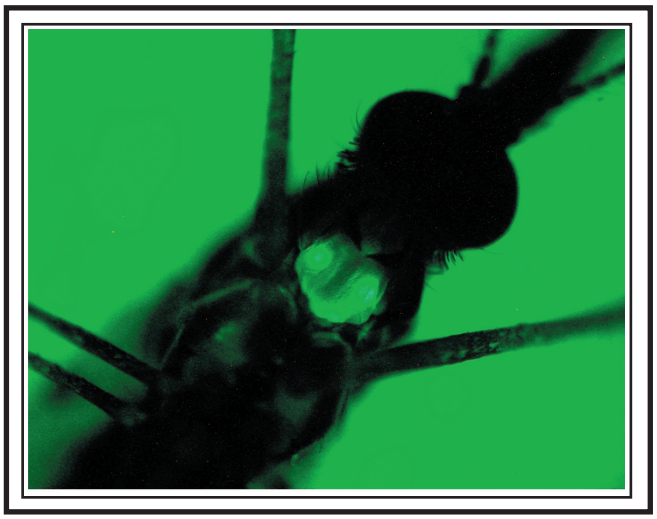
Thèse présentée pour obtenir le grade de

Docteur de l'Université Louis Pasteur de Strasbourg

Par

Stéphanie BLANDIN

Functional Analysis of Thioester-containing Proteins in Immune Responses of the Mosquito, *Anopheles gambiae*



Soutenue le 23 février 2004 devant la commission d'examen :

Prof. Fotis KAFATOS	Directeur de Thèse
Prof. Jules HOFFMANN	Directeur de Thèse
Prof. Bernard EHRESMANN	Rapporteur Interne
Prof. André CAPRON	Rapporteur Externe
Dr. Anne EPHRUSSI	Rapporteur Externe
Dr. Elena LEVASHINA	Examinateur

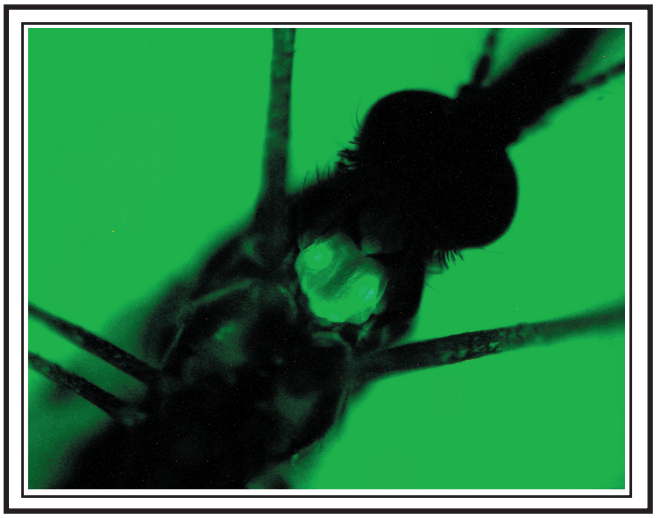
Thèse présentée pour obtenir le grade de

Docteur de l'Université Louis Pasteur de Strasbourg

Par

Stéphanie BLANDIN

Functional Analysis of Thioester-containing Proteins in Immune Responses of the Mosquito, *Anopheles gambiae*



Soutenue le 23 février 2004 devant la commission d'examen :

Prof. Fotis KAFATOS	Directeur de Thèse
Prof. Jules HOFFMANN	Directeur de Thèse
Prof. Bernard EHRESMANN	Rapporteur Interne
Prof. André CAPRON	Rapporteur Externe
Dr. Anne EPHRUSSI	Rapporteur Externe
Dr. Elena LEVASHINA	Examinateur

European Molecular Biology Laboratory
Meyerhofstrasse, 1
D-69117 Heidelberg
Germany

Tel: (+49) 6221.387.440
Fax: (+49) 6221.387.518

Institut de Biologie Moléculaire et Cellulaire
15 rue René Descartes
F-67087 Strasbourg Cedex

Cover picture: Green fluorescent parasites have invaded the mosquito salivary glands. The fluorescence is visible through the cuticule.

Abstract

The mosquito *Anopheles gambiae* is one of the most important vectors of human malaria in Africa. The causative agents of this disease are unicellular eukaryotic parasites of the genus *Plasmodium*. They are transmitted to humans by infected female mosquitoes when they take a blood meal. During its development in the mosquito, *Plasmodium* undergoes massive losses, which suggests that mosquitoes are able to mount a potent immune response and to limit parasite infection. However, the molecular mechanisms underlying parasite recognition and killing are not well understood. To address this question, we have chosen to study Thioester-containing proteins (TEPs) as this type of molecules is involved in pathogen recognition and activation of immune responses in other organisms. Indeed in vertebrates, members of this family comprise the universal protease inhibitors, α_2-macroglobulins, and the complement factors C3/C4/C5, which label pathogens and trigger their disposal through phagocytosis or cell lysis.

We identified a family of 15 *TEP* genes in the mosquito genome. To investigate the function of these and other genes, we adapted the RNA interference strategy to knockdown gene expression, either by transfection of cultured mosquito cells with double stranded RNA (dsRNA) or by direct injection of adult mosquitoes with dsRNA. We further established a set of phenotypic tests to analyze the effect of gene silencing on mosquito immune responses. In addition, we extended the *in vivo* knockdown by dsRNA injection to adult *D. melanogaster* where it represents a powerful method for rapid gene functional analysis and for rapid epistatic analysis.

Using these tools, we studied several aspects of the mosquito immune system, which could not have been addressed before. We provided the first functional evidence that an insect antimicrobial peptide plays an important role in the resistance to infections: Defensin1 is required for the mosquito antimicrobial defense against Gram+, but not against Gram- bacteria. Functional analysis of the mosquito Thioester-containing protein 1, TEP1, revealed that this molecule acts as an opsonin and promotes phagocytosis of bacteria by hemocytes. In addition, the binding of TEP1 to *Plasmodium* parasites mediates their killing. Our data provide the first molecular identification of a mosquito factor implicated in the establishment of vectorial capacity in *A. gambiae*. Further elucidation of the mechanisms of TEP1 activation and of the effector mechanisms triggered by TEP1 binding will be instrumental in our understanding of the mosquito immune responses, and especially of that against parasites.

Résumé

Le moustique *Anopheles gambiae* est l'un des principaux vecteurs du paludisme en Afrique. L'agent responsable de la maladie, *Plasmodium*, est un parasite eucaryote unicellulaire. Il est transmis à l'homme par un moustique femelle infecté lorsque celle-ci prend un repas de sang. Au cours de son développement chez le moustique, *Plasmodium* subit de nombreuses pertes, ce qui suggère que le moustique est capable de monter une réponse immunitaire et de limiter le développement du parasite. Cependant, les mécanismes moléculaires qui sous-tendent la reconnaissance et l'élimination des parasites ne sont pas connus. Afin d'aborder ce problème, nous avons choisi d'étudier les protéines à thioester (TEPs). En effet, des molécules de ce type sont impliquées dans la reconnaissance des pathogènes et dans l'activation des effecteurs de la réponse immunitaire chez d'autres organismes. Chez les vertébrés par exemple, on trouve dans cette famille les α2-macroglobulines, qui sont des inhibiteurs universels de protéases, et les facteurs C3/C4/C5 du complément, qui marquent les pathogènes et favorisent ainsi leur élimination par phagocytose ou par lyse cellulaire.

Nous avons identifié 15 gènes *TEP* dans le génome du moustique. Afin d'analyser leur fonction ainsi que celle d'autres gènes, nous avons adapté la technique d'interférence à ARN pour l'anophèle : l'expression d'un gène peut être inhibée par transfection d'ARN double brin (ARNdb) dans des cultures de cellules *in vitro*, ou par injection d'ARNdb dans des moustiques adultes. Nous avons ensuite mis en place un ensemble de tests phénotypiques qui nous ont permis d'étudier l'effet de ces knock-downs sur la réponse immunitaire du moustique face à divers pathogènes. En outre, nous avons élargi l'application de cette méthode à la drosophile adulte où le knock-down par injection d'ARNdb représente une méthode rapide pour l'étude fonctionnelle et épistatique des gènes.

L'utilisation de ces outils nous a permis d'analyser plusieurs aspects du système immunitaire du moustique, qui n'avaient pas pu être étudiés auparavant. D'une part, nous avons donné la première caractérisation fonctionnelle d'un peptide antimicrobien d'insecte : Defensin1 est nécessaire pour la résistance du moustique aux bactéries à Gram+, mais pas aux bactéries à Gram-. D'autre part, nous avons démontré que la protéine à thioester 1, TEP1 est une opsonine : le marquage des bactéries par TEP1 active leur phagocytose par les hémocytes, De plus, TEP1 est aussi capable de se fixer à la surface de *Plasmodium* et de déclencher son élimination. TEP1 est le premier facteur impliqué dans l'établissement de la capacité vectorielle à être identifié chez *A. gambiae*. L'analyse plus précise des mécanismes moléculaires qui activent TEP1 et de ceux qui sont déclenchés par la fixation de TEP1 nous apparaît essentielle à une meilleure compréhension du système immunitaire de l'anophèle, et en particulier de sa réponse antiparasitaire.

Acknowledgements

I would like to thank Pr. Fotis Kafatos for giving me the opportunity to work in his laboratory, in a scientific environment of very high quality and for his support. It has been (and still is!) an incredibly exciting time to work in the field of mosquito immunity! Thank you for having contributed so much to the happening of the mosquito revolution!

I am grateful to Pr. Jules Hoffmann for accepting to be my supervisor in France, and for his continuous support and interest in my work. Je tenais aussi à vous remercier pour votre confiance!

Prs. André Capron, Bernard Ehresmann and Dr. Anne Ephrussi accepted to judge my thesis work and I would like to thank them. Dr. Anne Ephrussi was also a member of my thesis advisory committee at the EMBL, as well as Dr. Jan Ellenberg. I would like to acknowledge them for their interest in my work and for their advice.

My deepest gratitude naturally goes to Dr. Elena Levashina, who has been my supervisor. I would like to thank her for her great support, her humanity and for her fairness. Lena, tu as su me faire bénéficier de ton expérience et de tes connaissances impressionnantes en toute simplicité, et c'est une qualité bien rare que j'admire beaucoup. Ton enthousiasme débordant pour la recherche et pour plein d'autres choses encore, est très communicatif! Merci pour ta bonne humeur! Merci pour ta confiance qui m'est infiniment précieuse!... Vraiment, je ne saurais comment exprimer toute ma reconnaissance pour ce que tu m'as donné, et en particulier, pour ton amitié!

This work has greatly benefited from diverse collaborations and I would like to thank the persons who have been involved in it, one way or another: The members of the TEP team, Shin-Hong Shiao and Luis Moita, with whom I have shared material, ideas, projects and fruitful discussions; Randy Clayton, whose expertise in transforming mosquitoes has been essential to the obtention of the TEP1 transgenic mosquitoes, also thanks for the nice discussions! Akira Goto, for his kindness and his determination in making the dsRNA knockdown work in flies. Special thanks to Aline Astrain and Ana Krstanovic, who came for a training period in the lab. I really appreciated your enthusiasm!

Thanks also to the past and present members of the Kafatos'lab: and especially to Albertino Danielli and mamoiselle Rosita Barrio for their steadfast kindness and cheerfulness, to Tasos Koutsos for his company in the lab, late in the evening and for long discussions about macs, lab, life… lab life, to Gareth Lycett for his British (mostly!) good humour and kindness, to Dolores Doherty who transformed the insectary from a boring desert into a crowded cheerful place and to the others for all the good times we have spent together. I very much enjoyed the young, dynamic and international atmosphere of the European Molecular Biology Laboratory. Special thanks to the French connection of the Gene Expression program, for unrestrained conversations.

I also would like to thank people from the IBMC and especially from Lena's lab, for their warm welcome each time I came to Strasbourg.

Many many thanks to all of you, with whom I have shared unforgettable moments in Heidelberg, but outside the lab! Tor Erik Rusten, Birgit Kerber, Michi Hannus and Katrin Flohr, Olaf Selchow, the founders of the Sat' Swim' club; David Hipfner and Dagmar Neubueser, Brenda Stride and Thomas Lemberger and their little Marie, and the others to keep the Sat' Swim' tradition going, for the nice diners and movie nights, and most importantly, for their precious friendship.

Mr. Doridot, mon cher professeur de biologie et de géologie du lycée! Si je n'en étais pas déjà convaincue, vous m'avez prouvé que la biologie était une science passionnante. Et puis, c'est aussi pendant vos cours que j'ai découvert le monde fascinant de l'immunologie. Au combien je je suis comblée de pratiquer cette science quotidiennement désormais! Merci de m'avoir soutenue et aidée à prendre les bonnes décisions quand il le fallait. Merci aussi pour votre amitié.

À ma famille, pour tout l'amour que vous m'avez apporté, pour votre soutien sans faille, pour vos encouragements et surtout, pour avoir su former et conserver ce petit nid douillet où il fait toujours bon se retrouver.

À François, parce que la vie est belle avec toi, tout simplement!

Tables and Figures

Abbreviations

TEP:	Thioester-containing protein
TE:	thioester
α2M:	alpha2-macroglobulin
C3, C4, C5:	complement factors C3, C4 and C5
MBL:	mannan-binding lectin
MASP:	MBL associated serine protease
MAC:	membrane attack complex
RNAi:	RNA interference
siRNA:	small interfering RNA
miRNA:	microRNA
dsRNA:	double stranded RNA
S:	susceptible (refers to the G3 strain of *A. gambiae*, susceptible to parasite infection)
R:	refractory (refers to the L3-5 strain of *A. gambiae,* refractory to parasite infection)
DaPKC:	*Drosophila* atypical protein kinas C

Table of Contents

Chapter 1

General Introduction

General Introduction

A number of arthropod species act as vectors of pathogenic organisms of medical importance. A common feature of these species is their habit to feed on blood, to support oogenesis or to fulfil other nutritional requirements. When feeding on blood, the arthropod vector can ingest and be infected by pathogens present in the host circulatory system or epidermis, and in subsequent blood meals, the vector can transmit these infectious agents to a new host. Among infectious diseases transmitted by arthropod vectors to humans, malaria is by far the most dreadful. It is caused by *Plasmodium* parasites and transmitted to humans by *Anopheles* mosquitoes.

1. Malaria in 2003

The social and economic burden of malaria

Malaria is the most widespread parasitic human disease in the world. It is currently endemic in Africa, South America and South-East Asia (Fig. 1A) and exerts its heaviest toll in Sub-Saharan Africa, where at least 90% of deaths from malaria occur. Every year, between 300 and 500 million cases of acute illness are reported, affecting nearly 10% of the global population. One to 3 million persons die as a result, mostly children under the age of five. Among adults, pregnant women and non-immune travellers are the main groups at risk. In addition to the heavy human burden, malaria also has devastating consequences on the economic al health of these countries (reviewed in Sachs and Malaney, 2002). The correlation between poverty and malaria is striking: the vast majority of malarious countries are also the ones with the lowest per capita gross domestic product (GDP) and the lowest rate of economic growth (Fig. 1B). Malaria and poverty somehow promote each other in a vicious circle. By affecting worker productivity and augmenting absenteeism, by increasing medical costs and premature mortality, by decreasing saving, foreign investment and tourism, malaria reduces incomes both at the household and at the macroeconomics levels. Reciprocally, poverty is a factor promoting malaria transmission as it reduces the expenditures on prevention, treatment and on government control programmes.

Control of malaria: a need for new strategies

The main elements in the transmission of malaria are: the causative agents, *Plasmodium* parasites; the insect vectors, *Anopheles* mosquitoes; the human hosts and the environment. Intervention on any of these can prevent spreading of the disease. Historically, the greatest successes came from programmes for vector control (Greenwood and Mutabingwa, 2002). In the first half of the 20th century, malaria was eliminated from the United States and from most of Europe as a result of changes in land use, agricultural practises and house construction and some targeted vector

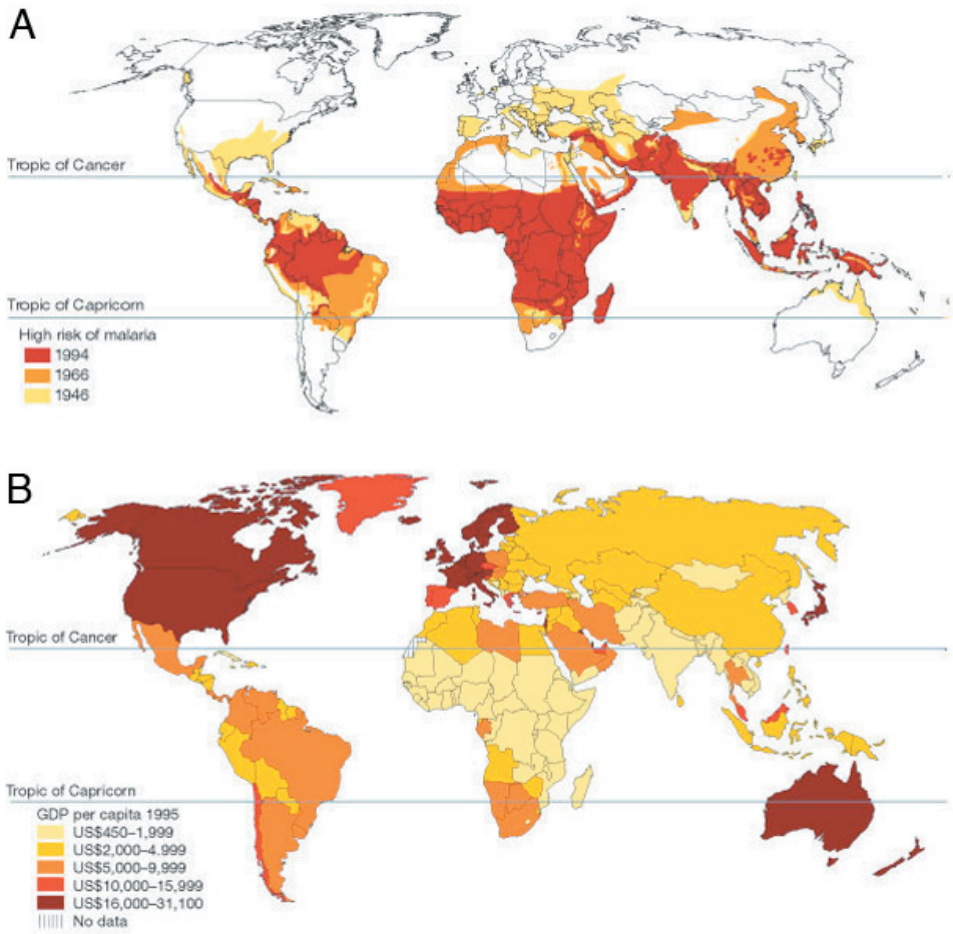


Figure 1. Correlation between malaria and poverty. **A**. Global distribution of malaria. In the last century, malaria has been successfully eradicated mostly from temperate regions as a result of changes in agricultural practices, house construction and targeted vector control efforts. However the most difficult part remains to be done: eradication of malaria from the tropics, where the climate is highly favourable to the development of mosquitoes and parasites. **B.** Global distribution of per capita gross domestic product in 1995, after adjusting to give parity of purchasing power. From (Sachs and Malaney, 2002).

control campaigns. The development of the highly effective insecticide DTT initiated a global eradication programme in the 1950s and 1960s, which has successfully eradicated or controlled the disease in countries with temperate climates and in some countries where malaria transmission was low or moderate (Fig. 1A). However this effort has not been sustained because of the costs of the programme, the opposition of many communities to repeated spraying of their houses and the emergence of resistance to DTT in mosquito populations. Additional factors have contributed to worsening the malaria burden, particularly in Africa (Greenwood and Mutabingwa, 2002). By far, the most important is the development by *Plasmodium falciparum* of resistance to cheap drugs such as chloroquine and sulphadoxine/pyrimethamine, and in some parts of Southeast Asia, to almost all antimalarial drugs. Other factors include poverty, political instability and collapse of health services, environment changes (such as the development of rice fields or the construction of small dams, that increase the number of breeding sites for mosquitoes), climatic changes (e.g.

floods, global warming), more frequent exposure of non-immune populations and emergence of HIV/AIDS. Alternative attempts to control malaria have focused on the prevention of the disease transmission to humans through vaccination. However, no effective vaccine against malaria has been reported thus far, mainly due to the fact that the parasite changes its antigenic structures so that vaccines rapidly become ineffective (reviewed in Richie and Saul, 2002). By the early 1990s, the international community began to realize that the malaria burden was becoming unbearably high, and in 1992, malaria control was re-established as a global health priority of the World Health Organization (WHO).

In 1998, the WHO launched the «Roll Back Malaria» partnership that is drawing together the main groups interested in malaria control and has committed itself to alleviate the burden of malaria by two by the year 2010. To achieve that goal, the WHO advocates several approaches which address the need to use existing, low-cost but under-used interventions to the fullest extend possible, and these are:

- insecticide-treated bed nets (ITN) to reduce the number of mosquito bites
- antimalarial drugs used in combination to delay the emergence of resistance
- intermittent treatment for malaria in pregnancy and infancy, to protect the two most vulnerable groups
- to improve access to treatment
- to improve the prediction and containment of epidemics

In addition, support for fundamental, applied and operational research aims to develop new technical tools, such as new drugs and vaccines, but also new strategies to control the vectorial capacity of mosquitoes. The past four years have witnessed the emergence of techniques that allow stable genetic transformation of mosquitoes. By genetic engineering, mosquitoes could be rendered refractory to parasite infection by the expression of an exogenous compound inhibiting parasite development (Ito et al., 2002), by the misexpression of a mosquito molecule essential to parasite development or by the fine-tuning of already existent mosquito immune defences. Although at present the ecological consequences of a release of genetically engineered mosquitoes, even when sterilized, are very difficult to foresee, it provides an attractive tool to modulate mosquito vectorial capacity. In any case, to design novel strategies and to assess the associated potential efficacy and risk, a more thorough understanding of mosquito-parasite interactions at the molecular level is required. In spite of more than 100 years of history, the research on *Anopheles-Plasmodium* interactions is still lacking important molecular insights into the finely-tuned host-parasite interactions.

2. Malaria parasite and mosquito vector: partners in crime

Historical perspective

As early as 95 BC, Lucretius hypothesized that swamp fever might result from a living organism. However, until the late 19th century, it was generally believed that malaria (from Italian "mal' aria") was caused by a poisonous vapour or miasma, released from swamps. The major breakthrough in the understanding of the disease etiology occurred in 1880, when the French physician Charles Louis Alphonse Laveran discovered the malaria parasite in blood samples of soldiers suffering from the swamp fever. Nearly twenty years later, the English physician Ronald Ross, encouraged by Sir Patrick Manson to investigate whether malaria parasites were transmitted by mosquitoes, discovered parasites on the midgut wall of a mosquito and went on to trace the entire lifecycle of bird malaria. Both Laveran and Ross were awarded the Nobel Prize in Medicine in 1907 and 1902, respectively. Further description of the parasite and mosquito species causing malaria was contributed by two Italian researchers. In 1885-85, Camillo Golgi, more famous for his pioneering work on the central nervous system and on cell compartments, discovered that different types of parasites cause different types of fever. These parasites were later regrouped in the genus *Plasmodium*. In 1898, Battista Grassi identified a species of *Anopheles* as a vector of human malaria, and proved that not all *Anopheles* species could transmit malaria to humans. In the following years, new *Plasmodium* species were identified, and their lifecycle characterized. Malaria is not a disease restricted to humans, it does also occur in other mammals as well as in birds. However a given species of *Plasmodium* can cause malaria in only one vertebrate species, and is usually transmitted by a limited number of Anopheline species.

Life cycle of Plasmodium in mosquitoes

Plasmodia are protozoan parasites that are transmitted by *Anopheles* vectors when a female mosquito takes an infected bloodmeal. To become infective to the next host, the parasite must undergo a complex developmental cycle in the mosquito (Fig. 2). Within the midgut, male and female gametocytes develop in the mosquito midgut into gametes that undergo fertilization within the next few hours. The newly formed zygotes transform into motile ookinetes that invade and cross the midgut epithelium between 24 and 48 hours post infection. Once they have reached the basal side of the epithelium, the parasites form protected capsules, called oocysts, between the midgut epithelium and the basal lamina. During the next 10 days, within each oocyst, a meiotic cycle followed by several rounds of mitosis, produce thousands of haploid sporozoites. Upon maturation, which takes place 14 – 16 days post infectious blood meal, sporozoites are released into the mosquito hemocoel, migrate to and invade the salivary glands. The parasite cycle in the mosquito is completed when the mosquito injects infective sporozoites into a new host.

Susceptibility and Refractoriness of mosquitoes to parasites

As we have seen in the previous paragraph, mosquitoes are not simple mechanical bridges between two hosts: malaria parasites have to undergo a series of growth and differentiation events within the mosquito to ultimately become infective to the next host. Because of the complexity of the interactions between the parasite and the mosquito vector, most parasites develop successfully only in a very narrow range of mosquito species and therefore, only a limited number of dedicated *Plasmodium – Anopheles* combinations cause malaria in a particular group of vertebrates. Moreover, when a mosquito species is permissive for a parasite species, transmission to the host might not occur due to the mosquito biting preferences. In addition to inter-species variations, the ability of mosquitoes to sustain parasite development (called vectorial capacity) varies substantially among individuals of the same species (reviewed in Vargas, 1949). Thus mosquito-parasite interactions constitute a critical aspect of disease transmission and represent potential targets for efforts to control malaria. The factors which determine why an individual mosquito will not become infected after it has received an infective meal are of two kinds (reviewed in Sinden, 2002). On one hand, the mosquito can lack an essential requirement to support the successful development of the parasite; examples include elements required for developmental induction such as xanthurenic acid, which activates gametogenesis in the mosquito midgut (Billker et al., 1998), nutrients or specific ligands or receptors necessary for the invasion of the midgut and salivary gland epithelia. Alternatively, mosquitoes may mount a potent immune response against parasites.

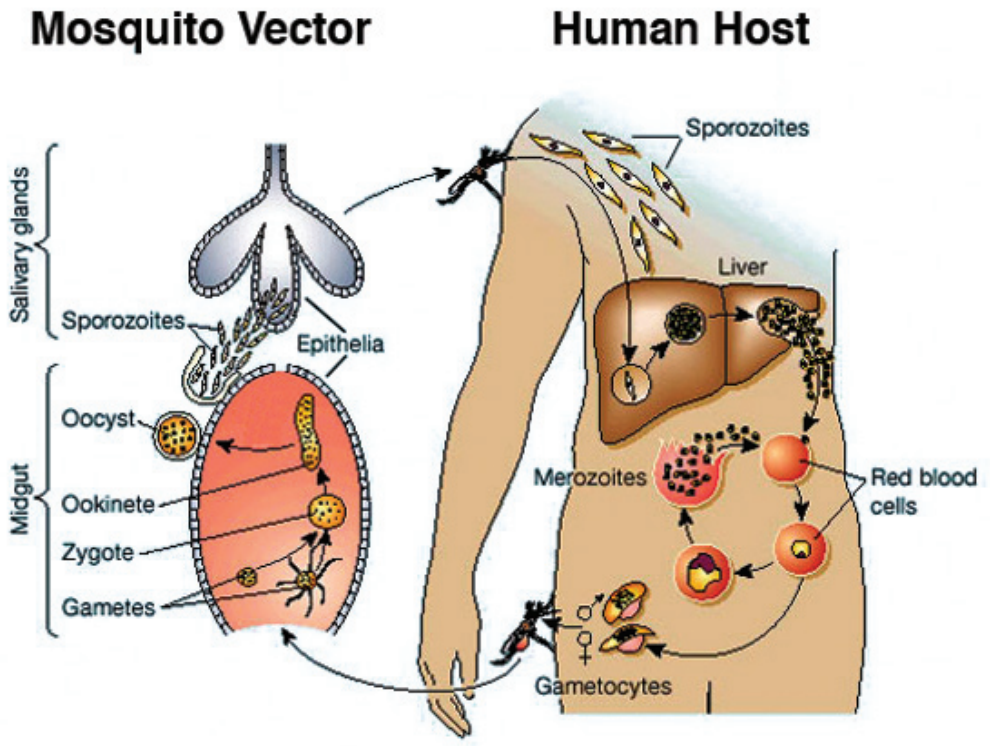


Figure 2. *Plasmodium* life cycle. The parasite develops in many diverse stages adapted to the invasion and multiplication in different cell types in the human host and mosquito vector.

Indeed, several lines of evidence suggest that mosquitoes have evolved potent defence mechanisms against malaria parasites. Even in susceptible mosquitoes, *Plasmodium* traverses several bottlenecks with massive losses during its development: the transitions between gametocytes and ookinetes, between ookinetes and oocysts, and between midgut sporozoites and salivary gland sporozoites (Fig. 3). The importance of the mosquito immune defences in limiting parasite loads is further supported by the fact that in two mosquito strains selected in the laboratory, parasites cannot develop: they are either killed while they traverse the midgut epithelium, illustrating to the extreme how efficient the mosquito response can be. One of these strains of *A. gambiae*, L3-5, completely aborts development of a number of malaria parasites, including the simian parasite *P. cynomolgi* B. Refractoriness is manifested by melanotic encapsulation of the ookinete (Fig. 4), after it completes its passage through the mosquito midgut (Collins et al., 1986). The genetic control of refractoriness appears to be complex. Not only it involves several quantitative trait loci, but also the relative contribution of each locus to oocyst encapsulation varies with the species of parasites (Zheng et al., 1997; Zheng et al., 2003). A different refractory mechanism resulting in complete lysis of ookinetes in the midgut was reported in a different infectious model: *A. gambiae* - *P. gallinaceum* (Vernick et al., 1995). The molecular mechanisms underlying

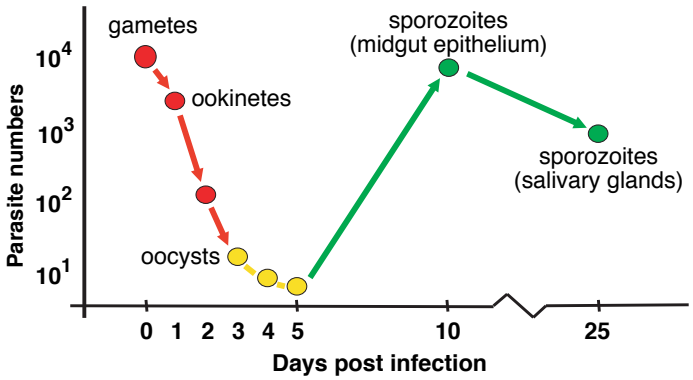


Figure 3. Parasite losses and amplification in the mosquito. Parasite numbers during the critical steps of transformation of gametes to ookinetes and of ookinetes to oocysts were quantified using real-time PCR and *P. berghei 28S* rRNA gene as a probe. The initial number of ingested parasites was estimated to be 10^4. Note the dramatic decrease in parasite numbers between 1 and 3 dpi (modified from Sinden, 1999).

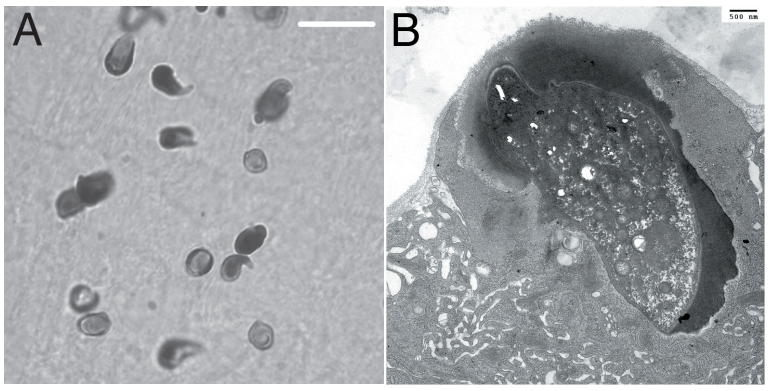


Figure 4. Melanization of parasites in the L3-5 refractory strain of *A. gambiae*. A. Melanized ookinetes are visible on the midgut of refractory L3-5 females 48 hours post infection. **B.** Electron micrograph of a melanized ookinete in the midgut of a L3-5 mosquito 32 hours post infection. Scale bars in μm: (A) 10; (B) 0.5. Electron micrograph by courtesy of Shin-Hong Shiao (IBMC, Strasbourg).

both types of refractoriness, and more generally mosquito immune defences, remained to be understood. The research in the field of mosquito immunity has greatly benefited from our knowledge of immune defences of vertebrates and of other arthropods, especially of another dipteran species, the fruitfly *Drosophila melanogaster*.

3. Insect immune responses: the fly paradigm

Antimicrobial defences in Drosophila melanogaster

D. melanogaster is a model of choice to study fundamental processes such as immunity, due to the exceptional experimental possibilities it offers, including a myriad of molecular and genetic tools. Fruitflies do resist challenges by various microorganisms remarkably well. For this, they rely solely on immune responses, which are based on the recognition of microorganisms by germline-encoded receptors and the rapid induction of effector mechanisms. This type of immune responses is evocative of innate immune responses in vertebrates (reviewed in Hoffmann, 2003). A short description of innate and adaptive immune systems is given in Box 1.

The very first line of defence is mechanical, aimed to avoid a direct contact with pathogens as much as possible. In this respect, fruitflies, as all Arthropods, are protected by their external rigid cuticule, which can function as a physical barrier between the internal milieu and the external world that contains pathogens. Internal epithelia, such as the digestive tracts and the trachea, are quite frequently in contact with microorganisms present in the food, in the air and are not shielded by the cuticule. All these surfaces locally produce antimicrobial peptides, which inhibit microbial growth (Ferrandon et al., 1998; Tzou et al., 2000). Antimicrobial peptides (AMPs) are small cationic peptides, very diverse in their structure and are believed to kill bacteria and fungi by damaging cell membranes (reviewed in Zasloff, 2002). They appear to be ubiquitous components of the innate immune defense arsenal used by both prokaryotic and eukaryotic organisms (reviewed in Zasloff, 2002). When the epithelial barriers are breached and tissues are infected, both cellular and humoral responses of the fruitfly are activated to fight invading pathogens. Cellular defences are mediated by hemocytes, the blood cells of insects, and comprise phagocytosis and cellular encapsulation (reviewed in Meister, 2004). Phagocytosis plays an important role in clearing bacteria, whereas larger pathogens that cannot be phagocytosed are encapsulated by hemocytes. Humoral reactions include the induction of proteolytic cascades in the hemolymph (the analogue of the vertebrate blood) leading to localized melanization and blood coagulation, which limit the spreading of invaders at the site of infection. The hallmark of the humoral response, however, is the rapid synthesis of antimicrobial peptides by the fat body (the analogue of the vertebrate liver). Although present on epithelial surfaces as described above, these peptides are normally absent from the hemolymph of naive flies. Microbial challenges induce the expression of the corresponding genes in fat body cells and the secretion of these peptides in the hemolymph (reviewed in Naitza and Ligoxygakis, 2004). Seven distinct inducible peptides (and their isoforms) have been characterized in *Drosophila* and their activities have

Box 1: Innate vs Adaptive Immunity

To resist infection by pathogens, metazoans have evolved diverse mechanisms of defense that can be classified in two types: Innate immunity, which is common to all metazoans and adaptive immunity, which is restricted to some 45000 vertebrate species.

The vast majority of metazoans rely solely on innate immune defenses. Recognition of microorganisms by the innate immune system is achieved by a limited number of germline-encoded receptors known as "pattern recognition receptors" (Janeway, 1989) that have conserved the memory of motives ("patterns") associated with pathogens throughout evolution. The principal function of these receptors is to trigger rapid effector mechanisms that include phagocytosis, activation of proteolytic cascades, and synthesis of antimicrobial peptides (Janeway and Medzhitov, 2002). These "ready-to-act" mechanisms are the front line of host defense against invading pathogens and act to limit the infection. In addition, the ability of the innate immune system to discriminate between self and non-self and to recognize broad classes of pathogens is essential for the induction of an appropriate adaptive response in vertebrates.

When the innate host defenses are bypassed, evaded or overwhelmed by pathogens, an adaptive response is enlisted in vertebrates. The adaptive immune system is able to generate an enormous variety of cells and molecules capable of specifically recognizing and eliminating an apparently limitless variety of foreign invaders. For this, it relies upon (i) the production of a large repertoire of receptors through somatic gene rearrangement to recognize a huge variety of antigens, and (ii) clonal expansion of antigen-specific effector cells that specifically target the pathogen with (iii) the generation of memory cells that can prevent re-infection with the same microorganism.

been determined *in vitro* (reviewed in Bulet et al., 1999): Drosomycin and Metchnikowin are predominantly antifungal, Attacins, Cecropins, Diptericins and Drosocin are active against Gram-negative bacteria and Defensins are active against Gram-positive bacteria.

Elegant genetic analyses of the induction of antimicrobial peptides have revealed the existence of two distinct pathways, both activating a different set of AMPs (reviewed in Hoffmann, 2003) (Fig. 5). The Toll pathway, originally found to be involved in development (Nusslein-Volhard and Wieschaus, 1980), is required for the induction of the antifungal peptide *Drosomycin* gene, as well as for the resistance to fungal and Gram+ bacterial infections (Lemaitre et al., 1996; Lemaitre et al., 1997). The induction of antibacterial AMPs and the resistance to Gram- bacteria rely predominantly on the Imd (Immune deficiency) pathway (Lemaitre et al., 1995). Our understanding of the cascades of signalling events triggered by microbial infection has progressed dramatically in the past years. Important missing pieces in the puzzle were the "pattern recognition receptors" that recognize infectious non-self and activate the Toll and Imd pathways. The preferential activation of the Toll pathway in response to fungi and Gram+ bacteria, and of the Imd pathway in response to Gram- bacterial infection had led to the hypothesis that flies possess pathogen-specific pattern recognition receptors. This hypothesis was confirmed by the recent characterization of the receptor complex for Gram+ bacteria that activates the Toll pathway. This complex comprises a Peptidoglycan-recognition protein, PGRP-SA, and a Gram negative bacteria binding protein, GNBP1 (Michel et al., 2001; Gobert et al., 2003). Full activation of the Imd pathway requires another receptor complex, a potential component of which, PGRP-LC, has been recently identified (Choe et al., 2002; Gottar et al., 2002; Ramet et al., 2002).

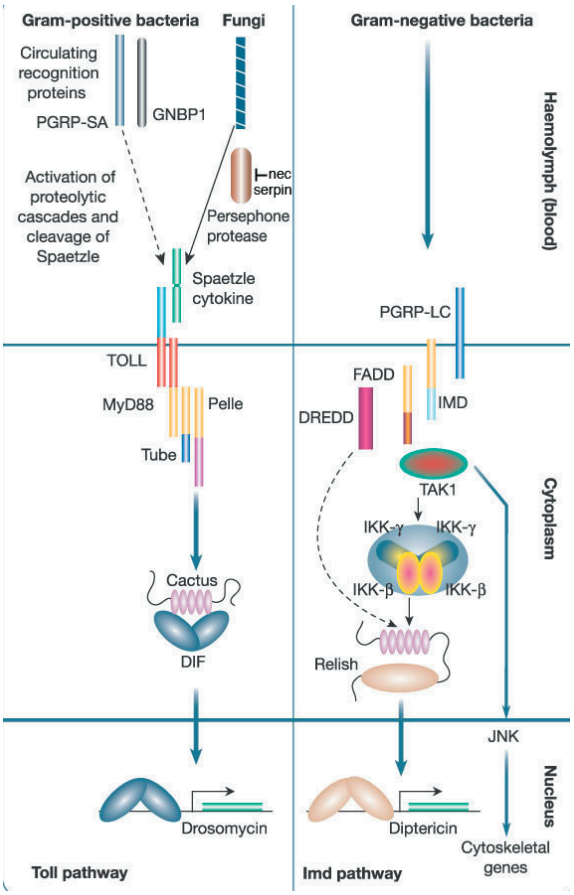


Figure 5. The Toll and Imd pathways in the control of expression of genes encoding antimicrobial peptides. (**Toll pathway**) Recognition of Gram+ bacteria and fungi is mediated by two separate extracellular branches that converge to the proteolytic cleavage of Spaetzle. The signal is transduced into the degradation of Cactus and the nuclear tranlocation of DIF, which activate antifungal and anti-Gram+ AMP genes. (**Imd pathway**) Gram- bacteria are most probably recognized by a receptor complex involving PGRP-LC, which activates two transduction cascades, leading either to the translocation of Relish into the nucleus and activation of anti Gram- AMPs, or to the activation of the expression of cytoskeletal genes through a JNK cascade. (Adapted from Hoffmann, 2003).

Therefore, it appears that the immune system of flies is able to sense infection, to discriminate between various classes of microorganisms and to induce the production of appropriate effector molecules.

Mosquito immunity: taking off?

Until recently our vision of the mosquito immune mechanisms had been shaped by the fly model. Similar types of cellular and humoral defenses had been described at the morphological level. In addition, a large collection of genes potentially implicated in immune responses of *A. gambiae* had been assembled, by cloning of homologues of immune genes of *Drosophila* and other insects, or of differentially expressed genes using substractive libraries. However, their role in mosquito immune responses in general, and especially in those directed against malaria parasites, remained to be confirmed by functional analysis. At the time we started this project, gene silencing by RNA interference (RNAi) had already been reported in several organisms, as well as in *Drosophila* cultured cells. Therefore, we decided to adapt this method to analyze gene function in *A. gambiae.* A short description of molecular mechanisms of RNA interference (reviewed in Denli and Hannon, 2003) is presented in Box 2.

The establishment of new techniques such as RNAi, germline transformation (Catteruccia et al., 2000b; Grossman et al., 2001), large-scale analysis of gene expression by microarrays coupled with the recent completion of the *A. gambiae* genome sequencing (Holt et al., 2002) has already begun to revolutionize the field of mosquito immunity. In a review we published recently (Blandin and Levashina, 2004a) and that can be found on pages 43-49, we summarize what has been learned from previous studies and reevaluate our knowledge on mosquito antiparasitic responses in the view of these recent advances.

4. The family of thioester-containing proteins

For a long time the molecular basis of pathogen recognition in mosquitoes, and also in flies, remained to be dissected. To address this question, we have chosen to study Thioester-containing proteins (TEPs) as putative molecules involved in pathogen recognition and activation of the innate immune responses. Indeed in vertebrates, members of this family play an important role in immune responses as components of the complement system, factors C3, C4 and C5, and as pan-protease inhibitors, α2-macroglobulins (α2Ms).

Complement factors in vertebrates

The complement system has been initially characterized in mammals. It comprises about 35 serum and cell surface molecules that react with one another in cascades to opsonize pathogens and to induce a series of inflammatory responses that help to fight infections. There are three distinct ways through which complement can be activated on pathogen surfaces: the classical pathway,

BOX2: RNA interference

RNA interference was first discovered in plants, and especially in petunias (van der Krol et al., 1990). To enhance a color already displayed by the plant, the pigmentation-encoding gene was cloned and several copies were re-introduced in the plant genome as transgenes to overexpress the gene of interest. However this operation quite often had the opposite effect, causing the formation of pigmentation-free patches in the plant, in which not only the transgene, but also the endogenous locus was silenced. This phenomenon, known in the plant field as Post-Transcriptional Gene Silencing (PTGS), remained unexplained for several years. In 1998, a similar fact, this time named RNA interference (RNAi), was described in *C. elegans*, where the injection of dsRNA in the worm efficiently disrupted the expression of the corresponding endogenous mRNA (Fire et al., 1998). Nowadays the term RNAi regroups all these phenomena as they rely on the same mechanisms. Indeed, although the sources of RNAi are diverse (expression of transgenes, viral RNAs, introduction of dsRNA, etc), they all trigger the same series of events that are catalyzed by homologous enzymes: (i) the formation of dsRNA (ii) which is then cleaved in small interfering RNAs (siRNAs) of 20-25 bp by Dicer ribonucleases. (iii) siRNAs trigger the specific degradation of target single-stranded RNA that contains perfectly complementary sequences, by the RNA-induced silencing complex (RISC).

More recently, it has been recently shown that the same mechanisms/enzymes also play an essential role in regulating gene expression via endogenous microRNAs (miRNAs). These are RNAs of small size encoded by endogenous genes and, as they are partially complementary on both ends, they form a hairpin structure. The annealed region is also recognized by Dicer ribonucleases and cleaved in small dsRNA fragments. miRNAs trigger the degradation of target RNAs that contain sequences perfectly complementary, whereas they block the translation of target RNAs that contain sequences imperfectly complementary. miRNAs are essential in the coordination of fundamental processes such as cell proliferation, cell death, stress resistance or control of development timing, via their ability to regulate the activity of mRNAs (reviewed in Ambros, 2003).

Recently, we have also learned that the RNAi influence extends beyond silencing mRNA to act directly on the genome and repress gene expression at the transcriptional level. DsRNA sharing sequence homology with promoter regions can induce gene silencing via *de novo* methylation of promoter sequences (Mette et al., 2000). In addition, in the fission yeast *Schizosaccharomyces pombe*, dsRNA was shown not only to trigger the degradation of target mRNA, but also to induce the formation of heterochromatin at the target locus via histone methylation (reviewed in Matzke and Matzke, 2003; Schramke and Allshire, 2003) .

The RNAi mechanism has been conserved throughout evolution, in plants, fungi, protozoans and metazoans. In addition to the role of miRNAs in endogenous gene regulation, RNAi might also represent a means of defence against exogenous pathogenic and endogenous parasitic nucleic acids (e.g. viral RNA and selfish DNA, respectively).

the lectin pathway and the alternative pathway. These activation pathways depend on different molecules for their initiation, but they all converge to generate the same set of effector molecules (Fig. 6, and refer to Janeway et al., 2001).

The classical pathway is generally initiated by the binding of the collectin C1q to antibody: antigen complexes, and thus represents a key link between the effector mechanisms of innate and adaptive immunity in vertebrates. However, C1q can also bind directly to the surface of certain pathogens. C1q is part of the C1 complex, with C1r and C1s. Binding of C1q to pathogen surfaces leads to the activation of an autocatalytic enzymatic activity in C1r; the active form of C1r then cleaves its associated C1s to generate an active serine protease. In turn, the active C1s protease cleaves C4 to produce a large fragment, C4b, which binds covalently to the pathogen surface, and a small fragment C4a. The attached C4b recruits C2 and makes it susceptible to cleavage by C1s, producing C2b which is itself an active serine protease, and C2a. The complex C4b,2b, bound to the pathogen surface, forms the C3-convertase.

The lectin pathway uses a protein very similar to C1q to trigger the complement cascade. This protein, the mannan-binding lectin, is a collectin, which recognizes mannose residues and several other sugars accessible on the surface of many pathogens. It forms a complex with MASP1 and MASP2 (MBL-associated serine proteases 1 and 2), which are closely homologous to C1r and C1s. The attachment of the MBL to a sugar allows MASP1 and MASP2 to activate C4 and C2 as in the classical pathway.

In contrast to the classical and lectin pathways, the alternative pathway does not depend on pathogen-binding proteins for its initiation. Instead, it is initiated through the spontaneous hydrolysis of C3, forming C3(H_2O). Factor B can bind to C3(H_2O), which allows the plasma protease factor D to cleave B into Ba and Bb, the latter remaining associated with C3(H_2O) to form the complex C3(H_2O),Bb. This complex is a fluid-phase C3-convertase that can cleave many C3 molecules into C3a and C3b. Much of the C3b formed this way is inactivated by hydrolysis, but some attaches to nearby surfaces of host cells or pathogens, where it can activate the cleavage of factor B by factor D and form the membrane-bound C3-convertase of the alternative pathway: C3b,Bb. Importantly, when C3b binds to the surface of host cells, several regulatory mechanisms prevent further complement activation from proceeding (see below).

The formation of C3-convertases is the point at which the three activation pathways converge to initiate the same subsequent events. The production of C3a and C3b by both C4b,2b and C3b,Bb is very efficient and constitutes an amplification step of the activation pathways. In addition, the alternative pathway can use the C3b molecules produced by the three pathways to bind factor B and create new C3-convertases.

The next step in the complement cascade is the generation of the C5-convertases via the binding of C3b to C4b,2b to yield C4b,2b,3b in the classical and lectin pathways, or to C3b,Bb to yield $C3b_2$,Bb in the alternative pathway. C5 is captured by the C5-convertases and is rendered susceptible to proteolytic cleavage by C2b or Bb, releasing C5a and C5b. C5b recruits complement factors C6, C7 and C8, which allows the complex to insert into the pathogen membrane and induces the polymerization of 10 to 16 molecules of C9 to form a pore in the membrane.

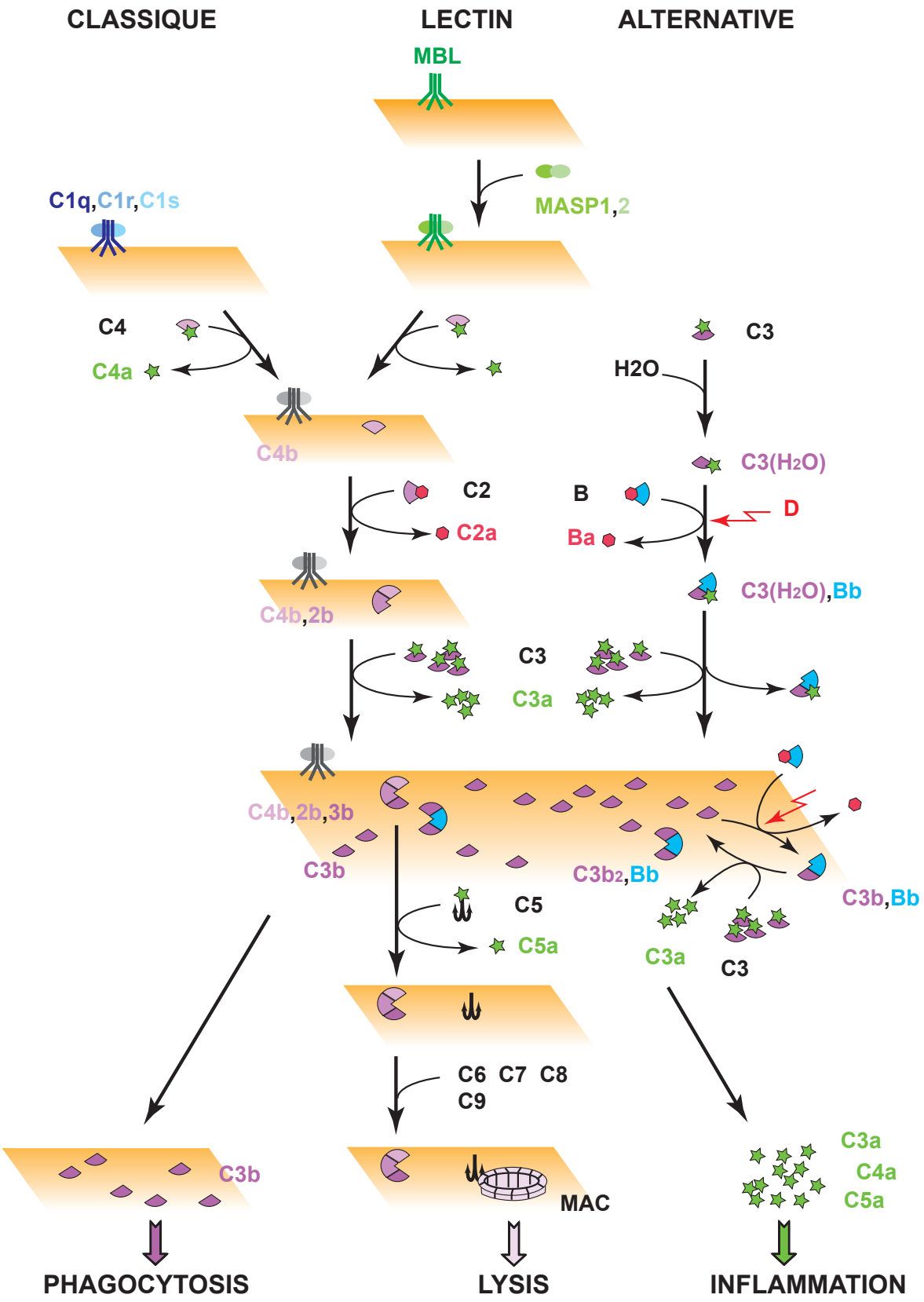


Figure 6. The complement system in action. For explanations, see the main text. Pathogen surfaces are represented in orange. Note that the regulatory mechanisms are not presented on this figure.

Four major functions have been reported for complement:

- The most important, and probably most ancient one, is to facilitate the uptake of pathogens by phagocytes. The many C3b molecules bound on pathogen surfaces promote their phagocytosis via complement receptors (CR) present on phagocytic cells (notably CR1, CR3 and CR4).
- The small fragments C3a, C4a and especially C5a, released upon cleavage of C3, C4 and C5, are known as anaphylatoxins. They are powerful chemoattractants to recruit phagocytes at the site of infection and to activate them. They also increase vascular permeability, hastening the movement of antigen presenting cells to the local lymph nodes, contributing to the prompt initiation of the adaptive immune response.
- The assembly of the membrane attack complex (MAC) allows the lysis of certain pathogens
- Finally, a breakdown product of C3b can bind to CR2, which is part of a B cell co-receptor that augments the signal received through the antigen-specific immunoglobulin receptor.

Complement activation is largely confined to the surface of a pathogen, and the spreading of activated molecules such as C3b and C4b is limited by their covalent attachment to molecules in the immediate vicinity of their site of activation on microbial surfaces via their highly reactive thioester bond (see below). If C3b and C4b do not rapidly form this bond, they are inactivated by hydrolysis. However inadvertent binding of activated complement factors to host cells or spontaneous activation of complement components in the plasma can occur and cause uncontrolled tissue damage but it is normally tightly regulated by several mechanisms which include:

- the inhibition of the protease activity of C1,
- the cleavage of C3b and C4b into inactive products,
- the dissociation of convertase complexes and the prevention of the MAC assembly.

Thus, complement is one of the major immune mechanisms by which pathogen recognition is converted into an effective host defence.

α2-macroglobulins

Most protease inhibitors are highly specific for target enzymes displaying similar specificities and catalytic mechanisms. In contrast, α2Ms inhibit a wide range of proteases and for this, they use a unique steric mechanism. α2Ms can be monomeric, dimeric when composed of two identical subunits and tetrameric when composed of four subunits. They are secreted under an inactive form. Proteolytic cleavage of the α2M induces a conformational change that entraps the attacking protease in a cage. It also exposes their highly reactive thioester bond that can mediate covalent binding to the protease. This does not affect the active site of the protease, which is still able to hydrolyse small substrates, but hinders larger substrates from reaching the active site. Once a protease has reacted with an α2M, the complex is rapidly cleared from the circulation via receptor-mediated endocytosis.

