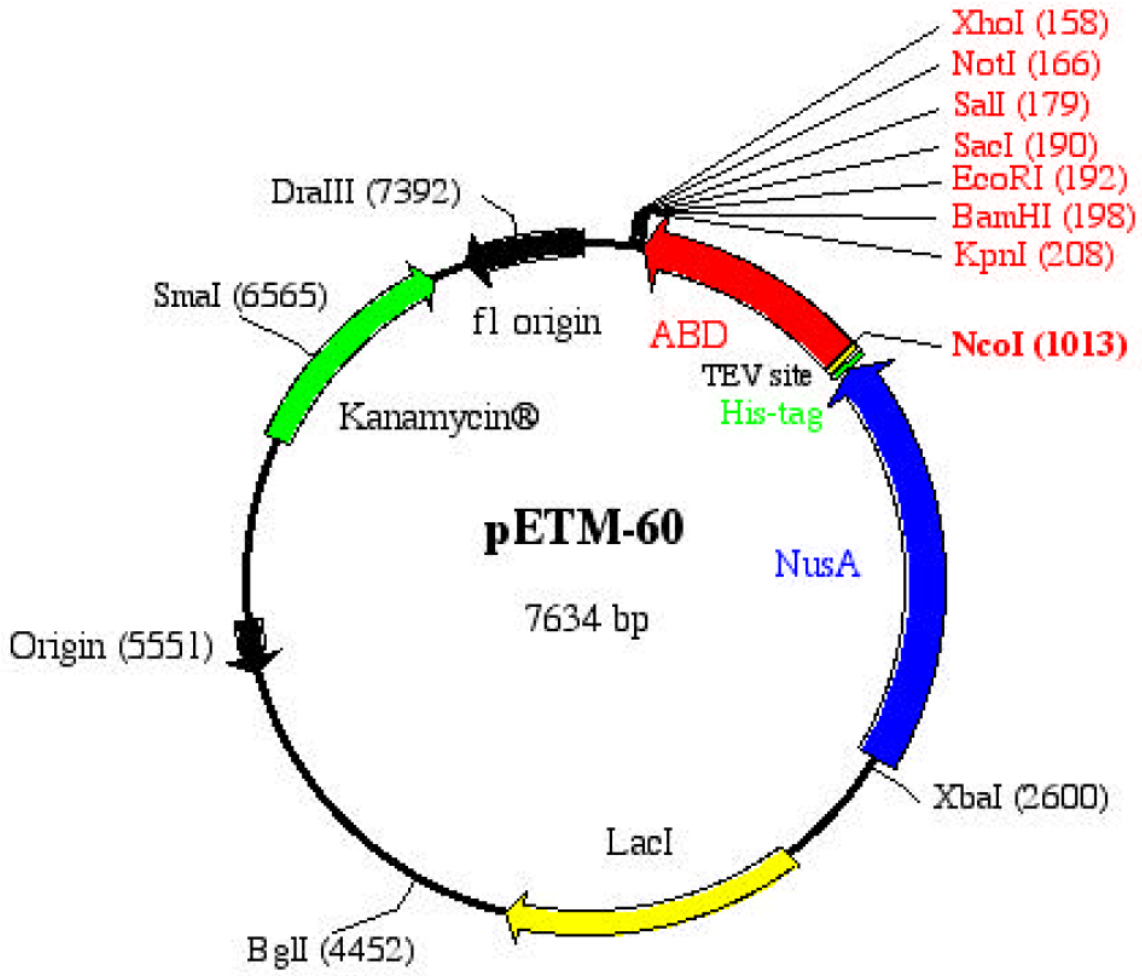


Fig. 9.4.3 The pETM-60 vector map.

```
     T7 promoter  -->      lac operator      XbaI
CGAAATTAATACGACTCACTATAGGGGAATTGTGAGCGGATAACAATTCCCCTCTAGAAA
GCTTTAATTATGCTGAGTGATATCCCCTTAACACTCGAATATTGTTAAGGGGAGATCTTT

                    RBS
TAATTTTGTTTAACTTTAAGAAGGAGATATACCATGAAAGAAATTTTGGCTGTAGTTGAA
ATTAAAACAAATTGAAATTCTTCCTCTATATGGTACTTTCTTTAAAACCGACATCAACTT
                                  MetLysGluIleLeuAlaValValGlu

                                                 His-tag
GCC...1446bp...GAAGCGACTAGTGGTTCTGGTCATCACCATCACCATCACTCCGCG
CGG....NusA....CTTCGCTGATCACCAAGACCAGTAGTGGTAGTGGTAGTGAGGCGC
Ala....482aa...GluAlaThrSerGlySerGlyHisHisHisHisHisHisSerAla

          TEV-site  NcoI
GGTGAGAATCTTTATTTTCAG GGCGCCATGGGCCAGATAGAG...777bp...GCGGCT
CCACTCTTAGAAATAAAAGTC CCGCGGTACCCGGTCTATCTC....ABD....CGCCGA
GlyGluAsnLeuTyrPheGln|GlyAlaMetGlyGlnIleGlu...259aa...AlaAla

              BamHI                SalI          NotI
         KpnI       EcoRI  SacI                 EagI   XhoI
AACAGGTGAGGTACCGGATCCGAATTCGAGCTCCGTCGACAAGCTTGCGGCCGCACTCGA
TTGTCCACTCCATGGCCTAGGCTTAAGCTCGAGGCAGCTGTTCGAACGCCGGCGTGAGCT
AsnArg***

          C-His-tag
GCACCACCACCACCACCACTGAGATCCGGCTGCTAACAAAGCCCGAAAGGAAGCTGAGTT
CGTGGTGGTGGTGGTGGTGACTCTAGGCCGACGATTGTTTCGGGCTTTCCTTCGACTCAA
 HisHisHisHisHisHis***
```

Single Cutters Listed by Site Order

158	XhoI	204	Asp718I	1046	SacII	3388	MluI
160	SciI	204	Acc65I	1046	SstII	3569	BstEII
166	XmaIII	208	KpnI	1077	SpeI	3599	ApaI
166	EagI	272	SplI	1324	BalI	3894	HpaI
166	NotI	272	SunI	1324	MscI	4452	BglI
166	Eco52I	804	AflII	1377	StuI	4470	MstI
179	SalI	859	Bsu36I	1756	RleAI	5234	Tth111I
188	Ecl136II	859	CvnI	2437	AatII	5373	SapI
188	EcoICRI	859	MstII	2494	Csp45I	5905	AlwNI
190	SstI	859	SauI	2494	AsuII	6563	XmaI
190	SacI	1013	NcoI	2600	XbaI	6565	SmaI
192	EcoRI	1044	Mlu113I	2863	SphI	7392	DraIII
198	BamHI						

Non Cutting Enzymes

AgeI	AhaIII	AscI	AvrII	Bsp1407I	CspI	DraI
Eam1105I	Eco31I	Eco72I	FseI	I-PpoI	MfeI	NdeI
NheI	PacI	PinAI	PmaCI	PmeI	PstI	ScaI
SfiI	SnaBI	SrfI	SwaI			

>**pETM-60;** Full length: 7634 bp
ATCCGGATATAGTTCCTCCTTTCAGCAAAAAACCCCTCAAGACCCGTTTAGAGGCCCCAAGGGGTTATGCT
AGTTATTGCTCAGCGGTGGCAGCAGCCAACTCAGCTTCCTTTCGGGCTTTGTTAGCAGCCGGATCTCAGTG
GTGGTGGTGGTGGTGCTCGAGTGCGGCCGCAAGCTTGTCGACGGAGCTCGAATTCGGATCCGGTACCTCAC
CTGTTAGCCGCTGTCTCGGCCTGCTCCGCGCCCGCAAAAGCGTGGTAGAAGCAAGAGACGTACGTCATGAT
GGCTCTTTCATCGGGTTTAGGGGTGTTCACGATGTCTTCAGCATCCAACATTTTAGGAATATCCAGGTGCT
TCTCAGCGATTTCCATAGCCAGGTTAATATTTCCTATGGGGTCATCCTTGTTAAGCTTTGAGTAGTCAATG
AGGTCAGGCCGGTGTCGGTGGATGAGGGCACAGAGTCCAAGGCCATCTTTCCAGCTAGTATGGAAGTTCTG
AATGTTCACATTTCTATAAGGAGCAGTTTTCCTCTGACACCAAAGCAGCAGACCTTCTTTGGCAGATGTTT
CTTCAACCGAAATATCCTGAATAGCAAAGCGAAGGATGATGGTCCAGATCATACCCAGGGTCATTTTCACA
TTGCCATCAACAATTTCTTCAGCGCCGATGGACACCAGTTTCACCCCTTTGCTGGCTATGTAATCCAAAGC
TTTGTTGACATTAGCAATTTTGTGGAACCGCATTTTTCCCCGGTCAGGTTTGGGCAGCCTTTCCCCTGAGA
TGACTTCCAAAAGCAGCATGAGCTTAAGGCCATTCCTGAAGTCTTCCTCGATGTTCTCAATCTGGGTGCCG
GCTTTCCTTAGGTGGGAGTTACACCAGGCAGTGAAGGTCTTCCTCTGCTGCTTCTTCCAGGCTGGGTCCAG
GAGCAGGTCGCGGTCCCACTCCTCCTCCTGGATCATGTACTCATCCTCGTCGTACACGTAGTTGTACTGCA
CGCCGGGCTCTATCTGGCCCATGGCGCCCTGAAAATAAAGATTCTCACCCGCGGAGTGATGGTGATGGTGA
TGACCAGAACCACTAGTCGCTTCGTCACCGAACCAGCAAATATTACGGGCAGCCATAATCAGTGCTCCGGC
TTTTTCGTCGGTCAACCCTTCGATATCAGCCAGATCATCAATGCCCTGTTCGGCGAGATCTTCCAGCGTAC
AAACGCCACGGGCGGCCAGTTTGAATGCCAAATCACGATCTACCCCTTCAAGGTTCAGCAGATCGTCAGCC
GGTTTGTTATCACCGAGGCTTTCTTCCTGGGCCTGTGCAATGGTGGCCAGTGCATTTTTAGCACGCTCGCG
CAGTGCTTCAACGGTCGGCTCATCAAGGCCTTCGATTTCCAACAGCTCTTTCATCGGCACATAGGCCAATT
CTTCCAGCGTCGAGAAGCCTTCTTCTACCAGAACAGTCGCGAAGTCTTCGTCGATGTCGAGATATTTGGTG