In spite of detailed characterization of their biochemical activities, the biological role of α2Ms is not well understood. The major function of α2Ms is probably to deal with proteases

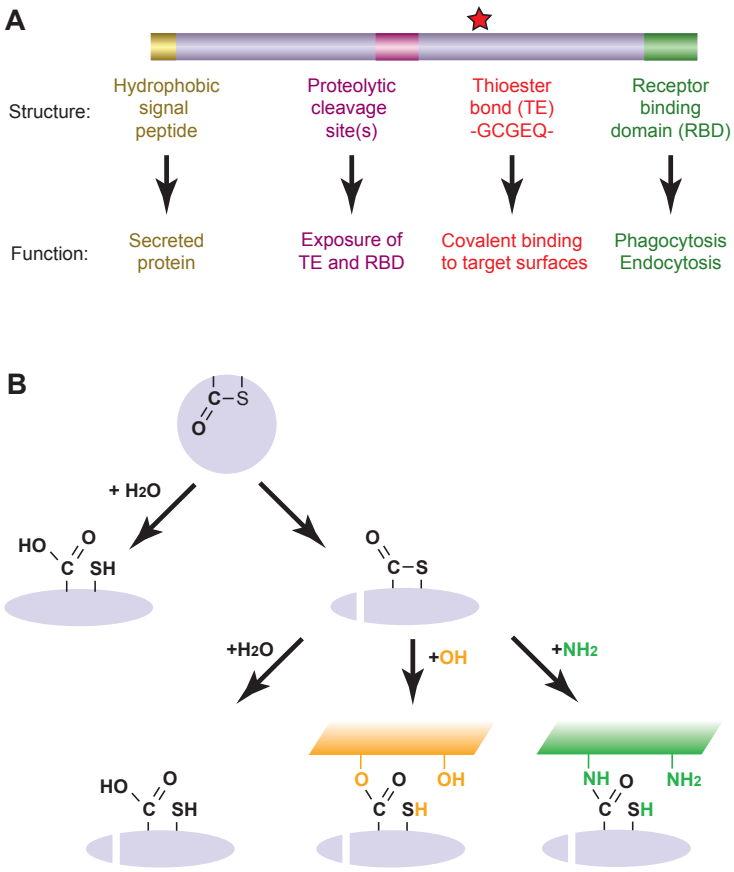


Figure 7. Hallmarks of TEPs and properties of their thioester bond. **A.** Prototype of a Thioester-containing protein. Attention, even with missing keys, your favorite protein might still belong to the family: e.g. complement factor C5 does not have the canonical thioester motif! **B.** Properties of the thioester bond. This highly reactive bond is burried inside the unactive TEP where it can react with small nucleophiles (including methylamine, H_2O). Proteolytic cleavage of the TEP induces a conformational change (symbolized by the change from round to oval) and exposes the thioester bond, which can then react with nearby amino or hydroxy groups. The highest concentration of hydroxyls in the local environment are carried by water molecules, therefore unreacted thioester bond are rapidely inactivated by hydrolysis, limiting the spread of these reactive molecules to more distant sites.

for which no specific inhibitors are present and enzymes that are released in excess of their specific inhibitor. This would include proteases released by tissue damage caused by pathogens and soluble or surface bound proteases produced by invading microorganisms. Consistently with this, several α2M inhibitory proteins were reported to be important virulence factors in bacteria (Kaminishi et al., 1995; Song et al., 2001). Moreover, in salmon stocks tolerant to *Cryptobia* infection, which causes high mortality rates in susceptible fish, the metalloprotease secreted by the parasite has been shown to be neutralized by the fish α2M (reviewed in Woo, 2003). On the other hand, α2Ms associate with a wide range of cytokines and growth factors and regulate their activity and distribution.

Four keys to recognize a thioester-containing protein

α2Ms and complement factors C3/C4/C5 are structurally quite close and probably derive from a common ancestor (reviewed in Dodds and Law, 1998). Several features are specific to the family (Fig. 7). These large proteins (about 1500 amino acids) (i) are secreted under an inactive form. (ii) To become functional, the proteins undergo proteolytic activation, which leads to major conformational changes. Note that complement factors are activated by specific convertases whereas a wide range of proteases can cleave α2Ms. (iii) TEPs bind covalently to nearby targets (pathogens for complement factors and proteases for α2Ms) through the highly conserved hyper-reactive thioester (TE) motif. (iv) Proteolytic activation and binding expose the C-terminal

domain of TEPs leading to clearance of labelled entities via receptor-mediated phagocytosis or endocytosis. Because of these structural and functional constrains, thioester-containing proteins share blocks of similarities that are distributed all over the molecule. Noteworthy, although the name of the family comes from it, some thioester-containing proteins lack the thioester motif. It is the case of complement factor C5 for instance, or of several dimeric or tetrameric α2Ms that cannot bind covalently to attacking proteases but can still trap them as they change conformation upon proteolytic activation.

Evolution of thioester-containing proteins in the animal kingdom

TEPs are not a privilege of vertebrates, members of the family have also been identified in more primitive Deuterostomes and in Protostomes, ranging from *C. elegans* to Arthropods. Components of the complement system have been characterized in all Deuterostomes (reviewed in Nonaka and Yoshizaki, 2004), whereas α2Ms are present in both Protostomes and Deuterostomes. In a review published in 2003 (Levashina et al., 2003), we summarize our knowledge of the TEPs of Protostome invertebrates, with a specific emphasis on the potential of insect models to dissect the primitive functions of this protein family. This review can be found on pages 21-42.

5. Objectives

The main aim of my project was the identification and characterization of thioester-containing proteins in the malaria vector, *Anopheles gambiae*. For that, I had to answer three questions:

Do mosquitoes have thioester-containing protein(s)?
We had good indications that the answer to this question would be yes. Indeed, members of this family had been characterized in Protostomes, including several Arthropods, such a horseshoe crab and a tick. In addition, Dr. M. Lagueux (IBMC, Strasbourg) had also identified 6 genes encoding thioester-containing proteins in the genome of *D. melanogaster*.

How can we analyze the function of these proteins?
Limitations inherent to mosquitoes hinder the traditional methods for gene function analysis, such as large mutagenesis screens, transgenic analysis, and therefore, no tools for functional analysis were available. We decided to adapt the RNAi strategy to knockdown gene expression in mosquitoes and use this technique to analyze mosquito gene function. Again, we had good hope to succeed as RNAi had already been reported in *Drosophila* cultured cells (Hammond et al., 2000).

Are mosquito TEPs involved in immune defences, and especially, in those against malaria parasites?
This was a challenging question! At that time, all characterized thioester-containing proteins in Protostomes had been showed to function as pan-protease inhibitors, they displayed more similarities to α2Ms than to complement factors at the sequence level and their biological function was unknown (Levashina et al., 2003). Moreover six *Teps* had been identified in the fruitfly genome, but no functional characterization had been reported. Some indirect indications from literature suggested that recognition of pathogens in mosquitoes could rely on complement-like reactions: (i) the fact that the activation of melanotic encapsulation reactions was limited to the surface of parasites and (ii) the work reported by Gorman et al, where they had shown that minimal surface characteristics are required for recognition of foreign bodies (Gorman et al., 1998).

As most of the results of our work have been published, in the following part of this manuscript I will present the publications, highlighting the motivations of the work that was conducted and summarizing the results. Finally in the last part of the manuscript, I will provide a synopsis of the obtained results and draw general conclusions.

Review 1

Thioester-containing Proteins of Protostomes

Elena A. LEVASHINA, Stéphanie BLANDIN,
Luis F. MOITA, Marie LAGUEUX, Fotis C. KAFATOS

From Infectious Disease: Innate Immunity
Edited by: R. A. B. Ezekowitz and J. A. Hoffmann
Humana Press Inc., Totowa, NJ

9

Thioester-Containing Proteins of Protostomes

Elena A. Levashina, Stephanie Blandin, Luis F. Moita, Marie Lagueux, and Fotis C. Kafatos

1. INTRODUCTION

The family of thioester-containing proteins (TEPs) appeared early in evolution: members of this family have been found in such diverse organisms as nematodes, insects, molluscs, fish, birds, and mammals *(1)*. They are characterized by homologous sequence features, including a unique intrachain β-cysteinyl-γ-glutamyl thioester, and a propensity for multiple conformationally sensitive binding interactions *(2)*. The presence of the highly reactive thioester bond renders the molecules unstable at elevated temperature and results in their autocatalytic fragmentation at the thioester site *(3,4)*. Moreover, when exposed, the thioester bond is readily hydrolyzed by water. To avoid precocious inactivation, the thioester in the native protein is protected by a shielded environment *(5,6)*. Proteolytic cleavage exposes a previously hidden thioester bond, which mediates covalent attachment through transacylation *(7)*. The reactivity associated with the thioester is one of the defining features of this protein family. Another important feature is the propensity for diverse conformationally sensitive interactions with other molecules. This includes covalent attachment to activating self and nonself surfaces (complement factors), covalent or noncovalent crosslinking to the attacking proteases [α_2-macroglobulins (α_2Ms)], interactions with receptors (complement factors and α_2Ms), and binding of cleavage-generated products to corresponding receptors (anaphylatoxins of complement factors). In addition, α_2Ms bind cytokines and growth factors and regulate their clearance and activity *(8,9)*.

Thioester-containing proteins are characterized by distinct structural forms. In this chapter we propose to use the following nomenclature. Complement factors C3, C4, and C5 are synthesized as single precursor molecules that are intracellularly processed into two- (C3 and C5) or three-*chain* (C4) molecules upon maturation. α_2Ms exist in a monomeric form (single-chain molecule), in a dimeric form, composed of identical *subunits,* and in a tetrameric form, composed of four *subunits*. Heat denaturation causes fragmentation of thioester-containing proteins, resulting in autocatalytic *fragments,* whereas proteolytic activation generates *cleavage products*.

The diverse biochemical activities of the TEPs suggest that they are involved in multiple biologic functions. Indeed, in addition to the well-studied role of the complement

From: *Infectious Disease: Innate Immunity*
Edited by: R. A. B. Ezekowitz and J. A. Hoffman © Humana Press Inc., Totowa, NJ

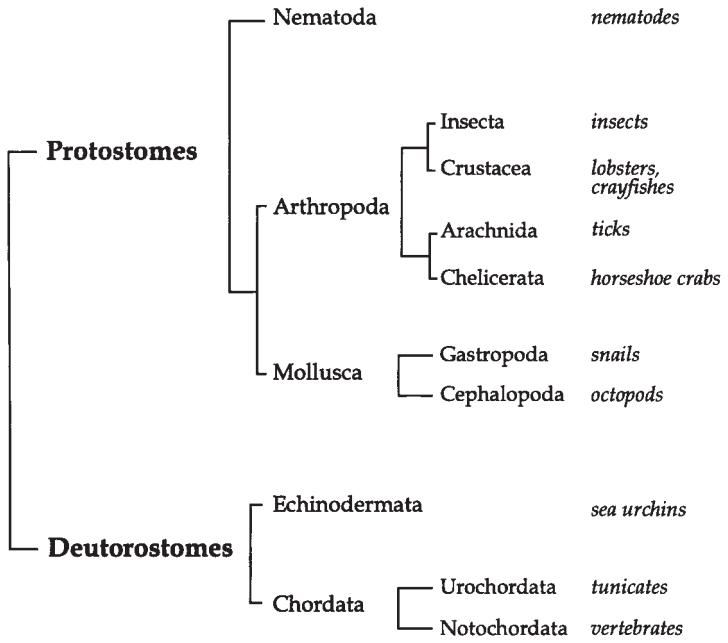


Fig. 1. Phylogenetic relationships between animal species discussed in this chapter.

factors in immune responses, it has been shown recently that C3 is expressed in the regenerating limb blastema cells of urodeles *(10),* suggesting that it plays a role in regenerative processes. Moreover, studies on C5 knockdown mice have demonstrated the importance of C5a in liver regeneration *(11).* Another report suggests that C3 is involved in fertilization, as it is associated with the extracellular matrix of eggs of an amphibian, *Bufo arenarum (12).*

Considerable evidence implicates α_2Ms as important players in several biologic processes, yet the actual mechanisms underlying their influence are mostly unsolved. Decreased concentrations of human α_2M occur in states associated with proteolytic problems, for example, pancreatitis *(13).* On the other hand, α_2Ms associate with a wide range of cytokines and growth factors and regulate their activity and distribution. Generation of mice deficient in either or both mouse α_2M and murinoglobulin-1 (MUG1), the single-chain protease inhibitor of the α_2M type, did not provide new insights into the protein function, as the knockdown mice are phenotypically normal under standard conditions *(14,15).* In experimental conditions, the α_2M knockdown mice were more sensitive to a diet-induced model of pancreatitis than wild-type mice, and this sensibility correlated with both antiprotease and cytokine-binding activities of α_2M. Other important biologic functions of α_2Ms may not have been revealed by these studies.

A thorough understanding of any biologic system requires detailed knowledge of its origin, evolution, and diversity. It is becoming increasingly clear that both complement factors and α_2Ms derive from a common ancestor, prototypes of which can be found in invertebrates going back to Nematodes. Several recent reviews describe primitive complement-like system in tunicates *(16,17)* and sea urchins *(18,19).* Here we summa-

rize current knowledge of the TEPs of protostome invertebrates, with a specific emphasis on the potential of insect models to dissect the primitive functions of this protein family.

2. THIOESTER-CONTAINING PROTEINS OF PROTOSTOME INVERTEBRATES

A spectacular advance in the biochemical methods of protein purification and characterization occurred in the early eighties. One of the challenges of that time was to identify proteins responsible for diverse biochemical activities, e.g., proteases and protease inhibitors, by purification of active protein fractions. Following the original discovery of α_2M antiprotease activity in mammals *(20,21),* this type of analysis was extended to invertebrate species *(22).* All the TEPs of the protostome invertebrates, except for insects, have been purified from protein fractions selected for their α_2M-like antiprotease activity (*see* Fig. 1 for phylogenetic relationships between studied species). Most of these α_2Ms are dimers of two identical subunits; however, tetramers are not a unique feature of vertebrate α_2Ms, since such molecules are present in two species of gastropod molluscs (Table 1). Interestingly, protease-inhibitory activities of protostomal α_2Ms vary in their requirements for a functional thioester bond: most of the α_2Ms from Arthropoda are sensitive to methylamine treatment, whereas in molluscs, methylamine does not interfere with the protease-inhibitory activities. To date, complete sequence information is available only for the *Limulus* α_2M, and the absence of molecular data hampers phylogenetic analysis of TEPs in invertebrates.

2.1. α_2-Macroglobulins in Chelicerata

A TEP of an invertebrate was first detected in the hemolymph of the American horseshoe crab, *Limulus polyphemus,* by Quigley and Armstrong *(23)* in 1983. Up to now it is the best biochemically characterized α_2M in protostomes. This protein is synthesized by hemocytes and secreted into the hemolymph at a concentration of 50 n*M (24),* whereas a minor portion is stored in the large granules of hemocytes, as demonstrated for the sister horseshoe crab species, *Tachypleus tridentatus (25).* Immune or physical stimulation causes blood cell degranulation and α_2M secretion *(26).* Like other TEPs, *Limulus* α_2M is a glycosylated protein with complex oligosaccharide chains *(25).* Its molecular mass is 185 kDa *(27),* and it circulates in the hemolymph mostly as a dimer of two disulphide-linked identical subunits *(28,29).* However, the existence of a minor tetrameric form that dissociates into dimers upon activation has been reported *(30).*

Comparison of the entire mature protein sequence of *Limulus* α_2M with human α_2M and C3 shows 31 and 23% similarities, respectively *(31).* Interestingly, phylogenetic analysis clusters this protein with the clade of α_2M from vertebrates, indicating that the protein has retained the main features of this class of pan-protease inhibitors (Fig. 2). The sequence of *Limulus* α_2M displays a highly conserved thioester motif *(25),* which is functional, as shown by the sensitivity of the native protein to autolytic fragmentation during heat denaturation: mild thermal treatment yielded fragments of 125 and 55 kDa *(32).* Treatment with methylamine, a small nucleophilic molecule that is used to inactivate thioester, rendered the protein resistant to heat denaturation. The interpreta-

Table 1
Structure, Biochemical Activities, and Biologic Functions of Thioester-Containing Proteins in Protostomal Invertebrates[a]

Phylum	Subphylum	Class	Species	Name	Stoichiometry and size (kDa)
Arthropoda	Chelicerata	Merostoma	*Limulus*	L α_2M	2 × 185 *(28)*
			Tachypleus		ND
		Arachnida	*Ornithodoros*	TAM	2 × (92 + 92) *(58)*
	Mandibulata	Crustacea	*Homarus*	Lobster α_2M	2 × 180 *(62)*
			Pacifistacus	P α_2M	2 × 190 *(63)*
			Astacus	*Astacus* α_2M	2 × 185 *(64)*
		Insecta	*Drosophila*	Tep1	ND
				Tep2	ND
				Tep3	ND
				Tep4	ND
				Tep6	ND
			Anopheles	aTEP-I *(31)*	165 *(31)*
Mollusca		Cephalopoda	*Octopus*	Octopus α_2M	2 × 180 *(67)*
		Gastropoda	*Biomphalaria*	Snail α_2M	4 × 200 *(69)*
			Helix	Hp α_2M	4 × 195 *(70)*

ND, not determined; **NA,** nonapplicable; **+TE,** the specified function of the protein requires the thioester bond; α_2M, α_2-macroglobulin TAM, tick α_2M. Numbers in parentheses indicate literature references.

tion that the thioester bond is involved in fragmentation was strengthened by the ability of *Limulus* α_2M to incorporate tritium-labeled glycerol *(28).* Surprisingly, this incorporation was 10-fold higher than for human α_2M but was comparable to the glycerol incorporation seen with human C3. These results were interpreted as indicating that the *Limulus* α_2M and human C3 have similar binding preferences. However, this conclusion is premature and requires direct comparison of binding efficiencies to both labeled glycerol and glycine. Preferential binding to [^{3}H]glycerol is characteristic for human C3, whereas human α_2M is more efficient in incorporating [^{3}H]glycine *(33).* This difference depends on a residue located at a site approximately 100 amino acids downstream of the thioester site *(34).* If this residue is a histidine, preferential binding to carboxyl groups is observed, whereas if it is an asparagine or an aspartic acid (as in many α_2Ms, including *Limulus* α_2M) formation of the amide bonds is favored.

The efficient binding to glycerol and the presence of methylamine-sensitive hemolytic activity in the hemolymph of *Limulus* led to the suggestion that *Limulus* α_2M might have dual properties, combining protease inhibitory and C3-like lytic activities *(28).* The story became more complicated when the inhibitory effect of methylamine treatment on the cytolytic activity was shown to be dependent on the experimental procedure *(35).* Later reports established that α_2M does not directly activate the hemolysis but modulates it through binding to another component of hemolymph, limulin *(36).* Limulin, a member of the pentraxin protein family, is necessary and sufficient to induce lysis of sheep red blood cells by an unknown mechanism

Thioester			Function			
Sequence	Functionality	Catalytic residue	Protease inhibitor	Opsonization	Protein binding	Expression patterns
+*(25)*	+*(32)*	N *(25)*	+TE *(27)*	ND	Limulin*(39)* Coagulin*(54)*	Hemocytes *(26)*
ND	ND	ND	ND	ND	ND	Hemocytes *(25)*
+*(58)*	+*(58)*	ND	+*(58)*	ND	ND	ND
+*(62)*	+*(62)*	ND	+TE *(62)*	ND	ND	ND
+ *(65)*	+ *(63)*	ND	+TE *(63)*	ND	ND	Hemocytes *(66)*
ND	+ *(64)*	ND	+TE *(64)*	ND	ND	ND
+ *(72)*	ND	H *(72)*	ND	ND	ND	Fat body, hemocytes *(72)*
+	ND	D	ND	ND	ND	ND
+	ND	E	ND	ND	ND	ND
+	ND	Y	ND	ND	ND	ND
–	NA	NA	ND	ND	ND	ND
+ *(31)*	+ *(31)*	H *(31)*	ND	+TE *(31)*	ND	Hemocytes *(31)*
+ *(67)*	+ *(67)*	ND	+ *(67)*	ND	ND	ND
ND	+ *(68)*	ND	+ *(68)*	ND	ND	ND
+ *(70)*	+ *(70)*	ND	+ *(70)*	ND	ND	ND

that depends on sialic acid binding *(37,38)*. Native, unreacted α_2M has no effect on the hemolytic activity, but the thioester-reacted forms of *Limulus* α_2M bind limulin and thus indirectly prevent hemolysis *(39)*. Although the role of such modulatory mechanisms in immunity is not clear, vertebrate α_2Ms reportedly bind cytokines and growth factors and modulate their activity *(9)*. Thus, binding of *Limulus* α_2M to a pentraxin may represent another ancient property of these proteins as modulators of biologically reactive molecules.

Historically, protease inhibitory activity was detected in the horseshoe crab hemolymph before α_2M was purified and characterized *(22,23)*. This activity was sensitive to mild acidification and to methylamine treatment and inhibited a wide range of proteolytic enzymes from mammals and bacteria, including trypsin, chymotrypsin, plasmin, elastase, subtilisin, and thermolysin. The inhibited proteases retained the ability to cleave small molecular weight substrates. All these properties are specific to the α_2M family and, consequently, the detected activity was attributed to the presence of an α_2M in the hemolymph. This has been confirmed by purification of the protein *(27,32)* and cloning of the corresponding gene *(25)*.

Inhibition is initiated when the protease cleaves α_2M at a defined domain, the bait region. This domain contains multiple cleavage sites for a wide variety of proteases. It is the most variable part in thioester-containing molecules, even from phylogenetically close species, suggesting the possibility of evolutionary pressure *(40)*. Reaction of *Limulus* α_2M with trypsin generates cleavage products of 100 and 85 kDa, correspond-

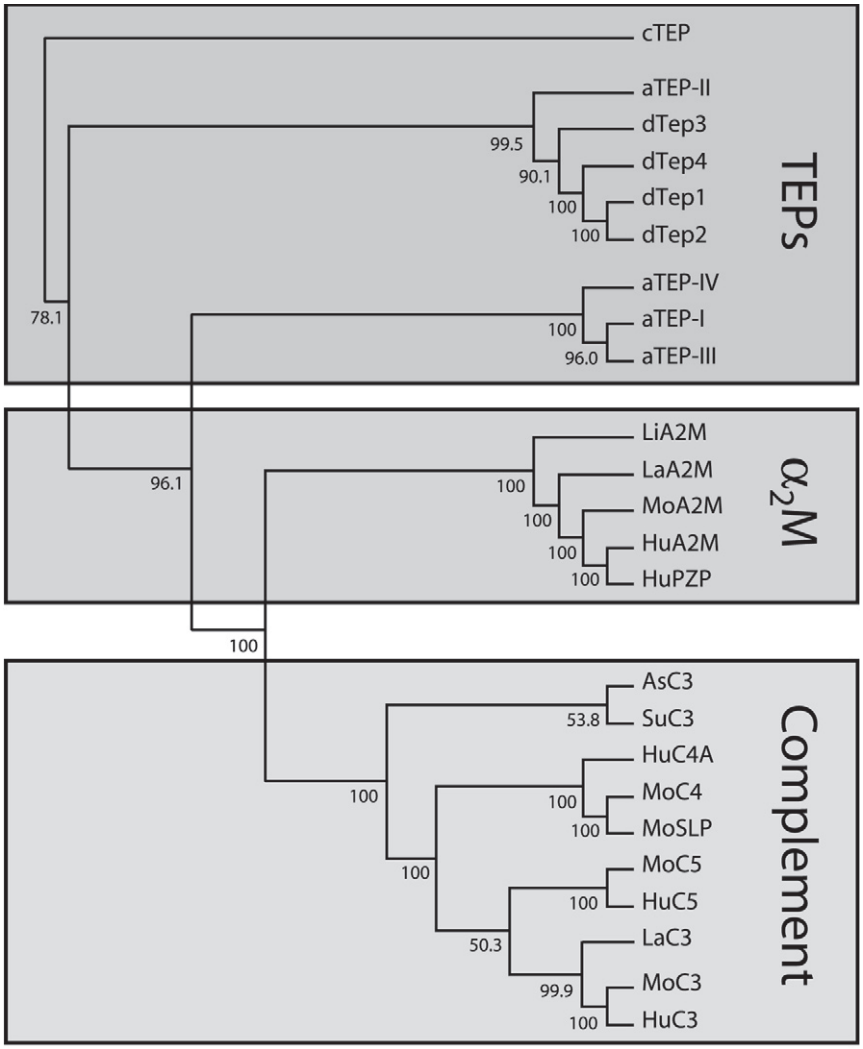


Fig. 2. Phylogenetic tree of thioester-containing proteins (TEPs). The tree is constructed using the neighbor-joining method based on alignment of the sequences using CLUSTAL_X. Numerals on the branches are bootstrap percentages to support the given partitioning. Clades of complement factors (Complement), α_2,-macroglobulins (α_2M), and insect/nematode TEPs are boxed.

ing to cleavage in the bait region, as well as three additional products with molecular masses greater than the native polypeptide *(41).* Inactive proteases are not bound by *Limulus* α_2M, indicating that it has to be proteolytically activated before it can bind and exert its inhibitory activity *(42).* The proteolytic activation of α_2M leads to its dramatic conformational change, causing the entrapment of the attacking protease *(21).* Entrapment may or may not involve covalent binding through the amide bond formed between the activated thioester and amino groups of lysyl residues in the protease *(40).* α_2M of *Limulus* entraps proteases mainly using the noncovalent mechanism. Interestingly, the "trapping" process is sensitive to methylamine; unlike the earlier cleavage of

α_2M in the bait region, the functional thioester crosslinks two α_2M subunits rather than the attacking protease. An isopeptide bond is formed between Lys254 of one α_2M subunit and the Glu residue of the thioester site in the second subunit *(43).* It is believed that intersubunit crosslinking stabilizes the altered conformational structure of the reacted α_2M *(44).* Only 1% of the attacking protease binds covalently to the *Limulus* α_2M *(28,41).* It is important to note that both trypsin and methylamine treatment result in extensive compaction of the molecule, as evidenced by electron microscopy *(28,29).* Such compaction results in a change of electrophoretic mobility of α_2Ms, a transition referred to as the slow to fast transformation *(45).*

Proteolytic activation and protease entrapment exposes a receptor binding domain of α_2M *(46).* This leads to receptor-mediated endocytosis of the α_2M-protease complexes from the circulation and their degradation in secondary lysosomes *(40).* Clearance of *Limulus* α_2M from the hemolymph was evidenced by injecting the fluorescein-labeled proteins into the heart lumen and measuring fluorescence in the plasma and in the hemocytes *(42).* Trypsin and α_2M-trypsin complexes were rapidly cleared from the hemolymph, in contrast to unreacted α_2M and the oxygen carrier hemocyanin, which remained in the circulation *(42).* Blood cells bound the fluorescein isothiocyanate-labeled trypsin-reacted α_2M, demonstrating their involvement in α_2M clearance.

A wealth of biochemical data suggests that α_2M may play an important role in proteolytic homeostasis of *Limulus.* Interestingly, this pan-protease inhibitor does not inhibit the key serine proteases, which activate the coagulation cascade of the hemolymph in both *Limulus (47)* and the closely related species, *Tachypleus tridentatus (25).* The coagulation cascade in the Japanese horseshoe crab has been well characterized at the molecular level *(48).* Constituents of microbial cell walls, such as lipopolysaccharides and β-1,3-glucans, activate a cascade involving four serine proteases that are normally kept in check by endogenous serpins *(49–51).* The proteolytic reaction culminates in the cleavage of soluble coagulogen into insoluble coagulin and the formation of a fibrillar extracellular clot *(52,53).* Surprisingly, *Limulus* α_2M was found associated with the coagulin clot in a Ca^{2+}-dependent fashion *(54).* The biologic role of these specific associations is not clear, but accumulation of the inhibitor may suppress degradation of the clot by microbial proteases *(54).* In addition, the coagulation cascade triggers the activation of another immune reaction—the prophenol oxidase cascade *(55).* Two clip-domain serine proteases of the coagulation cascade and an antimicrobial peptide, tachyplesin, convert hemocyanin to phenol oxidase, thus localizing the phenol oxidase activity to the site of infection *(56,57).* Therefore the coagulation system in the horseshoe crab appears to be quite a sophisticated system. Invading microorganisms are locally trapped into a gelatinous clot, where their escape is prevented through inhibition of microbial proteases by high concentrations of α_2M; antimicrobial peptides (big defensin, tachyplesins, anti-LPS factor) then break microbial walls; and, finally, the *champ de bataille* is isolated from the healthy surroundings by a melanotic barrier.

Horseshoe crabs represent a convenient model for the biochemical analysis of these proteins, because large volumes of plasma and other biologic materials are readily available for analysis. However, a deeper understanding of biologic function requires the analysis of specific mutant phenotypes, which is not readily applicable to these live fossils.

2.2. α_2-Macroglobulins in Arachnida

A recent report identified α_2M as one of the major hemolymph glycoproteins in the soft tick, *Ornithodoros moubata (58).* Although the complete protein sequence is unknown, partial sequencing confirmed that this protein is a new member of the family of TEPs; hence it was named TAM for tick α_2-macroglobulin. The structure of TAM is of particular interest, as it represents the first two-chain invertebrate α_2M in which two chains are held together by disulphide bridges. Such posttranslational processing of a processor protein is a specific feature of the complement proteins C3, C4, and C5 *(59).* However, a C3-like two-chain structure also exists in vertebrate α_2Ms from the plaice *Pleuronectes platessa (60)* and the carp *Cyprinus caprio (61).* The native 400-kDa TAM is a dimer of two two-chain subunits. In contrast to *Limulus* α_2M, the TAM subunits are associated by bonds that are sensitive to sodium dodecyl sulfate polyacrylamide gel electrophoresis even in nonreducing conditions *(58),* indicating that they are not disulphide or other covalent bridges. Purified TAM inhibits trypsin and thermolysin, and this activity is sensitive to methylamine treatment. The TAM-reacted trypsin is protected by steric interference against high molecular weight protease inhibitors, such as soybean trypsin inhibitor (SPI); it is still able to cleave low molecular weight substrates in the presence of the inhibitor. Further studies are needed to confirm the presence and functionality of the thioester bond in TAM and to clarify the mechanism of protease entrapment.

2.3. α_2-Macroglobulins in Crustacea

Although α_2M activity has been reported in several crustacean species *(44),* the proteins have been purified only from the hemolymph of the American lobster *Homarus americanus (62)* and the crayfish *Pacifastacus leniusculus (63)* and *Astacus astacus (64).* Although the full-length protein sequences are not available for these proteins, the N-terminal amino acid sequences are strikingly similar *(64).* In contrast, the similarity to the N-terminal sequence of the *Limulus* α_2M is less pronounced. All three proteins are dimers of two identical disulphide-linked subunits. Purified protein extracts inhibit trypsin proteolytic activity toward high molecular weight substrates but do not affect proteolysis of small molecular weight substrates. This inhibitory activity is sensitive to methylamine treatment. Consistently, the presence and functionality of the thioester bond was established by partial sequencing of the thioester-containing region in *Pacifastacus (65)* and *Homarus (62).* Heat denaturation results in autolytic fragmentation of α_2Ms in all three species. It is unknown whether these proteins bind covalently to proteases or form intersubunit disulphide bridges like *Limulus* α_2M. The defense and recognition reactions of arthropods are to a great extent carried out by the blood cells (hemocytes). In *Pacifastacus,* α_2M is almost exclusively present in the hemocytes *(66)* and is continuously secreted into the plasma reaching the concentration of about 0.2 mg/ml *(65).* Although a possible function of α_2M in regulating crayfish prophenol oxidase and/or clotting systems has been proposed *(65),* the role of these proteins in Crustacea remains to be established.

2.4. α_2-Macroglobulins in Mollusca

The first α_2M protein identified in molluscs was purified from the hemolymph of a cephalopod, *Octopus vulgaris (67).* The native glycoprotein forms a 360-kDa disul-

phide-bonded homodimer and inhibits three catalytic classes of proteases, serine-, metallo-, and cysteine proteases *(67).* Like other α_2Ms, the inhibitor does not interact with the active site of the proteases, but sterically shields it so that only protease inhibitors of molecular weight lower than 7 kDa are able to react with the encaged trypsin. Interestingly, proteolytic activation of octopus α_2M does not alter the apparent Stockes radius and consequently is based on a type of conformational change different from that in α_2M from *Limulus,* or on an entirely different mechanism. In addition, the protease entrapment does not require covalent binding *(67).* The functionality of the conserved thioester site was not addressed in this study, and thus the possibility that octopus α_2M does not have a reactive thioester bond remains open.

All the invertebrate thioester-containing molecules discussed above form predominantly homodimers, and only a minor fraction of *Limulus* α_2M was shown to exist as a tetramer prior to proteolytic activation *(30).* The first tetrameric α_2M in invertebrates was purified from the hemolymph of the tropical planorbid snail, *Biomphalaria glabrata (68,69).* The snail α_2M is similar to human α_2M in that the denatured molecule is composed of two subunits that form a disulphide-bonded dimer, whereas the native molecule consists of four subunits, each with a molecular mass of around 200 kDa *(69).* Proteolytic activation of the snail inhibitor produces two cleavage products of 100–105 kDa, indicating that the bait region is located near the middle of the subunit. The snail α_2M is a glycosylated protein that displays wide-range protease-inhibiting activity. It exhibits the characteristic methylamine-sensitive autocatalytic fragmentation of TEPs. However, as in the octopus, methylamine treatment does not induce dramatic conformational change and results in only a slight decrease in the protease inhibitory activity *(69).* However, the native molecule undergoes a slow to fast transition after complexing with trypsin, indicating that the snail α_2M uses a trapping mechanism similar to that of octopus α_2M, in which protease is encaged by conformational changes caused by cleavage of the bait region. The protease entrapment is independent of thioester activation and does not require covalent binding to the protease nor covalent crosslinking of the subunits.

Recently an α_2M has been purified from the hemolymph of another gastropod mollusc, *Helix pomatia (70).* The structure and inhibitory activities of the molecule are closely related to the above-discussed snail α_2M in that the native form consists of four subunits and shows protease inhibitory activity that is independent of the functional thioester bond. Thus the tetrameric forms of α_2M evolved quite early in evolution, and deeper characterization of these molecules could provide insight into the selective pressures that shaped their evolution.

3. THIOESTER-CONTAINING PROTEINS IN DIPTERAN INSECTS

Insects are able to mount a rapid and efficient response when confronting various microorganisms. This response is reminiscent of innate immune defenses of vertebrates and, because insects lack an adaptive immune response, it represents a valuable model to study the "ante-antibody" immunity *(71).* Characterization of immune responses in the fruitfly *Drosophila melanogaster* spearheads the research in the field owing to the exceptional experimental possibilities offered by this insect, including a myriad of molecular and genetic tools and the availability of the complete genome sequence.

Until recently, the major human malaria vector, the mosquito *Anopheles gambiae,* had few advantages, but in recent years this species has become an attractive insect model to study innate immunity. The forthcoming completion of the *A. gambiae* genome sequencing project will provide the first model of a parasitic disease in which full genome sequence information will be available for all three constituents: *Plasmodium falciparum* (the causative agent of human malaria), *Homo sapiens* (the host), and *A. gambiae* (the vector). The interest of this model is highlighted by the fact that the protozoan parasite has to evade two related but distinct immune systems, those of the vertebrate host and those of the mosquito. Unraveling the mechanisms that underlie mosquito-parasite interactions, which is itself of substantial interest, may provide a simplified model for understanding some aspects of host-parasite interactions.

All the TEPs discussed above were purified as active protein fractions that exhibited protease inhibitory activity. In insects, the TEPs were identified by molecular cloning *in silico (D. melanogaster),* or by homology cloning *(A. gambiae)* that was guided by the evolutionary conservation of the thioester region.

3.1. Thioester-Containing Proteins in Drosophila melanogaster

A BLAST search of the Berkeley *Drosophila* Genome Sequencing Project Database with the amino acid sequence of the α-chain of human complement factor C3 produced six hits. All of them showed significant sequence similarities and were named Teps for *t*hio*e*ster-containing *p*roteins. Four genes of six contain a conserved thioester motif, whereas *Tep5* and *Tep6* display a modified sequence (*72,73;* M.L., unpublished data). One gene, *Tep5,* is present only in genomic sequences and is not discussed here. Full-length sequences are now available for *Tep1, Tep2, Tep3,* and *Tep4 (72)* and were used in phylogenetic analysis (see below). Three specific protein features are characteristic of four *Drosophila* Teps: (1) a highly conserved region of 30 amino acid residues harboring a canonical thioester motif GCGEQ; (2) a distinctive cysteine signature encompassing 126 residues in the C-terminal part of the molecule; and (3) a highly variable central region of about 60 residues in length. Interestingly, in *Tep2* this variable region is encoded by five distinct exons, which produce distinct transcripts as a result of alternative splicing *(72).* Should *Tep2* code for a protease inhibitor, alternative splicing of the bait region may extend the number of inhibited proteases. It may also represent a novel mechanism for recognition of noxious structural patterns in the absence of the large repertoire of receptors of the adaptive immune response in vertebrates.

Interestingly, so far alternative splicing had not been reported in TEPs. Such a mechanism is often employed by insects to produce closely related proteins with distinct properties without dramatically enlarging the size of the genome. The best characterized example is a serine protease inhibitor (serpin) from a tobacco hornworm, *Manduca sexta,* in which alternative splicing generates 12 serpin isoforms, each of which has a distinct protease inhibitory activity *(74).* The distribution of the *Tep2* splice isoforms is unknown, but it will be of great interest to learn the temporal and tissue-specific patterns of particular isoforms. All Teps are located on the left arm of the second chromosome. Interestingly, *Tep2* and-*3* form a discrete cluster in which genes are oriented head to head and are separated by 1.5 kb of putative regulatory sequences (unpublished data). This clustering may result from a recent duplication-inversion event. Surprisingly, according to phylogenetic analysis of sequences, all inducible Tep proteins 1, 2,

and 4 (see below) cluster together, whereas the constitutively expressed Tep3 is somewhat divergent in sequence (Fig. 2).

TEPs are often acute-phase reactants *(40,75).* In *Drosophila,* the expression of *Tep1,-2,* and-*4* is strongly upregulated by bacterial challenge in larvae. During the adult stage, immune challenge markedly induces *Tep2* and-*4,* whereas the expression of *Tep1* is minor. To date, the expression pattern of *Tep1* in larvae is best characterized. This gene is mainly transcribed in the fat body after immune challenge, but it is also expressed in hemocytes in naive and challenged larvae. No expression of *Tep1* is detected in the *Drosophila l(2)mbn* and S2 Schneider cell lines.

The induction of immune responses in *Drosophila* is controlled by two distinct signaling pathways: the Toll pathway is primarily responsible for expression of the antifungal peptide gene *Drosomycin,* and the Imd pathway controls expression of most antibacterial peptide genes *(76,77).* Somewhat surprisingly, the inducible expression of *Tep1* is not controlled by either of these pathways *(72).* An indirect effect of the Toll pathway on *aTep1* expression is observed in *Toll* gain-of-function mutants and might be associated with the characteristic aggregation of blood cells into masses that tend to become melanized, a process referred to as melanotic tumor formation, in these mutants *(78).* It has been suggested that Toll-induced melanotic tumor formation is mediated by the JAK kinase Hopscotch *(79).* Remarkably, *Hopscotch* gain-of-function mutants constitutively express high levels of *Tep1.* Moreover, in *Hopscotch* loss-of-function mutants, the inducible expression of the gene dramatically levels off, suggesting a direct regulation of *Tep1* expression by the JAK pathway. Consistently, constitutive expression of the gene in *Toll* gain-of-function mutants is abolished in the *Hopscotch* loss-of-function background. It will be of interest to determine whether Toll-induced melanotic tumor formation is also abolished in the *Toll* gain-of-function/*Hopscotch* loss-of-function double mutants. If so, Tep1 may be involved in the aggregation of blood cells and the localized induction of melanization. Analysis of loss-of-function *Tep* mutants will provide crucial information on the role of this protein family in fly immunity and perhaps uncover new functions.

3.2. Thioester-Containing Proteins in Anopheles gambiae

A gene encoding a TEP, *aTEP-I,* has been cloned and characterized in detail in *A. gambiae (31).* Recent evidence suggests the presence of at least three more members of the family: aTEP-II (E.A.L., L.F.M., and S.B., unpublished data), aTEP-III *(80),* and aTEP-IV (IMCR-14) *(81).* Expression of all these molecules is transcriptionally upregulated by bacterial challenge and by the *Plasmodium berghei* infectious blood meal (*81;* E.A.L., unpublished data).

aTEP-I is a typical representative of TEPs in the mosquito *A. gambiae.* The amino acid sequence deduced from a cDNA clone codes for a protein of 150 kDa. It contains a signal peptide-like hydrophobic N-terminal segment characteristic of secreted proteins. The sequence contains the canonical thioester motif, which is followed 100 amino acids downstream by a catalytic histidine residue. The most C-terminal part displays a cysteine signature, which is characteristic of the *Drosophila* Teps (see above). The protein is glycosylated and secreted into the hemolymph by mosquito hemocytes. In the hemolymph, the glycosylated aTEP-I is present as a full-length form of approximately 165 kDa and as a smaller fragment of 80 kDa *(31),* indicating a constant low

level of proteolytic cleavage of the protein. Interestingly, the same type of cleavage can be induced by both wounding and bacterial challenge. It is unknown whether cleavage occurs during activation or subsequent inactivation of the molecule.

Biochemical analysis and protein purification in the mosquito are hampered by the small size of the animal and the nanoliter quantities of its hemolymph. To overcome these limitations, functional studies on aTEP-I were performed using a mosquito cell line established by H.-M. Mueller *(82,83).* Mosquito cells in vitro secrete aTEP-I into the conditioned medium, where it can be readily detected by affinity-purified rabbit polyclonal antibodies that recognize the full-length and the C-terminal fragment of the molecule *(31).* All evidence obtained so far suggests that aTEP-I is secreted as a single-chain molecule, but only purification of the native protein will provide an ultimate proof.

The functionality of the thioester bond in aTEP-I is supported by experiments on denaturation-dependent autocatalytic fragmentation. Heat treatment of the conditioned medium leads to the appearance of a new C-terminal fragment of 50 kDa, resulting from fragmentation of the molecule at the thioester site. Methylamine treatment completely prevents this autocatalytic fragmentation *(31).*

Cultured mosquito cells show phagocytic activity against latex beads and bacteria *(84).* Interestingly, this activity can be enhanced by pretreating Gram-negative *Escherichia coli* with the conditioned cell medium, before exposing the bacteria to the cells *(31).* Double-strand RNA knockdown of aTEP-I and methylamine treatment of the conditioned medium both decrease the efficiency of bacteria uptake by 50%. These experiments revealed for the first time that a thioester protein of a protostome has an ancient complement-like function of opsonising Gram-negative bacteria for phagocytosis *(31).* In the same experiments, the efficiency of phagocytosis of Gram-positive bacteria was shown to be very low and independent of opsonization by aTEP-I. Interestingly, aTEP-I opsonizes both Gram-positive and Gram-negative bacteria, binding to them in a thioester-dependent manner *(31).* Therefore, the low uptake of Gram-positive bacteria by the mosquito cells is not caused by inefficient opsonization by aTEP-I.

A similar phenomenon has been described for the opsonization and phagocytosis of group A streptococci by polymorphonuclear leukocytes (PMNs) *(85).* Streptococcal strains bearing M-surface antigens were totally resistant to phagocytosis by PMNs, whereas the M^- strains were susceptible in a C3-dependent fashion. Both bacterial strains were opsonized by C3b activated by the alternative pathway. It was proposed that in the M^+ strains, the rigid α-helical M-protein blocks or physically hinders the receptor binding site of attached C3b, making it inaccessible for interaction with the leukocyte receptors *(85).* It will be interesting to compare the phagocytic activity of the mosquito cells in culture with that of native hemocytes. This will allow us to discriminate between two broader explanations: (1) cultured cells may lack a component of the Gram-positive bacteria phagocytic receptor complex; or (2) the structure of the Gram-positive bacteria cell wall does not allow the opsonin to interact with the phagocytic receptors. In either case, the mosquito cell culture system provides a convenient model to address these questions.

Three additional genes encoding TEPs have been identified in *A. gambiae* thus far. In the phylogenetic tree, aTEP-I, aTEP-III, and aTEP-IV form a conspicuous cluster,

suggesting that this cluster may result from a recent duplication event (Fig. 2). Future studies on the structure and function of all four aTEPs will provide valuable information pertinent to the broader question of functional evolution of this protein family. It is already clear that the TEPs represent a multimember family in both the fruitfly and the mosquito, probably reflecting the evolutionary importance of this family in the biology of insects. It is unknown as yet whether each aTEP has distinct functions, or whether these primitive molecules share common functions, possibly performing both protease inhibition and opsonization.

Of special interest is the role of aTEPs in the interactions between *Plasmodium* and the mosquito. The malaria parasite performs only part of its life cycle in the mammalian host, where it reaches and invades hepatocytes, then enters the circulation, and multiplies in the red blood cells. The remaining development of the parasite, including its sexual cycle, obligatorily occurs in the mosquito. During its complex passage through both the vertebrate host and the invertebrate vector, it now appears that the parasite repeatedly faces attacks from complement-like systems that represent more than 400 million years of evolutionary diversification. Interestingly, mosquito stages of parasite development in vitro are less resistant to complement than vertebrate blood-stage parasites *(86,87),* indicating that some stage-specific factors may protect *Plasmodium* from the complement system in vivo. Dissection of the interactions between aTEPs and the parasite may provide new insights into the mechanisms of complement-parasite interactions.

5. CONCLUSIONS AND PERSPECTIVES

The mid-1980s evidenced an explosion of studies on the TEPs, especially on α_2Ms. In particular, several invertebrate species were analyzed for α_2M-like protease inhibitory activities, followed by purification and biochemical characterization of the relevant molecules. Animal size and the volume of hemolymph were important in selection of species for biochemical analysis. Big invertebrate species like the horseshoe crabs, crayfishes, octopuses, lobsters, and others discussed above were instrumental in demonstrating the presence of α_2M in protostomes and in revealing the amazing diversity of α_2M structural forms. Only dimeric α_2Ms were initially identified in invertebrates, leading to the belief that tetrameric forms are restricted to higher vertebrates. Recently tetrameric α_2Ms were characterized in two species of gastropod molluscs, and it is likely that insect TEPs are monomeric. Thus, monomeric, dimeric, and tetrameric TEPs might be widespread among animals. The protease inhibitory activity of tetrameric molecules appears to be less sensitive to methylamine than that of dimeric or putative monomeric molecules, in both invertebrate and vertebrate α_2Ms. The mechanism of protease entrapment in invertebrates has been characterized only for *Limulus* and represents a rather peculiar example, whereby the thioester forms a covalent bond with the opposite subunit of α_2M and not with the attacking protease.

Interestingly, not all thioester-containing molecules in invertebrates actually contain the thioester motif, e.g., *Drosophila* Tep6 lacks it. In vertebrates, complement factor C5 and ovostatin also lack this important protein signature, which would suggest that the thioester motif has been lost repeatedly during animal evolution. The overall picture suggests that the TEPs in the animal phyla undergo dramatic selective pressures that result in a set of polypeptides specific for each species. The nature of these selec-

tive pressures is unknown and can only be imagined when the functions of the proteins are more fully characterized.

Table 1 summarizes our present knowledge of the structure, biochemical activities, and biologic functions of TEPs in protostomal invertebrates. Obviously, this field of research is only at an early stage of development, and "not determined" dominates the table. The recent discovery of TEPs proteins in *Drosophila* will certainly play a key role in understanding the function(s) of this protein family. The immune-inducible expression of *Tep1, 2,* and-*4* suggests that *Drosophila* Teps may play a role in immune responses in the fruitfly. Mutant analysis may confirm this but may also reveal unexpected new functions. The mosquito represents a complementary model for functional analysis of this protein family. Cultured mosquito cells have provided a first quantitative phagocytic test and can potentially serve to dissect the molecules involved in regulation of phagocytosis. Another advantage of cell culture is the unlimited quantity of material available for biochemical studies; it will prove invaluable for purification of aTEPs and subsequent structure-functional analysis of the protein. Most importantly, the mosquito represents an elegant model for studying the role of TEPs during the response to parasitic infections. The malaria parasite faces the complement system and α_2Ms in the vertebrate host, as well as aTEPs in the mosquito, and is still able to complete its life cycle. Understanding the mechanisms of parasite evasion from the mosquito TEPs may extend our knowledge of parasite resistance to the mammalian complement system and to α_2M, with potential implications for the treatment of human disease.

Current data suggest that the TEPs in protostomes are mostly represented by α_2Ms (Table 1). However, this idea should be treated with caution, as most of the invertebrate proteins were purified as extracts that displayed α_2M-like protease inhibitory activity. Insect TEPs are the first example of a genomics-based approach for unbiased study of these proteins. Future systematic investigations on the invertebrate TEPs are very likely to break up the accepted view of a rigid separation of this family into complement factors and α_2M. This will open new perspectives for understanding the fascinating diverse and (probably) multifunctional TEP family.

ACKNOWLEDGMENTS

The authors thank Prof. Jules A. Hoffmann for critical reading of the manucript. Work in our laboratories has received support from the National Institutes of Health, grant 1PO1 AI44220, and from the European Commission Human Potential—Research Training Networks Programme, grant HPRN-CT-2000-00080.

REFERENCES

1. Nonaka M. Origin and evolution of the complement system. In: Du Pasquier L, Litman GW (eds.). Origin and Evolution of the Vertebrate Immune System, vol 248. New York: Springer-Verlag, 2000, pp. 37–50.
2. Chu CT, Pizzo SV. Alpha 2-macroglobulin, complement, and biologic defense: antigens, growth factors, microbial proteases, and receptor ligation. Lab Invest 1994;71:792–812.
3. Harpel PC, Hayes MB, Hugli TE. Heat-induced fragmentation of human alpha 2-macroglobulin. J Biol Chem 1979;254:8669–8678.

4. Howard JB. Methylamine reaction and denaturation-dependent fragmentation of complement component 3. Comparison with alpha2-macroglobulin. J Biol Chem 1980;255:7082–7084.
5. Salvesen GS, Sayers CA, Barrett AJ. Further characterization of the covalent linking reaction of alpha 2-macroglobulin. Biochem J 1981;195:453–461.
6. Sim RB, Sim E. Autolytic fragmentation of complement components C3 and C4 under denaturing conditions, a property shared with alpha 2-macroglobulin. Biochem J 1981;193:129–141.
7. Law SK, Dodds AW. The internal thioester and the covalent binding properties of the complement proteins C3 and C4. Protein Sci 1997;6:263–274.
8. Lysiak JJ, Hussaini IM, Webb DJ, et al. Alpha 2-macroglobulin functions as a cytokine carrier to induce nitric oxide synthesis and cause nitric oxide-dependent cytotoxicity in the RAW 264.7 macrophage cell line. J Biol Chem 1995;270:21919–21927.
9. Webb DJ, Wen J, Lysiak JJ, et al. Murine alpha-macroglobulins demonstrate divergent activities as neutralizers of transforming growth factor-beta and as inducers of nitric oxide synthesis. A possible mechanism for the endotoxin insensitivity of the alpha2-macroglobulin gene knock-out mouse. J Biol Chem 1996;271:24982–24988.
10. Del Rio-Tsonis K, Tsonis PA, Zarkadis IK, Tsagas AG, Lambris JD. Expression of the third component of complement, C3, in regenerating limb blastema cells of urodeles. J Immunol 1998;161:6819–6824.
11. Mastellos D, Papadimitriou JC, Franchini S, Tsonis PA, Lambris JD. A novel role of complement: mice deficient in the fifth component of complement (C5) exhibit impaired liver regeneration. J Immunol 2001;166:2479–2486.
12. Llanos RJ, Whitacre CM, Miceli DC. Potential involvement of C(3) complement factor in amphibian fertilization. Comp Biochem Physiol A Mol Integr Physiol 2000;127:29–38.
13. McMahon MJ, Bowen M, Mayer AD, Cooper EH. Relation of alpha 2-macroglobulin and other antiproteases to the clinical features of acute pancreatitis. Am J Surg 1984;147:164–170.
14. Umans L, Serneels L, Overbergh L, et al. Targeted inactivation of the mouse alpha 2-macroglobulin gene. J Biol Chem 1995;270:19778–19785.
15. Umans L, Serneels L, Overbergh L, Stas L, Van Leuven F. Alpha2-macroglobulin- and murinoglobulin-1-deficient mice. A mouse model for acute pancreatitis. Am J Pathol 1999;155:983–993.
16. Nonaka M. Origin and evolution of the complement system. Curr Top Microbiol Immunol 2000;248:37–50.
17. Nonaka M. Evolution of the complement system. Curr Opin Immunol 2001;13:69–73.
18. Smith LC, Azumi K, Nonaka M. Complement systems in invertebrates. The ancient alternative and lectin pathways. Immunopharmacology 1999;42:107–120.
19. Smith LC, Clow LA, Terwilliger DP. The ancestral complement system in sea urchins. Immunol Rev 2001;180:16–34.
20. Starkey PM, Barrett AJ. Inhibition by alpha-macroglobulin and other serum proteins. Biochem J 1973;131:823–831.
21. Barrett AJ, Starkey PM. The interaction of alpha 2-macroglobulin with proteinases. Characteristics and specificity of the reaction, and a hypothesis concerning its molecular mechanism. Biochem J 1973;133:709–724.
22. Starkey PM, Barrett AJ. Evolution of alpha 2-macroglobulin. The demonstration in a variety of vertebrate species of a protein resembling human alpha 2-macroglobulin. Biochem J 1982;205:91–95.
23. Quigley JP, Armstrong PB. An endopeptidase inhibitor, similar to mammalian alpha 2-macroglobulin, detected in the hemolymph of an invertebrate, *Limulus polyphemus.* J Biol Chem 1983;258:7903–7906.
24. Swarnakar S, Melchior R, Quigley JP. Regulation of the plasma cytolytic pathway of *Limulus polyphemus* α_2-macroglobulin. Biol Bull 1995;189:226–227.