AAGGTGTCGATCGCTGCGTGCGCTTCCGCCTGATGCTTAGCTTGCAGGTCGTCAACGGTCATCACGTTGAG
TTCCCAACCGCTCAGTTGCGAAGCCAGACGCACGTTCTGACCGTTACGGCCAATCGCCTGCGCCAGATTAC
CGGCTTCAACGGCGATGTCCATAGTGTGTTTATCTTCATCCACCACGATAGAAGCAACGTCTGCCGGTGCC
ATTGCGTTAATCACGAACTGCGCCGGGTTATCATCCCACAGGACGATATCGATACGCTCGCCACCCAGTTC
AGTAGACACCGCCTGAACACGCGCGCCACGCATACCTACGCAAGCACCTACCGGATCGATACGTTTATCGT
TGGTTTTCACCGCGATTTTCGCACGAGAACCCGGATCGCGAGCCGCTGCTTTAATTTCAATCACTTCTTCG
CCGATTTCTGGCACTTCAATACGGAACAGTTCGATCAGCATTTCCGGCTTGGAACGAGTGACGAACAGTTG
CGCGCCACGCGCTTCCGGGCGAACGGAATAGAGCACGCCACGAACGCGGTCGCCAGGGCGGAAGTTTTCAC
GCGGCAGCATATCTTCGCGCAGGATCACGGCTTCAGCGTTGTTGCCCAGATCCAGAGAGATGTTGTCGCGG
TTTACTTTTTTCACCACGCCGGTGATGATTTCACCTTCGTGTTCACGGAACTGATCAACCACCATCGCACG
TTCGGCTTCACGCACTTTCTGCACGATAACCTGTTTTGCCGTCTGGGTAGTGATACGGTCAAAGGTAACAG
ACTCAATCTGATCTTCAACGTAATCGCCCAGGTTCAGGCTTTCATCTTCATAACGTGCGGCTTCAAGGGTG
ATTTCCTTGGTCGGCTGGGTGACTTCATCAACAACTAACCAGCGACGGAAAGTGTCAAAATCACCGCTTTT
GCGATCGATCTGTACGCGGACGTCGATCTCTTGTTCATATTTTTTCTTTGTTGCTGTCGCCAGCGCGCTTT
CCAATGCTTCGAAAATCTTCTCGCGAGGTAGCGCCTTTTCATTGGATACGGCTTCAACTACAGCCAAAATT
TCTTTCATGGTATATCTCCTTCTTAAAGTTAAACAAAATTATTTCTAGAGGGGAATTGTTATCCGCTCACA
ATTCCCCTATAGTGAGTCGTATTAATTTCGCGGGATCGAGATCTCGATCCTCTACGCCGGACGCATCGTGG
CCGGCATCACCGGCGCCACAGGTGCGGTTGCTGGCGCCTATATCGCCGACATCACCGATGGGGAAGATCGG
GCTCGCCACTTCGGGCTCATGAGCGCTTGTTTCGGCGTGGGTATGGTGGCAGGCCCCGTGGCCGGGGGACT
GTTGGGCGCCATCTCCTTGCATGCACCATTCCTTGCGGCGGCGGTGCTCAACGGCCTCAACCTACTACTGG
GCTGCTTCCTAATGCAGGAGTCGCATAAGGGAGAGCGTCGAGATCCCGGACACCATCGAATGGCGCAAAAC
CTTTCGCGGTATGGCATGATAGCGCCCGGAAGAGAGTCAATTCAGGGTGGTGAATGTGAAACCAGTAACGT
TATACGATGTCGCAGAGTATGCCGGTGTCTCTTATCAGACCGTTTCCCGCGTGGTGAACCAGGCCAGCCAC
GTTTCTGCGAAAACGCGGGAAAAAGTGGAAGCGGCGATGGCGGAGCTGAATTACATTCCCAACCGCGTGGC
ACAACAACTGGCGGGCAAACAGTCGTTGCTGATTGGCGTTGCCACCTCCAGTCTGGCCCTGCACGCGCCGT
CGCAAATTGTCGCGGCGATTAAATCTCGCGCCGATCAACTGGGTGCCAGCGTGGTGGTGTCGATGGTAGAA
CGAAGCGGCGTCGAAGCCTGTAAAGCGGCGGTGCACAATCTTCTCGCGCAACGCGTCAGTGGGCTGATCAT
TAACTATCCGCTGGATGACCAGGATGCCATTGCTGTGGAAGCTGCCTGCACTAATGTTCCGGCGTTATTTC
TTGATGTCTCTGACCAGACACCCATCAACAGTATTATTTTCTCCCATGAAGACGGTACGCGACTGGGCGTG
GAGCATCTGGTCGCATTGGGTCACCAGCAAATCGCGCTGTTAGCGGGCCCATTAAGTTCTGTCTCGGCGCG
TCTGCGTCTGGCTGGCTGGCATAAATATCTCACTCGCAATCAAATTCAGCCGATAGCGGAACGGGAAGGCG
ACTGGAGTGCCATGTCCGGTTTTCAACAAACCATGCAAATGCTGAATGAGGGCATCGTTCCCACTGCGATG