25. Iwaki D, Kawabata S, Miura Y, et al. Molecular cloning of *Limulus* alpha 2-macroglobulin. Eur J Biochem 1996;242:822–31.
26. Armstrong PB, Quigley JP. The *Limulus* blood cell secretes α_2-macroglobulin when activated. Biol Bull 1990;178:137–143.
27. Quigley JP, Armstrong PB. A homologue of alpha 2-macroglobulin purified from the hemolymph of the horseshoe crab *Limulus polyphemus.* J Biol Chem 1985;260:12715–12719.
28. Enghild JJ, Thogersen IB, Salvesen G, et al. Alpha-macroglobulin from *Limulus polyphemus* exhibits proteinase inhibitory activity and participates in a hemolytic system. Biochemistry 1990;29:10070–10080.
29. Armstrong PB, Mangel WF, Wall JS, et al. Structure of alpha 2-macroglobulin from the arthropod *Limulus polyphemus.* J Biol Chem 1991;266:2526–2530.
30. Bowen ME, Armstrong PB, Quigley JP, Gettins PG. Comparison of *Limulus* alpha-macroglobulin with human alpha2-macroglobulin: thiol ester characterization, subunit organization, and conformational change. Arch Biochem Biophys 1997;337:191–201.
31. Levashina EA, Moita LF, Blandin S, et al. Conserved role of a complement-like protein in phagocytosis revealed by dsRNA knockout in cultured cells of the mosquito, *Anopheles gambiae.* Cell 2001;104:709–718.
32. Armstrong PB, Quigley JP. *Limulus* alpha 2-macroglobulin. First evidence in an invertebrate for a protein containing an internal thiol ester bond. Biochem J 1987;248:703–7.
33. Dodds AW, Law SK. Structural basis of the binding specificity of the thioester-containing proteins, C4, C3 and alpha-2-macroglobulin. Complement 1988;5:89–97.
34. Sepp A, Dodds AW, Anderson MJ, et al. Covalent binding properties of the human complement protein C4 and hydrolysis rate of the internal thioester upon activation. Protein Sci 1993;2:706–716.
35. Armstrong PB, Armstrong MT, Quigley JP. Involvement of alpha2-macroglobulin and C-reactive protein in a complement-like hemolytic system in the arthropod *Limulus polyphemus.* Mol Immunol 1993;30:929–934.
36. Armstrong PB, Melchior R, Swarnakar S, Quigley JP. Alpha2-macroglobulin does not function as a C3 homologue in the plasma hemolytic system of the American horseshoe crab, *Limulus.* Mol Immunol 1998;35:47–53.
37. Armstrong PB, Misquith S, Srimal S, Melchior R, Quigley JP. Identification of limulin as a major cytolytic protein in the plasma of the American horseshoe crab, *Limulus polyphemus.* Biol Bull 1994;187:227–228.
38. Armstrong PB, Swarnakar S, Srimal S, et al. A cytolytic function for a sialic acid-binding lectin that is a member of the pentraxin family of proteins. J Biol Chem 1996;271:14717–14721.
39. Swarnakar S, Asokan R, Quigley JP, Armstrong PB. Binding of alpha2-macroglobulin and limulin: regulation of the plasma haemolytic system of the American horseshoe crab, *Limulus.* Biochem J 2000;347:679–685.
40. Sottrup-Jensen L. Alpha-macroglobulins: structure, shape, and mechanism of proteinase complex formation. J Biol Chem 1989;264:11539–11542.
41. Quigley JP, Ikai A, Arakawa H, Osada T, Armstrong PB. Reaction of proteinases with alpha 2-macroglobulin from the American horseshoe crab, *Limulus.* J Biol Chem 1991;266:19426–19431.
42. Melchior R, Quigley JP, Armstrong PB. Alpha 2-macroglobulin-mediated clearance of proteases from the plasma of the American horseshoe crab, *Limulus polyphemus.* J Biol Chem 1995;270:13496–13502.
43. Dolmer K, Husted LB, Armstrong PB, Sottrup-Jensen L. Localisation of the major reactive lysine residue involved in the self-crosslinking of proteinase-activated *Limulus* alpha 2-macroglobulin. FEBS Lett 1996;393:37–40.

44. Armstrong PB, Quigley JP. Alpha2-macroglobulin: an evolutionarily conserved arm of the innate immune system. Dev Comp Immunol 1999;23:375–390.
45. Barrett AJ, Brown MA, Sayers CA. The electrophoretically 'slow' and 'fast' forms of the alpha2-macroglobulin molecule. Biochem J 1979;181:401–418.
46. Van Leuven F, Marynen P, Sottrup-Jensen L, Cassiman JJ, Van den Berghe H. The receptor-binding domain of human α_2-macroglobulin. Isolation after limited proteolysis with a bacterial proteinase. J Biol Chem 1986;261:11369–11373.
47. Armstrong PB, Levin J, Quigley JP. Role of endogenous protease inhibitors in the regulation of the blood clotting system of the horseshoe crab, *Limulus polyphemus.* Thromb Haemost 1984;52:117–120.
48. Iwanaga S, Kawabata S, Muta T. New types of clotting factors and defense molecules found in horseshoe crab hemolymph: their structures and functions. J Biochem (Tokyo) 1998;123:1–15.
49. Miura Y, Kawabata S, Iwanaga S. A *Limulus* intracellular coagulation inhibitor with characteristics of the serpin superfamily: purification, characterization and cDNA cloning. J Biol Chem 1994;269:542–547.
50. Miura Y, Kawabata S, Wakamiya Y, Nakamura T, Iwanaga S. A *Limulus* intracellular coagulation inhibitor type 2. Purification, characterization, cDNA cloning, and tissue localization. J Biol Chem 1995;270:558–565.
51. Agarwala KL, Kawabata S, Miura Y, Kuroki Y, Iwanaga S. *Limulus* intracellular coagulation inhibitor type 3. Purification, characterization, cDNA cloning, and tissue localization. J Biol Chem 1996;271:23768–23774.
52. Iwanaga S, Morita T, Miyata T, et al. The *Limulus* Coagulation System Is Sensitive to Bacterial Endotoxins. Heidelberg: Verlag Chemie, 1983, pp. 365–382.
53. Muta T, Seki N, Takaki Y, et al. Horseshoe crab factor G: a new heterodimeric serine protease zymogen sensitive to (1 $\rightarrow$ 3)-beta-D-glucan. Adv Exp Med Biol 1996;389:79–85.
54. Asokan R, Armstrong MT, Armstrong PB. Association of alpha2-macroglobulin with the coagulin clot in the American horseshoe crab, *Limulus polyphemus:* a potential role in stabilization from proteolysis. Biol Bull 2000;199:190–192.
55. Ashida M. The prophenoloxidase cascade in insect immunity. Res Immunol 1990;141:908–910.
56. Nagai T, Kawabata S. A link between blood coagulation and prophenol oxidase activation in arthropod host defense. J Biol Chem 2000;275:29264–29267.
57. Nagai T, Osaki T, Kawabata S. Functional conversion of hemocyanin to phenoloxidase by horseshoe crab antimicrobial peptides. J Biol Chem 2001;276:27166–27170.
58. Kopacek P, Weise C, Saravanan T, Vitova K, Grubhoffer L. Characterization of an alpha-macroglobulin-like glycoprotein isolated from the plasma of the soft tick *Ornithodoros moubata.* Eur J Biochem 2000;267:465–475.
59. Dodds AW, Law SK. The phylogeny and evolution of the thioester bond-containing proteins C3, C4 and alpha 2-macroglobulin. Immunol Rev 1998;166:15–26.
60. Starkey PM, Barrett AJ. Evolution of alpha 2-macroglobulin. The structure of a protein homologous with human alpha 2-macroglobulin from plaice (*Pleuronectes platessa* L.) plasma. Biochem J 1982;205:105–115.
61. Mutsuro J, Nakao M, Fujiki K, Yano T. Multiple forms of alpha2-macroglobulin from a bony fish, the common carp *(Cyprinus carpio):* striking sequence diversity in functional sites. Immunogenetics 2000;51:847–855.
62. Spycher SE, Arya S, Isenman DE, Painter RH. A functional, thioester-containing alpha 2-macroglobulin homologue isolated from the hemolymph of the American lobster *(Homarus americanus).* J Biol Chem 1987;262:14606–14611.
63. Hergenhahn HG, Hall M, Soderhall K. Purification and characterization of an alpha 2-macroglobulin-like proteinase inhibitor from plasma of the crayfish *Pacifastacus leniusculus.* Biochem J 1988;255:801–806.

64. Stocker W, Breit S, Sottrup-Jensen L, Zwilling R. Alpha2-macroglobulin from hemolymph of the freshwater crayfish *Astacus astacus.* Comp Biochem Physiol B 1991;98:501–509.
65. Hall M, Soderhall K, Sottrup-Jensen L. Amino acid sequence around the thiolester of alpha 2-macroglobulin from plasma of the crayfish, *Pacifastacus leniusculus.* FEBS Lett 1989;254:111–114.
66. Liang Z, Lindblad P, Beauvais A, et al. Crayfish alpha-macroglobulin and 76-kDa protein: their biosynthesis and subcellular localization of the 76-kDa protein. J Insect Physiol 1992;38:987–995.
67. Thogersen IB, Salvesen G, Brucato FH, Pizzo SV, Enghild JJ. Purification and characterization of an alpha-macroglobulin proteinase inhibitor from the mollusc *Octopus vulgaris.* Biochem J 1992;285:521–527.
68. Bender RC, Fryer SE, Bayne CJ. Proteinase inhibitory activity in the plasma of a mollusc: evidence for the presence of alpha-macroglobulin in *Biomphalaria glabrata.* Comp Biochem Physiol B 1992;102:821–824.
69. Bender RC, Bayne CJ. Purification and characterization of a tetrameric alpha-macroglobulin proteinase inhibitor from the gastropod mollusc *Biomphalaria glabrata.* Biochem J 1996;316:893–900.
70. Yigzaw Y, Gielens C, Preaux G. Isolation and characterization of an alpha-macroglobulin from the gastropod mollusc *Helix pomatia* with tetrameric structure and preserved activity after methylamine treatment. Biochim Biophys Acta 2001;1545:104–113.
71. Hoffmann JA, Kafatos FC, Janeway CA, Ezekowitz RA. Phylogenetic perspectives in innate immunity. Science 1999;284:1313–1318.
72. Lagueux M, Perrodou E, Levashina EA, Capovilla M, Hoffmann JA. Constitutive expression of a complement-like protein in *Toll* and *JAK* gain-of-function mutants of *Drosophila.* Proc Natl Acad Sci USA 2000;97:11427–11432.
73. Crowley TE, Hoey T, Liu JK, et al. A new factor related to TATA-binding protein has highly restricted expression patterns in *Drosophila.* Nature 1993;361:557–561.
74. Jiang H, Kanost MR. Characterization and functional analysis of 12 naturally occurring reactive site variants of serpin-1 from *Manduca sexta.* J Biol Chem 1997;272:1082–1087.
75. Volanakis JE. Transcriptional regulation of complement genes. Annu Rev Immunol 1995;13:277–305.
76. Anderson KV. Toll signaling pathways in the innate immune response. Curr Opin Immunol 2000;12:13–19.
77. Imler JL, Hoffmann JA. Signaling mechanisms in the antimicrobial host defense of *Drosophila.* Curr Opin Microbiol 2000;3:16–22.
78. Qiu P, Pan PC, Govind S. A role for the *Drosophila* Toll/Cactus pathway in larval hematopoiesis. Development 1998;125:1909–1920.
79. Mathey-Prevot B, Perrimon N. Mammalian and *Drosophila* blood: JAK of all trades? Cell 1998;92:697–700.
80. Dimopoulos G, Casavant TL, Chang S, et al. *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 USA 2000;97:6619–6624.
81. Oduol F, Xu J, Niare O, Natarajan R, Vernick KD. Genes identified by an expression screen of the vector mosquito *Anopheles gambiae* display differential molecular immune response to malaria parasites and bacteria. Proc Natl Acad Sci USA 2000;97:11397–11402.
82. Müller HM, Dimopoulos G, Blass C, Kafatos FC. A hemocyte-like cell line established from the malaria vector *Anopheles gambiae* expresses six prophenoloxidase genes. J Biol Chem 1999;274:11727–11735.
83. Catteruccia F, Nolan T, Blass C, et al. Toward *Anopheles* transformation: *Minos* element activity in anopheline cells and embryos. Proc Natl Acad Sci USA 2000;97:2157–2162.

64. Stocker W, Breit S, Sottrup-Jensen L, Zwilling R. Alpha2-macroglobulin from hemolymph of the freshwater crayfish *Astacus astacus.* Comp Biochem Physiol B 1991;98:501–509.
65. Hall M, Soderhall K, Sottrup-Jensen L. Amino acid sequence around the thiolester of alpha 2-macroglobulin from plasma of the crayfish, *Pacifastacus leniusculus.* FEBS Lett 1989;254:111–114.
66. Liang Z, Lindblad P, Beauvais A, et al. Crayfish alpha-macroglobulin and 76-kDa protein: their biosynthesis and subcellular localization of the 76-kDa protein. J Insect Physiol 1992;38:987–995.
67. Thogersen IB, Salvesen G, Brucato FH, Pizzo SV, Enghild JJ. Purification and characterization of an alpha-macroglobulin proteinase inhibitor from the mollusc *Octopus vulgaris.* Biochem J 1992;285:521–527.
68. Bender RC, Fryer SE, Bayne CJ. Proteinase inhibitory activity in the plasma of a mollusc: evidence for the presence of alpha-macroglobulin in *Biomphalaria glabrata.* Comp Biochem Physiol B 1992;102:821–824.
69. Bender RC, Bayne CJ. Purification and characterization of a tetrameric alpha-macroglobulin proteinase inhibitor from the gastropod mollusc *Biomphalaria glabrata.* Biochem J 1996;316:893–900.
70. Yigzaw Y, Gielens C, Preaux G. Isolation and characterization of an alpha-macroglobulin from the gastropod mollusc *Helix pomatia* with tetrameric structure and preserved activity after methylamine treatment. Biochim Biophys Acta 2001;1545:104–113.
71. Hoffmann JA, Kafatos FC, Janeway CA, Ezekowitz RA. Phylogenetic perspectives in innate immunity. Science 1999;284:1313–1318.
72. Lagueux M, Perrodou E, Levashina EA, Capovilla M, Hoffmann JA. Constitutive expression of a complement-like protein in *Toll* and *JAK* gain-of-function mutants of *Drosophila.* Proc Natl Acad Sci USA 2000;97:11427–11432.
73. Crowley TE, Hoey T, Liu JK, et al. A new factor related to TATA-binding protein has highly restricted expression patterns in *Drosophila.* Nature 1993;361:557–561.
74. Jiang H, Kanost MR. Characterization and functional analysis of 12 naturally occurring reactive site variants of serpin-1 from *Manduca sexta.* J Biol Chem 1997;272:1082–1087.
75. Volanakis JE. Transcriptional regulation of complement genes. Annu Rev Immunol 1995;13:277–305.
76. Anderson KV. Toll signaling pathways in the innate immune response. Curr Opin Immunol 2000;12:13–19.
77. Imler JL, Hoffmann JA. Signaling mechanisms in the antimicrobial host defense of *Drosophila.* Curr Opin Microbiol 2000;3:16–22.
78. Qiu P, Pan PC, Govind S. A role for the *Drosophila* Toll/Cactus pathway in larval hematopoiesis. Development 1998;125:1909–1920.
79. Mathey-Prevot B, Perrimon N. Mammalian and *Drosophila* blood: JAK of all trades? Cell 1998;92:697–700.
80. Dimopoulos G, Casavant TL, Chang S, et al. *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 USA 2000;97:6619–6624.
81. Oduol F, Xu J, Niare O, Natarajan R, Vernick KD. Genes identified by an expression screen of the vector mosquito *Anopheles gambiae* display differential molecular immune response to malaria parasites and bacteria. Proc Natl Acad Sci USA 2000;97:11397–11402.
82. Müller HM, Dimopoulos G, Blass C, Kafatos FC. A hemocyte-like cell line established from the malaria vector *Anopheles gambiae* expresses six prophenoloxidase genes. J Biol Chem 1999;274:11727–11735.
83. Catteruccia F, Nolan T, Blass C, et al. Toward *Anopheles* transformation: *Minos* element activity in anopheline cells and embryos. Proc Natl Acad Sci USA 2000;97:2157–2162.

84. Dimopoulos G, Müller HM, Kafatos FC. How does *Anopheles gambiae* kill malaria parasites? Parassitologia 1999;41:169–175.
85. Weis JJ, Law SK, Levine RP, Cleary PP. Resistance to phagocytosis by group A streptococci: failure of deposited complement opsonins to interact with cellular receptors. J Immunol 1985;134:500–505.
86. Touray MG, Warburg A, Laughinghouse A, Krettli AU, Miller LH. Developmentally regulated infectivity of malaria sporozoites for mosquito salivary glands and the vertebrate host. J Exp Med 1992;175:1607–1612.
87. Margos G, Navarette S, Butcher G, et al. Interaction between host complement and mosquito-midgut-stage *Plasmodium berghei.* Infect Immun 2001;69:5064–5071.

Review 2

Mosquito Immune Responses against Malaria Parasites

Stéphanie BLANDIN, Elena A. LEVASHINA

Current Opinion in Immunology 2004, 16: 16-20

ELSEVIER

Mosquito immune responses against malaria parasites

Stéphanie Blandin[1] and Elena A Levashina[1,2]

Anopheline mosquitoes are the major vectors of human malaria. Mosquito–parasite interactions are a critical aspect of disease transmission and a potential target for malaria control. Mosquitoes vary in their innate ability to support development of the malaria parasite, but the molecular mechanisms that determine vector competence are poorly understood. This area of research has been revolutionized by recent advances in the mosquito genome characterization and by the development of new tools for functional gene analysis.

Addresses
[1]European Molecular Biology Laboratory, Meyerhofstrasse 1, D-69117 Heidelberg, Germany
[2]Unité Propre de Recherche 9022 du Centre National de la Recherche Scientifique, Institut de Biologie Moléculaire et Cellulaire, 15 rue René Descartes, F-67084 Strasbourg Cedex, France
e-mail: E.Levashina@ibmc.u-strasbg.fr

Current Opinion in Immunology 2004, **16**:16–20

This review comes from a themed issue on
Innate immunity
Edited by Bruce Beutler and Jules Hoffmann

0952-7915/$ – see front matter

DOI 10.1016/j.coi.2003.11.010

Abbreviations

dpi	days post-infection
dsRNA	double-stranded RNA
hpi	hours post-infection
NOS	nitric oxide synthase
PCR	polymerase chain reaction
Pen	*Plasmodium* encapsulation locus
TEP	thioester-containing protein

Introduction

Malaria persists today as the most widespread and devastating protozoan disease of humans. It is one of the major causes of mortality of children under the age of five in sub-Saharan Africa. The causative agents of malaria are protozoan parasites of the genus *Plasmodium*, which are transmitted to humans when female *Anopheles* mosquitoes feed on blood. The parasite obligatorily performs the sexual stages of its life cycle in the mosquito. Shortly after the ingestion of an infected blood meal, male and female gametocytes develop in the mosquito midgut into gametes that undergo fertilization within the next few hours. The newly formed zygotes transform into motile ookinetes that invade and cross the midgut epithelium between 24 and 48 hours post-infection (hpi). When they reach the basal side of the epithelium the parasites form protected capsules, called oocysts. During the next 10 days, within each oocyst, a meiotic cycle followed by several rounds of mitosis produces thousands of haploid sporozoites. Upon maturation, which takes place 14–16 days post-infection (dpi), sporozoites are released into the mosquito hemocoel and then migrate to and invade the salivary glands. The parasite cycle in the mosquito is completed when the mosquito injects infective sporozoites into a new human host. In fact, out of 400 anopheline species worldwide only 40 are important malaria vectors. One of these, *Anopheles gambiae*, transmits *Plasmodium falciparum*, the parasite responsible for the cerebral form of the disease in much of sub-Saharan Africa.

During its differentiation in the mosquitoes, the parasite has to overcome three main bottlenecks where its development can be disrupted: the transition between gametocytes and ookinetes, between ookinetes and mature oocysts, and between midgut sporozoites and salivary gland sporozoites (Figure 1). The most dramatic losses occur during the ookinete invasion of the midgut 24–48 hpi. These losses are often attributed to efficient mosquito immune responses, which, in some cases, can completely block parasite development leading to total refractoriness, as observed in several genetically selected strains.

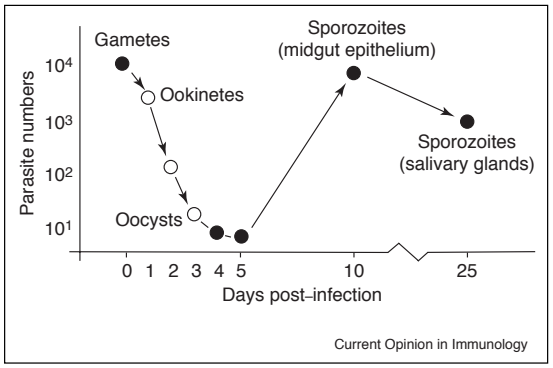
Several recent reviews provide detailed information on parasite development during gametogenesis, oocyst formation and sporozoite invasion [1–3,4••], and on basic mosquito immune responses [5–7]. Here, we have chosen to reevaluate our knowledge of mosquito immune responses against malaria parasites in the view of recent advances, namely the sequencing of the *A. gambiae* genome, and the first functional studies of mosquito immune genes by double-stranded RNA (dsRNA) knockout in adult mosquitoes.

Progress in understanding mosquito immune responses against parasites

Although mosquito immune responses have been studied intensively by the methods of molecular and cell biology [5,8], genes that are directly involved in mosquito–parasite interactions have not yet been conclusively identified.

Until recently, studies on the immune responses in vector mosquitoes were mostly based on gene transcriptional profiling during *Plasmodium* infection. Several approaches have been successfully applied to identify potential immune genes in mosquitoes. The first results came from a candidate gene approach. Mosquito homologs of immune genes of *Drosophila* or other insects were cloned using degenerate PCR techniques: five prophenoloxidase genes, a gene encoding the antimicrobial

Figure 1



Parasite losses and amplification in the mosquito. Parasite numbers during the critical steps of transformation of gametes to ookinetes and of ookinetes to oocysts were quantified using real-time PCR and *P. berghei 28S* rRNA gene as a probe. The initial number of ingested parasites was estimated to be 10^4. Note the dramatic decrease in parasite numbers 1–3 dpi (Modified from [33]).

peptide Defensin, the rel-like transcriptional activator Gambif1, the modular serine protease Sp22D and the serine protease inhibitor SRPN10 [5,9]. This investigation was followed by an unbiased differential display PCR approach and by the cloning of differentially expressed genes using subtractive cDNA libraries. Both of these approaches allowed the rapid identification of novel immune-inducible molecules, such as several serine proteases, serine protease inhibitors, an α_2-macroglobulin-like molecule, the novel antimicrobial peptide Gambicin, a mucin-like molecule and a Gram-negative bacteria-binding protein, in addition to some as yet uncharacterized genes [10–14]. Finally, large-scale expression profiling using cDNA microarrays identified at least 200 genes that are differentially expressed upon various types of immune stimulation, including parasite infection [15•]. The clustering of co-expressed genes revealed that the genes involved in the metabolism of the blood meal become induced at early timepoints after infectious feeding. The specific upregulation of immune genes begins around 20 hpi when the parasites are traversing the midgut epithelium. The upregulation of immune markers was detected not only locally in the midgut but also systemically in the fat body [16]. At present, little is known about the molecular mechanisms that trigger the induction of immune markers.

New insights into mosquito–parasite interactions resulted from cell biology studies of *P. berghei* midgut invasion in *A. stephensi*, and by an *in vitro* approach following the interactions of *P. gallinaceum* ookinete cultures with dissected midguts of *Aedes aegypti* [17,18]. Both studies suggest that mosquito midgut cells undergo extensive damage upon parasite invasion, which results in the apoptosis of invaded cells. A time bomb theory was proposed to explain both midgut cell apoptosis and parasite killing inside the epithelial cells. Reactive oxygen species, produced by a nitric oxide synthase (NOS), are suspected to play a major role in these processes [18,19]. *NOS* is transcriptionally induced after parasite infection. It is unclear whether *NOS* induction results in apoptosis, or whether the induction of apoptosis causes *NOS* upregulation. In a different study, the expression of two spliceforms of a serine protease inhibitor gene, *Spn10*, is also induced in *A. gambiae* midgut after parasite infection [9]. The functional relevance of *NOS* and *Spn10* induction is not understood at present, and further studies using dsRNA gene-specific silencing are needed. Indeed, some of the immune genes that are induced by parasite infections do not directly affect the efficiency of parasite transmission. For example, several lines of evidence suggested that the bactericidal peptide Defensin 1 might also be involved in limiting parasite development. In *A. gambiae* this gene is strongly induced at 24 hpi in the mosquito anterior midgut and in the fat body cells. However, the dsRNA silencing of *Defensin 1* does not affect the development of parasites, although it does confirm the expected role of this peptide in the resistance of mosquitoes to Gram-positive bacteria [20••]. Similarly, *TEP4*, a thioester-containing protein gene, is strongly induced by an infectious blood meal and after bacterial challenge [14]. Silencing of this gene by dsRNA has no effect on parasite development in the mosquito, suggesting that, similar to *Defensin*, this gene is not involved in the control of parasite loads in *A. gambiae* (S Blandin and EA Levashina, unpublished). Systematic functional analysis of parasite-inducible genes is required to improve and validate our view of mosquito immune responses to *Plasmodium* invasion.

Vector competence and refractoriness

Only a limited number of mosquito species are able to transmit malaria, and, in the majority of cases, a given mosquito species is an efficient vector of only a limited number of malarias. The ability of a vector to support the development of a given parasite is called vector competence, the mechanisms of which are poorly understood. The arrest in parasite development can be caused by the lack in the mosquito of a factor essential to support parasite survival, or by an active mosquito response that suppresses the passage of the parasite through the body tissues [21••]. Elegant recent studies on the development of two model malaria parasites (*P. berghei* and *P. gallinaceum*) in three mosquito species (*A. gambiae*, *A. stephensi* and *Ae. aegypti*) identified critical steps during the development of parasites in non-vector mosquito species. In fact, these steps correspond to the three bottlenecks described for the development of a parasite in a compatible vector [21••]. These observations suggest that the mechanisms of parasite killing are conserved between different mosquito species. Further studies are needed to

dissect molecular mechanisms that are essential for the parasite development.

Within the same susceptible species, mosquitoes display genetic variation in their susceptibility to parasites, and strains that do not support the development of various avian, rodent and simian malarias have been selected and studied, mostly morphologically. One strain of *A. gambiae* completely aborts the development of a simian malaria parasite *P. cynomolgi* B. Refractoriness, manifested by the melanotic encapsulation of ookinetes, is controlled by one major locus, *Pen1* (for *Plasmodium* encapsulation locus 1), and two additional loci, *Pen2* and *Pen3*, which have more minor effects on parasite melanization [22,23]. This strain is refractory to a large number of primate, rodent and avian malarias, but not to co-indigenous strains of human malarias, which successfully bypass the encapsulation response [22]. Another selected strain of *A. gambiae* efficiently eliminates *P. gallinaceum* by a novel lytic mechanism [24]. Only recently have studies on malaria resistance mechanisms and their significance for the biology of malaria transmission been initiated in natural populations of *A. gambiae* [25]. The first report, in which refractoriness is determined as a reduction in oocyst numbers 8 dpi, point to significant variation in permissiveness of mosquitoes for parasite development, and identify two genomic regions that affect this trait. The molecular mechanisms that underlie the refractory phenotypes and their respective importance for the transmission of malaria are not, as yet, understood.

Hemocytes and antiparasitic defense

In adult mosquitoes, hemocytes are involved in both cellular and humoral responses that are the hallmark of the insect immune system [26••]. Cellular responses include phagocytosis and cellular encapsulation. Phagocytosis plays an important role in the clearing of bacteria, especially during the early stages of infection. Larger pathogens that cannot be phagocytosed are encapsulated by hemocytes. However, no direct interaction between the parasite midgut stages (ookinetes and oocysts) and mosquito hemocytes has been reported to date, suggesting that the 'classical' cellular defenses are not involved in the mosquito response to the parasite early stages. Although one study reports phagocytosis of sporozoites by hemocytes [27], it remains unclear whether this is a general phenomenon, which may account either for major sporozoite losses, or for hemocyte scavenging of dead sporozoites.

Humoral immune responses include the synthesis of antimicrobial peptides by fat body cells and hemocytes. A recent review summarizes the properties of antimicrobial peptides in *A. gambiae* [7]; however, to date there is no functional evidence that these peptides are involved in the mosquito antiparasitic responses *in vivo*. Hemocytes also secrete molecules that are essential for the formation of basal membranes. Interestingly, some of these molecules (laminin, β-integrin 1) are found on the surface of parasites, suggesting that the components of the basal lamina interact with parasite surface proteins [28,29]. The precise role of these protein–protein interactions remains unclear. Another interesting molecule produced by hemocytes is TEP1. It belongs to a family of thioester-containing proteins (TEPs) that share significant similarities with the vertebrate complement factors C3, C4 and C5 and α_2-macroglobulins. TEP1 is cleaved shortly after septic injury. Attachment of the cleaved carboxy-terminal part of TEP1 to the surface of Gram-positive and Gram-negative bacteria through the conserved thioester site labels the bacteria for clearing by phagocytosis [30]. *TEP1* is moderately upregulated by bacterial and parasite infections. dsRNA knockout experiments in adult mosquitoes indicate that TEP1 plays an essential role in the mosquito antiparasitic response (Blandin *et al.*, unpublished). In susceptible mosquitoes, the knockout of *TEP1* results in a fivefold increase in the number of oocysts developing in the midgut, suggesting that parasite killing in mosquitoes is mediated by TEP1. The elucidation of the molecular mechanisms of TEP1 parasite killing will provide important insights in the evolution and function of the TEP family.

Conclusions

New tools have recently become available for the search of immune genes that play a role in the regulation of parasite load and, hence, vectorial capacity. The complete sequencing of the mosquito genome provides a wealth of information on putative immune genes that are homologous to components of the *Drosophila* and mammalian immune systems [31••]. Interestingly, the number of orthologous genes between the fruitfly and the mosquito is smaller than expected. Phylogenetic analysis clusters most immune genes into species-specific expansions, suggesting that these expansions appeared during the evolution of immune genes in response to distinct pathogenic environments in these two dipteran species [32]. Further functional analyses of *Anopheles*-specific genes will be informative about the mosquito immune response against malaria parasites.

With the knowledge of the mosquito genome, the availability of dsRNA technology and the determination of the precise timing of the mosquito immune responses, we are now in a position to identify mosquito immune factors/effectors that are responsible for parasite killing, and to dissect molecularly humoral and cellular immune responses during parasite infections in *A. gambiae*.

Acknowledgements

The authors are grateful to Jules A Hoffmann and Fotis C Kafatos for continuous support and to members of their laboratories for fruitful discussions. The work in the authors' laboratories has received support from the European Molecular Biology Laboratory, Centre National de la Recherche Scientifique, National Institutes of Health and European Union Research Training Network.

References and recommended reading

Papers of particular interest, published within the annual period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Menard R: **The journey of the malaria sporozoite through its hosts: two parasite proteins lead the way**. *Microbes Infect* 2000, **2**:633-642.

2. Shahabuddin M, Costero A: **Spatial distribution of factors that determine sporogonic development of malaria parasites in mosquitoes**. *Insect Biochem Mol Biol* 2001, **31**:231-240.

3. Han YS, Barillas-Mury C: **Implications of Time Bomb model of ookinete invasion of midgut cells**. *Insect Biochem Mol Biol* 2002, **32**:1311-1316.

4. •• Sinden RE: **Molecular interactions between *Plasmodium* and its insect vectors**. *Cell Microbiol* 2002, **4**:713-724.
This comprehensive review clearly defines the critical stages of parasite development in mosquitoes and gives an excellent overview of our understanding of mosquito–parasite interactions.

5. Dimopoulos G, Müller HM, Levashina EA, Kafatos FC: **Innate immune defense against malaria infection in the mosquito**. *Curr Opin Immunol* 2001, **13**:79-88.

6. Dimopoulos G, Kafatos FC, Waters AP, Sinden RE: **Malaria parasites and the anopheles mosquito**. *Chem Immunol* 2002, **80**:27-49.

7. Dimopoulos G: **Insect immunity and its implication in mosquito-malaria interactions**. *Cell Microbiol* 2003, **5**:3-14.

8. Barillas-Mury C, Wizel B, Soo Han Y: **Mosquito immune responses and malaria transmission: lessons from insect model systems and implications for vertebrate innate immunity and vaccine development**. *Insect Biochem Mol Biol* 2000, **30**:429-442.

9. Danielli A, Kafatos FC, Loukeris TG: **Cloning and characterization of four *Anopheles gambiae* serpin isoforms, differentially induced in the midgut by *Plasmodium berghei* invasion**. *J Biol Chem* 2003, **278**:4184-4193.

10. Dimopoulos G, Richman A, della Torre A, Kafatos FC, Louis C: **Identification and characterization of differentially expressed cDNAs of the vector mosquito, *Anopheles gambiae***. *Proc Natl Acad Sci USA* 1996, **93**:13066-13071.

11. Dimopoulos G, Richman A, Muller HM, Kafatos FC: **Molecular immune responses of the mosquito *Anopheles gambiae* to bacteria and malaria parasites**. *Proc Natl Acad Sci USA* 1997, **94**:11508-11513.

12. Vizioli J, Bulet P, Hoffmann JA, Kafatos FC, Müller HM, Dimopoulos G: **Gambicin: a novel immune responsive antimicrobial peptide from the malaria vector *Anopheles gambiae***. *Proc Natl Acad Sci USA* 2001, **98**:12630-12635.

13. Bonnet S, Prevot G, Jacques JC, Boudin C, Bourgouin C: **Transcripts of the malaria vector *Anopheles gambiae* that are differentially regulated in the midgut upon exposure to invasive stages of *Plasmodium falciparum***. *Cell Microbiol* 2001, **3**:449-458.

14. Oduol F, Xu J, Niare O, Natarajan R, Vernick KD: **Genes identified by an expression screen of the vector mosquito *Anopheles gambiae* display differential molecular immune response to malaria parasites and bacteria**. *Proc Natl Acad Sci USA* 2000, **97**:11397-11402.

15. • 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 USA* 2002, **99**:8814-8819.
The authors report the establishment of the microarray technology for *A. gambiae* and provide the first large-scale analysis of gene regulation upon immune challenges *in vitro* and *in vivo*.

16. Dimopoulos G, Seeley D, Wolf A, Kafatos FC: **Malaria infection of the mosquito *Anopheles gambiae* activates immune-responsive genes during critical transition stages of the parasite life cycle**. *EMBO J* 1998, **17**:6115-6123.

17. Zieler H, Dvorak JA: **Invasion *in vitro* of mosquito midgut cells by the malaria parasite proceeds by a conserved mechanism and results in death of the invaded midgut cells**. *Proc Natl Acad Sci USA* 2000, **97**:11516-11521.

18. Han YS, Thompson J, Kafatos FC, Barillas-Mury C: **Molecular interactions between *Anopheles stephensi* midgut cells and *Plasmodium berghei*: the time bomb theory of ookinete invasion of mosquitoes**. *EMBO J* 2000, **19**:6030-6040.

19. Luckhart S, Vodovotz Y, Cui L, Rosenberg R: **The mosquito *Anopheles stephensi* limits malaria parasite development with inducible synthesis of nitric oxide**. *Proc Natl Acad Sci USA* 1998, **95**:5700-5705.

20. •• Blandin S, Moita LF, Köcher T, Wilm M, Kafatos FC, Levashina EA: **Reverse genetics in the mosquito *Anopheles gambiae*: targeted disruption of the *Defensin* gene**. *EMBO Rep* 2002, **3**:852-856.
This is the first functional study of an immune gene by dsRNA silencing in adult mosquitoes. The knockout of the *Defensin* gene demonstrates that this antimicrobial peptide is required for resistance to Gram-positive bacteria, but does not affect the development of *Plasmodium in vivo*.

21. •• Alavi Y, Arai M, Mendoza J, Tufet-Bayona M, Sinha R, Fowler K, Billker O, Franke-Fayard B, Janse CJ, Waters A *et al.*: **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* 2003, **33**:933-943.
The authors compare the development of parasites in vector and non-vector mosquito species. The quantification of parasite losses identifies critical steps in mosquito–parasite interactions during which parasite development could be efficiently disrupted.

22. Collins FH, Sakai RK, Vernick KD, Paskewitz S, Seeley DC, Miller LH, Collins WE, Campbell CC, Gwadz RW: **Genetic selection of a *Plasmodium*-refractory strain of the malaria vector *Anopheles gambiae***. *Science* 1986, **234**:607-610.

23. Zheng L, Cornel AJ, Wang R, Erfle H, Voss H, Ansorge W, Kafatos FC, Collins FH: **Quantitative trait loci for refractoriness of *Anopheles gambiae* to *Plasmodium cynomolgi* B**. *Science* 1997, **276**:425-428.

24. Vernick KD, Fujioka H, Seeley DC, Tandler B, Aikawa M, Miller LH: ***Plasmodium gallinaceum*: a refractory mechanism of ookinete killing in the mosquito, *Anopheles gambiae***. *Exp Parasitol* 1995, **80**:583-595.

25. Niare O, Markianos K, Volz J, Oduol F, Toure A, Bagayoko M, Sangare D, Traore SF, Wang R, Blass C *et al.*: **Genetic loci affecting resistance to human malaria parasites in a West African mosquito vector population**. *Science* 2002, **298**:213-216.

26. •• Hoffmann JA, Reichhart JM: ***Drosophila* innate immunity: an evolutionary perspective**. *Nat Immunol* 2002, **3**:121-126.
This is an authoritative review of our state of knowledge of *Drosophila* immunity. The authors provide a clear overview of the effector mechanisms and intracellular signaling pathways that are activated by the recognition of infectious nonself, and discuss evolutionary perspectives of the antimicrobial defenses of the fruitfly.

27. Hillyer JF, Schmidt SL, Christensen BM: **Rapid phagocytosis and melanization of bacteria and *Plasmodium* sporozoites by hemocytes of the mosquito *Aedes aegypti***. *J Parasitol* 2003, **89**:62-69.

28. Dessens JT, Siden-Kiamos I, Mendoza J, Mahairaki V, Khater E, Vlachou D, Xu XJ, Kafatos FC, Louis C, Dimopoulos G *et al.*: **SOAP, a novel malaria ookinete protein involved in mosquito midgut invasion and oocyst development**. *Mol Microbiol* 2003, **49**:319-329.

29. Mahairaki V, Lycett G, Blass C, Louis C: **Beta-integrin of *Anopheles gambiae*: mRNA cloning and analysis of structure and expression**. *Insect Mol Biol* 2001, **10**:217-223.

30. Levashina EA, Moita LF, Blandin S, Vriend G, Lagueux M, Kafatos FC: **Conserved role of a complement-like protein in Phagocytosis revealed by dsRNA knockout in cultured cells of the mosquito, *Anopheles gambiae***. *Cell* 2001, **104**:709-718.

31. Christophides GK, Zdobnov E, Barillas-Mury C, Birney E, Blandin S,
•• Blass C, Brey PT, Collins FH, Danielli A, Dimopoulos G *et al.*: **Immunity-related genes and gene families in *Anopheles gambiae***. *Science* 2002, **298**:159-165.
The phylogenetic analysis of immune gene families in the two related dipteran species, *A. gambiae* and *Drosophila melanogaster*, revealed a marked deficit of orthologs and large species-specific expansions, possibly reflecting selection pressures from distinct pathogenic environments in the two species.

32. Zdobnov EM, von Mering C, Letunic I, Torrents D, Suyama M, Copley RR, Christophides GK, Thomasova D, Holt RA, Subramanian GM *et al.*: **Comparative genome and proteome analysis of *Anopheles gambiae* and *Drosophila melanogaster***. *Science* 2002, **298**:149-159.

33. Sinden RE: ***Plasmodium* differentiation in the mosquito**. *Parassitologia* 1999, **41**:139-148.

Chapter 2

Role of the Mosquito Thioester-containing Protein 1 in Phagocytosis

Role of the Mosquito Thioester-containing Protein 1 in Phagocytosis.

In mosquitoes as in other insects, different types of cellular and humoral defenses are activated to fight invading pathogens. Cellular responses have been described at the morphological level and include phagocytosis and cellular encapsulation of larger microorganisms. Humoral responses have been better characterized at the molecular level. They comprise a set of antimicrobial peptides (Defensins, Cecropins and Gambicin) that are secreted in the hemolymph upon immune challenge and display anti-bacterial and anti-fungal activities *in vitro* (Vizioli et al., 2000; Vizioli et al., 2001a; Vizioli et al., 2001b). In addition, activation of proteolytic cascades is postulated to cause melanotic encapsulation of pathogens and coagulation reactions at the site of infection. Several serine proteases and serpins potentially involved in these cascades have been identified (reviewed in Dimopoulos et al., 2001).

However by the time we started this project, the molecular basis of pathogen recognition remained to be dissected. The local activation of melanotic encapsulation reactions on the surface of parasites, and the minimal surface characteristics that are required for recognition of foreign bodies (Gorman et al., 1998) suggested that recognition could rely on a complement-like reaction. The identification of 6 thioester-containing proteins (*dTeps*) in the *Drosophila* genome (Lagueux et al., 2000) prompted us to search for homologs in mosquitoes. Having identified one mosquito TEP, we examined whether it could function as a complement-like molecule. Since the most ancient function of complement factors is believed to be the labelling of pathogens for phagocytosis, we developed a first approach *in vitro* to measure the binding and opsonizing properties of this mosquito TEP. For this, we made use of the mosquito hemocyte-like cell line 5.1* established in our laboratory by H.-M. Mueller (Catteruccia et al., 2000a) and developed a dsRNA knockdown method to silence gene expression in mosquito cells. Our results were published in 2001 (Levashina et al., 2001) and are presented on pages 57-68.

In this paper, we reported the identification and functional characterization of the first mosquito thioester-containing protein:

• We have cloned a novel Anopheles *thioester-containing protein, TEP1* (formerly known as *aTEP-I*), which shows significant sequence similarity to vertebrate complement factors and to the related family of α2Ms.

• The phylogenetic analysis of available sequences of thioester-containing proteins in animals, from *C. elegans* to humans, showed that insect and worm TEPs formed a new clade intermediate between the clade of α2Ms and the clade of complement factors. We propose that insect and worm TEPs represent a primitive but diversified group of TEP sequences, distinct from both complement factors and α2Ms.

• TEP1 is constitutively synthesized by mosquito hemocytes and secreted into the hemolymph. Bacterial challenge and *Plasmodium* infection further induce the expression of this gene.

• The proteolytic cleavage of TEP1 in the hemolymph is increased by both wounding and bacterial challenge. It is important to remember that proteolytic activation is required for complement factors and α2Ms first to expose their thioester bond and to bind covalently to nearby targets, and second, to expose their C-terminal domain leading to clearance of labelled entities via receptor-mediated phagocytosis or endocytosis. We speculate that TEP1 also requires proteolytic activation to become functional.

• We demonstrated the presence of an opsonizing factor in the conditioned medium of the 5.1* cell cultures. Indeed bacteria pre-exposed to conditioned medium were more readily uptaken by mosquito cells than when pre-exposed to fresh medium. Obviously, we next investigated whether TEP1 was this opsonizing factor.

• The conditioned medium of the 5.1* cell cultures resembled mosquito hemolymph in containing the full-length and the proteolytically cleaved forms of TEP1. We established the functionality of the thioester bond and showed that it can be chemically inactivated by methylamine treatment.

• The incubation of bacteria in normal or methylamine-treated condition medium allowed us to demonstrate that TEP1 binds to both Gram-positive and Gram-negative bacteria and that this binding is dependent on a functional thioester.

• Finally, and I was more particularly involved in this part of the work, we specifically depleted TEP1 from the condition medium by transfection of the mosquito cells with dsRNA directed against *TEP1*. Phagocytosis of Gram-negative bacteria pre-incubated in TEP1-depleted medium was reduced by 50% compared to control medium. In contrast, the efficiency of phagocytosis of Gram-positive bacteria was low and independent of opsonization by TEP1, suggesting that this cell line lacks an additional component promoting the uptake of Gram-positive bacteria.

In conclusion we demonstrated that a Protostome, the mosquito *A. gambiae*, expresses a thioester-containing protein that resembles complement in promoting phagocytosis. Our results, therefore, indicate that complement-like proteins have a much longer history than previously thought and can be traced back to dipteran insects. Furthermore we reported a successful application of dsRNA gene silencing using a cell culture model system to study gene function in mosquitoes, where, at that time, no *in vivo* genetic analysis could be performed.

Recently, the role of TEP1 in the phagocytosis of bacteria was assessed in adult mosquitoes using double-stranded RNA silencing (Moita et al, unpublished data). In these experiments, TEP1 was required for promotion of phagocytosis of both Gram-positive and Gram-negative bacteria *in vivo*, confirming that the absence of phagocytosis of Gram-positive bacteria is an artefact of the 5.1* cell line.

Article 1

Conserved Role of a Complement-like Protein in Phagocytosis Revealed by dsRNA Knockout in Cultured Cells of the Mosquito *Anopheles gambiae*

Elena A. LEVASHINA, Luis F. MOITA,
Stéphanie BLANDIN, Gert VRIEND, Marie LAGUEUX,
Fotis C. KAFATOS

Cell 2001, 104: 709-718

Cell, Vol. 104, 709–718, March 9, 2001, Copyright ©2001 by Cell Press

Conserved Role of a Complement-like Protein in Phagocytosis Revealed by dsRNA Knockout in Cultured Cells of the Mosquito, *Anopheles gambiae*

Elena A. Levashina,*§ Luis F. Moita,*§ Stephanie Blandin,* Gert Vriend,* Marie Lagueux,† and Fotis C. Kafatos*‡
*EMBL
1 Meyerhofstrasse
69117 Heidelberg
Germany
†UPR 9022 du CNRS
IBMC
15 rue René Descartes
67084 Strasbourg Cedex
France

Summary

We characterize a novel hemocyte-specific acute phase glycoprotein from the malaria vector, *Anopheles gambiae*. It shows substantial structural and functional similarities, including the highly conserved thioester motif, to both a central component of mammalian complement system, factor C3, and to a panprotease inhibitor, α_2-macroglobulin. Most importantly, this protein serves as a complement-like opsonin and promotes phagocytosis of some Gram-negative bacteria in a mosquito hemocyte-like cell line. Chemical inactivation by methylamine and depletion by double-stranded RNA knockout demonstrate that this function is dependent on the internal thioester bond. This evidence of a complement-like function in a protostome animal adds substantially to the accumulating evidence of a common ancestry of immune defenses in insects and vertebrates.

Introduction

Mosquitoes, like other insects, display powerful humoral and cellular defense reactions that help fight bacterial and parasite infections and are reminiscent of innate immune responses in vertebrates (Hoffmann et al., 1999). The humoral responses are best characterized, and a diverse set of immune-inducible genes has been identified. For example, antimicrobial peptides are produced after bacterial challenge, mainly by the mosquito fat body, and are secreted into the hemolymph (reviewed by Dimopoulos et al., 2001). Although the exact role of these genes in the antiparasitic response has not yet been elucidated, they may be involved in vector refractoriness to the parasite (Lowenberger et al., 1999). Cellular responses, which are not as well understood at the molecular level as the humoral defenses, include phagocytosis of bacteria and melanotic encapsulation of larger parasites (reviewed by Barillas-Mury et al., 2000).

‡To whom correspondence should be addressed (e-mail: kafatos@embl-heidelberg.de).
§These authors contributed equally to this work.

We are interested in the molecular mechanisms of recognition and activation of immune effector responses in the mosquito. The local activation of encapsulation reactions and the minimal surface characteristics that are required for recognition of foreign bodies (Gorman et al., 1998) suggest that recognition could rely on a complement-like reaction. In vertebrates, the complement system is a major effector system of innate immunity. It can be activated through three distinct pathways, classical, lectin, and alternative (reviewed by Volanakis, 1998), which intersect at the central component, C3. In all three pathways, proteolytic activation of C3 leads to covalent attachment of a C3 cleavage product through a thioester bond to the pathogen surface, followed by phagocytosis or cell lysis of pathogens. Recent cloning of C3-like molecules from an ascidian protochordate (Nonaka and Azumi, 1999) and a sea urchin (Al-Sharif et al., 1998), and the identification of related *t*hioester-containing *p*rotein (TEP) sequences in *Drosophila melanogaster* (Lagueux et al., 2000), prompted us to search for members of the same family in the most important insect vector of human malaria, *Anopheles gambiae*.

Here we report the identification and functional characterization of a novel *Anopheles* TEP (aTEP-I), which shows significant sequence similarity both to vertebrate complement factors and to the related family of animal α_2-macroglobulins (α_2M). aTEP-I is a hemocyte-specific acute phase glycoprotein, which is proteolytically processed in mosquito hemolymph shortly after septic injury. Studies on a hemocyte-like mosquito cell line reveal that aTEP-I can bind to *Escherichia coli* and *Staphylococcus aureus* and that this binding is dependent on a functional thioester bond. Finally, double-stranded RNA (dsRNA) knockout experiments demonstrate that this complement-like protein strongly promotes phagocytosis of three tested Gram-negative bacteria in a mosquito cell line. In contrast, phagocytosis of the Gram-positive *S. aureus* and two other bacterial species was consistently low and not significantly affected by the aTEP-I knockout, suggesting that additional molecules are required to promote uptake of Gram-positive bacteria.

Results

Cloning and Characterization of *aTEP-I*

To clone *TEP* sequences of *A. gambiae*, we designed degenerate primers corresponding to the highly conserved thioester site of the related genes of *D. melanogaster*. PCR experiments with a degenerate *TEP* primer and a T7 primer from the ZAP Express vector on a pool of cDNAs of the mosquito cell line 5.1* (Danielli et al., 2000) yielded a 1.6 kb 3′ sequence of a novel gene, which displayed sequence similarity with C3 and α_2M. 5′-RACE PCR amplified 1.5 kb of additional 5′ sequence, which was used as a probe to screen a 5′-enriched thoracic cDNA library (Arca et al., 1999). A 4.2 kb clone, *aTEP-I*, was isolated and contained a complete open reading frame of 1340 amino acids.

The translated sequence of aTEP-I was aligned with

sequences of representative thioester-containing proteins using the program CLUSTAL_X. The overall sequence similarity was approximately 22%. Furthermore, clustered matches of high identity were distributed along the molecule, leaving no doubt that aTEP-I is a member of the thioester-containing protein family. The predicted organization of aTEP-I is compared with that of human α_2M and C3 in Figure 1A. The aTEP-I sequence contains a signal peptide-like hydrophobic N-terminal segment suggesting that, like other TEPs, it is a secreted protein. The canonical thioester motif (GCGEQ) (red star) is located at a comparable position in all three molecules (residues 858–862 in aTEP-I).

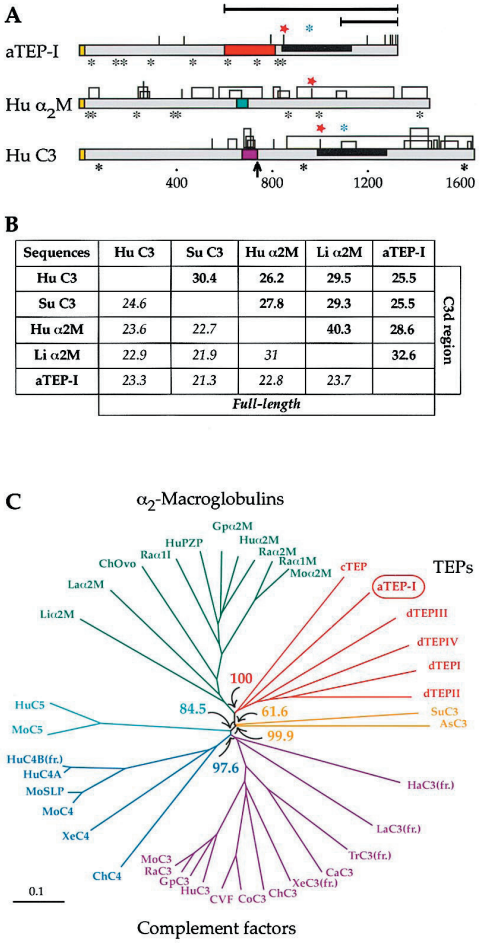
The thioester proteins show distinct preferences in their binding activities, related to a histidine residue about 100 amino acids downstream of the thioester site. In C3, His favors the formation of ester bonds with hydroxyl groups of carbohydrates whereas in α_2M, its absence leads to preferential formation of amide bonds with free amino groups of proteins (reviewed by Dodds and Law, 1998). Interestingly, in aTEP-I, a presumed catalytic His is found at the same site (position 972, blue asterisk).

A clear difference between C3 and α_2M is their post-synthetic intracellular processing. C3 is split into two chains, α and β, which are held together by disulfide bridges, while α_2Ms function as dimers or tetramers of intact chains (reviewed by Sottrup-Jensen, 1989). No clear-cut site (RRRR/RKKR) for this type of processing into two chains has been identified in aTEP-I, indicating that the molecule may be secreted as a single chain polypeptide. Subsequent limited extracellular proteolysis at centrally located sites activates the secreted TEPs. In human C3, a cleavage site (RS) (Figure 1A, black arrow) separates anaphylatoxin (purple) from the C-terminal effector region. Similarly, α_2M recognizes proteases through a particularly exposed peptide stretch (bait region, green). In aTEP-I, a region enriched in potential cleavage sites RX or KX (X = S, A, K, L, V) is indicated in red.

A cluster of six cysteine residues at the C terminus of aTEP-I (Figure 1A) forms a signature that is shared with *Drosophila* TEPs (Lagueux et al., 2000) but is absent from α_2M and complement factors. Otherwise, cysteines are not numerous in insect TEPs (ten in total in aTEP-I), and are not in conserved locations relative to the much more numerous cysteines of vertebrate TEPs. Potential N-glycosylation sites are found both in insect and vertebrate TEP sequences (Figure 1A, black asterisks). Neglecting potential posttranslational modifications, the calculated molecular mass of the secreted aTEP-I protein is 149.957 kDa.

Phylogenetic Analysis of aTEP-I and Related Thioester-Containing Proteins

The availability of insect TEPs from two different insect species and our detection of two nearly identical TEP sequences in the *Caenorhabditis elegans* WORMPEP database permitted us to perform a phylogenetic analysis of the animal TEP family across more than a billion years of evolution. The overall sequence similarity of aTEP-I to the C3, C4, and C5 complement factors and to α_2M ranged from 21 to 24%, while the similarity in



Sequences	Hu C3	Su C3	Hu α2M	Li α2M	aTEP-I
Hu C3		30.4	26.2	29.5	25.5
Su C3	*24.6*		27.8	29.3	25.5
Hu α2M	*23.6*	*22.7*		40.3	28.6
Li α2M	*22.9*	*21.9*	*31*		32.6
aTEP-I	*23.3*	*21.3*	*22.8*	*23.7*	
	Full-length				

C3d region

Figure 1. Sequence Comparison and Phylogenetic Tree of Thioester-Containing Proteins

(A) Schematic representation of thioester-containing proteins: *A. gambiae* aTEP-I (accession number AF291654) and human (Hu) C3 and α_2-macroglobulin. Thick marks indicate cysteines connected into the known disulfide bridges of C3 and α_2M. Red stars point to the internal β-cysteinyl-γ-glutamyl thioester sites. Blue asterisks show catalytic histidines and black asterisks are putative N-glycosylation sites. The color-filled segments indicate: yellow, signal peptide; red, a region containing putative protease cleavage sites for aTEP-I; green, the bait region of α_2M; violet, anaphylatoxin C3a; black half-filled segments, C3d region in C3 and its equivalent in aTEP-I. Black vertical arrow shows the C3 activation site. Numbers correspond to amino acid positions. Black horizontal bars show the aTEP-I fragments that were used as antigens for antibody production.

(B) Matrix of percentage amino acid similarity in the aligned full-length sequences (italics) and in the critical C3d-like region (bold) between representative members of the thioester-containing protein family.

(C) Unrooted phylogenetic tree of thioester-containing proteins. The tree was built by the neighbor-joining method based on the alignment of the sequences using CLUSTAL_X. Alignment and sequence information is accessible online. The insect/nematode TEP clade is colored red; the complement C3, C4, and C5 clades are violet, blue, and light blue, respectively; the α_2-macroglobulin clade is green; and the outbranching ascidian and sea urchin C3 factors are in orange. Bootstrap values of 1000 replicates (%) are for the nodes indicated by arrows.

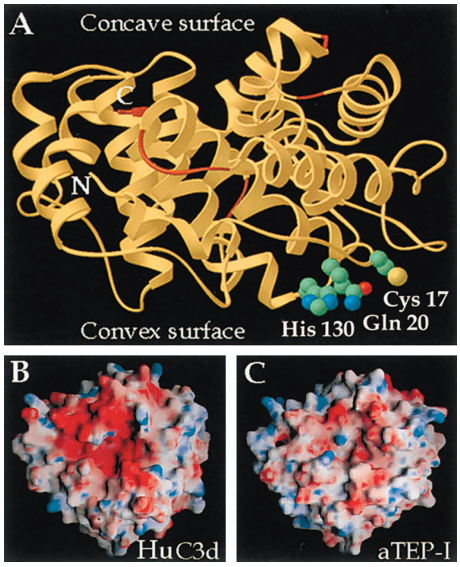


Figure 2. Modeling of the C3d-like Region of aTEP-I

(A) Ribbon model of aTEP-I. Areas that differ from the human C3d structure are colored in brown. The side chains of critical residues (Cys and Gln of the thioester site and a catalytic His) are shown as ball-and-stick models (C, green; N, blue; O, red; S, yellow).

(B and C) Comparison of electrostatic potential maps of concave surfaces between human C3d (B) and the aTEP-I model (C). Maps were built with GRASP (Nicholls et al., 1991). The potential range is from −10.011 (dark red) to +10.045 (dark blue).

the most conserved region (C3d-like, see below) ranged from 26 to 32% (Figure 1B).

Based on the full-length multisequence alignment built with CLUSTAL_X, we constructed an unrooted phylogenetic tree using the neighbor-joining method (Figure 1C). Insect TEPs formed a new clade in this family, intermediate between the clade of α_2M (from vertebrates and the horseshoe crab *Limulus*) and the clades of complement factors (vertebrate C3, C4, and C5, and the outlying ascidian and sea urchin C3). Interestingly, the *C. elegans* TEP clustered with the insect TEP clade, within which the four molecularly characterized dTEPs form deep branches. The same topology was observed for the phylogenetic tree based on the alignment of the C3d-like regions (data not shown). Thus, insect/worm TEPs may represent a primitive but diversified group of TEP sequences, distinct from both complement factors and α_2M.

Modeling of a Critical aTEP-I Region

Initial information about structural features of aTEP-I and their functional implications was gained by homology-based modeling of a region of aTEP-I corresponding to the thioester-containing C3d fragment of human C3 (Figure 1A, black half-filled segment). In the structure of human C3d (Nagar et al., 1998), the critical residues (Cys, Gln, and His) involved in the formation and catalytic activation of the thioester bond are clustered in and protrude from the convex surface (Figure 2A). The opposite, concave surface of C3d serves as the recognition site for complement receptor CR2. The C3d structure and the corresponding aTEP-I model are remarkably

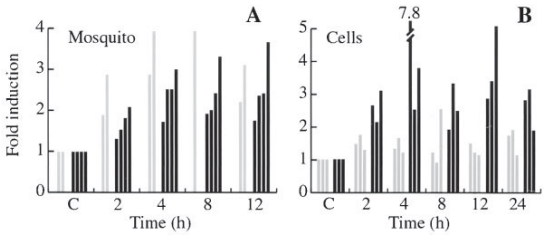


Figure 3. Immune-Inducible Transcription of *aTEP-I*

Semi-quantitative RT-PCR was performed and the levels of the aTEP-I transcript were normalized relative to the internal control transcript for ribosomal protein S7. Quantification was done by fluorimager using SYBR green fluorescent dye. (A) Immune challenge of nonfed (gray bars) and bloodfed (black bars) females with a live mix of *E. coli* and *M. luteus*. (B) Incubation of the 5.1* cell line with heat-killed bacteria: gray bars, *M. luteus*, black bars, *E. coli*. Results of independent experiments are shown as fold induction relative to the unchallenged control (C) at various time points after immune challenge (2 to 24 hr).

similar, except for the unmatched C terminus and four short nonhelical segments (brown in Figure 2A). In particular, the critical reactive and catalytic residues are similarly clustered at the convex surface. This and the very high conservation of surrounding residues (not shown) strongly suggest that aTEP-I is able to form a functional thioester bond very similar to C3. In contrast, the concave surfaces are different in electrostatic maps: the C3d surface (Figure 2B) shows an extended negatively charged acidic pocket, whereas the core of this surface in aTEP-I (Figure 2C) is hydrophobic and is surrounded by a hydrophilic region.

Transcriptional Profiles of *aTEP-I*

Thioester-containing proteins are often acute phase proteins (reviewed by Volanakis, 1995). The immune-responsiveness of *aTEP-I* was analyzed in female mosquitoes after poking with a mixture of live *E. coli* and *Micrococcus luteus*. Bacterial challenge rapidly induced transcription of *aTEP-I*, resulting in maximum transcript levels after 4–6 hr (Figure 3A). Within the adult, *aTEP-I* expression showed a pattern often associated with hemocytes and fat body: presence in head, thorax, and abdomen but virtual absence from the midgut (data not shown). Therefore, we investigated the expression of this gene in an immune-responsive, hemocyte-like cell line 5.1* established by H.-M. Müller (Catteruccia et al., 2000a). In this cell line, *aTEP-I* was expressed constitutively at a low level. Presence of heat-killed *E. coli* markedly induced gene expression, which reached a maximum of 4-fold induction at 4–8 hr of incubation (Figure 3B). In contrast, incubation with heat-killed *M. luteus* did not lead to substantial upregulation. These results indicate that *aTEP-I* is an immune-responsive gene, and that in cells in vitro it is responsive to *E. coli*, but not to *M. luteus*.

Protein Profiles of aTEP-I in Development and Immune Response

To detect the aTEP-I protein both biochemically and by immunolocalization, we raised and affinity-purified rabbit and rat polyclonal antibodies directed against a

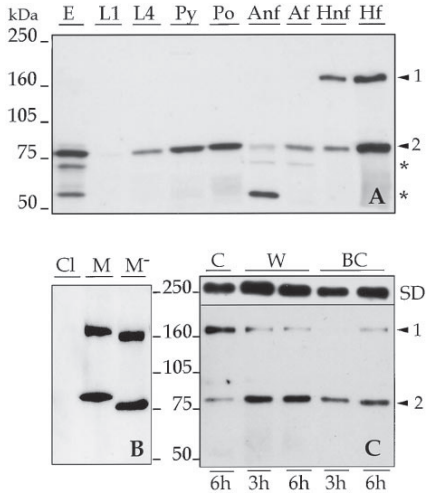


Figure 4. Immunoblotting Analysis of aTEP-I Using 6% SDS-PAGE

(A) Total protein extracts were from embryos (E), first instar larvae (L1), fourth instar larvae (L4), young white pupae (Py), old tanned pupae (Po), nonfed adult females (Anf), and bloodfed adult females (Af). Hemolymph was collected from ten nonfed females (Hnf) and ten bloodfed females (Hf). Total hemolymph extracts and 10–15 μg of total protein from the other extracts (quantified by the Bradford method) were analyzed. Arrowheads 1 and 2 point to putative full-length and cleaved aTEP-I bands, respectively, and asterisks indicate putative proteolytic fragments.

(B) The aTEP-I-positive bands were not detected in mosquito 5.1* cells (Cl) but were abundant in their conditioned medium (M). N-glycosidase F treatment of the conditioned medium (M^-) reduced the sizes of both aTEP-I-positive bands.

(C) Cleavage of aTEP-I protein in the hemolymph of bloodfed adults before (C) and after wounding (W) or bacterial challenge (BC). Hemolymph of ten females was collected at 3 and 6 hr directly into the loading buffer, immunoblotted with anti-aTEP-I rabbit antiserum, and revealed by anti-rabbit HRP-conjugated antibody. The blot was stripped and reprobed with a polyclonal antibody raised against the Sp22D modular protease (Danielli et al., 2000), which served as a loading control (SD). Molecular weight scale (50–250 kDa) is shown on the left.

GST-aTEP-I fusion protein representing two nested C-terminal 21 and 83 kDa polypeptides of aTEP-I (Figure 1A, black horizontal bars). In immunoblotting analysis, both anti-21 kDa and anti-83 kDa antibodies recognized a major protein of 80 kDa in mosquito extracts at all stages of development, as well as smaller moieties, two of which recurred in embryos and adults (arrowhead 2 and two asterisks in Figure 4A, respectively). We interpret these three bands as processed forms of aTEP-I resulting from proteolysis, after which the thioester bond is expected to be rapidly inactivated (reviewed by Dodds and Law, 1998). A bigger, approximately 165 kDa band (Figure 4A, arrowhead 1), interpretable as full-length glycosylated aTEP-I, was abundantly detected only in the hemolymph of adult mosquitoes and, weakly, in embryonic extracts. Interestingly, conditioned medium of 5.1* cell cultures resembled hemolymph in containing both the 165 and 80 kDa bands, while the cells themselves were essentially devoid of aTEP-I (Figure 4B). Treatment of the conditioned medium with N-glycosidase F led to reduction in size of both major bands (Figure 4B), confirming that aTEP-I is a glycosylated protein. Since the 80 kDa moiety is also glycosylated and detected by an antibody directed to the C-terminal region, some of the putative N-glycosylation sites in the vicinity of the thioester site (Figure 1A) are apparently functional. The same protein patterns were observed after SDS-PAGE both in reducing and nonreducing conditions (data not shown). We conclude that aTEP-I is secreted into the hemolymph as a single chain and that fragments generated by subsequent proteolytic cleavages are not connected by disulfide bonds.

We speculated that the proteolytic cleavage of aTEP-I might be inducible by microorganisms or by aseptic injury. To test this, bloodfed females were either challenged by injection of a bacterial mixture or wounded, and 3 or 6 hr later hemolymph was collected and analyzed by immunoblotting. Both wounding and bacterial challenge resulted in cleavage of full-length aTEP-I in hemolymph (Figure 4C). At later time points, we observed replenishment of the 165 kDa form (data not shown).

aTEP-I Is Selectively Expressed in Mosquito Hemocytes

The dispersed association of *aTEP-I* mRNA with head, thorax, and abdomen, the presence of the full-length protein primarily in hemolymph and in conditioned medium of a hemocyte-like cell line suggest that aTEP-I originates in hemocytes. This has been confirmed by immunostaining analysis of body walls: abdomens were dissected open and freed of the alimentary canal (whose epithelium does not itself express significant amounts of aTEP-I; data not shown). aTEP-I was mostly detected in hemocytes attached to tracheae (Figures 5A and 5B) or to the edges of fat body lobes; none was detected in the fat body cells themselves, in the nervous system, in muscles, or in ovaries. Interestingly, the aTEP-I-positive hemocytes (Figure 5C) coexpressed Sp22D (Danielli et al., 2000; Gorman et al., 2000), a hemocyte-specific serine protease (Figures 5D, 5E, and 5F). The only other cells positive for aTEP-I were the pericardial cells (Figure 5G). The putative role of these large binuclear cells in hemolymph filtration and clearance (Hoffmann, 1966; reviewed by Wigglesworth, 1970) suggests that the immunopositive signal may reflect uptake of inactivated aTEP-I rather than in situ synthesis of this protein. Again, aTEP-I and Sp22D were co-detected in the pericardial cells, albeit in different compartments (Figure 5H–5I). aTEP-I was also present in the form of cytoplasmic granules in a small fraction of 5.1* cells, up to 5% of the total, where it colocalized with Sp22D (Figures 5J–5L).

Denaturation-Dependent Fragmentation of aTEP-I

Structure modeling of aTEP-I predicted the functionality of the thioester bond. This internal bond makes known TEP proteins sensitive to denaturation, which causes a cleavage of the peptide bond between the Glu and Gln residues of the thioester site and consequent fragmentation of the molecule; the propensity to fragmentation is abolished by inactivation of the thioester bond through pretreatment with small nucleophilic molecules such as methylamine (MA) (Howard, 1980). Conditioned cell culture medium was collected and treated for 2 hr at 37°C

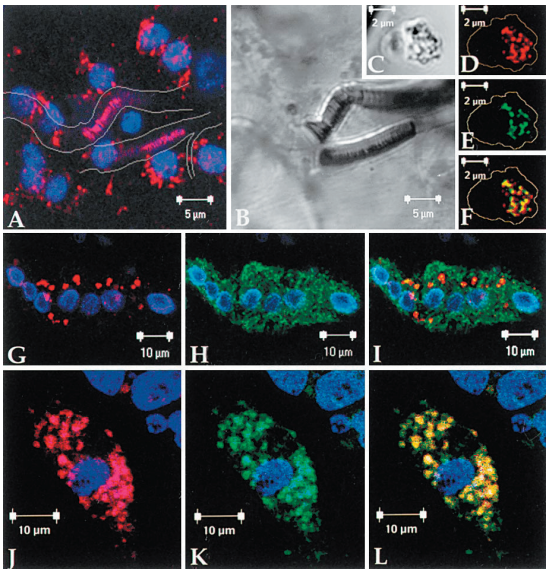


Figure 5. Immunolocalization of aTEP-I in Adult Tissues and a Cell Line

(A–I) Whole mount double staining of adult mosquito body wall showing location of aTEP-I, Sp22D, and nuclear histones (red, green, and blue, respectively). (A) Immunostaining of aTEP-I in hemocytes attached to tracheae, which are traced by thin lines. (B) Phase contrast image of the same field displaying the tracheae. (C) Phase contrast image of a mosquito hemocyte. Coexpression of aTEP-I (D) and Sp22D (E) in a mosquito hemocyte evidenced by merging the images (F). Constitutive presence of aTEP-I (G) and Sp22D (H) in pericardial cells. Note that the merged image (I) reveals that aTEP-I and Sp22D are in separate compartments. Double staining of 5.1* cells for aTEP-I (J) and for Sp22D (K) showing their colocalization in the merge image (L) in the 5.1* cells. Nucleic acids (blue) are colored with TO-PRO™-3.

with MA solution in PBS (0.2 M final concentration) or with PBS alone. Samples of control or MA-treated media were denatured at 80°C for 2 hr and analyzed by immunoblotting (Figure 6A). Consistent with our interpretation that the 80 kDa band is an inactive processed form lacking the thioester bond, this band was not affected by the treatment. In contrast, heat denaturation of the control sample led to appearance of a novel 50 kDa band with concomitant loss of the 165 kDa full-length protein. This shift required the thioester bond, as it was not observed after MA treatment. The estimated size of the novel 50 kDa band is in reasonable agreement with that predicted (54 kDa) for a denaturation-dependent fragmentation of aTEP-I at the thioester motif.

aTEP-I Binds to Bacteria through the Thioester Bond

In vertebrates, the thioester bond allows complement factors to bind covalently to target surfaces. To investigate whether this is also true for mosquitoes and, specifically, to test for binding of aTEP-I to bacteria, we pretreated aTEP-I containing conditioned medium with or without MA, and then incubated comparable *E. coli* or *S. aureus* preparations in these media. After a 5 min exposure to the media, the bacteria were washed and

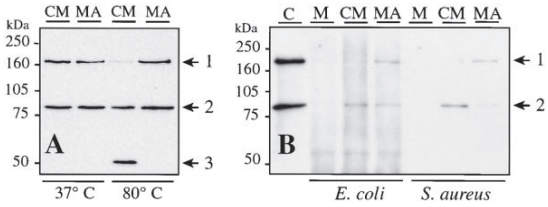


Figure 6. Functionality of the aTEP-I Thioester Site and Its Role in Binding to Bacteria

(A) Denaturation-dependent fragmentation and methylamine reaction of aTEP-I revealed by immunoblotting using 7% SDS-PAGE. Control (CM) and MA-treated (MA) samples of conditioned medium were incubated either at 37°C or at 80°C for 2 hr. Denaturation-induced fragmentation of the control sample occurred at the thioester site, resulting in degradation of full-length aTEP-I (arrow 1) and appearance of a novel 50 kDa band (arrow 3), distinct from the endogenously processed 80 kDa product (arrow 2).

(B) Cell-free conditioned medium of the 5.1* cell line was incubated with (MA) at 0.2 M final concentration or without (CM) methylamine for 2 hr. Fresh medium (M), CM, and MA-treated media were then mixed with the Gram-negative (*E. coli*) or the Gram-positive (*S. aureus*) bacteria for 5 min. The bacteria were centrifuged, washed in PBS, and denatured in 30 μl of reducing buffer at 68°C for 15 min. Aliquots of 15 μl from each sample including the initial conditioned medium (C) were fractionated by 7% SDS-PAGE and immunoblotted as in Figure 4.

their cell walls were mildly disrupted by heat denaturation in a reducing buffer. Bacterial pellets and extracted wall protein fractions were analyzed by immunoblotting (data not shown and Figure 6B). A major aTEP-I-positive signal corresponding to the 80 kDa cleaved fragment was detected both in *E. coli* and *S. aureus* cell wall extracts. Importantly, this signal was dramatically reduced in the extracts exposed to MA-treated medium. We conclude that binding of a processed form of aTEP-I to bacteria requires an intact thioester bond. In the same medium, a surprising faint signal corresponding to the full-length form was repeatedly detected.

aTEP-I Is Essential for Promotion of Phagocytosis of Gram-Negative Bacteria in a Mosquito Cell Line

We set up a phagocytic assay to examine whether binding of aTEP-I to bacteria can opsonize them for phagocytosis. We analyzed the propensity of mosquito 5.1* cells to phagocytose fluorescein-labeled bacteria, which had been pre-exposed to either fresh or conditioned culture media for 15 min. Opsonized bacteria were washed and presented to the mosquito cells in internalization medium. At 15 to 60 min, the mosquito cells were collected, treated with ethidium bromide, and plated on slides for fluorescence microscopy (Figures 7A and 7B). Staining with ethidium bromide helps to discriminate between engulfed (green) and surface-attached (orange) bacteria (Drevets and Campbell, 1991). Cells (n = 100) were counted in three different fields and the percentage of cells that contained one or more engulfed bacteria was used as the phagocytic index (PI). Bacteria preincubated with fresh culture medium were taken up rather slowly; the PI was only 10.6% at 15 min and reached 40.1% at 60 min, when it approached a plateau (Figure 7C and data not shown). Treatment of

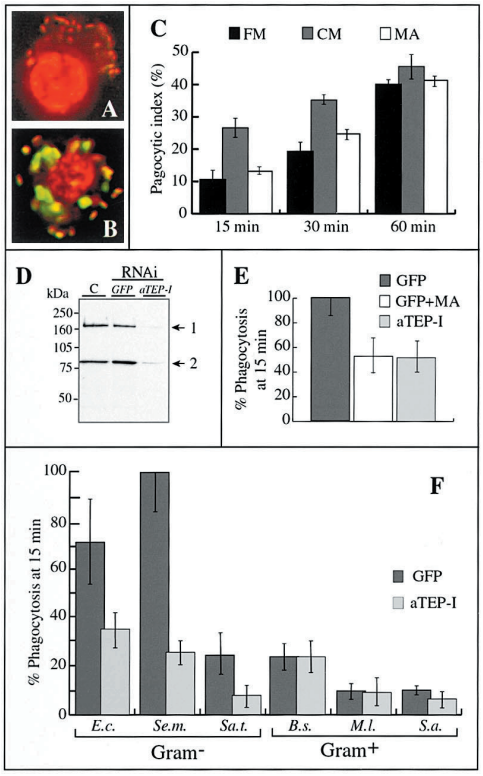


Figure 7. Phagocytosis of FITC-Labeled Bacteria by Mosquito 5.1* Cells

(A and B) Fluorescence microscopy of mosquito phagocytic cells. Treatment with ethidium bromide stains nucleic acids in red. Fluorescein-labeled *E. coli* attached to the cell surface emits in orange (A), whereas engulfed bacteria are colored in green (B).

(C) Phagocytic index (PI) showing the propensity of mosquito 5.1* cells to phagocytose fluorescein-labeled *E. coli* that had been pretreated with fresh medium (FM), conditioned medium (CM), and MA-treated conditioned medium (MA). Opsonized bacteria were washed and presented to naive 5.1* cells in internalization medium for 15, 30, and 60 min.

(D) Knockout of *aTEP-I* by dsRNA interference (RNAi) in mosquito cells. Full-length and processed aTEP-I (arrows 1 and 2) were detected by immunoblotting using 8% SDS-PAGE in control medium (C) and in conditioned medium of cells transfected with 500 ng of *GFP* dsRNA (*GFP*); it was absent from medium of cells transfected with *aTEP-I* dsRNA (*aTEP-I*). Molecular weight scale is on the left.

(E) Effect of *aTEP-I* RNAi knockout on phagocytosis of *E. coli* at 15 min of incubation. Bacterium was opsonized with conditioned medium of the ds(*GFP*) knockout experiments (GFP), with the same medium treated with MA (GFP+MA) and with conditioned medium of ds(*aTEP-I*) knockout (aTEP-I).

(F) Effect of the *aTEP-I* RNAi knockout on the phagocytosis of Gram-negative and Gram-positive bacteria after 15 min of incubation. Comparable preparations (see Experimental Procedures) of *E. coli* (*E.c.*), *Serratia marcescens* (*Se.m.*), *Salmonella typhimurium* (*Sa.t.*), *Bacillus subtilis* (*B.s.*), *Micrococcus luteus* (*M.l.*), and *Staphylococcus aureus* (*S.a.*) were opsonized with the ds(*GFP*) knockout and ds(*aTEP-I*) knockout conditioned media as described above. In (E) and (F), the maximal PI levels were expressed as 100% of phagocytosis. In all graphs vertical bars show mean ± SD (n = 3).

the cells with an inhibitor of actin polymerization, cytochalasin D, or incubation at 4°C drastically inhibited uptake (4 and 1% after 2 hr of incubation, respectively), confirming that the counted events represented phagocytosis rather than actin-independent endocytosis. Preincubation of bacteria with conditioned medium led to precocious phagocytosis (Figure 7C, PI of 27% at 15 min); this phenomenon was still evident at 30 min, but not at 60 min. Thus, the conditioned medium contains soluble factors secreted by 5.1* cells, which opsonize bacteria during the pretreatment and promote phagocytosis.

To address the potential role of mosquito TEPs in promotion of phagocytosis, we inactivated aTEP-I and potentially other aTEPs, present in the conditioned medium by MA treatment. The effect of this treatment on early phagocytosis was spectacular; at 15 min of incubation, the PI was induced only to the level of bacteria preincubated with fresh culture medium, and was 2-fold lower than after pretreatment with normal conditioned medium (Figure 7C). Again, this difference was abolished after 60 min of incubation. These results demonstrate an important role of mosquito TEPs in promoting early phagocytosis.

To specifically deplete aTEP-I from the conditioned medium, we used dsRNA interference (RNAi), which has been recently successfully applied to *Drosophila* cell cultures (Hammond et al., 2000) and is known to inhibit gene expression in a sequence-specific manner (reviewed in Bosher and Labouesse, 2000). Mosquito 5.1* cells were treated with dsRNA representing either a 640 bp 3′ fragment of *aTEP-I* [ds(*aTEP-I*)] or the full-length *GFP* gene [ds(*GFP*)] as a control. Immunoblotting confirmed that the level of aTEP-I was dramatically reduced in the ds(*aTEP-I*) medium, whereas it was present in the ds(*GFP*) medium at the control level (Figure 7D). The specific depletion of aTEP-I from the conditioned medium had a striking effect on phagocytosis of *E. coli* at 15 min of incubation; it reduced the PI to a level as low as that for the MA-treated conditioned medium of the ds(*GFP*) knockout (Figure 7E).

The observation that aTEP-I was essential for opsonization and promotion of phagocytosis of *E. coli* in the hemocyte-like mosquito cell line led us to examine opsonic properties of this protein in phagocytosis of other Gram-negative and Gram-positive bacteria. We followed the scheme described above and compared the PIs of 5.1* cells 15 min after incubation with bacteria opsonized with the conditioned media of ds(*GFP*) or ds(*aTEP-I*) knockout. The highest PI was detected for *Serratia marcescens* and we have expressed it as 100% of phagocytosis (Figure 7F). Phagocytosis of all three species of Gram-negative bacteria was reduced by 50%–75% when microorganisms were opsonized with aTEP-I-depleted medium. Strikingly, the 5.1* cells showed only low levels of phagocytic activity against three species of Gram-positive bacteria, and this activity was not affected by the aTEP-I knockout. These results suggest that aTEP-I promotes phagocytosis of some Gram-negative bacteria, but is not sufficient to promote the uptake by 5.1* cells of any of the three tested species of Gram-positive bacteria.

Discussion

The complement system was thought to be an exclusive hallmark of the host defense of vertebrates until C3-

like molecules were recently cloned in nonvertebrate deuterostomes, a sea urchin (Al-Sharif et al., 1998) and an ascidian (Nonaka and Azumi, 1999). Although thioester proteins had been described in protostomes (Armstrong and Quigley, 1987; Kopacek et al., 2000), they appeared to exhibit α_2-macroglobulin-like protease inhibitory activities rather than complement-like properties. In this report, we present the first functional evidence that a protostome, the mosquito *Anopheles gambiae,* expresses a thioester-containing protein that resembles complement in promoting phagocytosis. Our results, therefore, indicate that the complement-like proteins have a much longer history than previously thought and can be traced back to dipteran insects.

At the structural level, aTEP-I exhibits specific features of both C3 and α_2M, suggesting that it has retained ancient properties of both and may be a prototype of their common ancestor molecule. For example, our modeling studies point to a high conservation of key residues between C3 and aTEP-I on the convex surface of the respective C3d-like regions, although here the sequence similarity does not exceed 25.5%. In contrast, the corresponding region of α_2M from *Limulus*, which shows 32.6% of similarity to aTEP-I and 29.5% to human C3d, has structurally diverged and cannot be modeled successfully using the C3d structure as a template (data not shown). The structural conservation of the C3d-like region of C3 and aTEP-I probably reflects shared functional requirements for covalent attachment of these molecules to microbial surfaces. In keeping with this, we observed that the characteristic complement-like properties of aTEP-I, such as the attachment to bacteria and the subsequent promoting of precocious phagocytosis, are both dependent on the functional thioester bond.

As is the case for α_2M, aTEP-I is secreted as a single chain product. It is tempting to speculate that in these single chain molecules, the bait-like region is exposed in a way that makes it an easy target for a wide variety of activators. In contrast, the two- and three-chain factors C3, C4, and C5 rely on specific convertase complexes for their activation. We propose that activation of insect TEPs, unlike that of complement, does not require such convertase complex. This speculation is supported, albeit negatively, by an extensive search of the fruitfly genome that failed to find such specific complement-activating proteases as MASP or factors B and C2. Taking into account that such activating components of both the lectin and the alternative pathways exist in ascidians (reviewed by Nonaka, 2000), we suggest that these pathways first appeared in protochordate deuterostomes, and that insect TEPs are activated by endogenous proteases that are set free upon injury, or by proteases of pathogen origin. Indeed, our results imply that aTEP-I can be cleaved by endogenous proteases in the hemolymph after injury.

Although we have shown that aTEP-I binds to both *E. coli* and *S. aureus*, preliminary data indicate that it binds poorly to *M. luteus*; thus, the range of its binding specificity remains to be examined. A potential role of aTEP-I in host defense during *Plasmodium* infection was not addressed in this paper. However, preliminary data indicate that this gene responds to the presence of the parasite in the midgut by transcriptional upregulation. Further studies will clarify whether aTEP-I or other members of aTEP family are implicated in antiparasitic responses, since three additional members of the thioester family have been recently identified in *A. gambiae* (E. A. L., L. F. M., and S. B., unpublished data; Dimopoulos et al., 2000; Oduol et al., 2000).

We observed that in the absence of a functional thioester bond, full-length aTEP-I can bind, albeit weakly, to *E. coli* and to *S. aureus*. Similarly, several ELISA-based reports have shown binding of α_2M to pathogens (Araujo-Jorge et al., 1990). These results may indicate that early in evolution, this type of opsonin was able to bind noncovalently to microorganisms without proteolytic activation. We propose that later in evolution, such a primitive opsonization system adopted a proteolytic activation step, concomitant with covalent binding mediated by the thioester bond, and that this evolutionary step permitted a fine-tuning of pathogen recognition coupled to a high-affinity binding to the target.

The cell culture model system based on the hemocyte-like 5.1* cell line proved to be helpful in the establishment of a quantitative phagocytic test for the mosquito. The possibility to specifically inactivate the expression of *aTEP-I* by dsRNA interference demonstrated the importance of aTEP-I in opsonization for phagocytosis of the Gram-negative bacteria, *E. coli*, *S. marcescens*, and *S. typhimurium*. In contrast, the low level of phagocytosis of the Gram-positive bacteria, *B. subtilis*, *M. luteus*, and *S. aureus*, and its independence from the aTEP-I knockout, suggest that aTEP-I is not sufficient to promote phagocytosis of this latter type of bacteria in these cells. Thus, this in vitro cell culture model has provided a first insight into the complexity of the phagocytic system in mosquitoes and can potentially serve to further dissect the molecules involved in regulation of phagocytosis.

The bacteria binding and phagocytosis-promoting activities of aTEP-I clearly define this protein as an opsonin for some Gram-negative bacteria. Opsonization is thought to be the most conserved function of thioester-containing proteins (reviewed by Dodds and Law, 1998), and the discovery of an opsonizing insect TEP is important evidence favoring this concept. The predominant hemocyte origin of aTEP-I and the ability of this glycoprotein to promote phagocytosis by the hemocyte-like 5.1* cells highlight the role of hemocytes in the innate defense system of *Anopheles*. Several recent reports have pointed to the importance of insect hemocytes in providing interaction between cellular and humoral immune responses (Basset et al., 2000; Elrod-Erickson et al., 2000; Schneider and Shahabuddin, 2000). It is not clear whether aTEP-I is the only mosquito opsonin that promotes phagocytosis of Gram-negative bacteria. The effect of chemical inactivation or of RNAi depletion of aTEP-I is most striking at early time points of phagocytosis, which may suggest that additional opsonins can promote phagocytosis at later stages. An alternative explanation for the later leveling off of the differences between fresh and conditioned media may be the accumulation of newly synthesized aTEP-I, or, most likely, opsonin-independent internalization.

Finally, this report presents a successful application of double-stranded RNA interference using a cell culture model system to study gene function in mosquito, where

in vivo genetic analysis cannot be applied. This technique, in conjunction with subsequent validation by the very demanding method of genetic transformation of anopheline mosquitoes (Catteruccia et al., 2000b), opens up the possibility of functionally characterizing genes involved in complex phenotypes, such as the immune responses in vector insects.

Experimental Procedures

Biological Material

Mosquito colonies were maintained as described (Richman et al., 1997). Young adult nonfed and bloodfed females were injured with a thin needle (0.2 mm diameter) or infected by a needle dipped into a concentrated overnight culture of *Escherichia coli* strain 1106 and *Micrococcus luteus* strain A270. The Sua 5.1* cell line was cultured and incubated with heat-killed *E. coli* or *M. luteus* as described (Müller et al., 1999).

Cloning and Sequencing

Using a primer F2: 5′-GGITGYGGIGAGCAGAATATG-3′ corresponding to the thioester motif and a reverse primer T7: 5′-TAATACGACT CACTATAGGG-3′, a 1.6 kb PCR product was amplified from a lambda ZAP express (Stratagene) cell line cDNA library, cloned into pGEM-TA easy vector (Promega), and confirmed by sequencing. RACE-PCR using a specific primer: 5′-GTGCGGCCCGCTACGG TAGCG-3′ and the Marathon Kit (Clontech) produced a 1.5 kb fragment, which was used for screening a thoracic cDNA library (Arca et al., 1999). A clone of 4.2 kb was sequenced by Eurogentec (Herstal, Belgium) and analyzed using the SMART and BLAST programs.

Modeling of a Critical aTEP-I Region

The alignment of aTEP-I with C3d was performed by CLUSTAL_X and refined by introducing corresponding regions of other TEP sequences. The WHAT IF program (Vriend, 1990) was used for modeling as described (Chinea et al., 1995). Sequence identity between the template and the aTEP-I model was 25.5%. The model quality was checked with WHAT IF structure validation tools (Rodriguez et al., 1998). The sequence alignment and the coordinates of the model are available online.

Transcriptional Profiling by RT-PCR

Total RNA from adult mosquitoes and cultured cells was isolated with TRIZOL Reagent (Gibco BRL) according to the suppliers' instructions. First strand cDNA synthesis was performed as described (Müller et al., 1999). The aTEP-I primers 5′-GGAAATACCGGAGGA CAC-3′ and 5′-TGCTACCTTAAAGCGTCTG-3′ were used at 10 pmol per 25 μl PCR reactions to amplify a 409 bp fragment with a standard program (30 s at 94°C; 30 s at 56°C; 60 s at 72°C) for 25 cycles (cDNA from adults) or 32 cycles (cDNA from cells). The internal S7 control (Salazar et al., 1993) was amplified using 19 cycles of the same program. The linear range of all amplification reactions has been determined empirically. After electrophoresis on 1% agarose, gels were stained with the SYBR green dye (Molecular Probes) for 1 hr and analyzed with a fluorimager (Fuji).

Generation and Immunopurification of aTEP-I Antibody

Chimeric proteins composed of glutathione S-transferase (GST) and the C-terminal regions of aTEP-I were produced using the GST-aTEP-I vectors pLL2 and pLL4. For pLL2, a BamHI-XhoI 2351 bp fragment of *aTEP-I* cDNA was cloned into corresponding sites of the pGEX-3T expression vector (Frorath et al., 1992). For pLL4, a HindIII-XhoI 680 bp fragment was first subcloned into the pBluescript KS vector, and then recloned as a 705 bp BamHI-XhoI fragment into pGEX-3T. Fusion proteins were expressed in *E. coli* strain BL 21 and purified according to standard procedures (Pharmacia) on a glutathione Sepharose 4B column. For each protein, two rats and two rabbits were immunized with 50 μg protein per rat and 250 μg protein per rabbit using R-700 and R-730 adjuvants, respectively (RIBI Immunochem Research, Inc.). Rats were boosted every third week with 50 μg antigen until final bleed. Rabbits were boosted every fourth week with 250 μg antigen for 6 months. Identical protein profiles were detected by immunoblotting using antibodies directed against the pLL2 and pLL4 fusion proteins. Rabbit antibody against the pLL4 protein was affinity-purified using FPLC-purified antigen bound to activated CNBr Sepharose 4B beads as described (Harlow and Lane, 1988).

Immunoblotting

Protein extracts from adults, hemolymph collection, and immunoblotting were performed as described (Danielli et al., 2000). In all experiments, aTEP-I antiserum was used at 1:1,000 dilution. Bound antibodies were detected by an anti-rabbit IgG conjugated to horseradish peroxidase (Promega) at 1:40,000 using Western Blot Chemiluminescence Reagent Plus Kit (NEN™ Life Science Products). Protein loading and efficiency of protein transfer were monitored by blot staining with Indian ink (Harlow and Lane, 1988). Asparagine-linked glycan chains on aTEP-I were cleaved using the N-Glycosidase F Deglycosylation Kit (Roche Molecular Biochemicals) according to a standard protocol, with 5 μl of conditioned medium incubated for 15 min with or without N-glycosidase F (1 μl) at 37°C, followed by immunoblotting.

Immunostaining

Whole mount and cell line staining was performed as described (Danielli et al., 2000). Briefly, after blocking, the body walls were incubated overnight at 4°C with a mixture of primary antibodies [aTEP-I at 1:750, Sp22D at 1:1,000 and histone (MAB052, Chemicon) at 1:500] followed by 1 hr incubation with secondary antibodies (The Jackson Laboratory, 1:1,000). The samples were analyzed with a Zeiss LSM 510 confocal microscope. Cell line nuclei were colored with TO-PRO™-3 (Molecular Probes, 1:1,000) for 15 min.

Binding Assay

Conditioned media of 5.1* cell line were incubated with or without MA (0.2 M final concentration) for 2 hr at 37°C. Comparable preparations of chemically inactivated *E. coli* (K-12 strain) and *S. aureus* BioParticles®, fluorescein conjugate (Molecular Probes) were incubated for 5 min with the conditioned media, centrifuged, washed with PBS, and treated with 30 μl of denaturation buffer (Rosenbusch, 1974). After 15 min at 68°C the reactions were separated into soluble and nonsoluble fractions and analyzed by immunoblotting.

dsRNA Production and Interference

For the production of dsRNA, we modified the pBluescript-based pLL6 plasmid (a gift of T. Loukeris, EMBL, Heidelberg), which contained the full 721 bp *GFP* sequence between EcoRI and ApaI sites, by inserting a second T7 promoter in reverse orientation between the SstI and SstII. The resulting pLL6ds plasmid was used to produce a control *GFP* dsRNA. For producing *aTEP-I* dsRNA, the *GFP* fragment was exchanged for the 661 bp fragment of pLL4 using SmaI-KpnI sites of pLL6ds resulting in pLL17. Purified linearized plasmids served as templates for RNA synthesis using the MEGASCRIPT T7 transcription kit (Ambion). RNAs were isopropanol precipitated, quantified, and mixed for annealing in equal quantities by heating for 5 min at 95°C and cooling down to room temperature. Resulting dsRNAs were analyzed on agarose gel. In interference experiments, 500 ng of each dsRNA was transfected into a confluent culture of 5.1* cells using the Effectene Transfection Reagent (Qiagen) following the manufacturers' instructions. The conditioned media were exchanged for a fresh medium after 2 days. Samples for immunoblotting and phagocytic assay were collected 5 days later.

Phagocytic Assay

Cells and conditioned media (1- to 2-week-old) of the 5.1* mosquito confluent cell cultures (Catteruccia et al., 2000a) were used. Three bacterial species were comparable commercial preparations (Molecular Probes): BioParticles®, fluorescein conjugates *E. coli* (K-12 strain) and *S. aureus* (Molecular Probes) and *S. marcescens* (custom prepared by the manufacturer). Overnight cultures of *Salmonella typhimurium*, *Bacillus subtilis* (provided as cultures by P. Bulet, IBMC, Strasbourg), and *M. luteus* were prepared in a consistent manner. Bacteria were killed by a 30 min incubation with 4% formaldehyde (Polysciences, Inc.), washed in PBS, and fluorescein conjugated (Drevets and Campbell, 1991). All bacteria were incubated

with the conditioned media for 15 min, washed in PBS, and presented to mosquito cells (in a ratio of 20:1, respectively) in internalization medium consisting of 10 mM HEPES, 5 mM glucose in MEM (SEROMED®), pH 7.4. The phagocytic test was adapted from Drevets and Campbell (1991). All experiments were performed at least three times.

Acknowledgments

We thank Professor J. A. Hoffmann for initial support and constant interest in the work, A. Danielli and T. G. Loukeris for fruitful discussions, T. Gibson for help with the phylogenetic analysis, and H.-M. Müller for the 5.1* cell line. This work was supported by NIH program project (1PO1 AI44220) and European Commission Training and Mobility of Researchers Program (ERB4061PL95) grants. E. A. L. is a fellow of the European Molecular Biology Organization.

Received September 7, 2000; revised January 18, 2001.

References

Al-Sharif, W.Z., Sunyer, J.O., Lambris, J.D., and Smith, L.C. (1998). Sea urchin coelomocytes specifically express a homologue of the complement component C3. J. Immunol. *160*, 2983–2997.

Araujo-Jorge, T.C., de Meirelles, M. de N., and Isaac, L. (1990). *Trypanosoma cruzi*: killing and enhanced uptake by resident peritoneal macrophages treated with alpha-2-macroglobulin. Parasitol. Res. *76*, 545–552.

Arca, B., Lombardo, F., de Lara Capurro, M., della Torre, A., Dimopoulos, G., James, A.A., and Coluzzi, M. (1999). Trapping cDNAs encoding secreted proteins from the salivary glands of the malaria vector *Anopheles gambiae*. Proc. Natl. Acad. Sci. USA *96*, 1516–1521.

Armstrong, P.B., and Quigley, J.P. (1987). Limulus alpha 2-macroglobulin. First evidence in an invertebrate for a protein containing an internal thiol ester bond. Biochem. J. *248*, 703–707.

Barillas-Mury, C., Wizel, B., and Soo Han, Y. (2000). Mosquito immune responses and malaria transmission: lessons from insect model systems and implications for vertebrate innate immunity and vaccine development. Insect Biochem. Mol. Biol. *30*, 429–442.

Basset, A., Khush, R.S., Braun, A., Gardan, L., Boccard, F., Hoffmann, J.A., and Lemaitre, B. (2000). The phytopathogenic bacteria *Erwinia carotovora* infects *Drosophila* and activates an immune response. Proc. Natl. Acad. Sci. USA *97*, 3376–3381.

Bosher, J.M., and Labouesse, M. (2000). RNA interference: genetic wand and genetic watchdog. Nat. Cell Biol. *2*, E31–E36.

Catteruccia, F., Nolan, T., Blass, C., Müller, H.M., Crisanti, A., Kafatos, F.C., and Loukeris, T.G. (2000a). Toward *Anopheles* transformation: Minos element activity in anopheline cells and embryos. Proc. Natl. Acad. Sci. USA *97*, 2157–2162.

Catteruccia, F., Nolan, T., Loukeris, T.G., Blass, C., Savakis, C., Kafatos, F.C., and Crisanti, A. (2000b). Stable germline transformation of the malaria mosquito *Anopheles stephensi*. Nature *405*, 959–962.

Chinea, G., Padron, G., Hooft, R.W., Sander, C., and Vriend, G. (1995). The use of position-specific rotamers in model building by homology. Proteins *23*, 415–421.

Danielli, A., Loukeris, T.G., Lagueux, M., Müller, H.M., Richman, A., and Kafatos, F.C. (2000). A modular chitin-binding protease associated with hemocytes and hemolymph in the mosquito *Anopheles gambiae*. Proc. Natl. Acad. Sci. USA *97*, 7136–7141.

Dimopoulos, G., Casavant, T.L., Chang, S., Scheetz, T., Roberts, C., Donohue, M., Schultz, J., Benes, V., Bork, P., Ansorge, W., 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. USA *97*, 6619–6624.

Dimopoulos, G., Müller, H.M., Levashina, E.A., and Kafatos, F.C. (2001). Innate immune defense against malaria infection in the mosquito. Curr. Opin. Immunol. *13*, 79–88.

Dodds, A.W., and Law, S.K. (1998). The phylogeny and evolution of the thioester bond-containing proteins C3, C4 and alpha 2-macroglobulin. Immunol. Rev. *166*, 15–26.

Drevets, D.A., and Campbell, P.A. (1991). Macrophage phagocytosis: use of fluorescence microscopy to distinguish between extracellular and intracellular bacteria. J. Immunol. Methods *142*, 31–38.

Elrod-Erickson, M., Mishra, S., and Schneider, D. (2000). Interactions between the cellular and humoral immune responses in *Drosophila*. Curr. Biol. *10*, 781–784.

Frorath, B., Abney, C.C., Berthold, H., Scanarini, M., and Northemann, W. (1992). Production of recombinant rat interleukin-6 in *Escherichia coli* using a novel highly efficient expression vector pGEX-3T. Biotechniques *12*, 558–563.

Gorman, M.J., Schwartz, A.M., and Paskewitz, S. (1998). The role of surface characteristics in eliciting humoral encapsulation of foreign bodies in Plasmodium-refractory and -susceptible strains of *Anopheles gambiae*. J. Insect. Physiol. *44*, 947–954.

Gorman, M.J., Andreeva, O.V., and Paskewitz, S.M. (2000). Sp22D: a multidomain serine protease with a putative role in insect immunity. Gene *251*, 9–17.

Hammond, S.M., Bernstein, E., Beach, D., and Hannon, G.J. (2000). An RNA-directed nuclease mediates post-transcriptional gene silencing in *Drosophila* cells. Nature *404*, 293–296.

Harlow, E., and Lane, D. (1988). Antibodies: a Laboratory Manual (Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press).

Hoffmann, J.A. (1966). Etude ultrastructurale de l'absorption de saccharate de fer par les cellules pericardiales de *Locusta migratoria* (Orthoptere). C. R. Acad. Sci. Hebd. Seances Acad. Sci. D *262*, 1471–1496.

Hoffmann, J.A., Kafatos, F.C., Janeway, C.A., and Ezekowitz, R.A. (1999). Phylogenetic perspectives in innate immunity. Science *284*, 1313–1318.

Howard, J.B. (1980). Methylamine reaction and denaturation-dependent fragmentation of complement component 3. Comparison with alpha2-macroglobulin. J. Biol. Chem. *255*, 7082–7084.

Kopacek, P., Weise, C., Saravanan, T., Vitova, K., and Grubhoffer, L. (2000). Characterization of an alpha-macroglobulin-like glycoprotein isolated from the plasma of the soft tick *Ornithodoros moubata*. Eur. J. Biochem. *267*, 465–475.

Lagueux, M., Perrodou, E., Levashina, E.A., Capovilla, M., and Hoffmann, J.A. (2000). Constitutive expression of a complement-like protein in Toll and JAK gain-of-function mutants of *Drosophila*. Proc. Natl. Acad. Sci. USA *97*, 11427–11432.

Lowenberger, C.A., Kamal, S., Chiles, J., Paskewitz, S., Bulet, P., Hoffmann, J.A., and Christensen, B.M. (1999). Mosquito-*Plasmodium* interactions in response to immune activation of the vector. Exp. Parasitol. *91*, 59–69.

Müller, H.M., Dimopoulos, G., Blass, C., and Kafatos, F.C. (1999). A hemocyte-like cell line established from the malaria vector *Anopheles gambiae* expresses six prophenoloxidase genes. J. Biol. Chem. *274*, 11727–11735.

Nagar, B., Jones, R.G., Diefenbach, R.J., Isenman, D.E., and Rini, J.M. (1998). X-ray crystal structure of C3d: a C3 fragment and ligand for complement receptor *2*. Science *280*, 1277–1281.

Nicholls, A., Sharp, K.A., and Honig, B. (1991). Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. Proteins *11*, 281–296.

Nonaka, M. (2000). Origin and evolution of the complement system. In Origin and Evolution of the Vertebrate Immune System, L. Du Pasquier and G.W. Litman, eds. (Springer Verlag), pp. 37–50.

Nonaka, M., and Azumi, K. (1999). Opsonic complement system of the solitary ascidian, *Halocynthia roretzi*. Dev. Comp. Immunol. *23*, 421–427.

Oduol, F., Xu, J., Niare, O., Natarajan, R., and Vernick, K.D. (2000). Genes identified by an expression screen of the vector mosquito *Anopheles gambiae* display differential molecular immune response to malaria parasites and bacteria. Proc. Natl. Acad. Sci. USA *97*, 11397–11402.

Richman, A.M., Dimopoulos, G., Seeley, D., and Kafatos, F.C. (1997).

Plasmodium activates the innate immune response of *Anopheles gambiae* mosquitoes. EMBO J. *16*, 6114–6119.

Rodriguez, R., Chinea, G., Lopez, N., Pons, T., and Vriend, G. (1998). Homology modeling, model and software evaluation: three related resources. Bioinformatics *14*, 523–528.

Rosenbusch, J.P. (1974). Characterization of the major envelope protein from *Escherichia coli*. Regular arrangement on the peptidoglycan and unusual dodecyl sulfate binding. J. Biol. Chem. *249*, 8019–8029.

Salazar, C.E., Mills-Hamm, D., Kumar, V., and Collins, F.H. (1993). Sequence of a cDNA from the mosquito *Anopheles gambiae* encoding a homologue of human ribosomal protein S7. Nucleic Acids Res. *21*, 4147.

Schneider, D., and Shahabuddin, M. (2000). Malaria parasite development in a *Drosophila* model. Science *288*, 2376–2379.

Sottrup-Jensen, L. (1989). Alpha-macroglobulins: structure, shape, and mechanism of proteinase complex formation. J. Biol. Chem. *264*, 11539–11542.

Volanakis, J.E. (1995). Transcriptional regulation of complement genes. Annu. Rev. Immunol. *13*, 277–305.

Volanakis, J.E. (1998). Overview of the complement system. In The Human Complement System in Health and Disease, J.E. Volanakis, and M.M. Frank, eds. (New York: Marcel Dekker, Inc.), pp. 9–32.

Vriend, G. (1990). WHAT IF: a molecular modeling and drug design program. J. Mol. Graph. *8*, 52–56.

Wigglesworth, V.B. (1970). The pericardial cells of insects: analogue of the reticuloendothelial system. J Reticuloendothel. Soc. *7*, 208–216.

GenBank Accession Number

The GenBank accession number for the aTEP- I sequence reported in this paper is AF291654. For additional information, see http://www.embl-heidelberg.de/ExternalInfo/kafatos/publications/TEP.

Chapter 3

Phylogenetic Analyses of Thioester-containing Proteins

Phylogenetic Analyses of Thioester-containing Proteins...

... In insects.

A total of six *Tep* genes had been identified in *Drosophila*, and the overall number of *dTeps* reaches 10, as one of them, *dTep2*, encodes 5 alternative splice forms, a unique example in the TEP family (Lagueux et al., 2000). This suggested that the family of thioester-containing proteins might be represented by more than one member in dipterans. Hence we looked for additional *TEPs* in *A. gambiae*. In the absence of the genome sequence, we based our searches on homology cloning and on the analysis of the data from a pilot EST sequencing project (Dimopoulos et al., 2000) and from the cloning of differentially expressed genes (Oduol et al., 2000). This led us to identify three additional *TEPs*, *TEP2, TEP3* and *TEP4* (also named *Imcr14* in an independent study (Oduol et al., 2000)). To our great surprise, the sequencing of the mosquito genome (Holt et al., 2002) revealed the existence of a total of 19 *TEP* sequences. Noteworthy, the genome was sequenced from the *A. gambiae* PEST strain, which is a hybrid strain derived from crossing a laboratory colony originating from Nigeria with the offspring of field-collected mosquitoes from Kenya (Mukabayire et al., 1996). It was chosen partly because it lacked the large-scale chromosomal rearrangements that are typical of both wild populations and most other *A. gambiae* colonies. However, during the genome assembly, it became apparent that this strain possessed an unexpected amount of heterogeneity, posing problems for the automatic assembly of the genome (Holt et al., 2002). Indeed, some polymorphic sequence stretches, which were different enough (90-95% sequence identity) to be considered as distinct by automatic filters, appear as artificial duplications in the genome sequence. Although causing an overestimation of the number of genes, this heterogeneity also provides a rich depository of allelic sequences (reviewed in Craig et al., 2003). Our identification of the thioester-containing proteins of *A. gambiae* and their comparison with *Drosophila Teps* were part of a big collaborative effort led by Fotis Kafatos to annotate genes that are or might be involved in mosquito immune defences. The results were published in a common publication (Christophides et al., 2002) and can be found on pages 77-85.

Our analysis of the mosquito *TEPs* and their comparison with the fruitfly *Teps* led us to the following conclusions:

• 19 *TEP* sequences are present in the mosquito genome, including the 4 previously identified *TEPs*.

• Among them, we believe that 4 sequences, *TEP16-19*, are allelic forms of *TEP1, TEP5, TEP6* and *TEP8*, respectively. The potential functional relevance of these alleles will be discussed in Chapter 6. Therefore the *TEP* family in *A. gambiae* is likely to be represented by a total of 15 genes.

• At least nine (and maybe eleven) out the fifteen mosquito TEPs are devoid of the canonical thioester motif. Further functional analysis of mosquito TEPs which lack thioester might provide new insights into the importance of thioester motif and it requirements for the TEP function.

• The comparative analysis of the mosquito and fruitfly TEPs revealed a limited number of orthologs and the presence of large species-specific expansions, indicating that TEPs, in these two sister species, have a minimal number of common ancestors. All the *Anopheles*-specific *TEPs* are located in clusters at 39C-40B on the left arm of the third chromosome (Fig. 8). In addition, most of the insect TEPs that were reported to be inducible upon immune challenge are included within the species-specific expansions. Therefore, we speculate that these expansions might represent finely shaped evolution of TEPs in response to distinct pathogenic environments in the fruitfly and the mosquito.

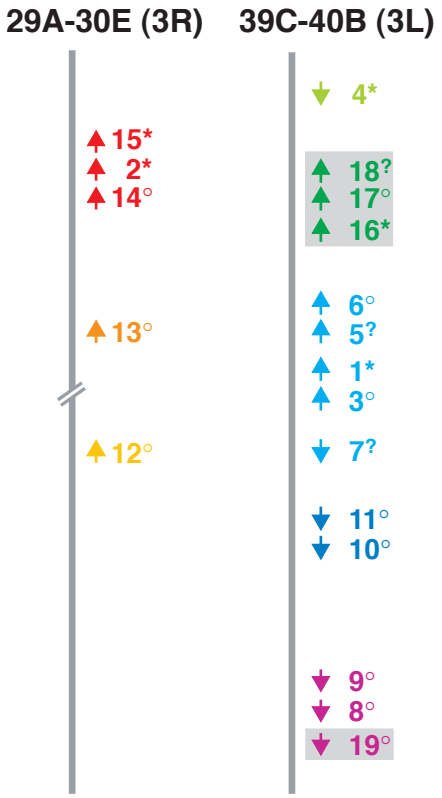


Figure 8. Localization of *Anopheles TEPs* on chromosome 3. The genomic organization of several genes changed since the initial release of the *A. gambiae* genome sequence. In this graph we represent the localization and orientation of the *TEP* genes as they are in the last version of the mosquito genome (v19.2a.1, 29/09/2003). Colors identify the different clusters of *TEPs* separated by at least 100 kb. *TEP16 – 19*, the putative allelic forms of *TEP1, 5, 6* and *8*, respectively, are shaded in grey. Superscript symbols after the names indicate that the thioester is (*) present, (º) absent or (?) unknown.