CTGGTTGCCAACGATCAGATGGCGCTGGGCGCAATGCGCGCCATTACCGAGTCCGGGCTGCGCGTTGGTGC
GGATATCTCGGTAGTGGGATACGACGATACCGAAGACAGCTCATGTTATATCCCGCCGTTAACCACCATCA
AACAGGATTTTCGCCTGCTGGGGCAAACCAGCGTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGGCGGTG
AAGGGCAATCAGCTGTTGCCCGTCTCACTGGTGAAAAGAAAAACCACCCTGGCGCCCAATACGCAAACCGC
CTCTCCCCGCGCGTTGGCCGATTCATTAATGCAGCTGGCACGACAGGTTTCCCGACTGGAAAGCGGGCAGT
GAGCGCAACGCAATTAATGTAAGTTAGCTCACTCATTAGGCACCGGGATCTCGACCGATGCCCTTGAGAGC
CTTCAACCCAGTCAGCTCCTTCCGGTGGGCGCGGGGCATGACTATCGTCGCCGCACTTATGACTGTCTTCT
TTATCATGCAACTCGTAGGACAGGTGCCGGCAGCGCTCTGGGTCATTTTCGGCGAGGACCGCTTTCGCTGG
AGCGCGACGATGATCGGCCTGTCGCTTGCGGTATTCGGAATCTTGCACGCCCTCGCTCAAGCCTTCGTCAC
TGGTCCCGCCACCAAACGTTTCGGCGAGAAGCAGGCCATTATCGCCGGCATGGCGGCCCCACGGGTGCGCA
TGATCGTGCTCCTGTCGTTGAGGACCCGGCTAGGCTGGCGGGGTTGCCTTACTGGTTAGCAGAATGAATCA
CCGATACGCGAGCGAACGTGAAGCGACTGCTGCTGCAAAACGTCTGCGACCTGAGCAACAACATGAATGGT
CTTCGGTTTCCGTGTTTCGTAAAGTCTGGAAACGCGGAAGTCAGCGCCCTGCACCATTATGTTCCGGATCT
GCATCGCAGGATGCTGCTGGCTACCCTGTGGAACACCTACATCTGTATTAACGAAGCGCTGGCATTGACCC
TGAGTGATTTTTCTCTGGTCCCGCCGCATCCATACCGCCAGTTGTTTACCCTCACAACGTTCCAGTAACCG
GGCATGTTCATCATCAGTAACCCGTATCGTGAGCATCCTCTCTCGTTTCATCGGTATCATTACCCCCATGA
ACAGAAATCCCCCTTACACGGAGGCATCAGTGACCAAACAGGAAAAAACCGCCCTTAACATGGCCCGCTTT
ATCAGAAGCCAGACATTAACGCTTCTGGAGAAACTCAACGAGCTGGACGCGGATGAACAGGCAGACATCTG
TGAATCGCTTCACGACCACGCTGATGAGCTTTACCGCAGCTGCCTCGCGCGTTTCGGTGATGACGGTGAAA
ACCTCTGACACATGCAGCTCCCGGAGACGGTCACAGCTTGTCTGTAAGCGGATGCCGGGAGCAGACAAGCC
CGTCAGGGCGCGTCAGCGGGTGTTGGCGGGTGTCGGGGCGCAGCCATGACCCAGTCACGTAGCGATAGCGG
AGTGTATACTGGCTTAACTATGCGGCATCAGAGCAGATTGTACTGAGAGTGCACCATATATGCGGTGTGAA
ATACCGCACAGATGCGTAAGGAGAAAATACCGCATCAGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCT
GCGCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAAT
CAGGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCG
TTGCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTG
GCGAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTC
CGACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCA
CGCTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCA
GCCCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCAC
TGGCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGG
TGGCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGG
AAAAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGC
AGCAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAG
TGGAACGAAAACTCACGTTAAGGGATTTTGGTCATGAACAATAAAACTGTCTGCTTACATAAACAGTAATA
CAAGGGGTGTTATGAGCCATATTCAACGGGAAACGTCTTGCTCTAGGCCGCGATTAAATTCCAACATGGAT
GCTGATTTATATGGGTATAAATGGGCTCGCGATAATGTCGGGCAATCAGGTGCGACAATCTATCGATTGTA
TGGGAAGCCCGATGCGCCAGAGTTGTTTCTGAAACATGGCAAAGGTAGCGTTGCCAATGATGTTACAGATG
AGATGGTCAGACTAAACTGGCTGACGGAATTTATGCCTCTTCCGACCATCAAGCATTTTATCCGTACTCCT
GATGATGCATGGTTACTCACCACTGCGATCCCCGGGAAAACAGCATTCCAGGTATTAGAAGAATATCCTGA
TTCAGGTGAAAATATTGTTGATGCGCTGGCAGTGTTCCTGCGCCGGTTGCATTCGATTCCTGTTTGTAATT
GTCCTTTTAACAGCGATCGCGTATTTCGTCTCGCTCAGGCGCAATCACGAATGAATAACGGTTTGGTTGAT
GCGAGTGATTTTGATGACGAGCGTAATGGCTGGCCTGTTGAACAAGTCTGGAAAGAAATGCATAAACTTTT
GCCATTCTCACCGGATTCAGTCGTCACTCATGGTGATTTCTCACTTGATAACCTTATTTTTGACGAGGGGA
AATTAATAGGTTGTATTGATGTTGGACGAGTCGGAATCGCAGACCGATACCAGGATCTTGCCATCCTATGG
AACTGCCTCGGTGAGTTTTCTCCTTCATTACAGAAACGGCTTTTTCAAAAATATGGTATTGATAATCCTGA
TATGAATAAATTGCAGTTTCATTTGATGCTCGATGAGTTTTTCTAAGAATTAATTCATGAGCGGATACATA
TTTGAATGTATTTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGAAAT
TGTAAACGTTAATATTTTGTTAAAATTCGCGTTAAATTTTTGTTAAATCAGCTCATTTTTTAACCAATAGG
CCGAAATCGGCAAAATCCCTTATAAATCAAAAGAATAGACCGAGATAGGGTTGAGTGTTGTTCCAGTTTGG
AACAAGAGTCCACTATTAAAGAACGTGGACTCCAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGG
CCCACTACGTGAACCATCACCCTAATCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATCGGAACC
CTAAAGGGAGCCCCCGATTTAGAGCTTGACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAAA

GCGAAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCACGCTGCGCGTAACCACCACACCCGCCGC
GCTTAATGCGCCGCTACAGGGCGCGTCCCATTCGCCA

pGEM®-T Easy (Promega) vector

Promoter:	T7/SP6
Selection:	Ampicillin
Source:	Promega® (www.promega.com)
notes:	

T7 RNA Polymerase transcription initiation site	1
SP6 RNA Polymerase transcription initiation site	141
T7 RNA Polymerase promoter	3002-6
SP6 RNA Polymerase promoter	136-158
multiple cloning site	10-128
lacZ start codon	180
lacoperon sequences	2839-2999, 166-395
lacoperator	100-216
b-lactamase coding region	1337-2197
phage f1 region	2383-2838
binding site of pUC/M13 Forward Sequencing Primer	2959-2975
binding site of pUC/M13 Reverse Sequencing Primer	176-192

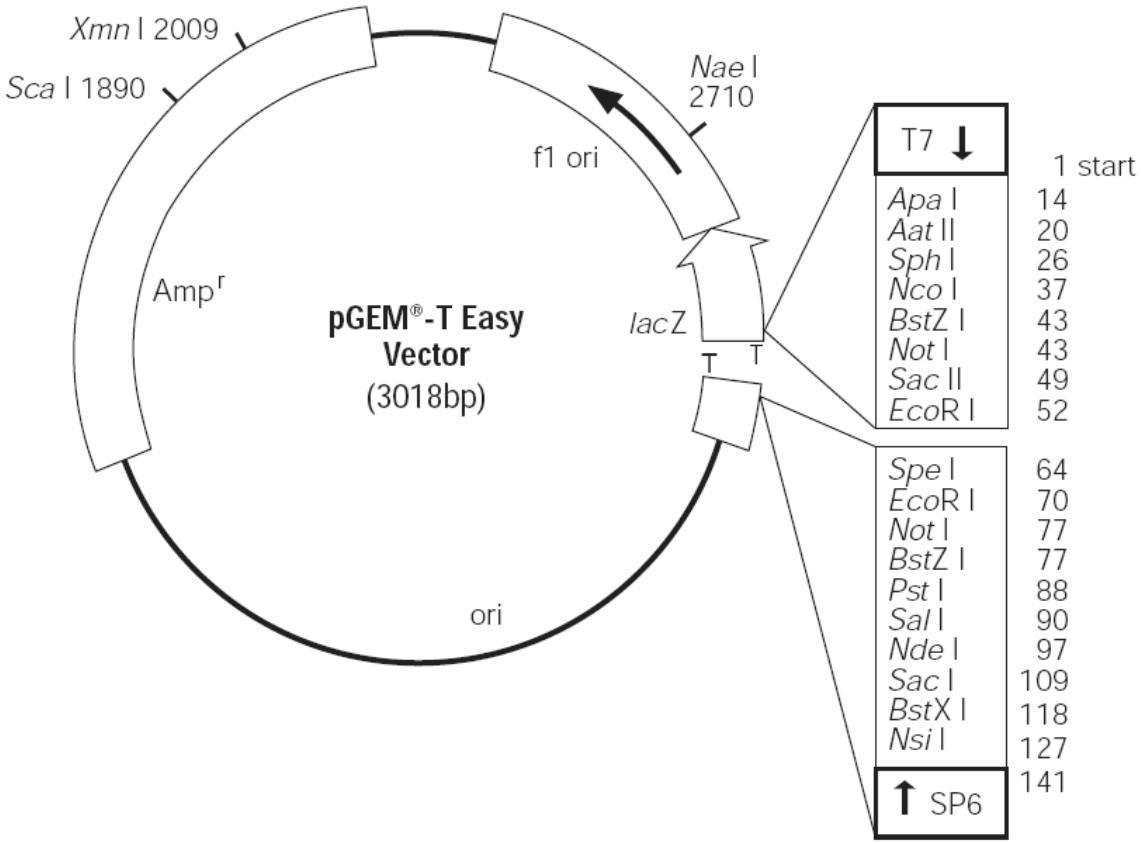


Fig. 9.4.4 The pGEM-T Easy vector map.

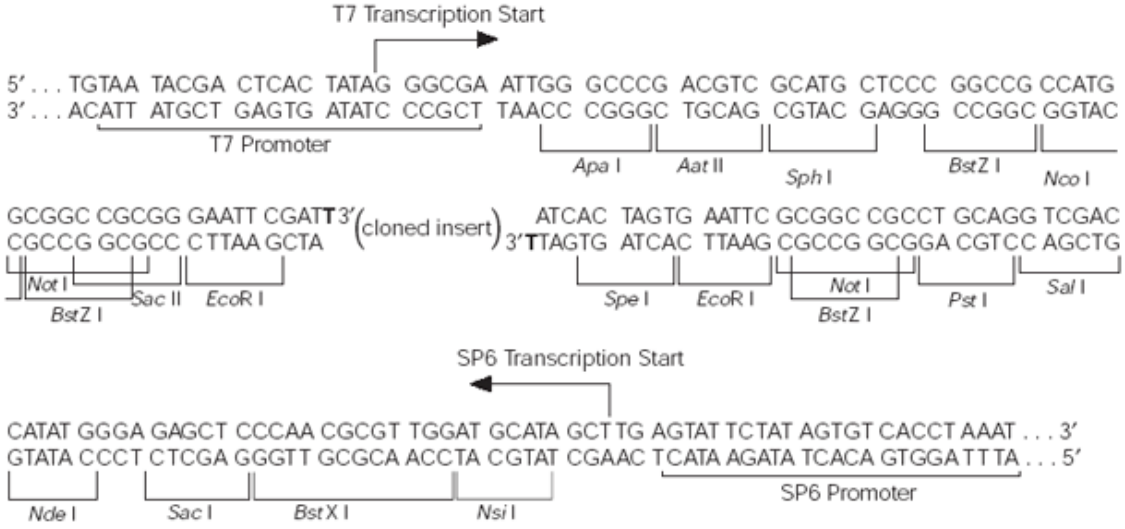


Fig. 9.4.5 The promoter and multiple cloning sequence of the pGEM®-T Easy vector. The top strand of the sequence shown corresponds to the RNA synthesized by T7 RNA Polymerase. The bottom strand corresponds to the RNA synthesized by SP6 RNA Polymerase.

>**pGEM-T Easy;** Full length: 3018 bp

GGGCGAATTGGGCCCGACGTCGCATGCTCCCGGCCGCCATGGCGGCCGCGGGAATTCGATATCACTAGTGA
ATTCGCGGCCGCCTGCAGGTCGACCATATGGGAGAGCTCCCAACGCGTTGGATGCATAGCTTGAGTATTCT
ATAGTGTCACCTAAATAGCTTGGCGTAATCATGGTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCA
CAATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAGCCTGGGGTGCCTAATGAGTGAGCTAACTC
ACATTAATTGCGTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCGTGCCAGCTGCATTAATGAAT
CGGCCAACGCGCGGGGAGAGGCGGTTTGCGTATTGGGCGCTCTTCCGCTTCCTCGCTCACTGACTCGCTGC
GCTCGGTCGTTCGGCTGCGGCGAGCGGTATCAGCTCACTCAAAGGCGGTAATACGGTTATCCACAGAATCA
GGGGATAACGCAGGAAAGAACATGTGAGCAAAAGGCCAGCAAAAGGCCAGGAACCGTAAAAAGGCCGCGTT
GCTGGCGTTTTTCCATAGGCTCCGCCCCCCTGACGAGCATCACAAAAATCGACGCTCAAGTCAGAGGTGGC
GAAACCCGACAGGACTATAAAGATACCAGGCGTTTCCCCCTGGAAGCTCCCTCGTGCGCTCTCCTGTTCCG
ACCCTGCCGCTTACCGGATACCTGTCCGCCTTTCTCCCTTCGGGAAGCGTGGCGCTTTCTCATAGCTCACG
CTGTAGGTATCTCAGTTCGGTGTAGGTCGTTCGCTCCAAGCTGGGCTGTGTGCACGAACCCCCCGTTCAGC
CCGACCGCTGCGCCTTATCCGGTAACTATCGTCTTGAGTCCAACCCGGTAAGACACGACTTATCGCCACTG
GCAGCAGCCACTGGTAACAGGATTAGCAGAGCGAGGTATGTAGGCGGTGCTACAGAGTTCTTGAAGTGGTG
GCCTAACTACGGCTACACTAGAAGGACAGTATTTGGTATCTGCGCTCTGCTGAAGCCAGTTACCTTCGGAA
AAAGAGTTGGTAGCTCTTGATCCGGCAAACAAACCACCGCTGGTAGCGGTGGTTTTTTTGTTTGCAAGCAG
CAGATTACGCGCAGAAAAAAAGGATCTCAAGAAGATCCTTTGATCTTTTCTACGGGGTCTGACGCTCAGTG
GAACGAAAACTCACGTTAAGGGATTTTGGTCATGAGATTATCAAAAAGGATCTTCACCTAGATCCTTTTAA
ATTAAAAATGAAGTTTTAAATCAATCTAAAGTATATATGAGTAAACTTGGTCTGACAGTTACCAATGCTTA
ATCAGTGAGGCACCTATCTCAGCGATCTGTCTATTTCGTTCATCCATAGTTGCCTGACTCCCCGTCGTGTA
GATAACTACGATACGGGAGGGCTTACCATCTGGCCCCAGTGCTGCAATGATACCGCGAGACCCACGCTCAC
CGGCTCCAGATTTATCAGCAATAAACCAGCCAGCCGGAAGGGCCGAGCGCAGAAGTGGTCCTGCAACTTTA
TCCGCCTCCATCCAGTCTATTAATTGTTGCCGGGAAGCTAGAGTAAGTAGTTCGCCAGTTAATAGTTTGCG
CAACGTTGTTGGCATTGCTACAGGCATCGTGGTGTCACGCTCGTCGTTTGGTATGGCTTCATTCAGCTCCG
GTTCCCAACGATCAAGGCGAGTTACATGATCCCCCATGTTGTGCAAAAAAGCGGTTAGCTCCTTCGGTCCT
CCGATCGTTGTCAGAAGTAAGTTGGCCGCAGTGTTATCACTCATGGTTATGGCAGCACTGCATAATTCTCT
TACTGTCATGCCATCCGTAAGATGCTTTTCTGTGACTGGTGAGTACTCAACCAAGTCATTCTGAGAATAGT
GTATGCGGCGACCGAGTTGCTCTTGCCCGGCGTCAATACGGGATAATACCGCGCCACATAGCAGAACTTTA
AAAGTGCTCATCATTGGAAAACGTTCTTCGGGGCGAAAACTCTCAAGGATCTTACCGCTGTTGAGATCCAG
TTCGATGTAACCCACTCGTGCACCCAACTGATCTTCAGCATCTTTTACTTTCACCAGCGTTTCTGGGTGAG

CAAAAACAGGAAGGCAAAATGCCGCAAAAAAGGGAATAAGGGCGACACGGAAATGTTGAATACTCATACTC
TTCCTTTTTCAATATTATTGAAGCATTTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATGTAT
TTAGAAAAATAAACAAATAGGGGTTCCGCGCACATTTCCCCGAAAAGTGCCACCTGTATGCGGTGTGAAAT
ACCGCACAGATGCGTAAGGAGAAAATACCGCATCAGGCGAAATTGTAAACGTTAATATTTTGTTAAAATTC
GCGTTAAATATTTGTTAAATCAGCTCATTTTTTAACCAATAGGCCGAAATCGGCAAAATCCCTTATAAATC
AAAAGAATAGACCGAGATAGGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTATTAAAGAACGTGG
ACTCCAACGTCAAAGGGCGAAAAACCGTCTATCAGGGCGATGGCCCACTACGTGAACCATCACCCAAATCA
AGTTTTTTGCGGTCGAGGTGCCGTAAAGCTCTAAATCGGAACCCTAAAGGGAGCCCCCGATTTAGAGCTTG
ACGGGGAAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAAAGCGAAAGGAGCGGGCGCTAGGGCGCTGG
CAAGTGTAGCGGTCACGCTGCGCGTAACCACCACACCCGCCGCGCTTAATGCGCCGCTACAGGGCGCGTCC
ATTCGCCATTCAGGCTGCGCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTCGCTATTACGCCAGCTG
GCGAAAGGGGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTGTAA
AACGACGGCCAGTGAATTGTAATACGACTCACTATA

pLL10

Promoter: T7
Selection: Kanamycin and Ampicillin
Origin: pLL17 plus pLL7 polylinker
Source: (Blandin, Moita et al. 2002)
notes: vector for RNAi; to sequence insert use M13 forward and M13 reverse primers