... In Metazoans.

The availability of the complete genome sequences of both *D. melanogaster* and *A. gambiae* provided us a unique opportunity to extend our initial phylogenetic analysis of animal thioester-containing proteins to the complete set of *Drosophila* and mosquito TEPs and to additional sequences presenting sufficient sequence similarity that we could find in databases. Results of this analysis are presented in a review we published in 2004 (Blandin and Levashina, 2004b) and can be found on pages 87-94.

The phylogenetic analysis of TEPs across the animal kingdom confirmed the previously demonstrated separation of thioester-containing proteins into three families: complement factors, α2Ms and invertebrate TEPs. We could not detect any α2M-like or any complement-like TEP in insects. This and the large species-specific expansions suggest that in insects TEPs undergo dramatic selection pressures that result in a set of polypeptides specific for each species.

In conclusion, in the fruitfly and in the mosquito, TEPs are represented by large families, most probably derived from species-specific duplications and diversification that gave rise to large species-specific expansions. We propose that these expansions reflect the adaptation of the two dipteran species to distinct pathogenic environments. Further functional analysis of these genes and especially of the *Anopheles*-specific *TEPs*, will most probably be informative about the mosquito immune response against malaria parasites.

Article 2

Immunity-Related Genes and Gene Families in *Anopheles gambiae*

George K. CHRISTOPHIDES, Evgeny ZDOBNOV,
Carolina BARILLAS-MURY, Ewan BIRNEY,
Stéphanie BLANDIN, et al

Science 2002, 298: 159-165

arms of the different species had already been noticed in the 1940s (*54*). Most of the interspecies rearrangements can be attributed to the occurrence of paracentric inversions (pericentric inversions degrade the integrity of the chromosomes). Additional processes such as simple or Robertsonian translocations (although occurring much less frequently than inversions in *Drosophila*) presumably would most easily explain major exchanges between chromosomal arms, which our analysis indicated. Finally, transposon-mediated rearrangements involving large chromosomal segments (*60*, *61*) could also have led to the extensive recombinations observed in our interspecies comparisons. The sequencing of additional insect genomes in the future will certainly help elucidate some of these evolutionary consequences.

References

1. M. W. Gaunt, M. A. Miles, *Mol. Biol. Evol.* **19**, 748 (2002).
2. D. K. Yeates, B. M. Wiegmann, *Annu. Rev. Entomol.* **44**, 397 (1999).
3. N. J. Besansky, J. R. Powell, *J. Med. Entomol.* **29**, 125 (1992).
4. M. Ashburner, in *Drosophila: A Laboratory Handbook* (Cold Spring Harbor Laboratory Press, Plainview, NY, 1989) p. 74.
5. R. A. Holt *et al.*, *Science* **298**, 129 (2002).
6. W. M. Fitch, *Syst. Zool.* **19**, 99 (1970).
7. R. L. Tatusov, E. V. Koonin, D. J. Lipman, *Science* **278**, 631 (1997).
8. P. Bork, *Genome Res.* **10**, 398 (2000).
9. E. S. Lander *et al.*, *Nature* **409**, 860 (2001).
10. S. Aparicio *et al.*, *Science* **297**, 1301 (2002).
11. M. Ashburner *et al.*, *Nature Genet.* **25**, 25 (2000).
12. G. K. Christophides *et al.*, *Science* **298**, 159 (2002).
13. A. T. Monnerat *et al.*, *Mem. Inst. Oswaldo Cruz* **97**, 589 (2002).
14. M. Affolter, T. Marty, M. A. Vigano, A. Jazwinska, *EMBO J.* **20**, 3298 (2001).
15. E. M. Zdobnov, R. Apweiler, *Bioinformatics* **17**, 847 (2001).
16. R. Apweiler *et al.*, *Bioinformatics* **16**, 1145 (2000).
17. A. Bateman *et al.*, *Nucleic Acids Res.* **30**, 276 (2002).
18. I. Letunic *et al.*, *Nucleic Acids Res.* **30**, 242 (2002).
19. S. R. Schmid, P. Linder, *Mol. Microbiol.* **6**, 283 (1992).
20. G. Dimopoulos *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **97**, 6619 (2000).
21. G. Dimopoulos *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 8814 (2002).
22. C. M. Adema, L. A. Hertel, R. D. Miller, E. S. Loker, *Proc. Natl. Acad. Sci. U.S.A.* **94**, 8691 (1997).
23. S. Gokudan *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **96**, 10086 (1999).
24. N. Kairies *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **98**, 13519 (2001).
25. H. Ranson *et al.*, *Science* **298**, 179 (2002).
26. S. M. Kanzok *et al.*, *Science* **291**, 643 (2001).
27. E. M. Zdobnov *et al.*, data not shown.
28. J. M. Ribeiro, J. G. Valenzuela, *J. Exp. Biol.* **202**, 809 (1999).
29. C. Barillas-Mury, unpublished results.
30. K. J. Schmid, D. Tautz, *Proc. Natl. Acad. Sci. U.S.A.* **94**, 9746 (1997).
31. P. Green *et al.*, *Science* **259**, 1711 (1993).
32. H. D. Youn, J. O. Liu, *Immunity* **13**, 85 (2000).
33. H. D. Youn, L. Sun, R. Prywes, J. O. Liu, *Science* **286**, 790 (1999).
34. A. J. Clark, K. Bloch, *J. Biol. Chem.* **234**, 2578 (1959).
35. R. H. Dadd, in *Comprehensive Insect Physiology, Biochemistry and Pharmacology*, L. I. Gilbert, Ed. (Pergamon, Oxford, 1985), vol. 4, pp. 313–390.
36. F. Lyko, *Trends Genet.* **17**, 169 (2001).
37. D. A. Petrov, *Trends Genet.* **17**, 23 (2001).
38. W. H. Li, T. Gojobori, M. Nei, *Nature* **292**, 237 (1981).
39. D. Torrents *et al.*, data not shown.
40. D. A. Petrov, E. R. Lozovskaya, D. L. Hartl, *Nature* **384**, 346 (1996).
41. A. Stoltzfus, J. M. Logsdon Jr., J. D. Palmer, W. F. Doolittle, *Proc. Natl. Acad. Sci. U.S.A.* **94**, 10739 (1997).
42. W. Gilbert, S. J. de Souza, M. Long, *Proc. Natl. Acad. Sci. U.S.A.* **94**, 7698 (1997).
43. I. B. Rogozin, J. Lyons-Weiler, E. V. Koonin, *Trends Genet.* **16**, 430 (2000).
44. D. Schmucker *et al.*, *Cell* **101**, 671 (2000).
45. I. Letunic, R. R. Copley, P. Bork, *Hum. Mol. Genet.* **11**, 1561 (2002).
46. E. L. George, M. B. Ober, C. P. Emerson Jr., *Mol. Cell. Biol.* **9**, 2957 (1989).
47. W. Dietmaier, S. Fabry, *Curr. Genet.* **26**, 497 (1994).
48. S. Aparicio *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **92**, 1684 (1995).
49. C. Hardison, *Trends Genet.* **16**, 369 (2000).
50. D. Thomasova *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 8179 (2002).
51. I. Bancroft, *Trends Genet.* **17**, 89 (2001).
52. S. Wong, G. Butler, K. H. Wolfe, *Proc. Natl. Acad. Sci. U.S.A.* **99**, 9272 (2002).
53. K. H. Wolfe, D. C. Shields, *Nature* **387**, 708 (1997).
54. H. J. Muller, in *The New Systematics*, J. H. Huxley, Ed. (Clarendon, Oxford, 1940), pp. 185–268.
55. V. N. Bolshakov *et al.*, *Genome Res.* **12**, 57 (2002).
56. R. L. Kelley *et al.*, *Cell* **98**, 513 (1999).
57. V. H. Meller, *Trends Cell Biol.* **10**, 54 (2000).
58. A. Pannuti, J. C. Lucchesi, *Curr. Opin. Genet. Dev.* **10**, 644 (2000).
59. C. Schutt, R. Nothiger, *Development* **127**, 667 (2000).
60. R. Paro, M. L. Goldberg, W. J. Gehring, *EMBO J.* **2**, 853 (1983).
61. P. Hatzopoulos, M. Monastirioti, G. Yannopoulos, C. Louis, *EMBO J.* **6**, 3091 (1987).
62. T. F. Smith, M. S. Waterman, *J. Mol. Biol.* **147**, 195 (1981).
63. J. E. Blair, K. Ikeo, T. Gojobori, S. B. Hedges, *Biomed. Central Evol. Biol.* **2**, 7 (2002).
64. A. M. Aguinaldo *et al.*, *Nature* **387**, 489 (1997).
65. J. O. Korbel, B. Snel, M. A. Huynen, P. Bork, *Trends Genet.* **18**, 158 (2002).
66. R. R. Sokal, F. J. Rohlf, *Biometry: The Principles and Practice of Statistics in Biological Research* (Freeman, New York, 1995).
67. A. N. Clements, *The Biology of Mosquitoes* (Chapman & Hall, London, 1992), vol. 1.
68. G. Wegener, *Experientia* **52**, 404 (1996).

Supporting Online Material

www.sciencemag.org/cgi/content/full/298/5591/149/DC1
Materials and Methods
Figs. S1 to S10
Tables S1 to S5
References

6 August 2002; accepted 6 September 2002

Immunity-Related Genes and Gene Families in *Anopheles gambiae*

George K. Christophides,[1]* Evgeny Zdobnov,[1]* Carolina Barillas-Mury,[2] Ewan Birney,[3] Stephanie Blandin,[1] Claudia Blass,[1] Paul T. Brey,[4] Frank H. Collins,[5] Alberto Danielli,[1] George Dimopoulos,[6] Charles Hetru,[7] Ngo T. Hoa,[8] Jules A. Hoffmann,[7] Stefan M. Kanzok,[8] Ivica Letunic,[1] Elena A. Levashina,[1] Thanasis G. Loukeris,[9] Gareth Lycett,[1] Stephan Meister,[1] Kristin Michel,[1] Luis F. Moita,[1] Hans-Michael Müller,[1] Mike A. Osta,[1] Susan M. Paskewitz,[10] Jean-Marc Reichhart,[7] Andrey Rzhetsky,[11] Laurent Troxler,[7] Kenneth D. Vernick,[12] Dina Vlachou,[1] Jennifer Volz,[1] Christian von Mering,[1] Jiannong Xu,[12] Liangbiao Zheng,[8] Peer Bork,[1] Fotis C. Kafatos[1]†

We have identified 242 *Anopheles gambiae* genes from 18 gene families implicated in innate immunity and have detected marked diversification relative to *Drosophila melanogaster*. Immune-related gene families involved in recognition, signal modulation, and effector systems show a marked deficit of orthologs and excessive gene expansions, possibly reflecting selection pressures from different pathogens encountered in these insects' very different life-styles. In contrast, the multifunctional Toll signal transduction pathway is substantially conserved, presumably because of counterselection for developmental stability. Representative expression profiles confirm that sequence diversification is accompanied by specific responses to different immune challenges. Alternative RNA splicing may also contribute to expansion of the immune repertoire.

Malaria transmission requires survival and development of the *Plasmodium* parasite in two invaded organisms: the human host and the mosquito vector. Interactions between the immune system of either organism with the parasite can hinder or even abort its development. The mosquito is known to mount robust immune reactions (*1*), accounting in part for the major parasite losses that occur within the vector. For example, melanotic encapsulation in a refractory strain of *A. gambiae*, the major vector of

human malaria, completely blocks parasite transmission (*2*).

The goal of this article is to describe potential molecular components and thus facilitate future in-depth analysis of the mosquito immune system's impact on the malaria parasite. This goal is best served by a comparative genomic analysis of the *Anopheles* and *Drosophila* immune systems. *Drosophila* is the best model system for the study of invertebrate immunity (*3*); it is a dipteran insect like the mosquito, and it also has a fully sequenced and extensively annotated genome, which has been compared with the *Anopheles* genome (*4*).

Insect immune reactions do not belong to adaptive immunity (which occurs only in chordates) but to the ancient defense system of innate immunity, which is relied upon by the vast majority of metazoans for dealing with invasive organisms, including pathogens and parasites. This system uses a wide range of gene families, some of which also have other physiological or developmental functions. It consists of both cellular and humoral responses, occurring first at the barrier epithelia (essentially the epidermis, gut, and tracheal respiratory organs of insects). Responses then become systemic, using the hemolymph-filled hemocoel, the open circulatory system of insects. Epithelial immunity is less well studied at present and occurs by direct interaction between epithelial cells and microorganisms. For malaria transmission, the key interaction is between the ookinete parasite stage and the midgut epithelial cells that it invades (*5*). In the systemic phase, key actors are the fat body (the insect's functional analog of liver and the main source of circulating immune-related components) and the hemocytes. The latter cells also engage in the cellular defenses of phagocytosis or encapsulation of larger invaders.

A Comparison of Immunity Gene Content in *Anopheles* and *Drosophila*

In this study, we have analyzed 18 mosquito gene families and a number of individual genes, for which comparative evidence from *Drosophila* and other organisms strongly suggested involvement in immune responses. We have used comparative bioinformatic analysis and manual annotation to characterize 242 *Anopheles* genes and relate them to 185 homologs from *Drosophila* (table S1). To facilitate future work, we named the mosquito genes systematically, using nomenclature rules that we propose for *Anopheles*, which are based largely on the HUGO nomenclature for the human genome.

The characterization and comparison of genes and families is summarized in Table 1. A basic conceptual framework of this analysis is that 1:1 orthologs correspond to well-conserved functions; orthologous groups (OG) represent functions that have begun to diversify; specific expansions (SE) represent major diversifications toward species-specific functions; and other genes (OT) represent genes that may have become highly specialized, or lost from the other species. A comparison of these global genome data against the immune genes is shown in Fig. 1A. In both species (and to a greater extent in *Anopheles*) we note that, relative to the genome as a whole (*4*), the immunity system has a deficit of 1:1 orthologs, contrasting

Table 1. Summary of potential immune components. Columns show gene numbers in orthologous pairs (1:1), total genes (TO), orthologous groups (OG), specific gene expansions (SE), and other homologs (OT). SCRB12 is not included in the analysis; CTL groups are not defined in *Drosophila*.

Family		*A. gambiae*				1:1	*D. melanogaster*			
		OT	SE	OG	TO		TO	OG	SE	OT
					Recognition					
PGRP	S	1	2	–	3	–	7	–	5	2
	L	–	–	1	4	3	6	2	–	1
TEP		2	10	2	15	1	6	1	4	–
GNBP	A	1	–	–	2	1	3	–	2	–
	B	–	4	–	4	–	–	–	–	–
SCR	A	1	–	–	5	4	5	–	–	1
	B	5	3	–	16	8	12	–	–	4
	C	–	–	–	1	1	4	–	2	1
CTL	MA	–	5	–	6	1	4	–	1	2
	GA	1	–	–	4	3	5	–	2	–
	SE	–	–	–	2	2	2	–	–	–
	O	5	–	–	10	5	24	–	16	3
GALE		–	5	–	8	3	5	–	–	2
FBN		3	52	–	57	2	13	–	11	–
					Modulation					
CLIP	A	2	6	–	10	2	11	–	7	2
	B	4	9	1	17	3	10	2	2	3
	C	3	2	2	7	–	5	1	2	2
	D	–	–	4	7	3	9	4	–	2
SRPN		1	–	8	10	1	17	6	5	5
IAP		1	2	–	6	3	4	–	–	1
					Signal transduction					
TOLL		–	2	4	10	4	9	2	2	1
MyD88		–	–	–	1	1	1	–	–	–
Tube		–	–	–	1	1	1	–	–	–
Pelle		–	–	–	1	1	1	–	–	–
Cactus		–	–	–	1	1	1	–	–	–
REL		–	–	–	2	2	3	–	–	1
Imd		–	–	–	1	1	1	–	–	–
STAT		–	–	2	2	–	1	1	–	–
					Effector molecules					
PPO		–	8	–	9	1	3	–	–	2
DEF		3	–	–	4	1	1	–	–	–
CEC		–	4	–	4	–	4	–	4	–
CASP	L	–	–	–	2	2	2	–	–	–
	S	2	–	8	10	–	5	3	–	2
SUM		35	114	32	242	61	185	22	65	37

[1]European Molecular Biology Laboratory, Meyerhofstrasse 1, D-69117 Heidelberg, Germany. [2]Department of Microbiology, Immunology and Pathology (MIP), Colorado State University, Fort Collins, CO 80523–1682, USA. [3]European Bioinformatics Institute–European Molecular Biology Laboratory, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SD, UK. [4]Unité de Biochimie et Biologie Moléculaire des Insectes, Institut Pasteur, 25 rue du Dr. Roux, 75724 Paris Cedex 15, France. [5]Center for Tropical Disease Research and Training, University of Notre Dame, Post Office Box 369, Notre Dame, IN 46556–0369, USA. [6]Department of Biological Sciences, Centre for Molecular Microbiology and Infection, Imperial College of Science, Technology and Medicine, London SW7 2AZ, UK. [7]Institut de Biologie Moléculaire et Cellulaire, Unité Propre de Recherche, 9022 du Centre National de la Recherche Scientifique, 15 rue Descartes, F67084 Strasbourg Cedex, France. [8]Yale University School of Medicine, Epidemiology and Public Health, 60 College Street, New Haven, CT 06520, USA. [9]Institute of Molecular Biology and Biotechnology–Foundation of Research and Technology Hellas, Vassilika Vouton, Post Office Box 1527, GR-711 10 Heraklion, Crete, Greece. [10]Department of Entomology, 237 Russell Lab, 1630 Linden Drive, University of Wisconsin, Madison, WI 53706, USA. [11]Columbia Genome Center and Department of Medical Informatics, Columbia University, Russ Berrie Medical Science Pavilion, 1150 St. Nicholas Avenue, New York, NY 10032, USA. [12]Department of Medical and Molecular Parasitology, New York University School of Medicine, 341 East 25th Street, Room 613, New York, NY 10010, USA.

*These authors contributed equally to this work.
†To whom correspondence should be addressed. E-mail: dg-office@embl-heidelberg.de

with an overabundance of specific gene expansions (Table 1). The same features are evident in the large immunity-related fibrinogen-domain (FBN) family, which we discuss in a companion comparative genomics paper as a prime example of gene family diversification (*4*). It would appear that in many immune families, orthologs are under pressure to diversify, or are lost, whereas certain immune genes reduplicate and then diversify. Our working hypothesis is that these prominent features reflect strong selective pressures to adjust and expand the innate immune repertoire in response to new challenges related to new ecological and physiological conditions; in the case of *Anopheles* the challenges include blood-borne infectious agents such as *Plasmodium*. When the immune genes are divided into four major categories (Fig. 1B) corresponding to the four major steps of the immune response, the ortholog deficit is greatest in the recognition, modulation, and effector categories; in contrast, the signal transduction category shows abundant 1:1 pairs and groups of orthologs, but minimal specific gene expansion.

Recognition of Infectious Nonself

In the terminology proposed by C. Janeway (*6*), innate immune responses begin when specialized, soluble or cell-bound pattern-recognition receptors (PRRs) recognize (and bind to) pathogen-associated molecular patterns (PAMPs) that are common in microorganisms but rare or absent in the responding species. PRRs can serve as opsonins facilitating phagocytosis; as receptors for signal transduction pathways that lead to synthesis of anti-pathogen effectors; and as initiators of clotting, melanization, or other protein modification cascades that are implicated in different steps of immunity. We have analyzed potential PRRs belonging to six gene families, two of which we will discuss in detail here and four primarily in the supplementary material.

Peptidoglycan Recognition Proteins (*PGRPs*). This family, distinguished by the PGRP domain (IPR002502), plays central and diverse roles in activating insect immune reactions. These include the melanization cascade, phagocytosis, and signal transduction pathways for production of anti–Gram-positive ($Gram^+$) and anti–Gram-negative ($Gram^-$) effectors (see supplementary material). We have identified seven distinct genes of this family in the *Anopheles* genome, of which three belong to the short (S) subfamily that encodes secreted proteins (*PGRPS1*, *S2*, and *S3*), while four belong to the long (L) subfamily (*PGRPLA*, *LB*, *LC*, and *LD*) encoding transmembrane or intracellular products. By comparison, *Drosophila* has 13 *PGRP* genes, six in the L subfamily (including the orthologs of the *Anopheles* L genes) and seven in the S subfamily (*7*).

Of special interest is the *Anopheles* chromosomal locus 21F (2L) encompassing two adjacent *PGRPLA* and *PGRPLC* genes within ~21 kb (Fig. 2A). The corresponding ~14-kb-long *Drosophila* locus at 67A8 (2R) includes an additional gene, *PGRP-LF*, an apparent product of species-specific tandem duplication. Except for *Drosophila PGRP-LF*, these genes have two or more PGRP domains, each domain encoded by two exons separated by introns at conserved positions (Fig. 2B). This gene architecture is compatible with alternative splicing, leading to proteins with alternative PGRP domains. Using a polymerase chain reaction–based approach on an adult cDNA library, we detected three main RNA isoforms (1, 2, and 3) from the *Anopheles PGRPLC* gene (see below); they carry alternative PGRP domains linked to a common backbone, which encodes a putative signal peptide and transmembrane domain. In *Drosophila*, isoforms of *PGRP-LC* are involved in the Imd signaling pathway and phagocytosis (*8–10*). The domains of this gene are more similar within a species than across species, indicating either that in the common ancestor this gene had one domain, which subsequently triplicated independently, or that a multidomain ancestral gene has followed concerted evolution after speciation (Fig. 2C). Similarly, the *PGRPLA* gene is represented in both species; however, the mosquito gene contains duplicated PGRP domains that are differentially spliced, leading to two distinct detected isoforms, *PGRPLA1* and *2* (Fig. 2A).

Microarray analysis (Fig. 2D) confirmed that some isoforms, which differ between species, are differentially regulated and functionally equivalent to gene expansions in other immunity gene families. After immune and oxidative challenges, the *Anopheles* isoform *PGRPLC2* is up-regulated by all four treatments tested, *PGRPLC1* by none, and *PGRPLC3* only by bacteria. Similarly, both *PGRPLA* isoforms respond to *Escherichia coli*, but additionally *PGRPLA1* responds to peptidoglycan (PGN) whereas *PGRPLA2* responds to *Staphylococcus aureus*. Taken together, these results suggest involvement of immunity signals in splice selection on transcripts of the *PGRPLA/C* gene cluster. Finally, the expression analysis revealed that *PGRPS1* is the only short *PGRP* to be induced by bacteria. *PGRPS2* is not up-regulated by these treatments, and *PGRPS3* is actually down-regulated by *S. aureus* and PGN.

Thioester-containing proteins (*TEPs*). This family is represented in many metazoa, from *Caenorhabditis elegans* to humans. It encodes proteins that play an important role in immune responses as part of the complement system and as the universal protease inhibitors, α2-macroglobulins. Recently, complement-like opsonin function for $Gram^-$ bacteria has been demonstrated (*11*) for the first member of this family studied in the mosquito, aTEP-I (now renamed TEP1). Another member of the family, TEP4, was shown to be up-regulated in *Plasmodium*-infected mosquitoes (*12*). A hallmark of the family is the conserved thioester (TE) motif.

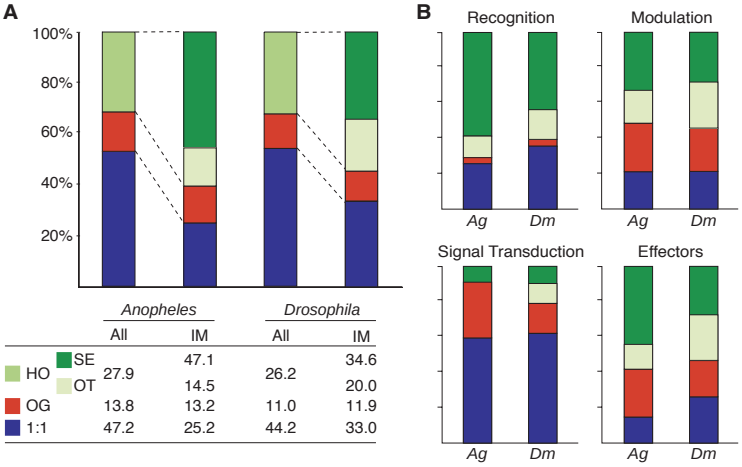


Fig. 1. Comparative analysis of immunity proteins in *Anopheles* and *Drosophila*, and comparison with the respective total proteomes (*4*). Proteins are divided into categories with their sizes shown as percentages. Category 1:1, orthologous pairs; OG, orthologous groups; HO, homologous proteins. The HO category is subdivided for the immunity studies as species-specific expansion (SE) and other homologs (OT). (**A**) Comparison of protein categories between whole genomes and immunity proteins. (**B**) Comparisons of protein categories in gene sets corresponding to the steps of recognition, modulation, signal transduction, and effectors.

After proteolytic activation, TEPs use TE for binding covalently to a nearby target, which is then cleared by phagocytic cells or destroyed by the membrane attack complex (MAC).

The *Drosophila* genome contains six *TEP* genes (*dTep*) (*13*). In strong contrast, after excluding putative haplotypes (designated *TEP16-19*), we have identified 15 *TEP* genes in the *Anopheles* genome (Fig. 3A and Table 1). Only a single 1:1 ortholog and one OG are shared. The majority of *TEP*s (4 in *Drosophila* and 10 in *Anopheles*) represent species-specific expansions, possibly permitting finely tuned responses to multiple pathogenic environments distinct in the two species. In addition, two dTeps and nine *Anopheles* TEPs lack the TE motif; as in vertebrate C5, the TE motif may not always be essential for the functions of insect TEPs.

A notable feature of the *Anopheles TEP* genes is arrangement in multiple chromosomal clusters (Fig. 3B). Genes that are either extensively diverged or resemble *Drosophila* most closely are all located at 29A-30E (3R). The two most similar genes (*TEP2*, *15*) are very close together, whereas the others (*TEP12*, *13*, *14*) are farther apart. Members of the major *Anopheles*-specific expansion are all located at 39C-40B (3L) in three clusters separated by 0.1 and 0.5 Mb. Close resemblance is evident between some genes in different clusters (e.g., *TEP5*, *7*, and *11*), as is a two-step specific expansion (*TEP8*, *9*, and *10*). The structural analysis of the *TEP* family is consistent with a model of sequential gene reduplications, potentially enabling diversified pathogen recognition.

Other recognition factors. We have analyzed four additional families associated with immune recognition in other species (see supplementary material). The Gram-Negative Binding Protein (GNBP) family includes members that are known to bind to Gram$^-$ bacteria, lipopolysaccharide (LPS), and β-1,3-glucan; to be involved in innate immune signaling in response to LPS (*14*); and to be up-regulated by immune challenge (*15*). In *Anopheles*, this family includes only one 1:1 ortholog and five other genes, four of which belong to a mosquito-specific subfamily derived from gene expansion (fig. S1). The multidomain scavenger receptor (SCR) family shows three disparate subfamilies (fig. S2) and is involved in immunity and development, recognizing multiple ligands and helping dispose of bacteria and apoptotic cells. Members of the large B subfamily are associated with uptake of multiple ligands, apoptotic corpses, and *Plasmodium*-infected erythrocytes; the fruit fly *croquemort* (*crq*) (*16*) is represented in the mosquito by a specific gene expansion. Two distinct carbohydrate-binding (lectin) families were also studied. C-type lectins (CTL), which bind to various sugars and LPS or are involved in cell adhesion, show prominent gene expansions (fig. S3). The Galectins (GALE) are associated with multiple functions, including apoptosis and innate immunity; in *Anopheles* several members (fig. S4) are induced by both bacteria and *Plasmodium* (*17*). Taken together (Fig. 1B), these six recognition families show great diversification by species-specific expansions and a deficit of 1:1 orthologs (less so in the case of SCR and CTLs).

Signal Modulation and Amplification

After recognition of infectious nonself, extracellular cascades of activating serine proteases and countervailing serine protease inhibitors (serpins) process the signal by either amplifying a strong "danger signal" or dampening false alarms (see also supplementary material). These modulatory families have a clear 1:1 ortholog deficit, but show increased numbers of OGs and only modest specific gene expansions.

The clip domain serine proteases (CLIPs) are characterized by the homonymous domain, a compact disulfide-bridged structure thought to regulate and localize the activity of the catalytic protease domain. One CLIP, Persephone (CG6367), acts to activate the Toll signaling cascade (*18*), whereas others

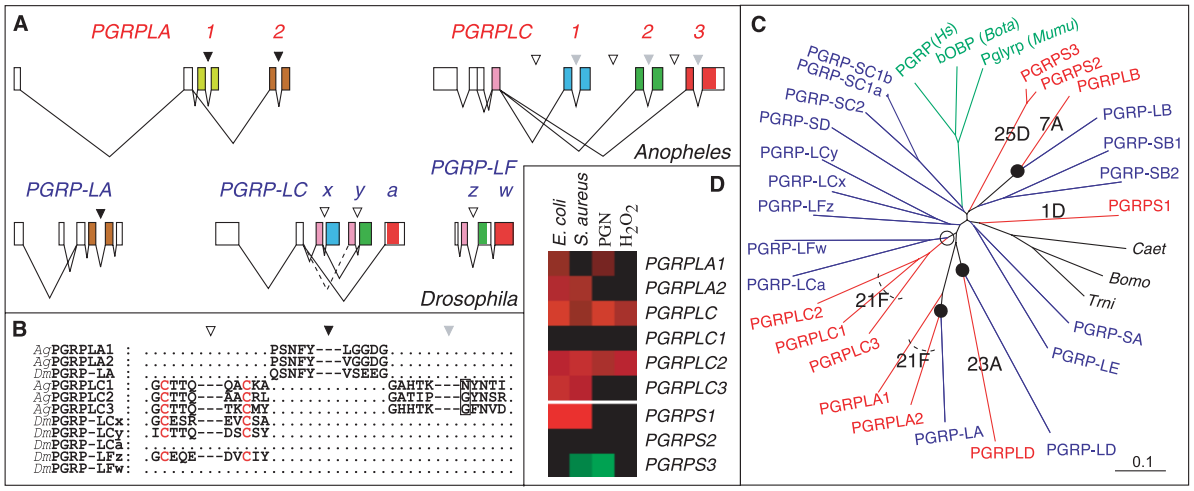


Fig. 2. Gene organization, transcriptional activity, and phylogenetic analysis of the *PGRP* gene family. (**A**) Exon/intron organization of the *Anopheles* 21F and *Drosophila* 67A8 PGRP loci. Exons coding for PGRP domains are colored. Arrowheads indicate introns as positioned in Fig. 2C. Numbers and letters designate PGRP domains included in alternative isoforms. (**B**) Intra- and interspecies conservation of introns in PGRP domains shown in 2A. Genes have maintained identical intron/exon boundaries, except that the *Drosophila* PGRP-LCa and -LFw domains lack introns possibly lost secondarily. *Anopheles* PGRP-LC has an additional exon in each of the PGRP domains. Amino acids encoded by codons spanning intron boundaries are boxed. A cysteine pairing conserved in almost all known PGRPs is highlighted in red; one is changed to a Tyr in the *Drosophila* PGRPSA of *semmelweis* mutants (*43*). (**C**) Phylogenetic analysis of the PGRP domains. In this and subsequent dendrograms, *Anopheles* genes/proteins are indicated as red branches, and *Drosophila* (blue), vertebrates (green), and invertebrates and common gene stems (black) are colored as shown; dots on nodes indicate orthologous pairs, and circles indicate orthologous groups. Numbers accompanying or grouping branches indicate chromosomal locations. (**D**) Expression profiles of PGRP isoforms in cultured cells challenged with *E. coli*, *S. aureus*, peptidoglycan (PGN), and H_2O_2. Color intensities indicate fold regulation relative to reference (naïve) cells (see Methods in supplemental material). Regulation values below 1.5-fold are masked.

are associated with immune effector cascades (e.g., the phenoloxidase cascade in Lepidoptera and hemolymph clotting in the horseshoe crab) or serve in development (e.g., Snake, Easter, and Stubble in *Drosophila*) (*19*). The *Anopheles* and *Drosophila* genomes encode 41 and 35 CLIPs, respectively, in four subfamilies (Fig. 4A). This apparent numerical conservation is deceptive, as only eight orthologous pairs and five OGs exist; numerical conservation appears to be the net effect of counterbalancing species-specific expansions. The developmental genes are conserved, unlike the single well-characterized *Drosophila* immune CLIP *Persephone*, which is not conserved in the mosquito.

Most serpins (SRPNs) are irreversible inhibitory substrates for proteases, often but not exclusively of the serine class. Noninhibitory serpins are less well characterized; some were shown to function in hormone transport or blood-pressure regulation. In mammals, serpins account for 10% of the plasma proteins and affect blood coagulation, fibrinolysis, phagocytosis, inflammation, microbial infection, and complement activation. The mosquito genome encodes 14 serpins, 10 of which are inhibitory. Again, gene expansions/losses result in species-specific diversification; only one orthologous pair and four OGs are evident (Fig. 4B). The *Drosophila* serpin encoded by the *nec* locus, which is a partner of the Persephone Clip-domain protease in the Toll-mediated antifungal response (*20*), also has no ortholog in the mosquito. The functions corresponding to Persephone and Nec must be served by independently evolved *Anopheles* CLIPs and SRPNs. The *Drosophila* serpin-27A (CG 11331), involved in control of melanization, forms an OG with three mosquito serpins, which constitute interesting potential modulators of prophenoloxidases (PPOs) (see below). In a separate study (*21*), we have determined that the mosquito SRPN10 (lacking a 1:1 ortholog, and initially named spi21F) is intracellular and has isoforms with distinct biochemical inhibitory specificities; thus, as in the *PGRPL* subfamily, alternative splicing augments SRPN diversification. Notably, one of these isoforms is greatly up-regulated in midgut cells during *Plasmodium* invasion.

Signal Transduction Pathways

Signal transduction pathways link recognition and amplification of the "danger" signal with transcriptional activation. In *Drosophila*, antimicrobial responses use two major signal transduction pathways, Toll and Imd (*3*), and at least in the mosquito a third pathway, STAT, is also involved (*22*). Here we will consider the well-characterized Toll pathway, which has both developmental and immune functions and engages many genes and families (Table 1 and supplementary material). Most steps in the pathway are served by well-conserved individual genes, presumably reflecting conservation of balanced functions. The signaling receptor family shows a modest level of diversification.

Anopheles has 11 *TOLL* genes (Fig. 4C), of which four (*TOLL 6*, *7*, *8*, and *9*) are unambiguous orthologs of *Drosophila* counterparts (*23*). However, orthologs of *Toll-2*, *-3*, and *-4* have not been detected in *Anopheles*, which shows instead a species-specific expansion (*TOLL 10* and *11*). Gene reduplication has also generated four mosquito genes—*TOLL 1A*, *1B*, *5A*, and *5B*—that together with the fruit fly *Toll-1* and *-5* genes form an interesting OG. The most parsimonious hypothesis is that single type 1 and 5 genes were ancestrally linked and that this pair reduplicated and translocated in the mosquito, forming the 1A/5A pair at chromosomal site 6C and the 1B/5B pair at 39C; in the fruit fly the ancestral pair separated to different locations. It remains to be determined whether the immune function of the *Drosophila* type 1 gene is ancestral and retained by both *Anopheles* 1A and 1B, and whether any of the type 5 genes have immune functions.

In *Drosophila*, Toll signal transduction is initiated by binding of a cleaved peptide ligand, Spaetzle, on the extracellular domain of Toll, the intracellular domain of which interacts with MyD88, Tube, and Pelle, probably forming a multimeric inactive protein kinase complex (*24*, *25*). Upon Spaetzle binding, Pelle phosphorylates (directly or indirectly) Toll, itself, and Cactus; Cactus phosphorylation causes release of the Rel transcription factors Dorsal and DIF, which translocate into the nucleus and activate numerous genes, including those encoding antifungal peptides (*26*). The intracellular pathway is intact in *A. gambiae*: We have identified single genes encoding orthologs of *MyD88*, *Tube*, *Pelle*, and *Cactus* (Table 1). Another Pelle-like domain is found in the COOH end of a predicted, unusually large, protein sequence whose NH_2-terminal part is homologous to Tube. The mosquito ortholog of *Dorsal*, *Gambif-1*, was identified previously (*27*), but surprisingly, no ortholog of *DIF* was found.

Effector Response Systems

After microbial recognition, signal modulation, and transduction, the transcriptional responses engage a large number of genes, including many with unknown function (*26*). However, three broad categories of effector systems are well recognized: antimicrobial peptides, the phenoloxidase-dependent melanization system, and the system of apoptosis-related genes. All three systems show a marked paucity of orthologs (Table 1).

Prophenoloxidases (*PPOs*): Melanization is an important immune response in insects and crustacea, and possibly in other arthropod classes. PPO proenzymes circulate through the hemolymph and, upon activation by clip domain proteases, catalyze key steps in the synthesis of melanin, thereby promoting cuticle sclerotization, wound healing, and melanotic encapsulation of pathogens (*28*, *29*); recently PPOs were also associated with hemolymph clotting (*30*). The genes show no signal peptide signature, suggesting that PPOs are released not by secretion but by rupture of hemocytes.

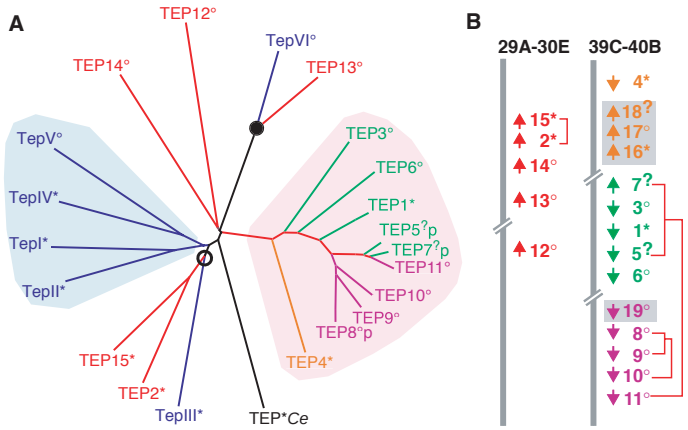


Fig. 3. Protein sequence comparison and chromosomal distribution of the *TEP* gene family. (**A**) Phylogenetic tree of complete sequence alignment. In this and subsequent figures, shading indicates gene expansions in *A. gambiae* (pink) and in *D. melanogaster* (blue). (**B**) *Anopheles* predicted *TEPs* (arrows) are physically located in four clusters and one isolated locus [identified by consistent colors in (A) and (B)]. Closest *Anopheles* paralogs are connected with brackets. Putative haplotypes are shaded in gray and are not discussed further. Superscript symbols after names indicate that the thioester motif is (*) present, (°) absent, or (?) unknown; p, partial sequence. Color scheme: blue, *D. melanogaster*; red, orange, green, and purple, *A. gambiae*; black, other invertebrates.

The *A. gambiae* genome encodes nine PPOs, threefold as many as the *Drosophila* genome. Six of the genes have been described (*31*–*33*) and numbered in order of discovery (*33*). Melanotic encapsulation of *Plasmodium* in refractory mosquitoes (*2*) suggests that antiparasitic defense may be one function of extra mosquito genes. Interestingly, the newly discovered *PPO9* is strongly induced in blood-fed *A. gambiae* (*34*) and may facilitate melanotic encapsulation. Other potential functions of extra genes may be to rapidly repair injuries endured by the swollen blood-fed mosquitoes, or by larvae living in swiftly running rainwater. The mosquito eggshell is also tanned after fertilization, and the adult mosquito cuticle and scales are more broadly melanized than those of *Drosophila*. Interestingly, most *Anopheles* PPO genes are part of a major expansion that may have occurred early in the mosquito lineage (Fig. 4D). Consistent with this hypothesis, all PPOs from other mosquitoes cluster with the *A. gambiae* genes. The sole exception is the *Anopheles PPO1* gene, which appears to be primitive; it clusters together with two members from *Drosophila* and one each from the fleshfly *Sarcophaga* and the beetle *Tenebrio*.

Other effector systems. Antimicrobial peptides (AMPs) are produced in the fat body, hemocytes, and epithelial tissues. Several hundred are now described, and their rapid evolution has been noted. The most important families are the widely distributed anti-Gram$^+$ insect defensins (DEF) and the predominantly anti-Gram$^-$ cecropins (CEC, in Diptera and Lepidoptera). Four *DEF* and four *CEC* genes exist in *Anopheles*, more numerous and more diverged than in *Drosophila*. Several other AMP families are specific to *Drosophila* but absent in *Anopheles*. Conversely, Gambicin (GAM1) (*35*) is mosquito specific. It appears that mosquitoes use few AMP families but may expand the spectrum of antibiotic activities, substantially diversifying both *DEF* and *CEC* sequences (see supplementary material).

The apoptotic machinery acts at three conceptually distinct levels: pro-apoptotic and anti-apoptotic regulators modulate the activity of initiator caspases (often by way of associated adaptor molecules), and these in turn activate effector caspases, the direct cell executioners. This system plays well-recognized roles in discarding unwanted cells during development but is also implicated in immunity. Intriguing evidence suggests that apoptosis and cell elimination is an important response of the *A. gambiae* midgut epithelium to *Plasmodium* invasion (*5, 21*). The initiator caspase DREDD, an important protease controlling morphogenic apoptosis, is also central to the *Drosophila* IMD (Immune Deficiency) pathway that resists Gram$^-$ bacterial infections (*36*). The number of caspase genes in *Anopheles* is somewhat higher than in humans (12 as compared to 11) and considerably higher than in *Drosophila* and *C. elegans* (7 and 3, respectively) (*37*). This overabundance is due to effector caspases, which have undergone a specific expansion, unlike initiator caspases (fig. S7). The negative regulators of caspases, IAPs (Inhibitor of Apoptosis Proteins), show both conservation (at least three orthologs with *Drosophila*) and specific expansion (two new IAPs) (fig. S8). The search for mosquito pro-apoptotic genes has been hampered by the extensive sequence diversification of the main players (*37*).

Diversified Gene Expression and Beyond

Immune gene sequence diversification suggests diversified functions. As a first step toward functional analysis, we evaluated the developmental regulation in whole mosquitoes, and in greater depth responsiveness to sterile injury or infections with bacteria (Gram$^+$ or Gram$^-$) and *Plasmodium*, for 24 representative mosquito genes belonging to 12 immunity families (Fig. 5), including one (*FBN*) described in a companion article (*4*). In immune-challenged mosquitoes, the expression profiles were specific to the gene and the particular challenge. By comparing sterile injury and bacterial infections of the mosquitoes, we determined that *E. coli* but not *S. aureus* specifically induces *GNBPB1*. Both types of bacteria induced *SRPN10* and *4* sequentially, whereas *SRPN9* was only induced late in *S. aureus* infection. *E. coli* induced *GAM1* late and robustly, whereas *S. aureus* induced *GAM1* only early and transiently. *CLIPB14* and *15* were up-regulated by both bacteria in a sustained manner, but *CLIPA6* was only induced transiently and modestly. Amongst six members of the *FBN*

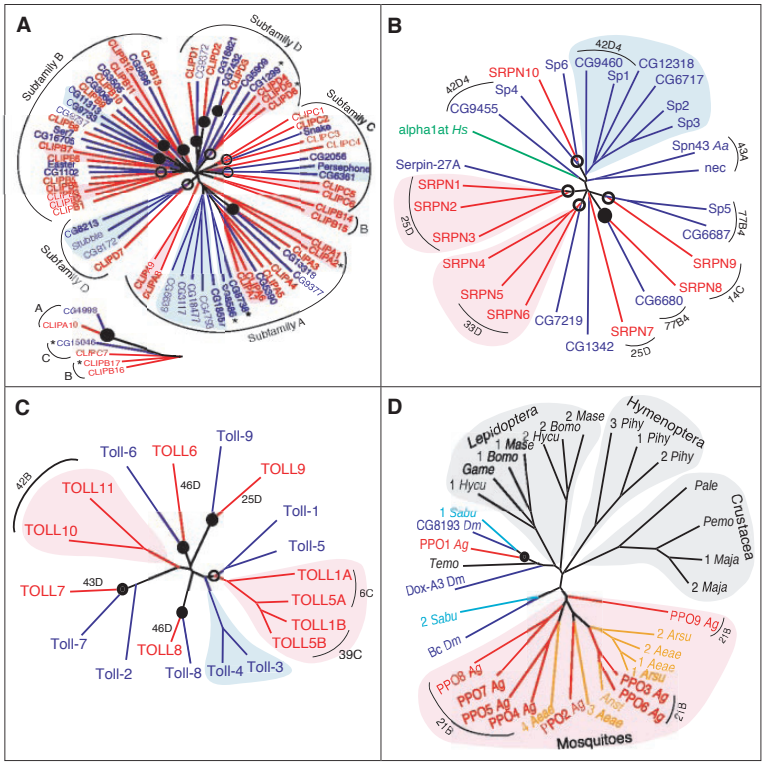


Fig. 4. Phylogenetic trees of CLIPs, SRPNs, TOLLs, and PPOs. (**A**) CLIP family. Tree was based on alignments that include the clip and serine protease domains. Proteins cluster into subfamilies A, B, C, and D and five hybrid sequences (shown at bottom of figure). Note: Long insertions within the clip domain were omitted for the orthologs CLIPA10 and CG4998 before sequence alignment; CG4914 was not aligned because of exceptionally arranged C residues in the clip domain. *, proteins containing two clip domains. (**B**) Inhibitory SRPNs. (**C**) TOLL family. (**D**) PPO family. Gray shading indicates groups of proteins from Hymenoptera, Lepidoptera, and Crustacea. Light blue branches indicate proteins from the dipteran, *Sarcophaga bullata*. For taxonomic abbreviations, see supplementary material.

family, only *FBN9* showed a sustained induction by bacteria. Finally, *TEP3* and *4* were strongly induced by both bacteria, possibly with different kinetics. Induction by sterile injury was rare, transient, and usually late (possibly in response to inadvertent infection of the wound, or cell damage; see late induction of *CASPL2* and *IAPB1* in Fig. 5); exceptionally, *PGRPLB* and *SCRBQ2* were induced similarly by sterile injury and bacteria.

During the life cycle of the parasite in the mosquito, six different genes (*FBN9*, *23*, and *CLIPB14*; *SRPN9*, *10*, and *4*) were activated primarily at 28 hours after infection of the mosquito, i.e., specifically when the midgut epithelium is invaded by ookinetes. In contrast, the parasite caused sustained induction of *PGRPLB*, *TEP4*, and *CLIP15* throughout its life cycle in the vector, suggesting an ongoing systemic rather than epithelial immune response. A delayed induction of *CEC1* and *GNBPA1* (which was not seen with bacteria) apparently represents reactions to the oocyst and sporozoite stages of the parasite.

The developmental profiles (for individual versus pooled stages) indicated stage-specific gene expression or up-regulation in the absence of a specific challenge: for example, *SRPN9* in pupae and *SRPN10* in early larvae, *CLIP15* primarily in early larvae, *CLIP6* in late larvae, and *CLIP14* and *PGRPLB* in adults. Developmental regulation of *FBN* family members was prominent, with three members expressed preferentially in embryos and early larvae, one in late larvae and pupae, and two in the adults. *FBN9*, which was strongly inducible both by bacteria and during *Plasmodium* penetration of the midgut, proved to be adult specific.

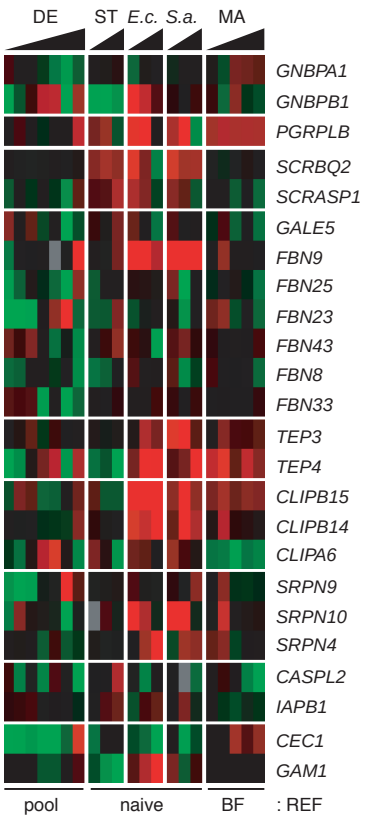


Fig. 5. Expression profiles of immunity gene family members. From left to right: Developmental (DE) expression profiles were examined at embryonic, 1st, 2nd, 3rd, 4th instar larval, pupal, and adult stages. Adult female mosquitoes were pricked with a sterile needle (ST) or infected with *E. coli* (*E. c.*) or *S. aureus* (*S. a.*), and assayed at 6, 12, and 24 hours after treatment. Mosquitoes were infected with malaria (MA), and expression profiles were examined at 24 hours, 28 hours, 6 days, 11 days, and 16 days after infection. Green- and red-colored data points of increasing intensity indicate up- and down-regulation relative to the reference (REF) samples, respectively. Regulation values below 1.5-fold are masked in black; gray represents missing points.

Concluding Remarks

The newly available genome sequence has created unprecedented opportunities for mosquito research. Genomic expression profiling will be facilitated by a consortium that is developing standarized whole-genome microarrays. Tools for reverse genetic analysis will be critically important. Hemocyte-like cell lines (*33*), coupled with in vitro transient and stable, transposable element–mediated transfection/transformation, are already in place (*38*). Germline transformation has been accomplished for both *A. stephensi* (*39*) and *A. gambiae* (*40*), and more sophisticated methodologies for gene disruption and conditional gain- and loss-of-function analysis are becoming available (*41*). Most recently, a convenient RNA interference–mediated approach for functional gene disruption by direct injection of double-stranded RNA has been developed (*42*). Phenotypic as well as genome-scale analysis of immune-related genes is now feasible for the malaria mosquito.

References and Notes

1. G. Dimopoulos, H. M. Muller, E. A. Levashina, F. C. Kafatos, *Curr. Opin. Immunol.* **13**, 79 (2001).
2. F. H. Collins *et al.*, *Science* **234**, 607 (1986).
3. J. A. Hoffmann, J.-M. Reichhart, *Nature Immunol.* **3**, 121 (2002).
4. E. M. Zdobnov, *Science* **298**, 149 (2002).
5. Y. S. Han, J. Thompson, F. C. Kafatos, C. Barillas-Mury, *EMBO J.* **19**, 6030 (2000).
6. C. A. Janeway Jr., R. Medzhitov, *Annu. Rev. Immunol.* **20**, 197 (2002).
7. T. Werner *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **97**, 13772 (2000).
8. K. M. Choe, T. Werner, S. Stoven, D. Hultmark, K. V. Anderson, *Science* **296**, 359 (2002).
9. M. Gottar *et al.*, *Nature* **416**, 640 (2002).
10. M. Ramet, P. Manfruelli, A. Pearson, B. Mathey-Prevot, R. A. Ezekowitz, *Nature* **416**, 644 (2002).
11. E. A. Levashina *et al.*, *Cell* **104**, 709 (2001).
12. F. Oduol, J. Xu, O. Niare, R. Natarajan, K. D. Vernick, *Proc. Natl. Acad. Sci. U.S.A.* **97**, 11397 (2000).
13. M. Lagueux, E. Perrodou, E. A. Levashina, M. Capovilla, J. A. Hoffmann, *Proc. Natl. Acad. Sci. U.S.A.* **97**, 11427 (2000).
14. Y. S. Kim *et al.*, *J. Biol. Chem.* **275**, 32721 (2000).
15. G. Dimopoulos, A. Richman, H. M. Muller, F. C. Kafatos, *Proc. Natl. Acad. Sci. U.S.A.* **94**, 11508 (1997).
16. N. C. Franc, J. L. Dimarcq, M. Lagueux, J. Hoffmann, R. A. Ezekowitz, *Immunity* **4**, 431 (1996).
17. G. Dimopoulos, D. Seeley, A. Wolf, F. C. Kafatos, *EMBO J.* **17**, 6115 (1998).
18. P. Ligoxygakis, N. Pelte, J. A. Hoffmann, J. M. Reichhart, *Science* **297**, 114 (2002).
19. H. Jiang, M. R. Kanost, *Insect Biochem. Mol. Biol.* **30**, 95 (2000).
20. E. A. Levashina *et al.*, *Science* **285**, 1917 (1999).
21. A. Danielli, unpublished data.
22. C. Barillas-Mury, Y. S. Han, D. Seeley, F. C. Kafatos, *EMBO J.* **18**, 959 (1999).
23. C. Luna, X. Wang, Y. Huang, J. Zhang, L. Zheng, *Insect Biochem. Mol. Biol.* **32**, 1171 (2002).
24. B. Shen, J. L. Manley, *Development* **125**, 4719 (1998).
25. S. Tauszig-Delamasure, H. Bilak, M. Capovilla, J. A. Hoffmann, J. L. Imler, *Nature Immunol.* **3**, 91 (2002).
26. E. De Gregorio, P. T. Spellman, P. Tzou, G. M. Rubin, B. Lemaitre, *EMBO J.* **21**, 2568 (2002).
27. C. Barillas-Mury *et al.*, *EMBO J.* **15**, 4691 (1996).
28. S. C. Lai, C. C. Chen, R. F. Hou, *J. Med. Entomol.* **39**, 266 (2002).
29. B. T. Beerntsen, A. A. James, B. M. Christensen, *Microbiol. Mol. Biol. Rev.* **64**, 115 (2000).
30. T. Nagai, S. Kawabata, *J. Biol. Chem.* **275**, 29264 (2000).
31. H. Jiang, Y. Wang, S. E. Korochkina, H. Benes, M. R. Kanost, *Insect Biochem. Mol. Biol.* **27**, 693 (1997).
32. W. J. Lee *et al.*, *Insect Mol. Biol.* **7**, 41 (1998).
33. H. M. Muller, G. Dimopoulos, C. Blass, F. C. Kafatos, *J. Biol. Chem.* **274**, 11727 (1999).
34. H. M. Muller, unpublished data.
35. J. Vizioli *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **98**, 12630 (2001).
36. F. Leulier, A. Rodriguez, R. S. Khush, J. M. Abrams, B. Lemaitre, *EMBO Rep.* **1**, 353 (2000).
37. S. Y. Vernooy *et al.*, *J. Cell Biol.* **150**, F69 (2000).
38. F. Catteruccia *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **97**, 6236 (2000).
39. F. Catteruccia *et al.*, *Nature* **405**, 959 (2000).
40. G. L. Grossman *et al.*, *Insect Mol. Biol.* **10**, 597 (2001).
41. G. Lycett, F. C. Kafatos, T. G. Loukeris, unpublished data.
42. S. Blandin, L. F. Moita, F. C. Kafatos, E. A. Levashina, *EMBO Rep.* **3**, 852 (2002).
43. T. Michel, J. M. Reichhart, J. A. Hoffmann, J. Royet, *Nature* **414**, 756 (2001).

44. We acknowledge the collaborative spirit of the *Anopheles* Genome Project and the importance of the genome sequence determination performed by the Celera Genomics, Genoscope, and TIGR teams. We are grateful for the support provided by the major funders: National Institutes of Health, the National Institute of Allergy and Infectious Disease (USA), National Science Foundation (USA), the French Ministry of Research, our institutions, and the European Union. G.K.C. and D.V. were supported by Marie-Curie fellowships. The long-standing support by Tropical Disease Research–TDR (WHO) and the John D. and Catherine T. MacArthur Foundation was also essential for the project. F.C.K. dedicates this article to the memory of Anne Gruner Schlumberger.

Supporting Online Material

www.sciencemag.org/cgi/content/full/298/5591/159/DC1
Methods
SOM Text
Figs. S1 to S10
Table S1
References and Notes

8 August 2002; accepted 3 September 2002

Review 3

Thioester-containing Proteins and Insect Immunity

Stéphanie BLANDIN, Elena A. LEVAHSHINA

Molecular Immunology 2004, 40: 903-908

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Molecular Immunology 40 (2004) 903–908
Molecular Immunology
www.elsevier.com/locate/molimm

Thioester-containing proteins and insect immunity

Stéphanie Blandin [a], Elena A. Levashina [a,b,*]

[a] *European Molecular Biology Laboratory, Meyerhofstrasse 1, D-69117 Heidelberg, Germany*
[b] *UPR 9022 du CNRS, Institut de Biologie Moléculaire et Cellulaire, 15 rue René Descartes, F-67084 Strasbourg Cedex, France*

Abstract

Here, we discuss the role of thioester-containing proteins in innate immune responses of insects. TEPs are represented by multi-member families both in the fruitfly, *Drosophila melanogaster*, and in the mosquito, *Anopheles gambiae*. Phylogenetic analysis of the family suggests that in these two dipteran species evolution of TEPs followed independent scenarios as a result of specific adaptation to distinct ecological environments. Research on these two relatively simple model systems, which lack adaptive immunity, may provide new insights into the evolutionary origins and functions of this important protein family.

Keywords: Thioester-containing proteins; Complement factors; Alpha2-macroglobulin; Innate immunity; *Anopheles gambiae*; *Drosophila melanogaster*

1. Introduction

Thioester-containing proteins (TEPs) appeared early in animal evolution: members of this family have been identified in such diverse organisms as nematodes, insects, molluscs, fish, birds and mammals (Nonaka, 2000). In vertebrates, these proteins play important roles in immune responses as components of the complement system, in the case of factors C3/C4/C5, or as universal protease inhibitors, in the case of α_2-macroglobulins (α2Ms). TEPs are characterized by homologous sequence features, including a unique intrachain β-cysteinyl-γ-glutamyl thioester bond, and a propensity for multiple conformationally sensitive binding interactions (Chu and Pizzo, 1994), which define their functional properties. The thioester bond mediates covalent attachment of the molecules to activating self and non-self surfaces (complement factors) and covalent or non-covalent cross-linking to attacking proteases (α2Ms). The thioester is highly reactive and is readily hydrolysed by water when exposed. To avoid precocious inactivation, it is protected by a shielded environment in the native protein (Salvesen et al., 1981; Sim and Sim, 1981). Activation of the molecule requires an exposure of the thioester bond, which is achieved through a proteolytic cleavage either by a specific "convertase" protease complex (complement factors) or by attacking proteases (α2Ms). The proteolytic activation of the complement factors generates two products, a small anaphylatoxin fragment without the thioester (C3a, C4a and C5a), and a larger fragment that binds covalently to the target surface via the hydrolysis of the thioester. Both cleavage products have signalling properties via specific interactions with receptors or effector molecules. Anaphylatoxins spread in the vicinity of complement activation and act as chemoattractants to recruit macrophages to the site of infection. Binding of the bigger fragment targets pathogens for phagocytosis or for destruction by lysis, mediated by a membrane attack complex. In the case of α2Ms, the proteolytic activation triggers a major rearrangement of the protein that has two important consequences. First it triggers the trapping of the attacking protease within the protein, therefore denying to the substrates access to the protease active site. And second, it exposes the C-terminal recognition domain that can then mediate the clearance of the protease-reacted α2M from the circulation by a receptor-mediated endocytosis.

Here, we summarise our current knowledge on the thioester-containing proteins in two dipteran insects. Recent completion of the genome sequencing of *Drosophila melanogaster* and *Anopheles gambiae* provides a unique opportunity to analyze phylogenetic relationships between complement factors, α2Ms and insect TEPs. The fruitfly *D. melanogaster* is a model of choice to study fundamental biological processes due to the exceptional experimental possibilities offered by this insect, including a myriad of molecular and genetic tools and the availability of the complete genome sequence. Recent advances in the development of tools for functional gene analysis and the completion

Abbreviations: TEP, thioester-containing protein; α2M, alpha2-macroglobulin; C3, C4 and C5, complement factors C3, C4 and C5, respectively

* Corresponding author. Tel.: +33-388-41-70-18; fax: +33-388-60-69-22.
E-mail address: e.levashina@ibmc.u-strasbg.fr (E.A. Levashina).

0161-5890/$ – see front matter © 2003 Elsevier Ltd. All rights reserved.
doi:10.1016/j.molimm.2003.10.010

of the genome sequencing validate *A. gambiae*, one of the major mosquito vectors of human malaria in Africa, as a powerful model to address mechanisms of vectorial capacity in medically important insect vectors. The interest of this model is highlighted by the fact that the protozoan parasite has to evade two related but distinct immune systems, those of the vertebrate host and of the mosquito. Insects are able to mount rapid and efficient responses, which are reminiscent of innate immune defences of vertebrates. Since insects lack an adaptive immune response, they represent a valuable model to dissect functions of thioester-containing proteins.

2. Thioester-containing proteins in *D. melanogaster*

Six TEP homologues have been identified in the *Drosophila* genome (Lagueux et al., 2000). Out of the six Teps, four genes contain a conserved thioester motif whereas *Tep5* and *Tep6* display a modified sequence (Crowley et al., 1993; Lagueux et al., 2000). One gene, *Tep5*, is present only in genomic sequences and may represent a pseudogene. Three features are characteristic of four *Drosophila Teps*: (i) a highly conserved region of 30 amino acid residues harbouring a canonical thioester motif GCGEQ; (ii) a distinctive cysteine signature encompassing 126 residues in the C-terminal part of the molecule; (iii) a highly variable central region of about 60 residues in length. Interestingly, in *Tep2* this variable region is encoded by five exons, which produce distinct transcripts as a result of alternative splicing (Lagueux et al., 2000). The alternative exons correspond to the bait region in α2Ms and the alternative splicing may serve to extend the repertoire of inhibited proteases. It may also represent a novel mechanism for recognition of noxious structural patterns in the absence of the large repertoire of

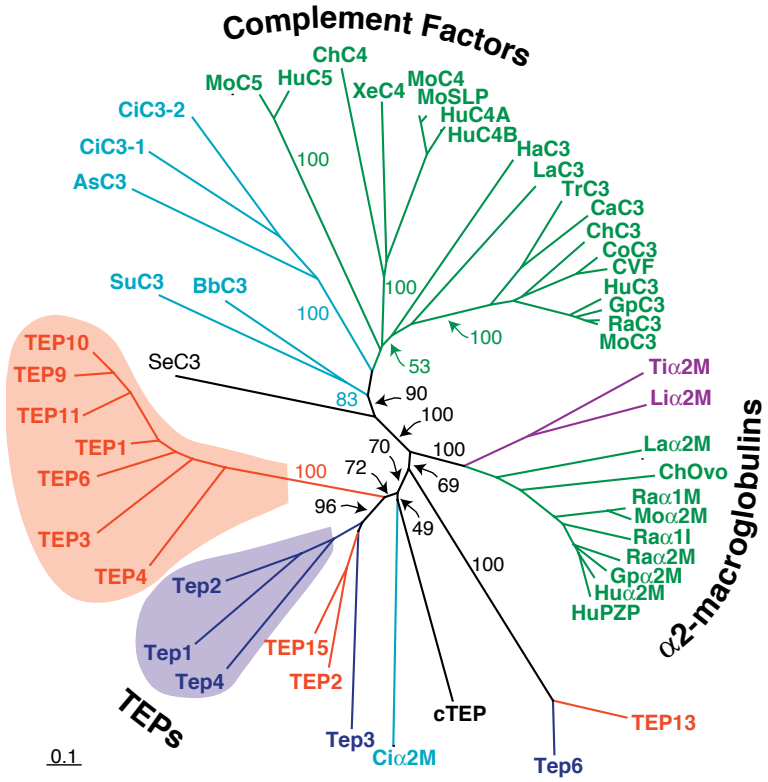


Fig. 1. Unrooted phylogenetic tree of thioester-containing proteins. The tree was constructed by the neighbour-joining method based on the sequence alignment generated with Clustal_X. Molecules of vertebrate species are in green and of primitive deuterostomes (*Strongylocentrotus purpuratus* (Su), *Branchiostoma belcheri* (Bb), *Halocynthia rorezi* (As), *Ciona intestinalis* (Ci)), in cyan. Proteins of *Drosophila melanogaster* (Tep) and *Anopheles gambiae* (TEP) are in blue and red respectively; other Arthropod proteins (*Limulus sp.* (Li) and *Ornithodoros moubata* (Ti)) in purple. Proteins of *Swiftia exserta* (SeC3) and *Caenorhabditis elegans* (cTEP) are in black. Bootstrap values of 1000 replicates (%) are indicated for some branches. Sequences, for which only partial information is available, or those that are too divergent from the majority of the sequences, are excluded from the analysis. The presence of partial (TEP5, 7 and 8) and highly divergent (TEP12, 14 and TEP5) sequences did not affect the general topology of the tree but modified the length of the branches. As most of the mosquito TEP sequences (except for TEP1-4) and *Drosophila* Tep5 are predictions from the genome sequences and have not yet been confirmed by cDNA cloning, this tree should be considered as preliminary. Species are as described in http://www.embl-heidelberg.de/ExternalInfo/kafatos/publications/TEP/.

receptors of the adaptive immune response in vertebrates. Until now alternative splicing had not been reported in thioester-containing proteins. Recently, a α2M from a soft tick, *Ornithodoros moubata*, was reported to be represented by four distinct isoforms as a result of alternative splicing of the similar bait-like region of this molecule (Saravanan et al., 2003). All *Drosophila Teps* are located on the left arm of the second chromosome, where *Tep2* and *3* form a discrete cluster in which genes are oriented head to head and are separated by 1.5 kb of putative regulatory sequences (reviewed in Levashina et al., 2002). This clustering may result from a recent duplication-inversion event. Thioester-containing proteins are often acute phase reactants (Sottrup-Jensen, 1989; Volanakis, 1995). In *Drosophila*, the expression of *Tep1*, *2* and *4* is strongly upregulated by bacterial challenge in larvae. During the adult stage, immune challenge markedly induces *Tep2* and *4*, whereas the expression of *Tep1* is minor (reviewed in Meister and Lagueux, 2003). To date, the expression pattern of *Tep1* in larvae is best characterized. This gene is mainly transcribed in the fat body after immune challenge, but it is also expressed in hemocytes in naive and challenged larvae. No expression of *Tep1* is detected in the *Drosophila l(2)mbn* and S2 Schneider cell lines (reviewed in Levashina et al., 2002). Surprisingly, according to phylogenetic analysis, all inducible Tep proteins 1, 2, and 4, cluster together in the *Drosophila*-specific expansion, whereas the constitutively expressed *Tep3* is somewhat divergent in sequence (Fig. 1). The induction of immune responses in *Drosophila* is controlled by two distinct signalling pathways: the Toll pathway is primarily responsible for expression of the antifungal peptide gene *Drosomycin*, whereas the Imd pathway controls expression of most antibacterial peptide genes (Anderson, 2000; Imler and Hoffmann, 2000). The inducible expression of *Tep1* is not controlled by either of these pathways (Lagueux et al., 2000). Indirect effect of the Toll pathway on the *Tep1* expression is observed in *Toll* gain-of-function mutants and might be associated with characteristic aggregation of blood cells into masses which tend to become melanized (Qiu et al., 1998). It has been suggested that the Toll-induced melanotic tumor formation is mediated by the JAK kinase, *Hopscotch* (Mathey-Prevot and Perrimon, 1998). Remarkably, *Hopscotch* gain-of-function mutants constitutively express high levels of *Tep1*. Moreover, in *Hopscotch* loss-of-function mutants the inducible expression of the gene dramatically levels off, suggesting a direct regulation of *Tep1* expression by the JAK pathway. Consistently, the constitutive expression of the gene in *Toll* gain-of-function mutants is abolished in the *Hopscotch* loss-of-function background.

3. Thioester-containing proteins in *A. gambiae*

The *A. gambiae* genome contains 19 homologues of *TEP* genes (Christophides et al., 2002). Among these, four gene pairs display characteristic features of haplotypes (*TEP1-TEP16*, *TEP5-TEP17*, *TEP6-TEP18* and *TEP8-TEP19*) and, therefore, might represent polymorphic variations rather than closely related but distinct genes. Consequently, the TEP family in *A. gambiae* is likely to be represented by a total of 15 *TEPs*. TEP1 is a typical thioester-containing protein of the mosquito and has been characterized in more detail (Levashina et al., 2001). This protein contains a signal peptide-like hydrophobic N-terminal segment characteristic of secreted proteins and the canonical thioester motif, which is followed 100 amino acids downstream by a catalytic histidine residue. The most C-terminal part displays a cysteine signature, characteristic of the *Drosophila* Teps (see above). The protein is glycosylated and secreted into the hemolymph by mosquito hemocytes. In the hemolymph, TEP1 is present as a full-length form of approximately 165 kDa and as a smaller fragment of 80 kDa (Levashina et al., 2001), indicating a constant low level of proteolytic cleavage of the protein. Interestingly, the same type of cleavage can be induced by both wounding and bacterial challenge. It is unknown whether cleavage occurs during activation or subsequent inactivation of the molecule. Functional studies of TEP1 were performed using a mosquito cell line, established by H.-M. Müeller (Catteruccia et al., 2000; Müller et al., 1999). Mosquito cells in vitro secrete TEP1 into the conditioned medium, where it is readily detected by affinity-purified rabbit polyclonal antibodies which recognise the full-length and the C-terminal fragment of the molecule (Levashina et al., 2001). The functionality of the thioester bond in TEP1 is supported by experiments on the denaturation-dependent autocatalytic fragmentation. Heat treatment of the conditioned medium leads to the appearance of a new C-terminal fragment of 50 kDa, resulting from fragmentation of the molecule at the thioester site. Methylamine treatment completely prevents this autocatalytic fragmentation (Levashina et al., 2001).

Mosquito cultured cells show phagocytic activity against latex beads and bacteria (Dimopoulos et al., 1999). Interestingly, this activity can be enhanced by pre-treating Gram-negative *Escherichia coli* with the conditioned cell medium, before exposing the bacteria to the cells (Levashina et al., 2001). Double-stranded RNA knockdown of *TEP1* or methylamine treatment of the conditioned medium both decrease the efficiency of bacteria uptake by 50%. These experiments demonstrate that a thioester protein of a protostome opsonises Gram-negative bacteria for phagocytosis and, therefore, displays an ancient complement-like function (Levashina et al., 2001). In contrast, the efficiency of phagocytosis of Gram-positive bacteria is low and independent of opsonization by TEP1. Interestingly, TEP1 binds to both Gram-positive and Gram-negative bacteria in a thioester-dependent manner (Levashina et al., 2001). Therefore, the low uptake of Gram-positive bacteria by the mosquito cultured cells is not due to inefficient opsonization by TEP1. Recently, the role of TEP1 in the phagocytosis of bacteria was assessed in adult mosquitoes using double-stranded RNA silencing. In these experiments,

TEP1 was required for promotion of phagocytosis of both Gram-positive and Gram-negative bacteria (Moita et al, in preparation), suggesting that the cultured cells used in the first experiments are deficient for a component of the Gram-positive bacteria phagocytic receptor complex.

Of special interest is the role of TEPs in the interactions between *Plasmodium* and the mosquito. The malaria parasite performs only part of its life cycle in the mammalian host, where it invades hepatocytes, then enters the circulation and multiplies in red blood cells. The remaining development of the parasite, including its sexual cycle, obligatorily occurs in the mosquito. After mating in the midgut lumen, parasites transform into motile ookinetes that cross the midgut epithelium. Between the epithelium and the basal lamina, ookinetes round up to form a kyst, called oocyst, where the parasite multiplies via mitosis. Ten days later, the oocyst bursts open and releases sporozoites that invade salivary glands and are transmitted during a subsequent blood meal. Recent double-stranded RNA knockdown experiments in adult mosquitoes indicate that TEP1 plays an essential role in the mosquito antiparasitic response (Blandin et al., in preparation). In susceptible mosquitoes, the knockdown of *TEP1* results in a five-fold increase in the number of oocysts developing on the midgut. Parasite killing in mosquitoes is mediated by direct binding of TEP1 to the surface of ookinetes. The role of the thioester in the binding of TEP1 to the parasite surface and the elucidation of the molecular mechanisms of TEP1 parasite killing will provide important insights in the evolution and function of the TEP family. During its complex passage through both the vertebrate host and the invertebrate vector, it now appears that the parasite repeatedly faces attacks from complement-like systems that represent more than 400 million years of evolutionary diversification.

The binding of TEP1 to different types of pathogens and the consequent promotion of phagocytosis of bacteria in the mosquito hemocoel or parasite killing in the midgut, are reminiscent of functions of vertebrate complement factors. It is not clear whether all mosquito TEPs are engaged in similar functions. Intriguingly, most mosquito TEPs have not conserved the thioester motif: only four have the thioester, nine proteins display modified sequences, and for two genes, we lack the sequence information in the thioester region. In vertebrates, TEPs with modified thioester regions serve as adaptors for initiation of the membrane attack complex, e.g. factor C5, or as tetrameric protease inhibitors. Further functional analysis of mosquito TEPs which lack thioester, might provide new insights into the importance of thioester motif and its requirements for the TEP function. As discussed above, TEPs are predominantly immune-inducible genes. In the mosquito, *TEP3* (S.B. and E.A.L., unpublished data) and *TEP4* (Oduol et al., 2000 and S.B. and E.A.L., unpublished data) are upregulated upon both bacterial challenge and parasite infection. Interestingly, our first functional data indicate that *TEP3* is indeed involved in the immune responses of the mosquito against both bacteria and parasites, whereas TEP4 plays an essential role in the defence against bacteria (S.B. and E.A.L., unpublished data). Although the functional analysis of the multi-member TEP family in mosquitoes is in its early days, it is clear that these proteins orchestrate distinct defence reactions against different pathogens, e.g. phagocytosis of bacteria and parasite killing.

4. Lessons and perspectives from the phylogenetic analysis of TEPs

The availability of the complete genome sequences from two dipteran species allowed us to refine our phylogenetic analysis of the TEP family across the animal kingdom (Levashina et al., 2001). Based on the full-length multisequence alignment generated with Clustal_X, we constructed an unrooted tree of the full-length sequences that were not too divergent in their local organisation using the neighbour-joining method (Fig. 1). We confirm the previously demonstrated separation of thioester-containing proteins into three families: complement factors, α_2-macroglobulins and invertebrate TEPs (Levashina et al., 2001 and Fig. 1). Clear separation into clades is observed for complement factors, where C3, C4 and C5 form prominent clades separated from the outlying deep branches representing primitive chordates. All these sequences, including the most divergent molecule from the coral *Swiftia exserta*, possess a characteristic C-terminal domain—one of the hallmarks of the complement factors. α_2Ms form a tight cluster which is suggestive of a slow rate of evolution of the represented sequences probably due to severe functional constraints on the structure of these pan-protease inhibitors. Interestingly, two molecules from the arthropods, the horseshoe crab *Limulus sp*. (Liα2M) and the soft tick *Ornithodoros moubata* (Tiα2M), form a subgroup in this subfamily. These two molecules are biochemically well characterized and possess inhibitory activity against a large spectrum of proteases (Kopacek et al., 2000; Quigley and Armstrong, 1985). Although insects belong to the same animal phylum, none of the mosquito or fruitfly TEPs falls into the α2M subfamily, instead they form the third unstable and the most divergent subfamily as evidenced by short internal branches with low bootstrap values. The overall topology of the tree suggests that the TEP subfamily is more closely related to α2Ms (note a relatively long separating branch between the complement factors and two other subfamilies) and the absence of the complement-specific C-terminal domain in both α2Ms and TEPs is consistent with this segregation.

The basis of the TEP subfamily displays a number of divergent outgroups, the position of which in the tree is supported by relatively low bootstrap values. Among these is the only clearcut orthologous group, which consists of *Drosophila Tep6* and *Anopheles TEP13*. Both molecules lack the thioester motive and their function in both insect species remains unknown. Another group, with neither bootstrap

nor phylogenetic support, includes the only TEP sequence from *Caenorhabditis elegans* (cTEP) and a recently identified α2M-like sequence from the ascidian *Ciona intestinalis*. Two striking features of the TEP subfamily are the presence of species-specific gene family expansions (Fig. 1, *Drosophila* TEPs are shaded in blue and *Anopheles* TEPs are shaded in red) and a minimal number of orthologs between these two sister insect species (Christophides et al., 2002). In addition to the unique orthologous group discussed above, there is one paralogous group, containing two mosquito TEPs (TEP2 and 15) and *Drosophila* Tep3 (Fig. 1). The minimal number of orthologs between *D. melanogaster* and *A. gambiae* indicates that TEPs in the fruitfly and in the mosquito have a limited number of common ancestors and that, most probably, in these species, TEPs have evolved independently. Indeed, the majority of insect TEPs represent species-specific expansions: the *Drosophila*-specific expansion includes all inducible Teps (Teps 1, 2 4 and 5), whereas the *Anopheles*-expansion includes all *TEP* genes located on the left arm of the third chromosome (Christophides et al., 2002). Interestingly, all mosquito *TEPs* that have been so far implicated in the immune defences in *A. gambiae* belong to this specific expansion. We suggest that species-specific expansions were finely shaped during the evolution of TEPs in response to distinct pathogenic environments in the two species. One may speculate that the fast species-specific evolution of insect TEPs has produced molecules with different functions. So far, only mosquito *TEP1* has been characterized in some detail. The structural analysis of *TEP1* suggests that this protein has conserved functional features of the complement factors (e.g. the structure of the C3d-like domain, the catalytic histidine and the thioester-mediated binding properties) (Levashina et al., 2001). Moreover, functional analysis based on loss-of-function phenotypes confirms the complement-like role of TEP1 in mosquito immune responses, such as promotion of the phagocytosis of bacteria and parasite killing. The pan-protease inhibitory activity has not yet been explored for insect TEPs due to technical limitations imposed by the availability of the biological material for the biochemical studies. Deeper characterization of loss-of-function phenotypes and the development of new tools for further biochemical analysis of insect TEPs will be instrumental for uncovering TEP functions and, eventually, for our understanding of this multifaceted family of proteins.

5. Conclusions

Compelling evidence suggests that complement factors and α_2-macroglobulins participate in inflammatory and immune responses. In addition to its important function in innate immunity, the central component of the complement system, factor C3, instructs the adaptive immune system to select appropriate antigens for humoral response, thus bridging the cellular and humoral arms of host defence. In contrast, the biological role of α2-macroglobulins is not well understood in spite of detailed characterizations of their biochemical activities. It has recently become apparent that new insights on the structure-function evolution of this protein family could be gained through the analysis of insect TEPs. The thioester-containing proteins represent a multi-member family both in the fruitfly and in the mosquito, probably reflecting the evolutionary importance of this family in the biology of insects. Interestingly, many thioester-containing molecules in insects lack the thioester motif but the function of these proteins remains unclear. The overall picture suggests that TEPs in the animal phyla undergo dramatic selective pressures that result in a set of polypeptides specific for each species. The nature of these selective pressures is unknown, and can only be imagined when the functions of the proteins are more fully characterized.

Acknowledgements

The authors are grateful to Professors Fotis Kafatos and Jules Hoffmann for continuous support and to Dr. Toby Gibson for help with the phylogenetic analysis. The work in the authors' laboratories has received support from the National Institutes of Health, Grant 1PO1 AI44220, and from the European Commission Human Potential—Research Training Networks Programme, Grant HPRN-CT-2000-00080.

References

Anderson, K.V., 2000. Toll signaling pathways in the innate immune response. Curr. Opin. Immunol. 12, 13–19.

Catteruccia, F., Nolan, T., Blass, C., Müller, H.M., Crisanti, A., Kafatos, F.C., Loukeris, T.G., 2000. Toward *Anopheles* transformation: *Minos* element activity in anopheline cells and embryos. Proc. Natl. Acad. Sci. U.S.A. 97, 2157–2162.

Christophides, G.K., Zdobnov, E., Barillas-Mury, C., Birney, E., Blandin, S., Blass, C., Brey, P.T., Collins, F.H., Danielli, A., Dimopoulos, G., Hetru, C., Hoa, N.T., Hoffmann, J.A., Kanzok, S.M., Letunic, I., Levashina, E.A., Loukeris, T.G., Lycett, G., Meister, S., Michel, K., Moita, L.F., Muller, H.M., Osta, M.A., Paskewitz, S.M., Reichhart, J.M., Rzhetsky, A., Troxler, L., Vernick, K.D., Vlachou, D., Volz, J., von Mering, C., Xu, J., Zheng, L., Bork, P., Kafatos, F.C., 2002. Immunity-related genes and gene families in *Anopheles gambiae*. Science 298, 159–165.

Chu, C.T., Pizzo, S.V., 1994. alpha 2-Macroglobulin, complement, and biologic defense: antigens, growth factors, microbial proteases, and receptor ligation. Lab. Invest. 71, 792–812.

Crowley, T.E., Hoey, T., Liu, J.K., Jan, Y.N., Jan, L.Y., Tjian, R., 1993. A new factor related to TATA-binding protein has highly restricted expression patterns in *Drosophila*. Nature 361, 557–561.

Dimopoulos, G., Müller, H.M., Kafatos, F.C., 1999. How does *Anopheles gambiae* kill malaria parasites? Parassitologia 41, 169–175.

Imler, J.L., Hoffmann, J.A., 2000. Signaling mechanisms in the antimicrobial host defense of *Drosophila*. Curr. Opin. Microbiol. 3, 16–22.

Kopacek, P., Weise, C., Saravanan, T., Vitova, K., Grubhoffer, L., 2000. Characterization of an alpha-macroglobulin-like glycoprotein isolated from the plasma of the soft tick *Ornithodoros moubata*. Eur. J. Biochem. 267, 465–475.

Lagueux, M., Perrodou, E., Levashina, E.A., Capovilla, M., Hoffmann, J.A., 2000. Constitutive expression of a complement-like protein in *Toll* and JAK gain-of-function mutants of *Drosophila*. Proc. Natl. Acad. Sci. U.S.A. 97, 11427–11432.

Levashina, E.A., Blandin, S., Moita, L.F., Lagueux, M., Kafatos, F.C., 2002. Thioester-containing proteins of protostomes. In: Ezekowitz, R.A.B., Hoffmann J.A. (Eds.), Infectious Disease: Innate Immunity, Humana Press Inc., Totowa, NJ, pp. 155–173.

Levashina, E.A., Moita, L.F., Blandin, S., Vriend, G., Lagueux, M., Kafatos, F.C., 2001. Conserved role of a complement-like protein in phagocytosis revealed by dsRNA knockout in cultured cells of the mosquito, *Anopheles gambiae*. Cell 104, 709–718.

Mathey-Prevot, B., Perrimon, N., 1998. Mammalian and *Drosophila* blood: JAK of all trades? Cell 92, 697–700.

Meister, M., Lagueux, M., 2003. *Drosophila* blood cells. Cell Microbiol. 5, 573–580.

Müller, H.M., Dimopoulos, G., Blass, C., Kafatos, F.C., 1999. A hemocyte-like cell line established from the malaria vector *Anopheles gambiae* expresses six prophenoloxidase genes. J. Biol. Chem. 274, 11727–11735.

Nonaka M., 2000. Origin and evolution of the complement system. In: Du Pasquier, L., Litman, G.W. (Eds.), Origin and Evolution of the Vertebrate Immune System, vol. 248, Springer-Verlag, New York, pp. 37–50.

Oduol, F., Xu, J., Niare, O., Natarajan, R., Vernick, K.D., 2000. Genes identified by an expression screen of the vector mosquito *Anopheles gambiae* display differential molecular immune response to malaria parasites and bacteria. Proc. Natl. Acad. Sci. U.S.A. 97, 11397–11402.

Qiu, P., Pan, P.C., Govind, S., 1998. A role for the *Drosophila* Toll/Cactus pathway in larval hematopoiesis. Development 125, 1909–1920.

Quigley, J.P., Armstrong, P.B., 1985. A homologue of alpha 2-macroglobulin purified from the hemolymph of the horseshoe crab Limulus polyphemus. J. Biol. Chem. 260, 12715–12719.

Salvesen, G.S., Sayers, C.A., Barrett, A.J., 1981. Further characterization of the covalent linking reaction of alpha 2-macroglobulin. Biochem. J. 195, 453–461.

Saravanan, T., Weise, C., Sojka, D., Kopacek, P., 2003. Molecular cloning, structure and bait region splice variants of alpha2-macroglobulin from the soft tick *Ornithodoros moubata*. Insect. Biochem. Mol. Biol. 33, 841–851.

Sim, R.B., Sim, E., 1981. Autolytic fragmentation of complement components C3 and C4 under denaturing conditions, a property shared with alpha 2-macroglobulin. Biochem. J. 193, 129–141.

Sottrup-Jensen, L., 1989. Alpha-macroglobulins: structure, shape, and mechanism of proteinase complex formation. J. Biol. Chem. 264, 11539–11542.

Volanakis, J.E., 1995. Transcriptional regulation of complement genes. Annu. Rev. Immunol. 13, 277–305.

Chapter 4

Reverse Genetics in Adult Mosquitoes

Reverse Genetics in Adult Mosquitoes

In the first chapter, we have shown that gene function can be assessed *in vitro* by dsRNA silencing in mosquito cells. Although this method can be useful to rapidly screen for genes involved in a particular process, *in vitro* results should be confirmed *in vivo*. In addition, some biological processes cannot be easily reproduced in cell culture, which is particularly true for the infection of mosquitoes by *Plasmodium* parasites (S.-H. Shiao, unpublished results). This prompted us to extend the dsRNA knockdown to adult mosquitoes for functional analysis *in vivo*. Several important technical questions were raised: In mosquito cells, we delivered dsRNA by lipotransfection, how should it be delivered in adult mosquitoes? When should it be delivered? Which degree of knockdown can we reach? And finally, how long does the knockdown last?

To answer these questions, we have chosen to silence the expression of the antimicrobial peptide gene, *Defensin1*, which had been well characterized in *A. gambiae*. In addition, no loss-of-function analysis of any insect antimicrobial peptide had been reported. Indeed, although antimicrobial peptides had been even extensively studied in *Drosophila* and were postulated to play an important role in the fruitfly defences against bacteria and fungi, no null mutant had been developed, mainly due to the small size of the genes that encode them, which dramatically decreases the chance to mutate these genes by EMS mutagenesis or by insertion of a P element.

In mosquitoes, Defensin1 is produced and secreted in the hemolymph by hemocytes and fat body cells (Richman et al., 1996). Defensin1 is active mainly against Gram-positive bacteria *in vitro* (Vizioli et al., 2001b). Several experiments suggested that it might also be involved in limiting parasite development. First, *Defensin1* is strongly upregulated in the anterior midgut and in the fat body upon parasite infection (Dimopoulos et al., 1998; Richman et al., 1997). A second indirect indication came from the fact that the injection of large quantities of exogenous Defensins from a fleshfly and a dragonfly in *Aedes aegyti* at specific time points after infection reduces oocyst density (Shahabuddin et al., 1998). We reasoned that if Defensin was important antibacterial and antimalarial factor, its absence would affect mosquito survival upon bacterial infection and result in higher parasite loads upon *Plasmodium* infection.

The functional analysis of Defensin1 was published in 2002 (Blandin et al., 2002), and can be found on pages 101-107. In this publication, we reported that:

• The injection of plain dsRNA directed against *Defensin1* in 1 to 2 day-old adult mosquitoes efficiently silences the expression of *Defensin1*. We followed the expression of *Defensin1* at the RNA level by Northern Blot and RT-PCR analyses and concluded that the injection of specific dsRNA inhibits the induction of *Defensin1* for as long as 22 days (this report and data not shown), therefore covering most of the life of an adult mosquito. We confirmed the efficiency of the knockdown at the protein level using MALDI-TOF analysis of hemolymph and anterior midgut extracts and demonstrated that dsRNA knockdown is efficient in the three distinct tissues tested (midgut, fat body and hemocytes).

• We then established a system to measure the effect of a gene knockdown on the survival of mosquitoes upon bacterial challenge. For this, we injected measured suspensions of different bacteria in control mosquitoes treated with an unrelated dsRNA and followed the survival of the mosquitoes for a week. These experiments allowed us first to select bacterial strains that were not excessively pathogenic to the mosquitoes, and second, to define the sublethal concentrations for infection with these bacteria.

• Using this survival test to analyze the *Defensin1* knockdown phenotype, we showed that this peptide is required for the mosquito antimicrobial defences against Gram-positive but not against Gram-negative bacteria, thus confirming the *in vitro* results.

• Finally, the knockout of *Defensin1* in *A. gambiae* had no effect on the development of *Plasmodium* midgut stages. Although our results demonstrate that endogenous Defensin1 does not act as a major antiparasitic factor *in vivo*, we do not argue the potential of antimicrobial peptides as antiparasitic factors when high quantities of these peptides are injected or when ectopically expressed.

In conclusion, we provided the first functional evidence that an insect antimicrobial peptide does play an important role in the resistance to pathogens. Evidently, other antimicrobial peptides, including those encoded by the three additional *Defensin* genes present in the mosquito genome, cannot substitute for Defensin1 in the mosquito defence against Gram-positive bacteria. The establishment of a simple *in vivo* dsRNA knockdown approach opens new perspectives for the study of mosquito gene function *in vivo* and provide a basis for systematic functional screens by targeted gene silencing.

ARTICLE 3

REVERSE GENETICS IN THE MOSQUITO *ANOPHELES GAMBIAE*: TARGETED DISRUPTION OF THE DEFENSIN GENE

Stéphanie BLANDIN, Luis F. MOITA,
Thomas KÖCHER, Matthias WILM, Fotis C. KAFATOS,
Elena A. LEVASHINA

EMBO Reports 2002, 3: 852-856

EMBO *reports*

scientific report

Reverse genetics in the mosquito *Anopheles gambiae*: targeted disruption of the *Defensin* gene

Stéphanie Blandin, Luis F. Moita, Thomas Köcher, Matthias Wilm, Fotis C. Kafatos & Elena A. Levashina+

European Molecular Biology Laboratory (EMBL), Meyerhofstrasse 1, D-69117 Heidelberg, Germany

Received May 24, 2002; revised July 18, 2002; accepted July 2002

***Anopheles gambiae*, the major vector of human malaria parasite, is an important insect model to study vector–parasite interactions. Here, we developed a simple *in vivo* double-stranded RNA (dsRNA) knockout approach to determine the function of the mosquito antimicrobial peptide gene *Defensin*. We injected dsRNA into adults and observed efficient and reproducible silencing of *Defensin*. Analysis of the knockdown phenotype revealed that this peptide is required for the mosquito antimicrobial defense against Gram-positive bacteria. In contrast, in mosquitoes infected by *Plasmodium berghei*, no loss of mosquito viability and no significant effect on the development and morphology of the parasite midgut stages were observed in the absence of Defensin. We conclude that this peptide is not a major antiparasitic factor in *A. gambiae in vivo*. Our results open new perspectives for the study of mosquito gene function *in vivo* and provide a basis for genome-scale systematic functional screens by targeted gene silencing.**

INTRODUCTION

Anopheles gambiae is the most important vector for *Plasmodium falciparum* malaria in sub-tropical Africa and thus a critical link in the transmission cycle of one of the most serious infectious diseases of humanity (Greenwood and Mutabingwa, 2002). In recent years, this mosquito has been studied intensively by the methods of molecular and cell biology, with a special emphasis on innate immune mechanisms possibly implicated in limiting the load of parasite transmission. Such studies promise to advance rapidly when the completion of the *A. gambiae* genome sequencing reveals the universe of genes upon which mosquito immunity resides (Hoffman *et al.*, 2002). Despite the existence of a good microsatellite-based genetic map (Zheng *et al.*, 1996) and robust techniques for germ-line transgenesis (Catteruccia *et al.*, 2000; Grossman *et al.*, 2001), inherent limitations of mosquito stock maintenance hinder the traditional methods for gene function analysis, such as large mutagenesis screens, fine-scale gene mapping and transgenic analysis. The double-stranded RNA (dsRNA) interference is potentially adaptable to systematic reverse genetic screens (Gonczy *et al.*, 2000), and we have shown recently that it can be used to assess gene function in cultured mosquito cells (Levashina *et al.*, 2001). Here, we demonstrate that dsRNA can also be used to disrupt essential gene function in the whole mosquito.

When dsRNA is taken up by cells, it is cleaved into small interfering fragments that can trigger specific degradation of the endogenous target mRNA (Zamore *et al.*, 2000). *Caenorhabditis elegans* is the only multicellular organism where the direct injection of dsRNA in the animal has been demonstrated to result in gene silencing throughout development (Fire *et al.*, 1998). In *Drosophila*, dsRNA injection in embryos is widely used for assessing gene function in early development, but in adult flies RNA interference is routinely mediated by the expression of hairpin dsRNA in transgenic strains (St Johnston, 2002), a technique that would be very difficult to apply systematically in mosquitoes. Instead, we have opted for a rapid and direct approach, intrathoracic injection of dsRNA in adult *A. gambiae*, to bring about efficient and reproducible silencing of the gene encoding the antimicrobial peptide, Defensin, which is encoded by a single gene in the *A. gambiae* genome.

So far, three antimicrobial peptides have been characterized in *A. gambiae*, Defensin, Cecropin and Gambicin, which are produced by the fat body and hemocytes and secreted into hemolymph upon immune challenge (Richman *et al.*, 1996; Vizioli *et al.*, 2000, 2001a). These polypeptides exhibit bactericidal and/or fungicidal activities *in vitro* and are thought to constitute the first line of defense against microbial infections

+Corresponding author. Tel: +49 6221 387 440; Fax: +49 6221 387 211; E-mail: elena.levashina@embl-heidelberg.de

scientific report

(reviewed in Dimopoulos *et al.*, 2001; Hoffmann and Reichhart, 2002). Related peptide families exist in vertebrates and are mostly expressed by epithelia and leukocytes, acting locally to limit bacterial infection (reviewed in Lehrer and Ganz, 2002). A knockout mouse lacking the Defensin-like peptide Cathelicidin has been shown recently to be susceptible to necrotic skin infections caused by Gram-positive Group A streptococci (Nizet *et al.*, 2001). In *Drosophila*, the inactivation of two major regulatory signalling pathways, *Toll* and *Imd*, has been used to turn off large groups of immune genes, including the antimicrobial peptide genes, and to associate them collectively with *in vivo* antifungal and antibacterial functions, respectively (Lemaitre *et al.*, 1996). However, no loss-of-function mutants for individual antimicrobial peptide genes have been reported as yet.

The expression of *Defensin* is predominantly induced in the mosquito fat body shortly after bacterial challenge. It is also induced locally in the midgut and salivary gland epithelia upon invasion by malaria parasites, suggesting that Defensin may have a broad role in the defense against both microbes and parasites (Richman *et al.*, 1996, 1997). This presumption was supported by *in vitro* tests of antiparasitic activity and by injection studies in *Aedes* mosquitoes infected by avian malaria (Shahabuddin *et al.*, 1998). However, rigorous conclusions about Defensin function *in vivo* require analysis by a loss-of-function approach in the intact mosquito.

RESULTS AND DISCUSSION

To knock down the *Defensin* gene expression, we injected 1- to 2-day-old females with dsRNA corresponding to the genes for either Defensin (*dsDEF*) or green fluorescent protein (*dsGFP*) as a control and allowed the mosquitoes to recover for 4 days. The dsRNA-treated mosquitoes were then challenged with *Escherichia coli*, and the presence/absence of the *Defensin* transcripts was monitored from day 1 to 8 by RNA blotting and RT–PCR (Figure 1). In six independent experiments, the injection of either Gram-negative *E. coli* or Gram-positive *Staphylococcus aureus* induced the expression of *Defensin* mRNA in non-treated control mosquitoes (Figure 1A; data not shown). We observed that the injection of *dsGFP* partially suppressed the ultimate level of *Defensin* induction after bacterial challenge (Figure 1A versus C), and therefore we used *dsGFP* mosquitoes as controls throughout this study. In the *dsDEF* mosquitoes, no *Defensin* mRNA of proper size was detected already at day 5 after dsRNA injection (Figure 1B and C, day 1 after bacterial challenge). Instead, a strong faster migrating signal was consistently present, accompanied by a faint signal migrating more slowly than *Defensin* mRNA. We interpret these signals as denatured and non-denatured dsRNA, respectively, as they are also exhibited by the input dsRNA (asterisks in Figure 1D). Signals corresponding to the input *dsGFP* were also detected using the *GFP* probe (bottom asterisk in Figure 1C and D), indicating that dsRNAs are stable for at least 12 days after injection in mosquitoes and therefore can provide a long-lasting inhibition of endogenous gene expression. In contrast to robust induction of *Defensin* in *dsGFP* mosquitoes (Figure 1E, days 1, 2 and 4), only traces of *Defensin* mRNA were detected in *dsDEF* mosquitoes using sensitive RT–PCR and primers corresponding to the 5′ and 3′ UTRs of the *Defensin* gene. We conclude that the injection of

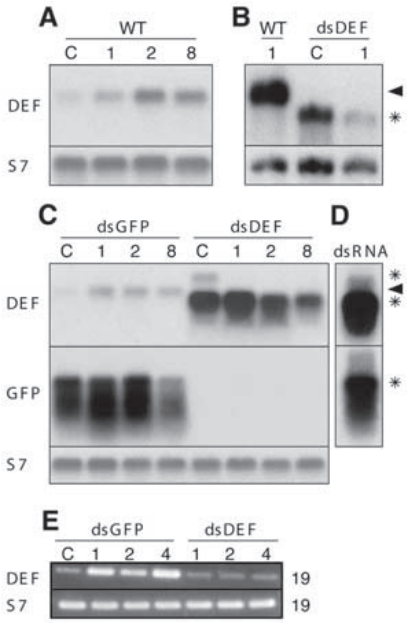


Fig. 1. RNA analysis of the *Defensin* gene knockout by dsRNA. RNA blots demonstrate that, over a period from 1 to 8 days (numbers above the images), *E. coli* challenge stably induces the expression of *Defensin* (*DEF*) mRNA (arrowheads) above the uninfected control level (C) in wild-type (WT) (**A** and **B**) and in *dsGFP*-treated, but not in *dsDEF*-treated, mosquitoes (B and **C**). In (**D**), the input *dsRNAs* match in size the signals (*DEF* and *GFP*) that are detected in dsRNA-injected mosquitoes (asterisks). (A), (C) and (D) were run on 1.2% agarose gels and (B) on an 1.4% agarose gel. The ribosomal protein *S7* transcript was used as a loading control. *GFP*, green fluorescent protein. (**E**) RT–PCR analysis of the *DEF* (19 cycles) gene expression in *dsGFP*- and *dsDEF*-treated mosquitoes before (C) and after (days 1, 2 and 4) bacterial challenge. The expression of the ribosomal protein gene *S7* (19 cycles) served as control.

specific dsRNA successfully inhibits induction of *Defensin* after bacterial challenge.

The efficacy of the *Defensin* knockout was further validated at the polypeptide level by matrix-assisted laser desorption ionization time-of-flight (MALDI-TOF) mass spectometry (MS) of hemolymph samples (Figure 2A). Although the *dsDEF* and *dsGFP* peptide profiles were mostly comparable, the latter showed a strong peak at 4136.2 Da, which was undetectable in the *dsDEF* sample and corresponded to singly protonated *A. gambiae* Defensin bearing three disulfide bridges (calculated mass 4136.7 Da) (Vizioli *et al.*, 2001b). This identification was confirmed by sequencing of the reduced peptide (data not shown). We also followed the *Defensin* knockout in the epithelium of the anterior midgut 24 h after infectious bloodmeal. The *dsGFP* and *dsDEF* midgut extracts again differed by a prominent Defensin peak at 4136.9 Da, which was present only in the midgut cells of *dsGFP* mosquitoes (Figure 2B). Thus, the injection of dsRNA in adult mosquitoes disrupts the expression of the targeted gene at both the RNA and polypeptide levels. Our data demonstrate that dsRNA knockout is efficient in the three different cell types

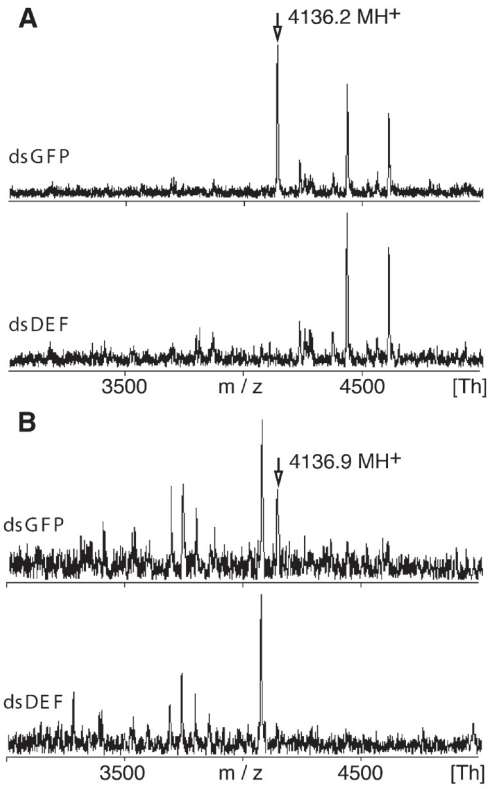


Fig. 2. MALDI-TOF MS analysis of the *Defensin* gene knockout by dsRNA. (**A**) Mosquitoes were challenged with a mixture of *E. coli* and *S. aureus*, and 36 h later the presence of Defensin peptide was detected in the hemolymph of control *dsGFP*, but not of *dsDEF*, mosquitoes. (**B**) Defensin is present in the anterior midgut 24 h after infectious bloodmeal in control *dsGFP*, but not in *dsDEF*, mosquitoes. The peaks corresponding to Defensin, and their molecular weights are indicated by arrows.

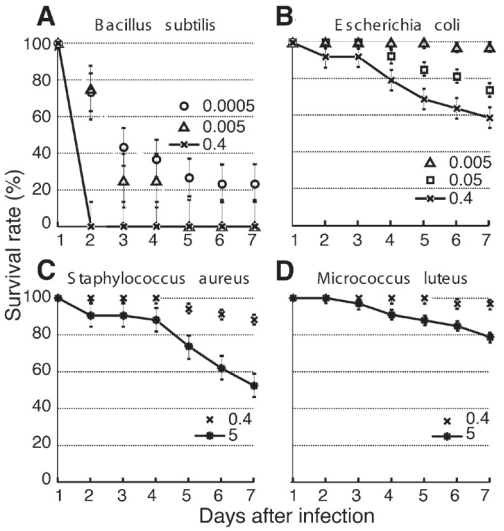


Fig. 3. Lethality of control *dsGFP*-treated mosquitoes after infection with different doses of bacteria. Survival rates (%) are presented for mosquitoes infected with (**B**) the Gram-negative bacterium *E. coli* or the Gram-positive bacteria (**A**) *B. subtilis*, (**C**) *S. aureus* and (**D**) *M. luteus*. The bacterial concentrations are expressed as optical densities (OD) of suspensions at 600 nm: OD_{600} = 0.0005 (open circles), OD_{600} = 0.005 (open triangles), OD_{600} = 0.05 (open squares), OD_{600} = 0.4 (crosses) and OD_{600} = 5 (asterisks). Each experiment was performed with 50 mosquitoes for each bacterial species, and the results shown are representative of three independent experiments. Standard errors are indicated by bars.

tested (midgut, fat body and hemocytes), which originate from two distinct cell lineages: endoderm and mesoderm.

The efficacy, effectiveness and reproducibility of the knockout allowed us to determine the function and specificity of Defensin *in vivo*. Because of *in vitro* indications that Defensin is particularly potent against Gram-positive bacteria (Vizioli *et al.*, 2001b), we focused on three members of this class, *S. aureus*, *Micrococcus luteus* and *Bacillus subtilis*, as well as on a Gram-negative species, *E. coli*. We injected measured bacterial suspensions of each species into control *dsGFP A. gambiae* females and followed the survival of the mosquitoes for 7 days (Figure 3). The bacteria exhibited different degrees of pathogenicity that were independent of Gram classification. *Bacillus subtilis* rapidly killed *A. gambiae* even at concentrations as low as 85 bacteria per mosquito (Figure 3A). The next most efficient pathogen, *E. coli*, caused comparable mosquito lethality only when injected in 10-fold higher numbers; *S. aureus* was substantially less pathogenic, and *M. luteus* was the least effective (Figure 3B–D). Differential susceptibility to bacteria has also been reported in *Drosophila melanogaster*, where, surprisingly, *S. aureus*, as well as *B. subtilis*, cause rapid death, while *M. luteus* is again only weakly pathogenic (Tzou *et al.*, 2002). Comparative analysis of bacterial pathogenicity may be fruitful in pinpointing specificities of the immune response in related dipteran species.

These experiments allowed us to define the sublethal concentrations for infection with the three bacterial species other than *B. subtilis*. We injected the selected concentrations of bacteria in both *dsDEF* and control *dsGFP* females and followed their respective survival rates over a period of 7 days (Figure 4). The profiles did not differ significantly in the case of *E. coli*- and mock-injected mosquitoes. In contrast, the *dsDEF* mosquitoes were modestly susceptible to *M. luteus* (20% lethality) and highly susceptible to *S. aureus* (80% lethality). This is the first demonstration *in vivo* that a single endogenous immune peptide, Defensin, is necessary for the resistance of a mosquito to two Gram-positive bacteria and not to a Gram-negative species of bacteria. The plateau in mortality that is seen in *dsDEF* mosquitoes 5 days after infection by *M. luteus* suggests that additional immune factors or cellular processes may be implicated in the clearance of this mild pathogen.

We next examined the potential role of Defensin in the immune response of *A. gambiae* to the rodent malaria parasite, *Plasmodium berghei*. *Defensin* is constitutively expressed in the anterior midgut epithelium and is further induced by malaria infection

scientific report

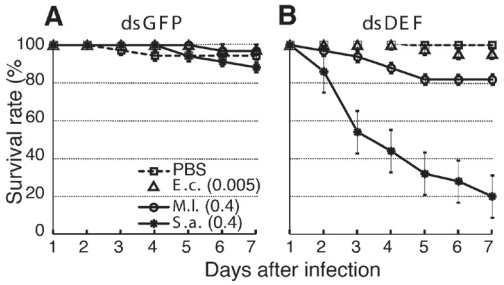


Fig. 4. *Defensin* knockdown mosquitoes are susceptible to Gram-positive, but not to Gram-negative, bacteria. The survival rates (%) of (**A**) *dsGFP*-treated mosquitoes and (**B**) *dsDEF*-treated mosquitoes after injection of PBS (open squares), *E. coli* (OD_{600} = 0.005, open triangles), *M. luteus* (OD_{600} = 0.4, open circles) or *S. aureus* (OD_{600} = 0.4, asterisks) are shown. Each experiment was performed with 50 mosquitoes for each bacterial species, and the results shown are representative of three independent experiments. Standard errors are indicated by bars.

(Richman *et al.*, 1997; Vizioli *et al.*, 2001b). In *Anopheles aegypti*, the injection of high doses of related Defensins from a dragonfly and a fleshfly at specific time points after an infectious bloodmeal interfered with the development of the midgut stages of the parasite (Shahabuddin *et al.*, 1998). These studies prompted us to monitor parasite numbers during *P. berghei* infections in *dsGFP* and *dsDEF* mosquitoes, using as a readout the number of oocysts per midgut 10–12 days after infection. We reasoned that, if endogenous Defensin acts antiparasitically, its absence during development of the *Plasmodium* midgut stages would remove a constraint and result in higher parasite loads. We observed that the mean number of oocysts per midgut was unchanged or even slightly lower in *dsDEF* than in *dsGFP* mosquitoes (Table I). Moreover, parasite-infected *dsDEF* and *dsGFP* mosquitoes showed no significant differences in mosquito viability, ookinete/oocyst morphology or the frequency distribution of oocyst numbers. We conclude that malaria-induced endogenous *A. gambiae* Defensin does not act as a significant antiparasitic factor *in vivo*.

In conclusion, the simple and convenient dsRNA technique that we describe here efficiently disrupts gene function in distinct tissues of adult mosquitoes. As a proof of principle, we have used this method to delimit phenotypically the *in vivo* function of a single immune peptide against different types of infections, showing it to play an important role in the resistance to Gram-positive bacteria. Evidently, antimicrobial peptides that would be unaffected by *dsDEF* injection cannot substitute for Defensin. Recently, we have successfully applied dsRNA to knock out 20 additional immune genes in *A. gambiae* (data not shown), thus confirming the general validity of this method in the study of immune responses in the mosquito. With a reverse genetics method now in hand, it should be possible to conduct systematic functional genomic analysis in this major vector of human malaria.

METHODS

Mosquito colony. *Anopheles gambiae* strain G3 was reared as described previously (Richman *et al.*, 1996).