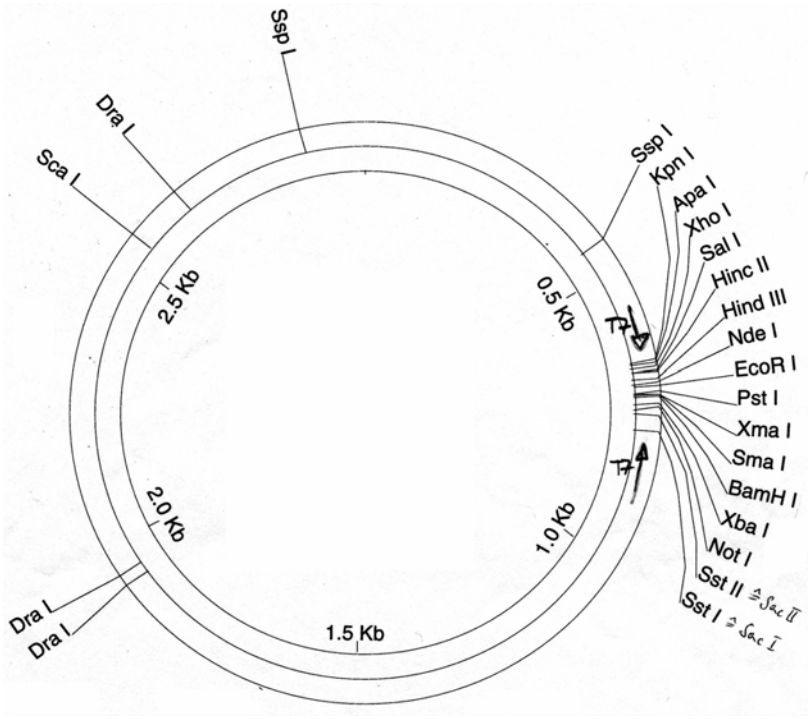


Fig. 9.4.6 The pLL10 vector map.

pLL10 enzymes	compatible with	Common sequence
Sst II (Sac II)	none	-
Not I	none	-
Xba I	Nhe I, Spe I	CTAG
Bam HI	Bcl I, Bgl II	GATC
Sma I / Hinc II	Ssp I, Sca I, Pvu II, Eco RV, Nae I, Hpa I, Dra I	blunt end
Pst I	Nsi I	TGCA
Eco RI	Apo I, Mfe I	AATT
Nde I	none	-
Hind III	none	-
Hinc II	Cf Sma I	blunt end
Sal I	none	-
Xho I	none	-
Apa I	none	-
Kpn I	none	-

Fig. 9.4.7 The pLL10 restriction enzymes and compatible enzymes.

pUAST

Promoter:	UAS/GAL4
Selection:	Ampicillin
Origin:	pCaSpeR3
Source:	(Brand and Perrimon 1993)
notes:	pUAST consists of five tandemly arrayed optimized GAL4 binding sites (red) followed by the hsp70 TATA box and transcriptional start (blue), a polylinker (green) containing unique restriction sites for EcoRI, Bg/II, NotI, Xho, KpnI and XbaI and the SV40 small t intron and polyadenylation site. These features are included in a P-element vector (pCaSpeR3) containing the P element ends (P3' and P5') and the white gene which acts as a marker for successful incorporation into the Drosophila genome.

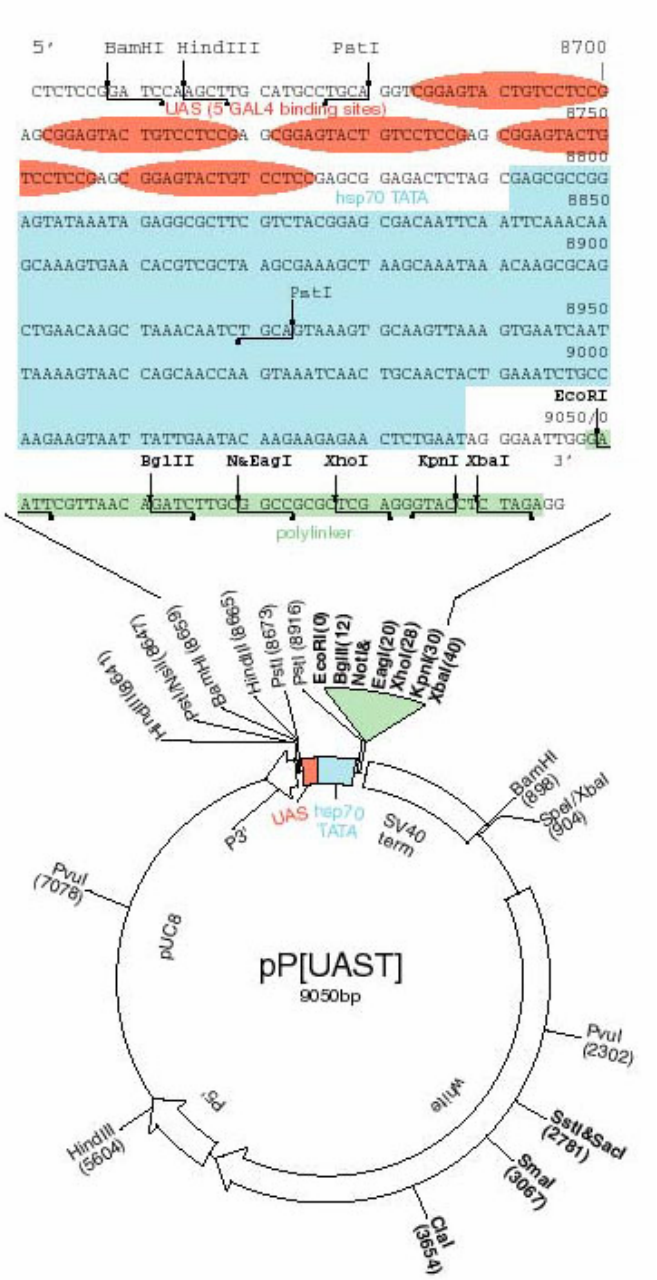


Fig. 9.4.6 The pUAST vector map.