Double-stranded RNA preparation and injection in mosquitoes. dsRNAs were produced as described previously using the plasmids pLL6ds for control *dsGFP* (Levashina *et al.*, 2001) and pLL80 for *dsDEF*. pLL80 was constructed in two steps. *Defensin* cDNA of 404 bp was PCR-amplified using *dfn a* and *dfn b* primers (Richman *et al.*, 1996) and cloned into pCR2.1-TOPO (Invitrogen), resulting in pLL79. The 515 bp *Hin*dIII–*Xba*I fragment of pLL79 was then subcloned between the two T7 promoters of pLL10. Sense and antisense RNAs were synthesized using the T7 Ambion kit, annealed in water and stored as dsRNAs at –80°C until use. A nano-injector (Nanoject, Drummond) was used to introduce 69 nl of dsRNAs (1 mg/ml) in water in the thorax of CO_2-anesthetized mosquito females, which were then allowed to recover for 4 days.

RNA analysis. Total RNA was extracted from 15 mosquitoes with TRIzol Reagent (Invitrogen) and separated by electrophoresis. Two different conditions were used: to ascertain the absence of endogenous *Defensin* transcripts, the electrophoresis were performed for 6 h using 1.2% agarose gels (Figure 1B); clear signals corresponding to input dsRNAs were detected using 1.4% agarose gels and 4 h migration time (Figure 1C and D). Separated total RNAs were transferred to a nylon membrane (Hybond-N^+, Amersham Pharmacia Biotech). Blots were hybridized sequentially with the radioactively-labeled probes [Ready-To-Go DNA labeling beads (dCTP), Amersham Pharmacia Biotech]: *Defensin* (pLL79 insert), *S7* (Salazar *et al.*, 1993) and *GFP* (pLL6 insert).

For RT–PCR analysis, *Defensin*-specific primers were selected that do not overlap with the dsRNA used for the knockout: 5′-UTR primer, 5′-AAC TCC AGC CAA GCT AAA GC-3′; and 3′-UTR primer, 5′-GAA TTA AGC CTG TGT TGT AAA C-3′. Total RNA was extracted as above from whole mosquitoes at the indicated time points. S7-specific primers and the RT–PCR

Table I. Survival of *P. berghei* oocysts in the *dsGFP*- and *dsDEF*-treated mosquitoes

Experiment	*dsRNA*	Number of midguts	Oocyst number per midgut			Mean ± SE[a](oocysts per midgut)
			0	1–29	>30	
1	*dsGFP*	19	5	9	5	16.74 ± 3.7
	dsDEF	24	7	14	3	10.75 ± 3.1
2	*dsGFP*	13	4	9	0	2.77 ± 0.6
	dsDEF	25	7	18	0	2.48 ± 0.4

[a]SE, standard error.

scientific report

conditions were as described previously (Richman *et al.*, 1996). For amplification of both *S7* and *Defensin*, 19 PCR cycles were used.

Mass spectrometry. The hemolymph of 40 mosquitoes was collected by centrifugation of decapitated females (3000 r.p.m., 5 min) and used in 1/10 dilution in acidified water. Five dissected anterior midguts were homogenized in 0.5% trifluoracetic acid, 50% acetonitrile; and extracts were cleared by centrifugation. The samples were analyzed without further purification in a modified thin-layer preparation (Vorm *et al.*, 1994). MALDI-TOF MS analysis was performed on a Bruker Biflex (Bremen, Germany) mass spectrometer in linear positive mode using delayed extraction. Tandem MS with a nano-electrospray source mounted on a Micromass QTOF1 (Manchester, UK) mass spectrometer was used for peptide sequencing.

Bacterial challenge and mosquito survival. GFP-expressing *E. coli* OP-50 was a gift from J.J. Ewbank (INSERM, Marseille-Luminy, France). *Bacillus subtilis, M. luteus* and *S. luteus* were kind gifts from P. Bulet (IBMC, Strasbourg, France). Bacteria were cultured to OD_{600} = 0.4, pelleted, washed and resuspended in phosphate-buffered saline (PBS) to indicated concentrations. The number of bacteria injected was estimated by the plating of appropriate aliquots of bacterial suspension on LB plates. Mosquitoes were anesthetized with CO_2, injected into the thorax with 69 nl of the bacterial suspension or PBS for controls and allowed to recover. Mosquitoes that died within 24 h of injection were not considered in the analysis. Dead mosquitoes were daily counted and removed over a period of 7 days. The results shown here are representative of at least three independent experiments, each carried out with 50 mosquitoes per tested group.

Parasite infections and oocyst counting. Parasite infections were performed essentially as described previously (Richman *et al.*, 1997). Briefly, *P. berghei* parasites were passaged in CD1 mice, and parasitemia was determined from Giemsa-stained blood films. For each experiment, *dsGFP* and *dsDEF* mosquitoes were fed on the same infected mouse. Mosquito midguts were dissected 10–12 days later, fixed and DAPI-stained. Morphology was examined, and the numbers of oocysts were counted using a UV-light fluorescent microscope (Zeiss).

ACKNOWLEDGEMENTS

We thank D. Doherty for help with the mosquito colony and J. Ewbank and P. Bulet for bacterial strains. This work was supported by EMBL, the National Institutes of Health and the European Commission.

REFERENCES

Catteruccia, F., Nolan, T., Loukeris, T.G., Blass, C., Savakis, C., Kafatos, F.C. and Crisanti, A. (2000) Stable germline transformation of the malaria mosquito *Anopheles stephensi*. *Nature*, **405**, 959–962.

Dimopoulos, G., Müller, H.M., Levashina, E.A. and Kafatos, F.C. (2001) Innate immune defense against malaria infection in the mosquito. *Curr. Opin. Immunol.*, **13**, 79–88.

Fire, A., Xu, S., Montgomery, M.K., Kostas, S.A., Driver, S.E. and Mello, C.C. (1998) Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature*, **391**, 806–811.

Gonczy, P. *et al.* (2000) Functional genomic analysis of cell division in *C. elegans* using RNAi of genes on chromosome III. *Nature*, **408**, 331–336.

Greenwood, B. and Mutabingwa, T. (2002) Malaria in 2002. *Nature*, **415**, 670–672.

Grossman, G.L., Rafferty, C.S., Clayton, J.R., Stevens, T.K., Mukabayire, O. and Benedict, M.Q. (2001) Germline transformation of the malaria vector, *Anopheles gambiae*, with the piggyBac transposable element. *Insect Mol. Biol.*, **10**, 597–604.

Hoffman, S.L., Subramanian, G.M., Collins, F.H. and Venter, J.C. (2002) *Plasmodium*, human and *Anopheles* genomics and malaria. *Nature*, **415**, 702–709.

Hoffmann, J.A. and Reichhart, J.M. (2002) *Drosophila* innate immunity: an evolutionary perspective. *Nat. Immunol.*, **3**, 121–126.

Lehrer, R.I. and Ganz, T. (2002) Defensins of vertebrate animals. *Curr. Opin. Immunol.*, **14**, 96–102.

Lemaitre, B., Nicolas, E., Michaut, L., Reichhart, J.M. and Hoffmann, J.A. (1996) The dorsoventral regulatory gene cassette *spatzle/Toll/cactus* controls the potent antifungal response in *Drosophila* adults. *Cell*, **86**, 973–983.

Levashina, E.A., Moita, L.F., Blandin, S., Vriend, G., Lagueux, M. and Kafatos, F.C. (2001) Conserved role of a complement-like protein in phagocytosis revealed by dsRNA knockout in cultured cells of the mosquito, *Anopheles gambiae*. *Cell*, **104**, 709–718.

Nizet, V. *et al.* (2001) Innate antimicrobial peptide protects the skin from invasive bacterial infection. *Nature*, **414**, 454–457.

Richman, A.M., Bulet, P., Hetru, C., Barillas-Mury, C., Hoffmann, J.A. and Kafalos, F.C. (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**, 203–210.

Richman, A.M., Dimopoulos, G., Seeley, D. and Kafatos, F.C. (1997) *Plasmodium* activates the innate immune response of *Anopheles gambiae* mosquitoes. *EMBO J.*, **16**, 6114–6119.

Salazar, C.E., Mills-Hamm, D., Kumar, V. and Collins, F.H. (1993) Sequence of a cDNA from the mosquito *Anopheles gambiae* encoding a homologue of human ribosomal protein S7. *Nucleic Acids Res.*, **21**, 4147.

Shahabuddin, M., Fields, I., Bulet, P., Hoffmann, J.A. and Miller, L.H. (1998) *Plasmodium gallinaceum*: differential killing of some mosquito stages of the parasite by insect defensin. *Exp. Parasitol.*, **89**, 103–112.

St Johnston, R. (2002) The art and design of genetic screens: *Drosophila melanogaster*. *Nat. Rev. Genet.*, **3**, 176–188.

Tzou, P., Reichhart, J.M. and Lemaitre, B. (2002) Constitutive expression of a single antimicrobial peptide can restore wild-type resistance to infection in immunodeficient *Drosophila* mutants. *Proc. Natl Acad. Sci. USA*, **99**, 2152–2157.

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

Vizioli, J., Bulet, P., Hoffmann, J.A., Kafatos, F.C., Muller, H.M. and Dimopoulos, G. (2001a) Gambicin: a novel immune responsive antimicrobial peptide from the malaria vector *Anopheles gambiae*. *Proc. Natl Acad. Sci. USA*, **98**, 12630–12635.

Vizioli, J., Richman, A.M., Uttenweiler-Joseph, S., Blass, C. and Bulet, P. (2001b) The defensin peptide of the malaria vector mosquito *Anopheles gambiae*: antimicrobial activities and expression in adult mosquitoes. *Insect Biochem. Mol. Biol.*, **31**, 241–248.

Vorm, O., Roepstorff, P. and Mann, M. (1994) Improved resolution and very high sensitivity in MALDI TOF of matrix surfaces made by fast evaporation. *Anal. Chem.*, **66**, 3281–3287.

Zamore, P.D., Tuschl, T., Sharp, P.A. and Bartel, D.P. (2000) RNAi: double-stranded RNA directs the ATP-dependent cleavage of mRNA at 21 to 23 nucleotide intervals. *Cell*, **101**, 25–33.

Zheng, L., Benedict, M.Q., Cornel, A.J., Collins, F.H. and Kafatos, F.C. (1996) An integrated genetic map of the African human malaria vector mosquito, *Anopheles gambiae*. *Genetics*, **143**, 941–952.

DOI: 10.1093/embo-reports/kvf180

Chapter 5

dsRNA Silencing in Adult *Drosophila melanogaster*

dsRNA silencing in Adult *Drosophila. melanogaster*

In *Drosophila* adults, gene silencing is mainly achieved by establishment of transgenic strains expressing hairpin constructs against the targeted genes (reviewed in St Johnston, 2002). This method requires a few months for construction of the transgene and generation of transgenic flies, and therefore, is not really adapted to large-scale reverse genetic screens (reviewed in Adams and Sekelsky, 2002). Although Dzitoyeva and coworkers reported in 2001 that the injection of specific dsRNA in adult flies disrupted the expression of a *LacZ* transgene in the central nervous system of the fruitfly (Dzitoyeva et al., 2001), no other attempt to use this rapid method had been described, suggesting that further improvements were required. Hence, we undertook to establish the conditions for dsRNA injection in adult flies to silence the expression of endogenous genes in a rapid, efficient and reproducible manner.

As a proof of principle, we decided to silence the expression of components of the *Toll* signalling cascade, representing positive and negative regulators that encode secreted, intracellular or transmembrane proteins which loss-of-function phenotypes had already been well characterized in mutant flies (reviewed in Hoffmann, 2003, and see Fig. 5). In addition, two novel components of the Toll pathway had been identified in a cell culture system: the atypical protein kinase C (DaPKC) and Ref(2)P (Avila et al., 2002). However, analysis of their function using classical genetics has been hindered by the fact that mutations in *DaPKC* are embryonic lethal (Wodarz et al., 2000) and *Ref(2)P*-mutants males are sterile (Gay and Contamine, 1993). To place these components in the Toll pathway, we developed a dsRNA-based epistatic gene analysis.

Our results were published in 2003 (Goto et al., 2003), and the article can be found on pages 113-119:

• Our first result was striking: we demonstrated that the direct injection of dsRNA into adult flies is efficient to disrupt gene expression only in flies that are at least 5 day-old, and that younger flies are not sensitive to dsRNA treatment. Noteworthy, we did not observe a similar age-dependent efficiency of gene silencing in *A. gambiae*, where the injection of specific dsRNA in newly hatched (1 to 2 day-old) adults fully inhibits the induction of *Defensin1* (see Chapter 4). At present, it is unclear why such a difference exists between these two insect species.

• We next examined the duration of dsRNA knockdown and showed that it lasts for at least a week.

• In the standard conditions we used to analyze dsRNA knockdown phenotypes, i.e. injection of dsRNA in 6 day-old flies and analysis of phenotypes 4 days later, we reported that dsRNA silencing is not equal for all components of the Toll pathway: the expression of some genes is more efficiently disrupted than that of others. Although we cannot exclude incomplete gene silencing by dsRNA treatment, we believe that these differences rather come from the slow turn-over or the post-translational regulation of some proteins.

• Finally, we confirmed *in vitro* data placing the atypical protein kinase C (DaPKC) and Ref(2)P in the Toll pathway via epistatic analyses using the injection of a mixture of two dsRNAs.

In the *Drosophila* model, where many genetic tools for functional analysis are already available, we believe that injection of dsRNA represents a powerful method for rapid functional analysis of genes in adult flies, and particularly of those that are essential for fly development, and for large-scale reverse genetic screens. Co-injections of dsRNAs can be used for rapid epistatic analysis of double knockdowns.

ARTICLE 4

SILENCING OF TOLL PATHWAY COMPONENTS BY DIRECT INJECTION OF dsRNA INTO *DROSOPHILA* ADULT FLIES

Akira GOTO, Stéphanie BLANDIN, Julien ROYET, Jean-Marc REICHHART, Elena A. LEVASHINA

Nucleic Acids Research 2003, 31: 6619-6623

Nucleic Acids Research, 2003, Vol. 31, No. 22 ***6619–6623***
DOI: 10.1093/nar/gkg852

Silencing of Toll pathway components by direct injection of double-stranded RNA into *Drosophila* adult flies

Akira Goto[1], Stéphanie Blandin[2], Julien Royet[1], Jean-Marc Reichhart[1] and Elena A. Levashina[1,2,*]

[1]UPR 9022 du Centre National de la Recherche Scientifique, Institut de Biologie Moléculaire et Cellulaire, 15 rue René Descartes, F-67084 Strasbourg Cedex, France and [2]European Molecular Biology Laboratory (EMBL), Meyerhofstrasse 1, D-69117 Heidelberg, Germany

Received July 3, 2003; Revised and Accepted September 23, 2003

ABSTRACT

Double-stranded RNA (dsRNA) gene interference is an efficient method to silence gene expression in a sequence-specific manner. Here we show that the direct injection of dsRNA can be used in adult *Drosophila* flies to disrupt function of endogenous genes *in vivo*. As a proof of principle, we have used this method to silence components of a major signaling cascade, the Toll pathway, which controls fruit fly resistance to fungal and Gram-positive bacterial infections. We demonstrate that the knockout is efficient only if dsRNA is injected in 4- or more day-old flies and that it lasts for at least 1 week. Furthermore, we report dsRNA-based epistatic gene analysis via injection of a mixture of two dsRNAs and propose that injection of dsRNA represents a powerful method for rapid functional analysis of genes in *Drosophila melanogaster* adults, particularly of those whose mutations are lethal during development.

INTRODUCTION

Drosophila melanogaster has been widely used as a model organism to identify genes that are involved in diverse biological processes (1). The publication of the fly genome created a great need for new approaches to do the reverse, namely to disrupt the functions of specific genes whose sequences are known (2). RNA interference (RNAi) is one of these tools (reviewed in 3). In fly embryos, injection of double-stranded RNA (dsRNA) is routinely used to analyze gene function but it is inherently limited to early development (4,5). Although direct injection of dsRNA in adults was previously reported in *Caenorhabditis elegans* (6), *Anopheles gambiae* (7) and *D.melanogaster* (8,9), gene silencing in *Drosophila* adults is mainly achieved by establishment of transgenic strains expressing hairpin constructs against the targeted genes (1).

Here we establish an efficient dsRNA silencing method by direct injection into adult flies of dsRNA for components of the Toll signaling pathway. This pathway was initially characterized for its role in dorso-ventral development in embryos (reviewed in 10). It was later found that it also exerts an immune function in adult flies (11–13). In this case, Toll is activated after immune challenge by a cleaved form of the polypeptide Spaetzle, leading to the cytoplasmic dissociation of the NF-κB family member Dif from its inhibitor, the IκB-like protein Cactus. This is followed by nuclear translocation of Dif, which controls expression of many genes including that encoding the antifungal peptide Drosomycin (Drs). In the absence of immune challenge, cleavage of Spaetzle is inhibited by the serpin Spn43Ac. Two novel components of this pathway have been recently identified in a cell-culture system: the atypical protein kinase C (DaPKC) and Ref(2)P (14). However, analysis of their function using classical genetics is difficult as mutations in *DaPKC* are embryonic lethal (15) and *Ref(2)P*-mutant males are sterile (16). We used direct injection of dsRNA into adult flies to place these components in the Toll pathway *in vivo*. For that, we developed dsRNA-based epistatic gene analysis via injection of a mixture of two dsRNAs. Our results demonstrate that injection of dsRNA represents a novel powerful method for rapid *in vivo* functional analysis of genes in *D.melanogaster* adults, particularly those for which mutations are lethal during development.

MATERIALS AND METHODS

Fly stocks and microbial infection

DD1 flies, carrying two reporter transgenes *P[w$^+$mC Dipt::LacZ = pDipt-LacZ]* and *P[w$^+$mC Drom::GFP = pDrs-GFP S65T]* (17), were grown on standard medium at 25°C. Four days after dsRNA injection, flies were pricked with a thin tungsten needle previously dipped into a concentrated culture of Gram-positive bacteria, *Micrococcus luteus*. For natural fungal infection, anesthetized flies were manually shaken for 30 s on a Petri dish containing a sporulating culture of *Beauvaria bassiana* and kept at 25°C for 48 h (18).

*To whom correspondence should be addressed. Tel: +33 388 417 018; Fax: +33 388 606 922; Email: e.levashina@ibmc.u-strasbg.fr

6620 *Nucleic Acids Research, 2003, Vol. 31, No. 22*

dsRNA preparation and injection into adult flies

Templates for the preparation of dsRNA were PCR-derived fragments sandwiched by two T7 promoter sequences (TAA TAC GAC TCA CTA TAG GGA GAC CAC). The amplified cDNA fragments for each gene were as follows: *green fluorescent protein* (*GFP*: nucleotides 35–736, GenBank accession no. L29345), *Cactus* (nucleotides 1041–1582, GenBank accession no. L04964), *dorsal-related immunity factor* (*Dif*: nucleotides 1921–2550, GenBank accession no. L29015), *serpin 43Ac* (*Spn43Ac*: nucleotides 57–1473, GenBank accession no. AJ245444), *Toll* (nucleotides 3289–3866, GenBank accession no. M19969), *Drosomycin* (*Drs*: nucleotides 28–260, GenBank accession no. X75595), *Drosophila atypical protein kinase C* (*DaPKC*: nucleotides 301–900, GenBank accession no. AF288482) and *Ref(2)P* (nucleotides 348–655, GenBank accession no. X69829). As a negative dsRNA control we used *odd-paired* (*opa*: nucleotides 135–1156, GenBank accession no. S78339). Single-stranded RNAs were synthesized by using the MEGAscript T7 transcription kit (Ambion, Austin, TX) according to instructions. Annealed dsRNAs were ethanol precipitated and dissolved in injection buffer (0.1 mM sodium phosphate, pH 6.8; 5 mM KCl). Using a nano-injector (NANOJECT II; Drummond), 32.2 nl of dsRNA (3 mg/ml) were injected in the thorax of CO_2-anesthetized adult females which were allowed to recover for 4 days before bacterial infection. In double knockdown experiments, 32.2 nl of a 1:1 mix of dsRNAs (3 mg/ml) were injected.

RNA preparation, northern blotting and RT–PCR

Total RNA from 20–25 flies was extracted by TRIzol Reagent (Invitrogen). Twenty micrograms of total RNA were separated by 1% agarose/formaldehyde gel electrophoresis, blotted to nylon membrane (Schleicher & Schuell) and hybridized sequentially with the ^{32}P-labeled cDNA probes for *GFP*, *Drs* and *ribosomal protein 49* (*rp49*) [Ready-To-Go DNA labeling beads (dCTP); Amersham Pharmacia Biotech]. The *rp49* was used as an internal loading control. Signal quantification was done by a Fuji Film BAS 2000 Image Analyzer (Fuji Photo Film Co., Tokyo, Japan).

For reverse transcription, 10 µg of total RNA were reverse transcribed with Superscript II RNase H$^-$ reverse transcriptase (Gibco BRL). Single-stranded cDNAs of different dilution were amplified by PCR using recombinant *Taq* DNA polymerase (Invitrogen). Primers were made to amplify endogenous *Toll* (nucleotides 3289–3866, GenBank accession no. M19969) for 30 cycles, *Dif* (nucleotides 1921–2550, GenBank accession no. L29015) for 35 cycles and *Actin 5C* (nucleotides 213–1205, NCBI accession no. NM-167053) for 25 cycles. *Actin 5C* was used as an internal control.

RESULTS AND DISCUSSION

Efficiency of dsRNA silencing in *Drosophila* adults is age specific

To analyze the efficiency of gene silencing by direct injection of dsRNA into adult flies, we prepared dsRNAs corresponding to several components of the Toll pathway. As a read-out of Toll activation we have used DD1 flies, which carry two reporter transgenes, *Drosomycin-GFP* (*Drs-GFP*) and *Diptericin-Lac Z*,

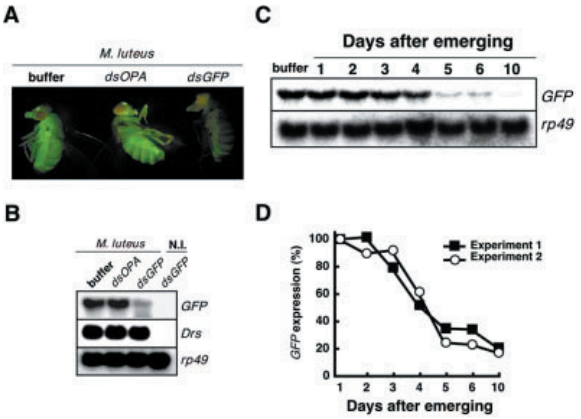


Figure 1. dsRNA silencing of the *Drs-GFP* transgene. Six-day-old DD1 females were injected with dsRNA against the *Drs-GFP* transgene (*dsGFP*), the immune-unrelated gene *oda* (*dsOPA*) or injection buffer (buffer) and 4 days later were either infected with Gram-positive bacteria (*M.luteus*) or left unchallenged as a control (N.I.). (**A**) GFP expression was analyzed 1 day later using fluorescence microscopy. (**B**) The expression of *Drs-GFP* (*GFP*), *Drs* and ribosomal protein gene *rp49* as a loading control, was monitored by northern blotting. (**C**) *dsGFP* was injected into 1–10-day-old flies and the level of *Drs-GFP* mRNA after *M.luteus* infection was monitored. (**D**) The graph represents relative percentages of *Drs-GFP* expression normalized by *rp49* in two independent experiments. 100% of expression corresponds to *Drs-GFP* expression in buffer-injected flies.

on the X chromosome (17). In this study, we monitored the expression of a *Drs-GFP* reporter transgene. We first injected into the thorax of females dsRNAs directed against an inducible *GFP* reporter gene (*dsGFP*) or against the unrelated gene *opa* (*dsOPA*). The injected flies were allowed to recover for 4 days and then challenged with Gram-positive bacteria, *M.luteus*, to induce the expression of the reporter *Drs-GFP* gene (17). We monitored the expression of GFP 1 day after bacterial challenge, when the maximal level of fluorescence is observed (11). *dsGFP* flies displayed a strong reduction in *Drs-GFP* expression compared with both buffer and *dsOPA* control flies (Fig. 1A). The efficiency of the knockout was confirmed at the mRNA level (Fig. 1B). We observed that the *M.luteus*-induced expression of *Drs-GFP* was similar in buffer- and *dsOPA*-injected flies, suggesting that dsRNA *per se* does not interfere with the expression of antimicrobial peptides. Interestingly, when we used flies of different ages, we noticed a high heterogeneity in *Drs-GFP* silencing. To determine the age dependency of gene silencing, we injected *dsGFP* into 1–10-day-old flies and monitored the *Drs-GFP* expression after bacterial challenge (Fig. 1C and D). We could not detect any significant decrease in *Drs-GFP* expression in flies injected with *dsGFP* earlier than 4 days after eclosion. Silencing became more efficient as flies grew older and was at its maximum at days 5 and 6 (5-fold reduction in *Drs-GFP* expression). The same age-dependent efficiency of the knockout was observed for other dsRNAs: *dsDRS*, *dsTOLL* and *dsSPN43Ac* (data not shown). This effect had not been detected in the mosquito *A.gambiae*, where the injection of specific dsRNA in newly hatched adults fully inhibited the induction of the antimicrobial peptide gene *Defensin* (7). Differential age-dependent efficiency of dsRNA knockout may reflect physiological differences between these two insect species.

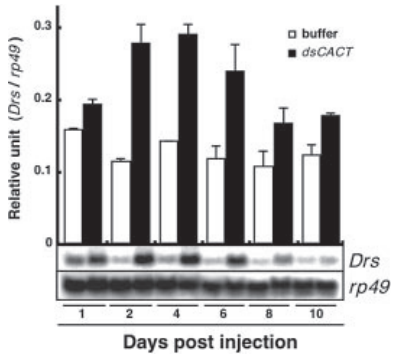


Figure 2. Time course of *Drs* expression in *dsCACT*-injected flies. DD1 females were injected with buffer or *dsCACT* and *Drs* expression was monitored over 10 days. The graph represents results of two independent experiments and shows relative *Drs* expression normalized by *rp49*. The lower panel shows representative northern blotting data for *Drs* and *rp49*.

Duration of the effect of gene silencing

We next examined the duration of dsRNA knockout efficiency by targeting *Cactus*, a gene encoding the ankyrin-repeat protein that negatively regulates the Toll pathway in adult flies (19,20). Loss-of-function mutations in *Cactus* lead to a signal-independent *Drs* expression (11). Therefore, to test the duration of dsRNA knockout in flies, we recorded *Drs* expression in DD1 flies over a period of 10 days after *dsCACT* injection (Fig. 2). *Drs* expression in *dsCACT* reached a maximum 2–4 days post-dsRNA injection, and gradually decreased on the following days. One day after injection we observed a significant level of *Drs* expression in buffer-treated flies, which we attribute to the septic injury caused by the injection (Fig. 2). These results allowed us to determine the following standard conditions to analyze dsRNA knockout phenotypes: injection of dsRNA into 6-day-old flies and analysis of phenotypes 4 days later.

dsRNA silencing of components of the Toll signaling pathway

We extended our study to other components of the Toll pathway. The serine protease inhibitor (serpin) Spn43Ac is a negative regulator of the extracellular proteolytic cascade, which regulates Toll activation, and mutations in *Spn43Ac* lead to a challenge-independent expression of *Drs* (21). Flies injected with *dsSPN43Ac* displayed phenotypes similar to *Spn43Ac* mutants: constitutive activation of *Drs-GFP* and *Drs* (Fig. 3A and B) and necrotic patches (Fig. 3A, arrow) (21,22). In contrast to *Spn43Ac* mutants, which die within 4 days after hatching, we did not observe any lethality phenotype for the *dsSPN43Ac*-injected flies, indicating that this lethality does not directly result from the constitutive activation of the Toll pathway. These results demonstrate that this extracellular regulator of the Toll pathway was efficiently silenced by dsRNA injection.

As both Cactus and Spn43Ac are negative regulators of the Toll pathway, we further attempted to silence positive regulators, namely *Toll* and *Dif*, as well as the *Drs* gene itself. Flies were injected with specific dsRNAs or buffer as a control, challenged with *M.luteus* and *Drs* induction was

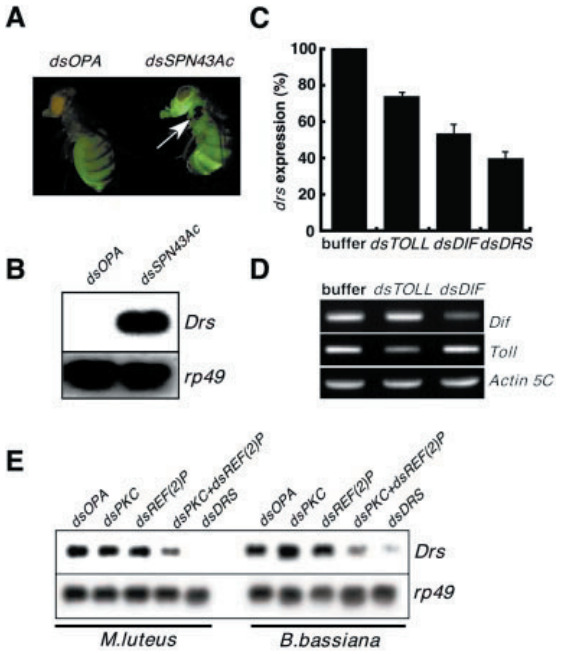


Figure 3. dsRNA silencing of other components of Toll pathway. Four days after *dsOPA* or *dsSPN43Ac* injection into 6-day-old DD1 females. Efficiency of silencing was analyzed by fluorescence microscopy (**A**) or by northern blotting using probes specific for *Drs* and ribosomal protein *rp49* (**B**). The white arrow in (A) points to a necrotic spot appearing after *dsSPN43Ac* injection. (**C**) Level of *Drs* expression 1 day after *M.luteus* challenge of buffer-, *dsTOLL*-, *dsDIF*- or *dsDRS*-injected flies. The graph shows relative percentages of *Drs* expression normalized by *rp49*. (**D**) RT–PCR analysis of dsRNA silencing of *Dif* (*dsDIF*) and *Toll* (*dsTOLL*). *Actin 5C* was used as an internal control. (**E**) Four days after *dsOPA*, *dsPKC*, *dsREF(2)P*, *dsPKC+dsREF(2)P* or *dsDRS* injection, flies were infected with the Gram-positive bacterium *M.luteus* or the fungus *B.bassiana*. The expression of *Drs* and *rp49* was analyzed by northern blotting 1 day after infection with *M.luteus* and 2 days after infection with *B.bassiana.*

monitored. All dsRNA-treated flies showed a decrease in *Drs* expression compared with control flies: 26, 47 and 62% for *dsTOLL*, *dsDIF* and *dsDRS*, respectively (Fig. 3C). The sequence-specific silencing of *Toll* and *Dif* was confirmed by RT–PCR (Fig. 3D). However, this semi-quantitative method did not allow us to determine precisely the efficiency of the knockouts, leaving open the possibility that the differences in the remaining activity of *Drs* are due to incomplete gene silencing.

We next extended our analysis to two new components of the Toll pathway, the atypical protein kinase C (DaPKC) and Ref(2)P, which were recently implicated in the Toll signaling pathway by *in vitro* experiments with a *Drosophila* cell line (14). DaPKC is involved in regulation of many basic developmental functions in *Drosophila* and loss-of-function mutants of this gene are embryonic lethal (15). The *Ref(2)P* gene product is required for male fertility and, later in development, mediates fly resistance to a sigma rhabdovirus (16,23). We first analyzed the effect of the knockout of these genes on the expression of the *Drs* gene after infection with *M.luteus* or *B.bassiana*. Single knockouts of either molecule did not significantly affect *Drs* expression. However, simultaneous silencing of both genes resulted in a reduction of *Drs*

6622 Nucleic Acids Research, 2003, Vol. 31, No. 22

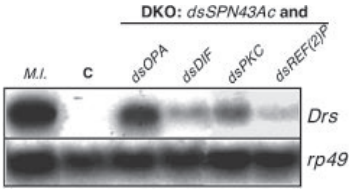


Figure 4. dsRNA-based epistatic analysis in adult flies. *Drs* expression in double knockouts (DKO) was monitored 4 days after concomitant injection of dsRNA corresponding to the serpin *Spn43Ac* gene (*dsSPN43Ac*) in combination with dsRNA against either a control unrelated gene (*dsOPA*), *Dif* (*dsDIF*), the atypical protein kinase C (*dsPKC*) or *Ref(2)P* [*dsREF(2)P*]. Levels of *Drs* expression in the double knockouts were compared with those of buffer-injected flies before (C) or after challenge with *M.luteus* (*M.l.*).

expression after infection by both microorganisms. These results suggest that the single knockouts of either *DaPKC* or *Ref(2)P* genes are not sufficient to block *Drs* induction after infection and that both molecules are required for full pathogen-induced *Drs* expression (Fig. 3E). Our data confirm the results of the earlier *in vitro* studies that place these genes in the Toll pathway.

Epistatic gene analysis using dsRNA

Epistatic analysis provides a powerful tool to dissect signaling pathways. We attempted to use dsRNA silencing to perform such an analysis *in vivo*. For that, we first injected flies with a mixture of *dsSPN43Ac* and *dsDIF*, targeting the negative regulator gene *Spn43Ac* and the downstream transcription factor gene *Dif*. We observed that the challenge-independent induction of *Drs* and the *necrotic* phenotype in the *dsSPN43Ac*-treated flies were abolished by the co-injection of *dsDIF* but not by the co-injection of the *dsOPA* control (Fig. 4 and data not shown).

We then used co-injection of *dsSPN43Ac* and *dsPKC* or *dsREF(2)P* to place these genes in the Toll pathway *in vivo*. Co-injection of *dsSPN43Ac* with either *dsPKC* or *dsREF(2)P* dramatically reduced the level of *Drs-GFP* expression (Fig. 4). However, we noticed that the knockout of *DaPKC* was not as efficient in counterbalancing the *dsSPN43Ac* phenotype as that of *REF(2)P*. These results demonstrate that in the context of the *dsSPN43Ac* knockdown, *DaPKC* and *Ref(2)P* are required for full induction of the Toll pathway *in vivo*. It is unclear why single gene silencing is sufficient to decrease *Drs* expression in the *dsSPN43Ac* background, but not after bacterial or fungal infection (Fig. 4). Since the level of challenge-independent *Drs* expression in *dsSPN43Ac*-treated flies is 40–50% weaker than after infection with Gram-positive bacteria (data not shown), it is conceivable that *dsSPN43Ac*-treated flies represent a more sensitized background for the analysis of *PKC* and *Ref(2)P* genes by dsRNA. In addition, bacterial and fungal infections are known to induce multiple signaling cascades, which can compensate for silencing of each, but not of both, of these genes. In either way, our results show that dsRNA co-injections can be used to dissect signaling pathways to gain a rapid insight into gene function that can be followed by a more laborious analysis of double mutant phenotypes by means of forward genetics.

In conclusion, we demonstrate that the direct injection of dsRNA into adult flies efficiently disrupts gene function, provided that the flies are at least 5 days old, and this effect lasts for at least 10 days. We successfully silenced components of the Toll signaling cascade, representing positive and negative regulators that encode secreted, intracellular or transmembrane proteins. Our results show that dsRNA injection into adult flies allows the analysis of knockout phenotypes for genes that are essential for *Drosophila* development and, as such, are not amenable for forward genetics. We also demonstrate that dsRNA co-injections can be used for rapid epistatic analysis of double knockouts. Finally, our finding that the degree of silencing is dependent on the age of the fly provides important information for development of dsRNA technology in other insect species, in particular vectors of insect-borne diseases.

ACKNOWLEDGEMENTS

We thank Professors Jules A. Hoffmann and Fotis C. Kafatos for encouragement and critical reading of the manuscript. This work was supported by CNRS, the NIH (1PO1 AI44220), the European Commission (HPRN-2000-00080) and Entomed (Strasbourg).

REFERENCES

1. St Johnston,D. (2002) The art and design of genetic screens: *Drosophila melanogaster*. *Nature Rev. Genet.*, **3**, 176–188.
2. Adams,M.D. and Sekelsky,J.J. (2002) From sequence to phenotype: reverse genetics in *Drosophila melanogaster*. *Nature Rev. Genet.*, **3**, 189–198.
3. Hannon,G.J. (2002) RNA interference. *Nature*, **418**, 244–251.
4. Kennerdell,J.R., Yamaguchi,S. and Carthew,R.W. (2002) RNAi is activated during *Drosophila* oocyte maturation in a manner dependent on aubergine and spindle-E. *Genes Dev.*, **16**, 1884–1889.
5. Schmid,A., Schindelholz,B. and Zinn,K. (2002) Combinatorial RNAi: a method for evaluating the functions of gene families in *Drosophila*. *Trends Neurosci.*, **25**, 71–74.
6. Fire,A., Xu,S., Montgomery,M., Kastas,S.A., Driver,S.E. and Mello,C.C. (1998) Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature*, **391**, 806–811.
7. Blandin,S., Moita,L.F., Kocher,T., Wilm,M., Kafatos,F.C. and Levashina,E.A. (2002) Reverse genetics in mosquito *Anopheles gambiae*: targeted disruption of the defensin gene. *EMBO Rep.*, **9**, 852–856.
8. Dzitoyeva,S., Dimitrijevic,N. and Manev,H. (2001) Intra-abdominal injection of double-stranded RNA into anesthetized adult *Drosophila* triggers RNA interference in the central nervous system. *Mol. Psychiatry*, **6**, 665–670.
9. Dzitoyeva,S., Dimitrijevic,N. and Manev,H. (2003) γ-Aminobutyric acid B receptor 1 mediates behavior-imparing actions of alcohol in *Drosophila*: adult RNA interference and pharmacological evidence. *Proc. Natl Acad. Sci. USA*, **100**, 5485–5490.
10. Belvin,M.P. and Anderson,K.V. (1996) A conserved signaling pathway: the *Drosophila* toll-dorsal pathway. *Annu. Rev. Cell. Dev. Biol.*, **12**, 393–416.
11. Lemaitre,B., Nicolas,E., Michaut,L., Reichhart,J.M. and Hoffmann,J.A. (1996) The dorsoventral regulatory gene cassette spatzle/Toll/cactus controls the potent antifungal response in *Drosophila* adults. *Cell*, **86**, 973–983.
12. Hoffmann,J.A. and Reichhart,J.A. (2002) *Drosophila* innate immunity: an evolutionary perspective. *Nature Immunol.*, **3**, 121–126.
13. Tzou,P., De Gregorio,E. and Lemaitre,B. (2002) How *Drosophila* combats microbial infection: a model to study innate immunity and host–pathogen interactions. *Curr. Opin. Microbiol.*, **5**, 102–110.
14. Avila,A., Silverman,N., Diaz-Meco,M.T. and Moscat,J. (2002) The *Drosophila* atypical protein kinase C-Ref(2)P complex constitutes a

conserved module for signaling in the Toll pathway. *Mol. Cell Biol.*, **22**, 8787–8795.
15. Wodarz,A., Ramrath,A., Grimm,A. and Knust,E. (2000) *Drosophila* atypical protein kinase C associates with Bazooka and controls polarity of epithelia and neuroblasts. *J. Cell Biol.*, **150**, 1361–1374.
16. Gay,P. and Contamine,D. (1993) Study of the ref(2)P locus of *Drosophila melanogaster*. II. Genetic studies of the 37DF region. *Mol. Gen. Genet.*, **239**, 361–370.
17. Jung,A., Criqui,M.C., Rutschmann,S., Hoffmann,J.A. and Ferrandon,D.A. (2001) A microfluoremeter assay to measure the expression of β-galactosidase and GFP reporter genes in single *Drosophila* flies. *Biotechniques*, **30**, 594–601.
18. Tauszig-Delamasure,S., Bilak,H., Capovilla,M., Hoffmann,J.A. and Imler,J.L. (2002) *Drosophila* MyD88 is required for the response to fungal and Gram-positive bacterial infections. *Nature Immunol.*, **3**, 91–97.
19. Rutschmann,S., Jung,A.C., Hetru,C., Reichhart,J.M., Hoffmann,J.A. and Ferrandon,D. (2000) The Rel protein DIF mediates the antifungal but not the antibacterial host defense in *Drosophila*. *Immunity*, **12**, 569–580.
20. Tatei,K. and Levine,M. (1995) Specificity of Rel-inhibitor interactions in *Drosophila* embryos. *Mol. Cell Biol.*, **15**, 3627–3634.
21. Levashina,E.A., Langley,E., Green,C., Gubb,D., Ashburner,M., Hoffmann,J.A. and Reichhart,J.M. (1999) Constitutive activation of Toll-mediated antifungal defense in serpin-deficient *Drosophila*. *Science*, **285**, 1917–1919.
22. Green,C., Levashina,E., McKimmie,C., Dafforn,T., Reichhart,J.M. and Gubb,D. (2000) The necrotic gene in *Drosophila* corresponds to one of a cluster of three serpin transcripts mapping at 43A1.2. *Genetics*, **156**, 1117–1127.
23. Wyers,F., Petitjean,A.M., Dru,P., Gay,P. and Contamine,D. (1995) Localization of domains within the *Drosophila* Ref(2)P protein involved in the intracellular control of sigma rhabdovirus multiplication. *J. Virol.*, **69**, 4463–4470.

Chapter 6

TEP1 is a Determinant of Vectorial Capacity in *Anopheles gambiae*

TEP1 is a Determinant of Vectorial Capacity in *A. gambiae*

With all the tools in hand: an efficient dsRNA gene silencing in adult mosquitoes and several assays to analyze knockdown phenotypes, we could finally achieve what we had always been aiming to, i.e. to assess the role of thioester-containing proteins in the mosquito antiparasitic defense, and especially, in the recognition of parasites and activation of effector responses. The first obvious candidate for this analysis was TEP1, which, as we have shown in Chapter 2, resembles its vertebrate complement homologues in recognizing bacteria and labelling them for destruction.

As a laboratory model, we have chosen the infection of *A. gambiae* by *P. berghei*, a rodent malaria parasite, to study mosquito – parasite interactions in a simple and safe system. In addition, to examine the biological principles of mosquito refractoriness, we compared two strains of *A. gambiae*: the susceptible G3 strain and the refractory L3-5 strain, which kills and melanizes many different parasites, including *P. berghei*, after they have traversed the midgut epithelium. Noteworthy, our observation of parasite development in mosquitoes by optical microscopy was greatly facilitated by the use of a GFP-expressing parasite provided by B. Frankle-Fayard (LUMC, Leiden) (Franke-Fayard et al., 2004).

Our results are reported in a manuscript which has recently been accepted for publication in Cell (Blandin et al., 2004c), and that can be found on pages 125-136. We demonstrated that:

• Parasites suffer significant losses during midgut invasion, not only in refractory (R) mosquitoes, but also in fully susceptible (S) mosquitoes. The kinetics of parasite killing and the observation of bubble-like projections in close proximity to dying ookinetes led us to propose (i) that ookinetes are lyzed in the basal labyrinth of the midgut epithelium and that this accounts for major parasite losses associated with midgut invasion in both strains, and (ii) that in contrast to what was generally believed, melanization in R mosquitoes does not a kill parasites, but occurs after parasite death, most likely to isolate dead parasites from surrounding mosquito tissues.

• TEP1 binds to the surface of ookinetes, after they have crossed the midgut epithelium. Interestingly, we observed that TEP1 binding is temporally correlated with the appearance of morphologically degenerate ookinetes. Moreover the vast majority of TEP1-decorated ookinetes were dead. In S mosquitoes, we also detected TEP1 on the surface of oocysts. Later

parasite stages, sporozoites, which have to cross the hemocoel on their way to the salivary glands, are not recognized by TEP1.

• The binding of TEP1 to the ookinetes and oocysts is associated with two peaks of transcriptional upregulation, possibly to replenish the level of TEP1 circulating in the hemolymph.

• The silencing *TEP1* by dsRNA injection results in a five-fold increase in parasite survival in S mosquitoes. In R mosquitoes, *TEP1* knockdown both increases parasite numbers and completely abolishes their melanization, thus converting refractory mosquitoes into susceptible. Our results clearly demonstrate that TEP1 is an essential mosquito factor implicated in parasite killing.

• Two major differences exist between the S and R strains. The first is related to the kinetics of parasite death, which proceeds faster in R mosquitoes, and to the efficiency of ookinete killing, which reaches 100% in R mosquitoes. The second difference pertains to the disposal of dead parasites: lysis in S mosquitoes and lysis and melanization in the encapsulating refractory strain. Parallel to the first difference, we observed that TEP1 binding to and killing of parasites was also faster and more efficient in R than in S mosquitoes.

We examined several reasons to explain why TEP1 was more efficient and faster to bind to ookinetes in R than in S mosquitoes. This may result from:

• A differential pattern of *TEP1* expression in these two strains. However we showed that *TEP1* expression is upregulated in the same manner in S and R mosquitoes upon *Plasmodium* infection.

• Intrinsic properties of different alleles of *TEP1*. As mentioned in Chapter 3, we had proposed that *TEP1* and *TEP16* are two allelic forms of the same gene. Using primers specific for *TEP1* or *TEP16*, we showed that *TEP1*, renamed *TEP1s*, is specifically associated with the susceptible strain, whereas *TEP16*, renamed *TEP1r*, is detected only in refractory mosquitoes. The comparative analysis of the two alleles revealed that most of the substitutions are clustered in the C3d-like region, which contains the thioester site and, in complement factors, directly binds to substrates. Noteworthy, we could evidence a similar segregation of the *TEP1* alleles in another set of refractory and susceptible strains, independently selected in the group of Hilary Hurd (Keele University). We speculate that the polymorphism detected between TEP1r and TEP1s might account, at least partially, for the more efficient parasite killing in R mosquitoes.

In conclusion, we demonstrated that the hemocyte-specific TEP1 is directly involved in the control of parasite loads in mosquitoes and identified ookinetes as a parasite stage sensitive to TEP1-dependent killing. These results document the important role of mosquito immune responses, especially those orchestrated by hemocytes, in the establishment of vectorial capacity and transmission of malaria.

Article 5

Complement-like Protein TEP1 is a Determinant of Vectorial Capacity in the Malaria Vector *Anopheles gambiae*

Stéphanie BLANDIN, Shin-Hong SHIAO,
Luis F. MOITA, Chris J. JANSE, Andrew P. WATERS,
Fotis C. KAFATOS, Elena A. LEVASHINA

Cell 2004, 116:661-670

Cell, Vol. 116, 661–670, March 5, 2004, Copyright ©2004 by Cell Press

Complement-Like Protein TEP1 Is a Determinant of Vectorial Capacity in the Malaria Vector *Anopheles gambiae*

Stephanie Blandin,[1] Shin-Hong Shiao,[2] Luis F. Moita,[1,4] Chris J. Janse,[3] Andrew P. Waters,[3] Fotis C. Kafatos,[1] and Elena A. Levashina[1,2,*]
[1]European Molecular Biology Laboratory
Meyerhofstrasse 1
69117 Heidelberg
Germany
[2]UPR 9022 du CNRS
Institut de Biologie Moléculaire et Cellulaire
15 rue René Descartes
67084 Strasbourg Cedex
France
[3]Department of Parasitology
Leiden University Medical Centre
Postbus 9600 RC
Leiden
The Netherlands

Summary

***Anopheles* mosquitoes are major vectors of human malaria in Africa. Large variation exists in the ability of mosquitoes to serve as vectors and to transmit malaria parasites, but the molecular mechanisms that determine vectorial capacity remain poorly understood. We report that the hemocyte-specific complement-like protein TEP1 from the mosquito *Anopheles gambiae* binds to and mediates killing of midgut stages of the rodent malaria parasite *Plasmodium berghei*. The dsRNA knockdown of *TEP1* in adults completely abolishes melanotic refractoriness in a genetically selected refractory strain. Moreover, in susceptible mosquitoes this knockdown increases the number of developing parasites. Our results suggest that the TEP1-dependent parasite killing is followed by a TEP1-independent clearance of dead parasites by lysis and/or melanization. Further elucidation of the molecular mechanisms of TEP1-mediated parasite killing will be of great importance for our understanding of the principles of vectorial capacity in insects.**

Introduction

Human malaria is one of the most devastating diseases, especially in Sub-Saharan Africa. Every year, 300 to 500 million people suffer from this disease and more than one million die, mostly children under the age of 5 (Snow et al., 1999). The causative agents of malaria are protozoan parasites of the genus *Plasmodium*, which are transmitted to an *Anopheles* mosquito vector when the female mosquito takes an infected bloodmeal. To become infective to the next host, the parasite must undergo a complex developmental cycle in the mosquito. Within the insect midgut, *Plasmodium* gametocytes are rapidly activated to produce gametes. Fertilization leads to the formation of zygotes which, 16–20 hr later, transform into motile ookinetes. Approximately 24 hr after an infectious bloodmeal, the ookinetes invade and cross the midgut epithelium, reaching the basal side of the midgut where they form protected capsules called oocysts. In the next 10 days, within each oocyst, a meiotic cycle and several ongoing rounds of mitosis produce thousands of haploid sporozoites. After maturation (14–16 days after the infectious bloodmeal), sporozoites are released into the mosquito hemocoel and migrate to the salivary glands which they invade. The parasite cycle in the mosquito is completed when this vector injects infective sporozoites into a new vertebrate host during a subsequent blood feeding.

Most mosquito species are not permissive for *Plasmodium* parasites and only a limited number of dedicated *Plasmodium–Anopheles* combinations cause malaria in a particular group of vertebrates. Thus, mosquito-parasite interactions constitute a critical aspect of disease transmission and are a potential target for efforts to control malaria. *A. gambiae* is the most important vector of human malaria in much of Africa. We have chosen as a laboratory model the infection of *A. gambiae* by *P. berghei*, a rodent malaria parasite, to study mosquito-parasite interactions and thus comprehend the biological principles of mosquito antiparasitic responses in a simple and safe system.

It has long been known that within the same species, mosquitoes display genetic variation in their susceptibility to parasites. Strains refractory to various malarias have been selected and studied, mostly morphologically (reviewed in Vargas, 1949). In the present study, we made use of a refractory strain of *A. gambiae* (R), which completely aborts development of a number of malaria parasites, including the simian parasite *P. cynomolgi* B. Refractoriness is manifested by melanotic encapsulation of the ookinete, after it completes its passage through the mosquito midgut (Collins et al., 1986). The genetic control of refractoriness appears to be complex. It not only involves several quantitative trait loci, but also the relative contribution of each locus to oocyst encapsulation varies with the species of parasites (Zheng et al., 1997, 2003). A different refractory mechanism resulting in complete lysis of *P. gallinaceum* ookinetes in the midgut was reported in *A. gambiae* (Vernick et al., 1995). To date, the molecular mechanisms underlying both types of refractoriness, and more generally parasite recognition and killing, are not well understood.

We are interested in the mechanisms of parasite recognition and have chosen a family of *t*hioester-containing *p*roteins (TEPs) as potential candidate recognition molecules. In vertebrates, members of this family comprise the universal protease inhibitors α_2-macroglobulins, and the complement factors C3/C4/C5, which are involved in labeling pathogens and triggering their disposal through phagocytosis or cell lysis. A family of 19 TEPs has been identified in the genome of *A. gambiae* (Christophides et al., 2002) and one of these, *TEP1*, has

*Correspondence: e.levashina@ibmc.u-strasbg.fr
[4]Present address: Center for Immunology and Inflammatory Diseases, Massachusetts General Hospital, 149 13th Street, Charlestown, Massachusetts 02129.

Cell
662

been studied in detail (Levashina et al., 2001). TEP1 is an acute phase glycoprotein secreted by mosquito hemocytes into the hemolymph. Similarly to its vertebrate complement homologs, it is cleaved shortly after septic injury. Moreover, the cleaved C-terminal part of the protein binds to the surfaces of gram-positive and gram-negative bacteria through the conserved thioester bond and labels gram-negative bacteria for clearance by phagocytosis in vitro.

The properties of TEP1 to recognize microorganisms and target them for destruction led us to investigate the possibility, and obtain molecular evidence, that TEP1 is one of the mosquito factors that determine vectorial capacity in *A. gambiae*. In this study, we make use of a GFP-expressing strain of *P. berghei* and demonstrate that parasite killing in both susceptible and refractory mosquitoes is mediated by direct binding of the TEP1 protein to the surface of ookinetes. The essential role of TEP1 in parasite killing is further supported by double-stranded RNA (dsRNA) knockdown experiments: in susceptible mosquitoes, the knockdown of *TEP1* results in a 5-fold increase in the number of oocysts developing on the midgut, and in the refractory strain the knockdown completely abolishes parasite melanization, thus converting refractory mosquitoes into susceptible. We propose a model for TEP1 function, stressing the differences between susceptible and refractory mosquitoes regarding the kinetics of TEP1 binding, parasite killing, and clearance. Furthermore, we show that *TEP1* is encoded by two distinct alleles, *TEP1s* and *TEP1r* that appear to be specific to the susceptible and the melanotically encapsulating refractory strain, respectively. These results document the important role of mosquito immune responses, especially those orchestrated by hemocytes, in the establishment of vectorial capacity and transmission of malaria. Further elucidation of the molecular mechanisms of TEP1 killing will provide important insights toward development of antimalarial strategies.