CURRICULUM VITAE

STEPHAN MEISTER

PERSONAL DETAILS

Imperial College London,
Sir Alexander Fleming Building,
South Kensington Campus
London, SW7 2AZ
UK

born 07th June 1975

Phone: (+44) 7962 604481

E-mail: stephan.meister@embl.de
s.meister@imperial.ac.uk

Nationality: German

RESEARCH EXPERIENCE

2001 – present
PhD project

European Molecular Biology Laboratory (EMBL), Heidelberg, Germany
& Imperial College, London, UK
Prof. Fotis C. Kafatos and Dr. George K. Christophides
"The role of PGRP proteins in innate immunity pathways in the malaria vector *Anopheles gambiae.*"

- **Meister S** et al, "The *Anopheles gambiae* PGRPLC gene cluster"; in preparation
- **Meister S** et al, PhD Thesis, EMBL/University of Heidelberg, 2006
- Koutsos, AC,… **Meister S** et al, "*Lifecycle Transcriptomics of the Malaria Mosquito* Anopheles gambiae *and Comparison with the Fruitfly* Drosophila melanogaster" submitted
- **Meister S** et al., *Proc Natl Acad Sci U S A. 2005* Aug 9;102(32):11420-5.
- **Meister S**, et al., *Int J Parasitol. 2004* Dec;34(13-14):1473-82.
- Kumar S,…**Meister S**, et al., *Proc Natl Acad Sci U S A. 2003* Nov 25; 100(24): 14139-44.
- Christophides GK,…**Meister S** et al., *Science 2002* Oct 4;298(5591):159-65.
- Holt RA,… **Meister S** et al., *Science 2002* Oct 4;298(5591):129-49.
- Dimopoulos G,…**Meister S** et al., *Proc Natl Acad Sci U S A. 2002* Jun 25;99(13):8814-9.

1999 – 2000
Independent study

Duke University, Durham, NC, USA
& Universität Würzburg, Würzburg, Germany
Hubert O. Amrein, Ph.D. and Prof. Dr. Martin Heisenberg
"Identification and characterization of the putative odorant receptor gene family in *Drosophila melanogaster*. "

- Dunipace L,…**Meister S**, *Curr Biol. 2001* Jun 5;11(11):822-35.
- **Meister S,** Diploma Thesis, University of Würzburg, 2001.

1999
Independent study

Universität Würzburg, Würzburg, Germany
Prof. Dr. Martin Heisenberg
"Variability of neuron specific GAL4 lines driving GFP-expression in the *Drosophila melanogaster* brain."

EDUCATION

2005–present	Imperial College (EMBL PhD, continued)	London, UK
2001–2005	European Molecular Biology Laboratory (EMBL PhD)	Heidelberg, Germany
1999–2000	Duke University (foreign exchange program)	Durham, NC, USA
1996–2001	Bayrische Julius-Maximilians Universität MSc equivalent: "Diplom Biologe"; grade: very good.	Würzburg, Germany

1987–1994	Hohenlohe Gymnasium Öhringen (High School)	Öhringen, Germany
1981–1987	Deutsche Schule Bombay (Elementary & High School)	Mumbai, India

CONFERENCES & COURSES

2006	2nd Annual BioMalPar Conference on the Biology and Pathology of the Malaria Parasite, Heidelberg, Germany *Poster*: "The Role of PGRP Proteins in Innate Immunity Pathways in the Malaria Vector Anopheles *gambiae*."
2005	7th International EMBL PhD Student Symposium: Biology at Work – a journey through applied life sciences, Heidelberg, Germany.
2005	6th International EMBL PhD Student Symposium:: Animal models – tips and tricks from nature, Rome, Italy.
2004	1st Finnish National PhD Symposium: Interaction & Networks in Biology, Hämeenlinna, Finnland.
2004	Biology of Disease Vectors Course, Colorado State University, Fort Collins, Colorado, USA. *Poster:* "The *Anopheles* gambiae Peptidoglycan Recognition Protein LC (PGRPLC) in Innate Immunity and Malaria".
2003	**Organizer**. 4th International EMBL PhD Student Symposium: A life of Encounters - Recognition in Biology, Heidelberg, Germany.
2003	EMBO Workshop on Pattern Recognition Proteins and Receptors, Trest, Czech Republic. *Poster*: "PGRPs in the malaria mosquito *Anopheles gambiae*".

TEACHING EXPERIENCE

2006	Supervision of Imperial College Undergraduate student Louise Downs 6 week final year project. (May-June 2006)
2006	Supervision of Imperial College MSc student Joanna Waldock's 13 week project "The Toll pathway – effect of immune signaling on fungal and *Plasmodium berghei* infection in *Anopheles gambiae* mosquitoes." (Feb-April 2006)

OTHER EXPERIENCE

2003–2005	EMBL **First Aid Helper**.
2004–2005	**PhD Student representative** on EMBL Staff Association & EMBL **Safety Committee**.
2004–2005	EMBL BSAC **Diving Club Member** and **Assistant Instructor**.
1995–1996	Backpacking Trip through South-East Asia.
1994–1995	Community Service at „Psychosoziale Beratungsstelle in Öhringen" (Counseling Center for Substance Abuse), Öhringen, Germany.

IMPERIAL COLLEGE LONDON, DIVISION OF CELL & MOLECULAR BIOLOGY • SIR ALEXANDER FLEMING BUILDING • SOUTH KENSINGTON CAMPUS • LONDON • SW7 2AZ • UK
PHONE: (+44) 7962 604481 • E-MAIL: S.MEISTER@IMPERIAL.AC.UK

SKILLS & COMPETENCES

Language skills: **German** (native speaker), **English** (excellent written & spoken), French (basic comprehension).

Excellent **computer skills**: Windows XP, MS Office, Adobe Illustrator, Adobe Photoshop, Adobe Acrobat, DNAStar, Artemis, GeneSpring and various other bioinformatics tools.

Clean driving record, licensed for cars, trucks up to 7.5 tons and motorcycles.

INTERESTS

Martial arts, salsa, scuba diving, computers.

FELLOWSHIPS

"EMBL predoctoral fellowship" for PhD studies from 2002 to 2006.

"Integriertes Auslandsstudium" (IAS) fellowship of the German Academic Exchange Service (DAAD) for attendance of Duke University Graduate School from 1999 to 2000.

REFERENCES AVAILABLE UPON REQUEST

IMPERIAL COLLEGE LONDON, DIVISION OF CELL & MOLECULAR BIOLOGY • SIR ALEXANDER FLEMING BUILDING • SOUTH KENSINGTON CAMPUS • LONDON • SW7 2AZ • UK
PHONE: (+44) 7962 604481 • E-MAIL: S.MEISTER@IMPERIAL.AC.UK