Results

Binding of TEP1 to *P. berghei*

In vertebrates, complement factors bind covalently to target surfaces, opsonizing them for phagocytosis or initiating the formation of a lytic membrane attack complex (for review, see Carroll and Fischer, 1997; Law and Dodds, 1997). TEP1 binds to the surface of both gram+ and gram− bacteria in a thioester-dependent manner and promotes phagocytosis of bacteria in vitro and in vivo (Levashina et al., 2001; L.F.M., S.B., F.C.K., and E.A.L., unpublished data). To investigate whether TEP1 also recognizes *Plasmodium* parasites, we first performed immunohistochemical analysis of susceptible (S) mosquito midgut tissues infected with *P. berghei*, using rabbit polyclonal antibodies directed against the C-terminal fragment of TEP1 (Levashina et al., 2001). *A. gambiae* females of the S strain were infected with *P. berghei*, dissected at selected time points and analyzed by confocal microscopy. No TEP1-positive signal was observed in the midgut epithelial cells at any time. Instead, specific expression of *TEP1* was detected in the mosquito hemocytes attached to the midgut and Malpighian tubules. TEP1 staining in hemocytes was more pronounced at 24 and 48 hr postinfection (hpi) as compared to uninfected controls (Levashina et al., 2001 and data not shown). Importantly, starting at 24 hpi, TEP1 appeared on the surface of some ookinetes (Figures 1A–1G), as evidenced by colocalization of the signals for TEP1 and for the ookinete surface protein P28 (Simonetti et al., 1993). The number of ookinetes positive for TEP1 increased with time and attained a maximum at approximately 48 hpi (data not shown), when the surviving parasites have completed their migration through the midgut epithelium and rest in the extracellular space, between the membrane folds of the basal labyrinth of the midgut cells, facing the basal lamina. We have never observed TEP1 labeling of all ookinetes in the susceptible strain. Between 22 and 48 hpi, parasites smaller than normal and strongly labeled with TEP1 were often detected. These parasites were usually associated with P28-positive blobs and lacked nuclear staining (Figures

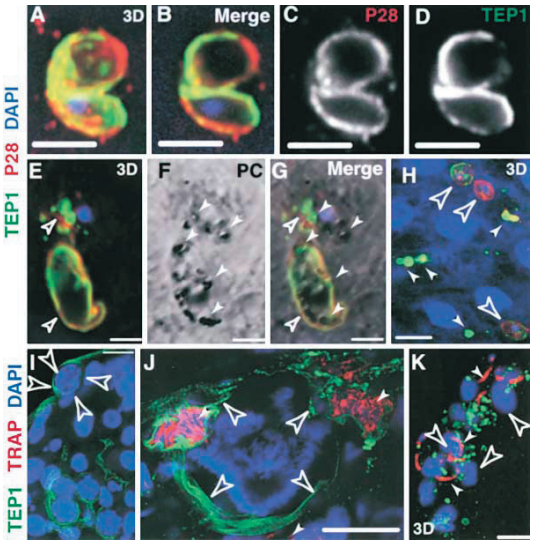


Figure 1. Binding of TEP1 to *P. berghei* Parasites in Susceptible Mosquitoes

Confocal sections (B-D, I–J) and 3D reconstructions (A, E–H, and K) of midgut tissues from susceptible mosquitoes at different time points after infection. Triple staining showing TEP1 in green, a parasite surface protein in red (P28 in A–H or TRAP in J and K) and nucleic acids in blue.

(A–D) TEP1 was detected on the surface of ookinetes 22 hpi. The colocalization of P28 (C, red channel) and TEP1 (D, green channel) is evidenced in the merged image (B).

(E–G) TEP1-labeled ookinetes display a lytic phenotype 24 hpi. Bubbles are of parasite origin as evidenced by the presence of P28 (E, open arrowheads) and hemozoine granules (F, phase contrast image, filled arrowheads) in the merged image (G).

(H) Two days pi, ookinetes (filled arrowheads) are heavily labeled with TEP1, whereas young oocysts (open arrowheads) display weak or no opsonization.

(I–K) Later in development, oocysts are covered by TEP1 (open arrowheads in I and J at 11 and 16 dpi, respectively). Note that sporozoites (filled arrowheads in J and K) are already visible in matured oocysts (J). No TEP1 is detected on sporozoites (K). Rare evidence of contact between mosquito hemocytes (open arrowheads) and sporozoites (K). Scale bars in μm: (A–G), 2; (H–J), 10; (K), 5.

1E–1H). This morphology suggested that these parasites might have been killed during the invasion process.

We have also examined whether TEP1 binds to later stages of parasites. Shortly after reaching the basal lamina, ookinetes round up and transform into young oocysts. At 48 hpi, nondeveloping and presumably dead, ookinetes were heavily labeled with TEP1, whereas developing oocysts showed only weak and patchy TEP1 labeling (Figure 1H, filled and outlined arrowheads, respectively). However, at 11 days postinfection (dpi) well-developed oocysts were covered by abundant TEP1-positive material, and the morphology suggested that TEP1 was associated with components of the basal lamina in which the oocysts are embedded (Figures 1I and 1J). In contrast, sporozoites that develop from oocysts and are stained with specific antibodies against the *t*hrombospondin-*r*elated *a*nonymous *p*rotein (TRAP) never displayed TEP1 staining on the surface, indicating that TEP1 does not directly interact with this haploid form of the parasite (Figures 1J and 1K). Even in exceptional cases when clusters of mosquito hemocytes were in contact with sporozoites, no colocalization was detected between the TEP1 and TRAP signals (Figure 1K).

We next examined TEP1 binding to *P. berghei* in the refractory (R) strain. As in S mosquitoes, TEP1 was found to bind to ookinetes in a time-dependent manner, but clearly with faster kinetics: most of the parasites were labeled with TEP1 as early as 24 hpi (Figures 2C and 3E). The binding of TEP1 to ookinetes was further confirmed by electron microscopy (Figures 2A and 2B). It often correlated with perturbations in the P28 signal on ookinete surface, parasite blebbing, condensation and degeneration of nuclei and, in extreme cases, localization of P28 inside the degenerated parasites (Figures 2D–2G). At later time points, all detectable parasites in the R mosquitoes were melanized and, therefore, opaque. The early and extensive TEP1 binding to ookinetes in the basal labyrinth correlates with ookinete melanization, which begins at the same time and location (24–36 hpi). We explored the potential role of TEP1 in melanization in more detail by confocal microscopy using specific antibodies against TEP1 and against the conserved copper binding domain of PPO6 (Müller et al., 1999). A PPO-positive signal was detected in the cytoplasm of hemocytes but not in the midgut cells, indicating that PPOs, like TEP1, are produced by the mosquito blood cells and are released into the hemolymph (data not shown). The hemocytes often coexpress PPOs and TEP1 (data not shown). Importantly, the PPO signal was detected only on the surface of TEP1-labeled ookinetes and yet, no precise colocalization between TEP1 and PPOs was observed (Figures 2F–2I). This suggests that TEP1 is associated with ookinete melanization but is not directly involved in the tethering of the PPO complex on the parasite surface. The PPO-positive layer was later surrounded by a second layer of TEP1, reminiscent of the TEP1 deposition on oocysts described for S mosquitoes (compare Figures 2H and 1I).

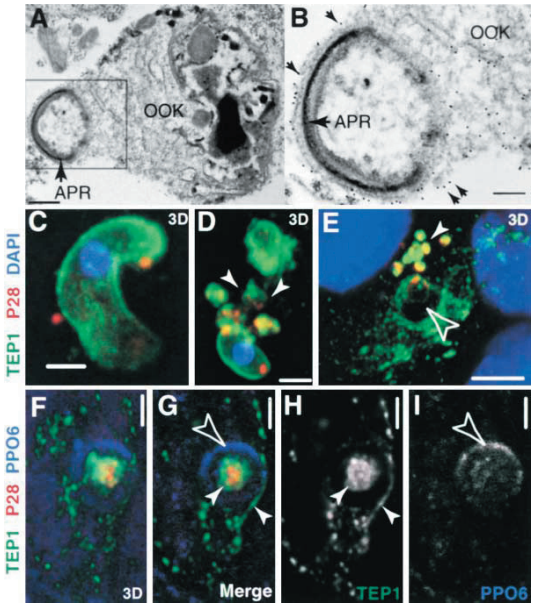


Figure 2. Binding of TEP1 to *P. berghei* Parasites in Refractory Mosquitoes

(A) The electron micrograph (30 hpi) shows the apical part of an ookinete (OOK), and its intracellular apical polar ring (APR).

(B) A close up view of the boxed area in (A). Binding of TEP1 to ookinetes is evidenced by gold particles on the parasite surface (arrows).

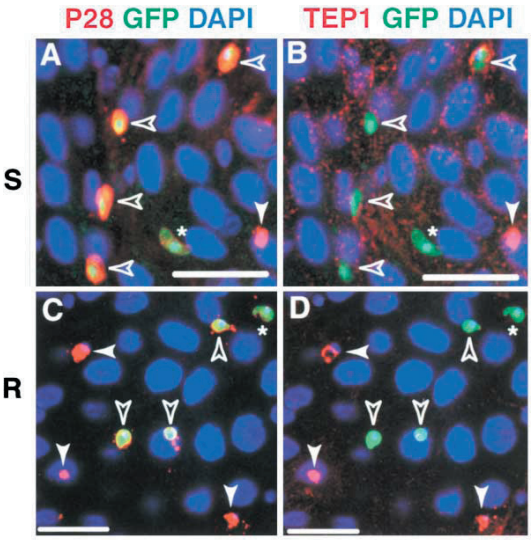
(C–F) 3D reconstructions and (G–I) confocal sections of midgut tissues at different time points after infection. Triple staining as in Figure 1 with P28 in red. Ookinetes labeled with TEP1 (C, 24 hpi) display a lytic phenotype as in S mosquitoes (compare small arrowheads in D, E, 32 hpi, and Figure 1E). Melanized ookinetes are covered with a second layer of TEP1 (E, big arrowhead, and data not shown).

(F–I) The parasite is surrounded by concentric layers of TEP1, PPO, and TEP1 (44 hpi). Note the absence of colocalization between TEP1 (H, green channel, small arrowhead) and PPOs (I, blue channel, open arrowhead) in the merged image (G). Scale bars in μm: (A), 0.5; (B), 0.2; (C, D, F–I), 2; (E), 5.

Parasite Losses during Midgut Invasion

Our results indicate that TEP1 is able to recognize and bind to specific developmental stages of *P. berghei*. Potentially, this binding might serve two opposing functions: (1) help parasite invasion or, (2) limit parasite development. To distinguish between these two hypotheses, we monitored quantitatively the development of *P. berghei* and the kinetics of TEP1 binding in both S and R mosquitoes, using a transgenic parasite strain $PbGFP_{CON}$ that expresses GFP under control of the *eIF1* gene promoter throughout its life cycle in the mosquito (Alavi et al., 2003). Female mosquitoes were blood-fed on $PbGFP_{CON}$ infected mice and dissected 24 and 32 hr later for analysis by multichannel confocal microscopy. In these experiments, in addition to GFP fluorescence, we used P28 antibodies to detect parasites and TEP1 antibodies to follow their opsonization.

Surprisingly, the only universal (although occasionally weak) marker for the parasites was P28 staining (Figures 3A and 3C). Only a minority of ookinetes expressed GFP at 24 hpi, both in S and especially in R mosquitoes (Figure 3E, 38% and 14%, respectively). The number of GFP-positive ookinetes decreased with time and, at 32 hpi, it was down to 20% and 3% in S and R mosquitoes,



E

	GFP+		TEP1+		GFP/TEP1+		Melanized	
	S	R	S	R	S	R	S	R
24 h	37.9	13.9	57.2	81.1	5.0	1.4	0	3.6
32 h	20.1	3.1	75.9	85.7	4.0	2.0	0	9.1

Figure 3. Parasite Killing in the Mosquito Midgut

Double staining of midgut tissues of susceptible (A–B) and refractory (C–D) mosquitoes 24 hpi with the GFP-parasite: (A, C): P28 in red; GFP in green; (B, D): TEP1 in red; GFP in green. Nucleic acids are in blue. Scale bars are equal to 20 μm. Parasites showing a strong P28 staining and GFP fluorescence are indicated with open arrowheads. Filled arrowheads point to GFP-negative parasites. Parasites displaying a weak P28 staining are indicated with asterisks. Most of the TEP1-labeled ookinetes have lost their GFP staining (B and D, filled arrowheads). Table (E) shows percentages of ookinetes positive for GFP (GFP+), for TEP1 (TEP1+), for both (GFP/TEP1+) or melanized, in S and R mosquitoes at 24 and 32 hpi. The results of three independent experiments were highly consistent (1000–3000 ookinetes were counted for each time point and mosquito strain) and one representative experiment is shown.

respectively. Conversely, TEP1-decorated parasites were always a majority, their numbers increased over time, and they were always more numerous in R than in S mosquitoes: at 24 hpi 81% versus 57%, respectively, and at 32 hpi 86% versus 76% (Figure 3E). Notable observations were that GFP-negative ookinetes were heavily labeled with TEP1 (Figures 3B and 3D, filled arrowheads), and that many of them had also lost nuclear staining, suggesting that the lack of GFP fluorescence reflects parasite death. Only very rarely did TEP1-positive ookinetes show GFP fluorescence (Figure 3E, no higher than 5%); possibly, these double positives represent an early stage in a process that begins with TEP1 binding and proceeds with loss of GFP fluorescence, nuclear disintegration, and parasite elimination. The key observation was that binding of TEP1, loss of parasite fluorescence, and parasite death proceeded in parallel, faster in R than in S mosquitoes. All but about 5% of the parasites were killed by 32 hpi in the refractory strain (as indicated by lack of GFP fluorescence or melanization), whereas in the susceptible strain approximately 24% of the parasites survived.

TEP1 Limits Parasite Development

If TEP1 binding is essential for parasite killing, its absence should result in higher oocyst numbers in the mosquito midguts. The in vivo expression of *TEP1* was silenced by injecting dsRNA corresponding to *TEP1* into the thorax of young susceptible females; controls received dsRNA of unrelated genes. The efficacy of the knockdown (KD) was confirmed by immunoblotting: the 165 kDa full-length and 80 kDa cleaved forms of TEP1 (Levashina et al., 2001) were detected in hemolymph samples of control but not *dsTEP1*-treated mosquitoes (Figure 4A). Four days after dsRNA injection, mosquitoes were fed on GFP-parasite infected mice, and 24 hr later immunohistochemistry further confirmed the efficacy of the KD, as no TEP1 labeling of parasites was detected in *dsTEP1*-treated mosquitoes (compare Figures 4B and 4C). The absence of TEP1 was correlated with substantially higher levels of ookinete survival evidenced by GFP expression (see below). Careful analysis detected a number of parasites that were negative for GFP as well as TEP1, suggesting that more than one mechanism controls parasite loads in *A. gambiae* (Figure 4C, arrowheads).

In control experiments, mosquitoes were injected with either *dsGFP* or *dsLacZ* RNAs; the respective mean numbers of fluorescent oocysts were indistinguishable (data not shown), indicating that the expression of a parasite-borne gene (*GFP* in this case) cannot be silenced when dsRNA is delivered to mosquitoes. In three independent experiments performed on susceptible mosquitoes, we observed a 5-fold increase in the mean number of oocysts in *dsTEP1* mosquitoes as compared to *dsLacZ* controls, as well as a uniquely large class of superinfected mosquitoes with >300 oocysts per midgut (Figures 4D, 4E, and 4P). These results clearly demonstrate that *A. gambiae* senses the malaria parasites and actively limits their development in a TEP1-dependent manner, even in the susceptible strain. The parasites that remained GFP-labeled successfully developed into oocysts and produced infectious sporozoites that efficiently invaded the mosquito salivary glands (Figures 4F and 4G).

We have also silenced *TEP1* expression in R mosquitoes by injecting L3-5 females with *dsTEP1*, and with *dsGFP* or *dsLacZ* as controls. The efficacy of the knockdown was again confirmed by immunoblotting (Figure 4A) and by immunohistochemical analysis (Figures 4H and 4I), which demonstrated that *dsTEP1* completely depleted TEP1 from the hemolymph, resulting in the absence of parasite opsonization. Melanized ookinetes and young oocysts were detected only in the midguts of control mosquitoes starting from 24 hpi and persisting throughout the lifetime, demonstrating that after melanization parasites remain in the midgut tissues until the end of mosquito life (Figures 4J, 4K, and data not shown). Strikingly, we did not find any single melanized parasite in *dsTEP1*-treated R mosquitoes; instead, green fluorescent oocysts developed normally and released sporozoites (Figures 4L–4N). As all parasites are melanized in control R mosquitoes, we compared the number of sur-

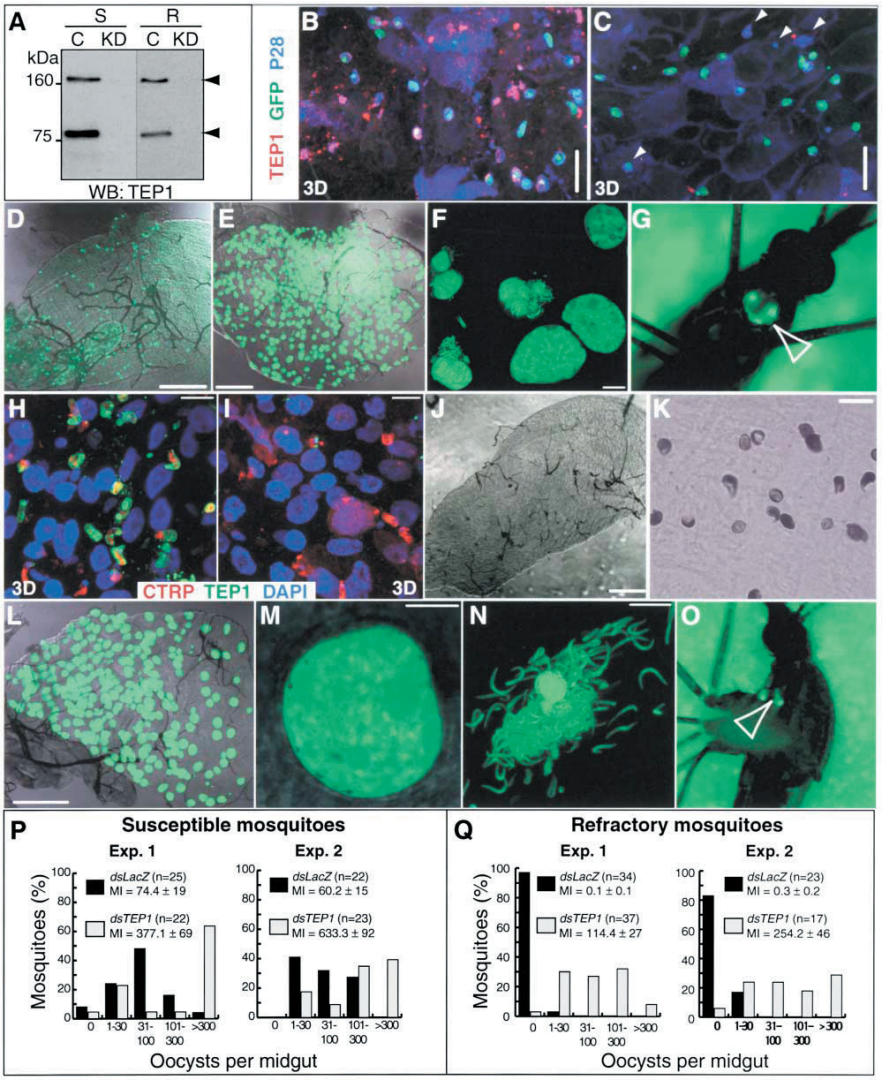


Figure 4. Role of TEP1 in the Mosquito Antiparasitic Response

(A) dsRNA knockdown of *TEP1* in adult susceptible (S) and refractory (R) females. The full-length and processed forms of TEP1 (arrowheads) were detected by immunoblotting using 7% SDS PAGE in the hemolymph of control mosquitoes treated with *dsGFP* (S and R, C) but not in the hemolymph of *dsTEP1*-treated mosquitoes (S and R, KD).

(B–O) Fluorescence microscopy of midguts and salivary glands of susceptible (B–G) and refractory (H–O) mosquitoes infected with the GFP-parasite. Phase contrast and GFP fluorescence are merged in (D–E) and (J–M), only the GFP fluorescence is displayed in (F–G) and (N–O).

(B–C) Double staining of midgut tissues of S mosquitoes treated with *dsLacZ* (B) and *dsTEP1* (C) 32 hpi showing the absence of TEP1 in *dsTEP1* mosquitoes: TEP1 in red; GFP in green; P28 in blue. Note the presence of parasites that do not express GFP in *dsTEP1* mosquitoes (C, arrowheads).

(D–G) Development of GFP-parasites in *dsLacZ* (D) and *dsTEP1* (E–G) mosquitoes. Note higher infection rates in *dsTEP1* midguts 11 dpi (E). In *dsTEP1* mosquitoes, oocysts develop normally (F, 11 dpi) and produce infective sporozoites that invade salivary glands (arrowhead in G, 21 dpi).

(H–I) Triple staining of midgut tissues of R mosquitoes 24 hpi (CTRP in red; TEP1 in green; nucleic acids in blue) demonstrating the absence of TEP1 in *dsTEP1*-treated mosquitoes.

(J–K) Ookinetes melanization in refractory mosquitoes 48 hpi.

(L–O) The knockdown of *TEP1* in R mosquitoes completely abolishes the melanization phenotype. Ookinetes successfully transform into oocysts (L, M, 10 dpi) and further mature into sporozoites (N, 10 dpi) that invade salivary glands (arrowhead in O, 21 dpi).

(P–Q) Frequency distribution of oocysts in mosquito midguts after dsRNA knockdown in susceptible (P) and refractory (Q) mosquitoes. Control (*dsLacZ*) and *dsTEP1*-treated (*dsTEP1*) mosquitoes were infected with the GFP-parasite, dissected 10 days later, and the number of oocysts on each midgut was counted. Results of two out of four independent experiments for each mosquito strain are shown. n, number of mosquitoes per experiment, MI, mean intensity of infection plus/minus standard error.

Scale bars in μm: (D, E, J, L), 200; (B, C, F, M, N), 20; (H, I, K), 10.

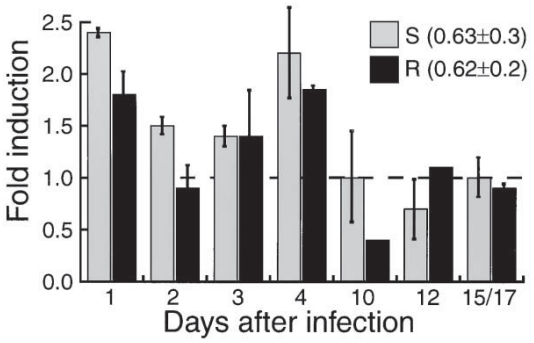


Figure 5. Inducibility of *TEP1* Expression upon *P. berghei* Infection

Quantitative real-time PCR using specific primers and probes for *TEP1s* in susceptible (S) and *TEP1r* in refractory (R) mosquitoes. The levels of *TEP1* transcript at various time points after infection were normalized to the internal control transcript for ribosomal protein S7. Results are shown as fold induction relative to the 6 hpi control. Absolute values for this time point are shown in brackets in the upper right. For most of the time points, experiments were performed 3 to 4 times (error bars indicate standard errors). Results for 10 and 12 dpi in R mosquitoes are from one experiment.

viving parasites in *dsTEP1* R mosquitoes to that of control S mosquitoes. In three independent experiments, the number of developing parasites in R mosquitoes after *dsTEP1* treatment was 2- to 3-fold higher than in *dsLacZ* S mosquitoes. However, in contrast to the results of *dsTEP1* knockdown in S mosquitoes, it never reached the 5-fold increase and superinfected mosquitoes with >300 oocysts per midgut did not represent a major class in *dsTEP1* R mosquitoes (Figures 4P and 4Q), indicating that the melanotic refractoriness is a complex phenomenon which requires coordinated action of a number of genes, including *TEP1*. Mature sporozoites successfully invaded the salivary glands (Figure 4O) and were infective to a mouse (data not shown). To demonstrate that the surviving parasites did not represent a revertant clone that escaped melanization, we fed a new batch of naive (not *dsTEP1*-treated) R mosquitoes on the same mouse that had been infected with the surviving parasites. All parasites in these mosquitoes were melanized (data not shown), indicating that the normal development of parasites in the knockdown R mosquitoes was indeed due to silencing of a single gene, *TEP1*.

Expression of *TEP1* Is Upregulated after Parasite Infection

Distinct kinetics and efficiency of parasite killing in the S and R strains may result from a differential pattern of *TEP1* expression in these two strains. We compared the transcriptional profiles of *TEP1* by quantitative real-time PCR in whole S and R females at selected time points after infection (Figure 5). In these experiments, the expression level of *TEP1* was normalized to that of the gene encoding the ribosomal protein rpS7. At 6 hpi, no parasite-specific interactions take place and mosquitoes react mostly to the physiological changes induced by a bloodmeal (Dimopoulos et al., 2002). Therefore, the normalized expression of *TEP1* at 6 hpi was used as a reference to determine any later induction of *TEP1*.

Transcription of *TEP1* followed an overall similar pattern in S and R mosquitoes. The relative level of TEP1 upregulation in S mosquitoes was somewhat higher than in R, judging by the absolute levels of expression at 6 hpi that were similar in these two strains (0.63 ± 0.3 and 0.62 ± 0.2, respectively). The expression of *TEP1* was rapidly upregulated at 24 hpi, when it reached a maximum of 2.5- and 1.8-fold in the S and R strains, respectively. This was followed by a temporary depression in transcript abundance and then by a second peak of 2.2- and 1.8-fold induction at 4 dpi in S and R mosquitoes, respectively. Thus, the peak upregulation of *TEP1* expression coincided with two important steps in the development of *P. berghei* in *A. gambiae*: the passage through the columnar cells of the midgut epithelium (24 hpi) and the establishment of oocysts (beginning on day 3–4 pi, when the oocysts become surrounded by the basal lamina; Meis et al., 1989). No upregulation of *TEP1* expression was detected at the time points associated with maturation, release, and massive migration of sporozoites from the midgut oocysts to the salivary glands (Figure 5, days 10, 12, and 15/17 pi). These results suggest that *TEP1* upregulation is a mosquito response to the presence of *Plasmodium* during key early steps of infection, both in the S and R strains.

Genomic Polymorphism of the *TEP1* Gene Correlates with Susceptible and Refractory Strains of *A. gambiae*

We have previously proposed that *TEP1* and *TEP16* are two allelic forms of the same gene and that initial annotation of *TEP16* as a distinct gene was an artifact of the automatic genome assembly (Christophides et al., 2002). To explore whether *TEP1* polymorphism could be correlated with the S and R phenotypes, *TEP1*- and *TEP16*-specific PCR primers were designed to follow the presence of the two alleles in these strains. Primers specific for *TEP1* amplified the expected fragment in S but not in R mosquitoes. Conversely, *TEP16*-specific primers produced an amplicon in R but not S mosquitoes (Figure 6D). The results of reciprocal crosses between S and R mosquitoes confirmed that these strains are homozygous for one of the two alternative alleles, *TEP1* or *TEP16*, respectively: in contrast to the homozygous parents (S and R), individual mosquitoes of the F_1 progeny were positive for both forms (Figure 6D). These results suggest that *TEP1* and *TEP16* are two alleles of the same gene, which by precedence will be called *TEP1*. We propose to rename the allele associated with the S strain, *TEP1s* (formerly *TEP1*) and the allele associated with the R strain, *TEP1r* (formerly *TEP16*).

The overall identity and similarity between the TEP1s and TEP1r deduced proteins are 92% and 95%, respectively. Interestingly, the differences between these two isoforms are very unevenly distributed and are mostly concentrated in one region (Figure 6A region IV, dots for single aa substitutions and triangles for clusters of at least 5 substitutions per 50 residues).

The highly polymorphic region IV (Figure 6A) corresponds to the functionally important C3d-like domain, the structure of which was previously analyzed in TEP1

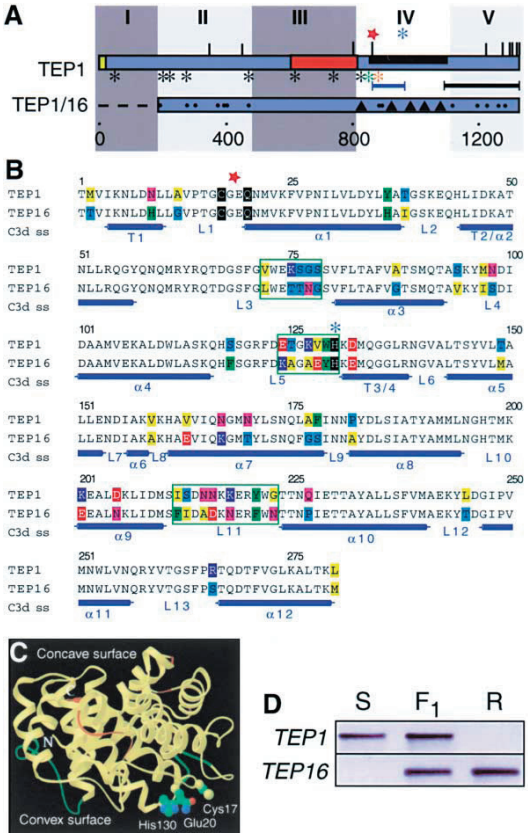


Figure 6. Sequence and Structure Comparison of Two Allelic Forms of *TEP1*

(A) Schematic representation of TEP1 structure (TEP1) and sequence comparison between TEP1 and TEP16 (TEP1/16). Vertical bars indicate cysteines, the red star points to the internal thioester site and the blue asterisk to the catalytic histidine. Black, green, and orange asterisks represent conserved, TEP1-specific and TEP16-specific putative glycosylation sites, respectively. The color-filled fragments of TEP1 indicate: yellow, signal peptide; red, a region containing putative protease cleavage sites; black half-filled segment, the C3d-like region. The black horizontal bar shows the TEP1 fragment that was used to produce polyclonal antibodies and for dsRNA synthesis, the blue horizontal bar shows the fragment amplified by PCR using *TEP1* and *TEP16*-specific primers. The TEP1/TEP16 scheme shows major differences: circles for single modifications and triangles for clusters of at least 5 modifications per 50 amino acids (aa). Shaded boxes (I–V) represent regions differing in the extent of sequence conservation. The absent N-terminal sequence in TEP16 is indicated by a dashed line. Numbers correspond to aa positions in TEP1.

(B) Sequence comparison of the C3d-like region of TEP1 and TEP16. Differences between the two sequences are colored according to the aa properties. Residues of the active site are shaded in black and are indicated with a red star and blue asterisk, as in (A). Extended clusters of modifications are boxed in green. Secondary structures (turns, T, α helices, α and loops, L) are indicated below (C3d ss). Numbers correspond to aa positions in the C3d-like domain.

(C) Modeled 3D structure of the C3d-like region of TEP1 (adapted from Levashina et al., 2001). Differences between TEP1 and human C3d are indicated in brown, between TEP1 and TEP16, in green. The side chains of the active site residues are shown as ball-and-stick models.(D) PCR amplification of *TEP1* and *TEP16* in the susceptible and refractory strains. Genomic DNA from parents (S and R) and progeny (F1) of reciprocal crosses between susceptible G3 (S) and refractory L3-5 (R) mosquitoes was amplified using *TEP1*- and *TEP16*-specific primers.

by homology-based modeling (Levashina et al., 2001). The key TEP1 features are conserved in both allelic forms: (1) the canonical thioester (TE) motif (red star); (2) the catalytic histidine (blue asterisk); and (3) the cysteine residues, including a cluster of 6 cysteines at the C terminus (vertical bars). According to the model, the majority of modifications were observed in three loops (L) between putative α helices and turns (α and T blue cylinders, respectively) located on the convex side of the molecule, with loops 3 and 5 nested in close proximity to the thioester site (Figures 6B and 6C, green boxes and green loops, respectively). It is conceivable that substitutions in the thioester environment might affect the reactivity of the thioester bond leading to the different kinetics of TEP1 binding to parasites that we have detected in the S and R strains.

Discussion

In this study, we demonstrate that the knockdown of the *TEP1* gene is sufficient to enhance substantially the number of oocysts in susceptible mosquitoes, and also to convert refractory into highly susceptible mosquitoes. We used a combination of morphological, immunochemical, and fluorescence markers to demonstrate that parasites suffer significant losses even in fully susceptible mosquitoes. A substantial fraction of the ookinetes is killed between 24 and 48 hpi, after they have accomplished the passage through the columnar cells of the midgut and lie in the basal labyrinth facing the basal lamina. Live ookinetes that express the *eIF1*-driven *GFP* transgene have smooth, regular edges, distinct nuclei, and generally, an even distribution of the P28 surface protein. In contrast, dead or degenerating ookinetes, identified by their complete loss of *GFP* expression, are characterized by an irregular surface, the diminution of nuclear size or even nuclear loss. They often disintegrate, as evidenced by the appearance of bubble-like projections, patchy surface, or intracellular distribution of P28. Association of ookinete death with extensive blebbing is suggestive of lytic destruction. A similar lysis phenotype of *P. gallinaceum* ookinetes was reported in *A. gambiae* by Vernick and coworkers (1995), although in this case lysis was only observed in the selected refractory strain, unlike in our experiments that identified lysis of *P. berghei* parasites in susceptible as well as refractory mosquitoes. Our observations parallel earlier morphological studies on *Culex pipiens* and *P. cathemerium* by Huff in 1934 (reviewed in Vargas, 1949), where normal and degenerate ookinetes were reported in susceptible and insusceptible mosquitoes alike. Thus, we propose that ookinete lysis in the basal labyrinth of the midgut epithelium accounts for major parasite losses associated with midgut invasion, and that it represents a general mechanism of parasite destruction even in mosquitoes that are conventionally described as susceptible.

Huff and others postulated the existence of mosquito

Cell
668

hereditary factor(s) that determine(s) whether or not an individual mosquito will become infected after it has received an infective meal and has been kept under conditions favorable to the development of the parasite (reviewed in Vargas, 1949). Here, we identify TEP1 as one of these factors.

Several lines of evidence support the conclusion that TEP1 is a mosquito factor implicated in parasite killing and, therefore, in establishment of vector capacity in *A. gambiae*. First, TEP1 binds to the surface of ookinetes after they cross the midgut epithelium, as well as to the surface of oocysts but not of sporozoites. Second, this binding is associated with two respective peaks of transcriptional upregulation, possibly to replenish the level of circulating TEP1. Third, TEP1 binding is temporally correlated with the appearance of morphologically degenerate ookinetes. Fourth, the vast majority of TEP1-decorated ookinetes do not express the vital fluorescent marker, GFP. Some do, however, and some TEP1-positive ookinetes also have a regular shape and normal P28 surface covering. These observations suggest that TEP1 binds first and that loss of GFP fluorescence and morphological abnormalities ensue. The fifth and most stringent evidence comes from the dsRNA knockdown experiments. Indeed, *TEP1* silencing results in a 5-fold increase in parasite survival in S mosquitoes. In R mosquitoes, *TEP1* knockdown both increases parasite numbers and completely abolishes their melanization.

Taken together our results lead us to propose the following two-step model for immune responses of *A. gambiae* to *P. berghei* (Figure 7). In the first step, after crossing the midgut epithelium, parasites come in contact (within the basal labyrinth) with soluble hemolymph components, but not with hemocytes. One of the hemolymph components, TEP1, recognizes and binds to the ookinetes, causing them to die by an as yet unknown mechanism. The second step is the disposal of dead parasites by either lysis or, in the case of R mosquitoes, lysis and melanization. The model predicts that the processes of disposal of dead parasites (e.g., melanotic encapsulation and lysis) are controlled by genes other than *TEP1*. It is notable that TEP1 binding to and killing of parasites occurs in both R and S mosquitoes, albeit more slowly in the latter. It is tempting to speculate that TEP1 may function as a complement-like factor, with the covalent binding of its C-terminal part recruiting the formation of a structure similar to the membrane attack complex in mammals, and directing killing of the parasite. Detection of a small number of dead ookinetes that are not labeled with TEP1 in the knockdown experiments suggests that *TEP1* is not the only mosquito gene that controls parasite development and that other genes and gene cascades act in concert to keep the parasite load low during infection.

We stress two major differences between S and R strains. The first is related to the kinetics and efficiency of ookinete killing: approximately 80% of the ookinetes are killed in S mosquitoes by 32 hpi, whereas 100% of the parasites are killed by the same time in the R strain, 80% of which are dead already at 24 hpi. The second difference pertains to the disposal of dead parasites: lysis in S mosquitoes and lysis and melanization in the encapsulating refractory strain.

So far we did not address the role of the thioester in

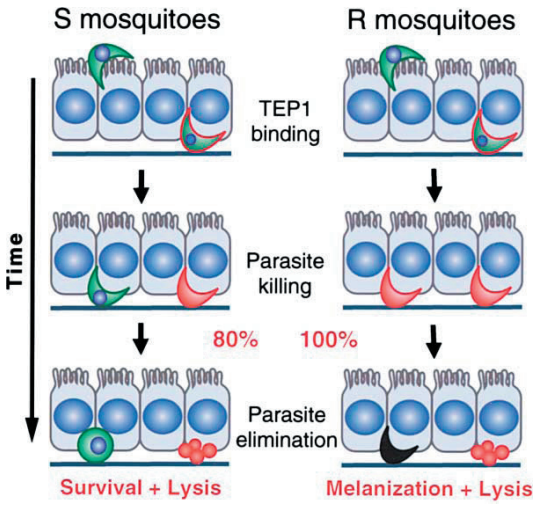


Figure 7. Model for the Role of TEP1 in the Antiparasitic Response of Susceptible (S) and Refractory (R) Mosquitoes

Banana-shaped ookinetes (green) traverse the mosquito midgut epithelium (represented as a cell layer, with microvilli on the lumen side) to reach the basal lamina (thick line) where they become labeled with TEP1 (red line around ookinetes). This labeling targets parasites for killing. Dead parasites (red) loose their nucleus (blue), and are eliminated. Two major differences are observed between susceptible and refractory mosquitoes. First, all parasites are killed in R mosquitoes whereas in S mosquitoes, 20% of parasites survive and transform into oocysts (round and green). Second, in S mosquitoes, dead parasites are disposed by lysis (red balls), whereas in R mosquitoes both lysis and melanization (black) are observed.

parasite binding. Further mutant analysis of the efficiency of TEP1 binding to the parasites will provide important information for structure-function analysis of this and other thioester-containing proteins. We have earlier reported proteolytic cleavage of TEP1 in the hemolymph of wounded or bacteria-infected mosquitoes and binding of the cleaved C-terminal fragment of TEP1 to bacteria in a thioester-dependent manner (Levashina et al., 2001). At present, it is unclear whether the same proteolytic activation is required for binding of TEP1 to ookinetes or whether the same C-terminal cleavage product is observed on the surface of the parasites. We are currently developing monoclonal antibodies against the N-terminal fragment of TEP1, aiming to use double staining with antibodies against the N- and C-terminal domains to answer this question.

Intriguingly, our data indicate that S and R strains exhibit an allelic polymorphism for *TEP1*: *TEP1r* is associated with the refractory L3-5 strain, whereas *TEP1s* is detected only in the susceptible G3 strain. The absence of *TEP1r* in the G3 strain, from which the L3-5 strain was selected, suggests that this allele might have been lost during breeding in our mosquito colony. It would be of interest to examine existent G3 colonies for preservation of the *TEP1r* allele in other laboratories. The comparative sequence analysis of the two alleles reveals that most of the substitutions are clustered in the C3d-like region, which contains the thioester site and, in complement factors, directly binds to substrates. In the

structural model of this region, the three loops enriched in modifications are all located on the convex surface and in close proximity to the thioester active site. The striking accumulation of polymorphisms in the thioester region might indicate its importance for the protein function. In this case, polymorphisms between *TEP1r* and *TEP1s* might affect either the reactivity of the thioester bond and/or the kinetics of protein binding to the substrate, and might account, at least partially, for the more efficient TEP1 binding and parasite killing in R mosquitoes. However, an important note of caution is that this cannot be directly demonstrated by analysis of F1 crosses. Indeed melanotic refractoriness is a complex genetic trait that has long been known to be controlled by several different genes (Zheng et al., 1997). Furthermore, it recently became clear that the complexity is even higher than previously suspected, especially as the importance of several mapped loci is dependent on the species of parasite used for infections (Zheng et al., 2003). In addition, both parental R (L3-5) and S (G3) strains are not genetically homogeneous (Zheng et al., 2003). Although we have performed reciprocal RxS crosses and observed that the transheterozygotes (*TEP1r/TEP1s*) are intermediate between the R and S phenotypes (data not shown), this does not in itself prove that the difference of the strains results from the TEP1 polymorphisms. Transgenic studies will be necessary to test critically whether the expression of the *TEP1r* allele in the genetic background of S mosquitoes is sufficient to accelerate the kinetics of TEP1 binding to parasites and to augment parasite killing.

The laboratory model of infection that we have used, *A. gambiae* and *P. berghei*, is not encountered in the field. Future studies will focus in part on the role of TEP1 in the immune response of *A. gambiae* to its natural parasites. Our preliminary results indicate that TEP1 recognizes and binds to *P. falciparum* ookinetes (S.B., E.A.L. and R.E. Sinden, unpublished data), suggesting that TEP1 binding is not limited to *P. berghei*. Moreover, molecular analysis of *A. gambiae* populations points to the existence of the *TEP1s* and *TEP1r* alleles in malaria-endemic regions of West Africa (S.B., E.A.L. and D. Fontenille, unpublished data). Further studies on the genetics of natural populations will be required to examine whether the distribution of *TEP1* alleles correlates with the incidence of mosquito refractoriness in Africa.

Experimental Procedures

Mosquito Colonies and Parasite Infections

The susceptible G3 and refractory L3-5 colonies were maintained as described previously (Richman et al., 1997). *P. berghei* (ANKA strain) clones 2.34 and transgenic 259cl2 (Frankle Fayard et al., submitted) were passaged in CD1 mice and parasitemia was determined from blood films stained with Diff-Quik I (Dade Behring). In most of the infection experiments, the prevalence of infection was examined 24 hpi, and was about 100%.

Polymorphism Detection by PCR

Genomic DNA was prepared from the parents and progeny of reciprocal crosses between susceptible G3 and refractory L3-5 mosquitoes. Specific primers for *TEP1:* Fwd: 5′-AAAGCTGTTGCGTCA-3′ Rev: 5′-ATAGTTCATTCCGTTTTGGATTACCA-3′ and for *TEP16*: Fwd as for *TEP1* Rev: 5′-CCTCTGCGTGCTTTGCTT-3′ were used at 0.3 μM to amplify fragments of 372 and 349 bp, respectively, with a standard program (30 s at 94°C; 30 s at 52°C; 45 s at 72°C) for 40 cycles.

Transcription Profiling by Real-Time PCR

At selected time points after infection, total RNA from at least 15 fed females was extracted with TRIzol reagent (Invitrogen) and reverse transcribed. Specific TaqMan primers and probes were designed using the Primer Express software (Applied Biosystems). *TEP1:* Fwd: 5′-AAAGCTGTTGCGTCA-3′ Rev: 5′-TTCTCCCACACACCAAACGAA-3′ and Probe: 5′-FAM-CCAGATGCGTTACCGCCAGACAGATG-TAMRA-3′; *TEP16:* Fwd as for *TEP1*, Rev 5′-ATTAGTAGTCTCCCACAAACCA AAT-3′ and Probe 5′-FAM-CCAGATGCGCTACCGTCAGACGGATG-TAMRA-3′; *S7*: Fwd 5′-AGCAGCTACAGCACTTGATTATTGG-3′, Rev 5′-GATATTTTTAACGGCTTTTCTGCGT-3′ and Probe 5′-FAM-CCC GATTTCTCCGATCTTTCACATTCCA-TAMRA-3′. The PCR reactions were assembled and run in the ABI PRIS 7000 Sequence Detection System (Applied Biosystems) following the manufacturer's instructions.

Immunostainings for Confocal and Electron Microscopy

Immunostainings were performed essentially as described (Danielli et al., 2000). For confocal analysis, after blocking, midguts and abdomen walls were incubated overnight at 4°C with a mixture of primary antibodies (rabbit polyclonal antibody against TEP1 at 1:350, mouse monoclonal antibody against *P. berghei* P28 (gift from R. Sinden) at 1:1000, mouse polyclonal antibody against *P. berghei* TRAP (gift from S. Naitza and A. Chrisanti) at 1:300, rat polyclonal antibody against PPO6 (gift from H.-M. Mueller) at 1:500) followed by 1 hr incubation with secondary antibodies (The Jackson Laboratory, 1:2000). Cell nuclei were colored with DAPI (Roche Applied Science, 1 ng/mL). Samples were mounted using the ProLong Antifade Kit (Molecular Probes) and analyzed under a Zeiss LSM 510 confocal microscope. For electron microscopy analysis, midguts were dissected 30 hr postinfection and fixed in 1% paraformaldehyde, 0.5% glutaraldehyde in 0.1 M phosphate buffer [pH 7.2] for 1 hr. After permeabilization and blocking, midguts were incubated overnight at 4°C with rabbit polyclonal antibody against TEP1 at 1:100, followed by 1 hr incubation with a 10 nm gold particle-conjugated secondary antibody (British BioCell International, 1:60). The samples were then dehydrated, embedded in epoxy resin and sectioned for examination with an electron microscope (BioTwin, Phillips).

Double-Stranded RNA Knockdown

dsRNAs were produced as described previously using the plasmids pLL6ds for *dsGFP*, pLL17 for *dsTEP1* (Levashina et al., 2001), and pLL100 for *dsLacZ*. The 816 bp ClaI-BamHI fragment of pC4 (Thummel and Pirrotta, 1991) was cloned between the two T7 promoters of pLL10, resulting in pLL100. Adult females were injected with dsRNA (3 mg/mL) and allowed to recover for 4 days before infection (Blandin et al., 2002) or before collecting hemolymph for immunoblotting (Levashina et al., 2001). The specificity of the knockdown was examined using TEP1-specific antibodies as well as rabbit polyclonal antibodies against other members of TEP family (TEP2, 3 and 4).

Infection Intensity and Mean Oocyst Numbers in Mosquitoes

For each experiment, mosquitoes were blood-fed on a parasite-infected mouse. Mosquito midguts were dissected 6–11 days later and numbers of oocysts were counted using a Zeiss fluorescence microscope.

Acknowledgments

We are grateful to Jules A. Hoffmann for continuous support and interest in this work. The authors are indebted to B. Franke-Fayard for sharing of the GFP-parasite clone before publication. We thank R.E. Sinden for P28 antibodies; S. Naitza and A. Crisanti for TRAP antibodies; H.-M. Mueller for PPO6 antibodies; D. Zachary and G. Griffiths for help with electron microscopy; and P. Aloy and R. Russell for help with the model of the C3d region. We are grateful to D. Doherty for mosquito rearing. This work received financial support from EMBL, CNRS, NIH, and EU-RTN.

Received: September 30, 2003
Revised: December 17, 2003
Accepted: January 14, 2004
Published: March 4, 2004

References

Alavi, Y., Arai, M., Mendoza, J., Tufet-Bayona, M., Sinha, R., Fowler, K., Billker, O., Franke-Fayard, B., Janse, C.J., Waters, A., and Sinden, R.E. (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*, 933–943.

Blandin, S., Moita, L.F., Köcher, T., Wilm, M., Kafatos, F.C., and Levashina, E.A. (2002). Reverse genetics in the mosquito *Anopheles gambiae*: targeted disruption of the *Defensin* gene. EMBO Rep. *3*, 852–856.

Carroll, M.C., and Fischer, M.B. (1997). Complement and the immune response. Curr. Opin. Immunol. *9*, 64–69.

Christophides, G.K., Zdobnov, E., Barillas-Mury, C., Birney, E., Blandin, S., Blass, C., Brey, P.T., Collins, F.H., Danielli, A., Dimopoulos, G., et al. (2002). Immunity-related genes and gene families in *Anopheles gambiae*. Science *298*, 159–165.

Collins, F.H., Sakai, R.K., Vernick, K.D., Paskewitz, S., Seeley, D.C., Miller, L.H., Collins, W.E., Campbell, C.C., and Gwadz, R.W. (1986). Genetic selection of a *Plasmodium*-refractory strain of the malaria vector *Anopheles gambiae*. Science *234*, 607–610.

Danielli, A., Loukeris, T.G., Lagueux, M., Müller, H.M., Richman, A., and Kafatos, F.C. (2000). A modular chitin-binding protease associated with hemocytes and hemolymph in the mosquito *Anopheles gambiae*. Proc. Natl. Acad. Sci. USA *97*, 7136–7141.

Dimopoulos, G., Christophides, G.K., Meister, S., Schultz, J., White, K.P., Barillas-Mury, C., and Kafatos, F.C. (2002). Genome expression analysis of *Anopheles gambiae*: responses to injury, bacterial challenge, and malaria infection. Proc. Natl. Acad. Sci. USA *99*, 8814–8819.

Law, S.K., and Dodds, A.W. (1997). The internal thioester and the covalent binding properties of the complement proteins C3 and C4. Protein Sci. *6*, 263–274.

Levashina, E.A., Moita, L.F., Blandin, S., Vriend, G., Lagueux, M., and Kafatos, F.C. (2001). Conserved role of a complement-like protein in phagocytosis revealed by dsRNA knockout in cultured cells of the mosquito, *Anopheles gambiae*. Cell *104*, 709–718.

Meis, J.F., Pool, G., van Gemert, G.J., Lensen, A.H., Ponnudurai, T., and Meuwissen, J.H. (1989). *Plasmodium falciparum* ookinetes migrate intercellularly through *Anopheles stephensi* midgut epithelium. Parasitol. Res. *76*, 13–19.

Müller, H.M., Dimopoulos, G., Blass, C., and Kafatos, F.C. (1999). A hemocyte-like cell line established from the malaria vector *Anopheles gambiae* expresses six prophenoloxidase genes. J. Biol. Chem. *274*, 11727–11735.

Richman, A.M., Dimopoulos, G., Seeley, D., and Kafatos, F.C. (1997). *Plasmodium* activates the innate immune response of *Anopheles gambiae* mosquitoes. EMBO J. *16*, 6114–6119.

Simonetti, A.B., Billingsley, P.F., Winger, L.A., and Sinden, R.E. (1993). Kinetics of expression of two major *Plasmodium berghei* antigens in the mosquito vector, *Anopheles stephensi*. J. Eukaryot. Microbiol. *40*, 569–576.

Snow, R.W., Craig, M., Deichmann, U., and Marsh, K. (1999). Estimating mortality, morbidity and disability due to malaria among Africa's non-pregnant population. Bull. World Health Organ. *77*, 624–640.

Thummel, C.S., and Pirrotta, V. (1991). Technical notes: New pCasPer P-element vectors. Dros. Info. Service *71*, 150.

Vargas, L. (1949). Culicine and aedine mosquitoes and the malaria infections of lower animals. In Malariology, M.F. Boyd, ed. (Philadelphia: W.B. Saunders Company), pp. 526–538.

Vernick, K.D., Fujioka, H., Seeley, D.C., Tandler, B., Aikawa, M., and Miller, L.H. (1995). *Plasmodium gallinaceum*: a refractory mechanism of ookinete killing in the mosquito, *Anopheles gambiae*. Exp. Parasitol. *80*, 583–595.

Zheng, L., Cornel, A.J., Wang, R., Erfle, H., Voss, H., Ansorge, W., Kafatos, F.C., and Collins, F.H. (1997). Quantitative trait loci for refractoriness of *Anopheles gambiae* to *Plasmodium cynomolgi* B. Science *276*, 425–428.

Zheng, L., Wang, S., Romans, P., Zhao, H., Luna, C., and Benedict, M.Q. (2003). Quantitative trait loci in *Anopheles gambiae* controlling the encapsulation response against *Plasmodium cynomolgi* Ceylon. BMC Genet. *4*, 16.

Chapter 7

General Discussion

General Discussion

The main aim of my project was to understand the role of thioester-containing proteins in the immune response of *Anopheles gambiae*, and especially in the antiparasitic defence of this malaria vector mosquito. For this, we had defined three questions:

- Do mosquitoes have thioester-containing protein(s)?
- How can we analyze the function of these proteins?
- Are mosquito TEPs involved in immune defences, and especially, in those against malaria parasites?

Identification of the family of TEPs *in* A. gambiae.

Analysis of the genome sequence revealed the existence of 15 *TEP* genes and 4 putative allelic sequences, suggesting that in the mosquito, as in the fruitfly, *TEPs* are represented by large families. Another characteristic feature of insect TEP families is the clustering of molecules into species-specific expansions, which, most probably, derived from species-specific duplications and diversification. We propose that these expansions reflect the adaptation of the two dipteran species to distinct pathogenic environments. Further systematic functional analysis of these genes is required for elucidation of the importance of the wide TEP repertoire and its potential role in immune responses of insects.

Establishment of tools for functional analysis.

We have developed rapid and convenient techniques for silencing gene expression in a sequence-specific manner and for analysis of the resulting phenotypes. In a mosquito cell culture system, gene expression can be efficiently disrupted by transfecting cells with dsRNA. Cells in culture represent a convenient system for analysis of such processes as phagocytosis. However, many complex phenomena could not be addressed *in vitro*. We developed a protocol for efficient gene silencing in different immune-responsive tissues by simple injection of dsRNA in adult mosquitoes. We have designed a set of phenotypic assays to analyze the effect of gene knockdown on different aspects of mosquito immunity:

- survival of mosquitoes upon challenge with Gram+ and Gram- bacteria,
- phagocytosis of Gram+ and Gram- bacteria *in vivo* (developed by Luis Moita and Rui Wang, unpublished data),
- development of parasite midgut stages in susceptible mosquitoes
- parasite melanization in refractory mosquitoes.

Currently this methodology is used in the laboratory in the functional screen of more than 100 immune-inducible genes.

Furthermore, we extended the *in vivo* knockdown by dsRNA injection to adult *D. melanogaster* where it represents a powerful method for rapid gene functional analysis and rapid epistatic analysis. Our finding that the degree of silencing is dependent on the fly age provides important information for the development of dsRNA technology in other insect species.

Role of TEP1 and Defensin1 in mosquito immune defences.

Using dsRNA silencing in adult mosquitoes, we provided the first functional evidence that an insect antimicrobial peptide does play an important role in the resistance to infections: Defensin1 is required for the mosquito antimicrobial defence against Gram+, but not to Gram- bacteria; it does not act as a significant antiparasitic factor.

Functional analysis of TEP1 allowed us to demonstrate (i) that the binding of TEP1 to Gram-positive and Gram-negative bacteria promotes their phagocytosis by hemocytes and (ii) that TEP1 binds to and mediates killing of midgut stages of *P. berghei*. The binding of TEP1 to different types of pathogens and the consequent promotion of phagocytosis of bacteria in the mosquito hemocoel or parasite killing in the midgut, are reminiscent of functions of complement factors in vertebrates.

Importantly, our studies on the role of TEP1 upon parasite invasion demonstrate (i) that mosquitoes actively defend themselves against *Plasmodium* parasites, (ii) that the dramatic losses undergone by parasites in the first days after infection are essentially due to this efficacious immune response and (iii) identify TEP1 as an important actor of this immune response.

Taken together, our results presented in chapter 6 led us to propose the following two-step model for immune responses of *A. gambiae* to *P. berghei* (Fig. 9). In the first step, after crossing the midgut epithelium, parasites come in contact (within the basal labyrinth) with soluble hemolymph components, but not with hemocytes. One of the hemolymph components, TEP1, recognizes and binds to the ookinetes, causing them to die by an as yet unknown mechanism. The second step is the disposal of dead parasites by either lysis or, in the case of refractory mosquitoes, lysis and melanization. The model predicts that the processes of disposal of dead parasites (i.e. melanotic encapsulation and lysis) are controlled by genes other than *TEP1*. It is notable that TEP1 binding to and killing of parasites occurs in both R and S mosquitoes, albeit more slowly in the latter. We have observed that the two *TEP1* alleles, *TEP1s* and *TEP1r*, were specifically associated with the S and R strains, respectively. We speculate that the polymorphism of the *TEP1* locus could account, at least partially, for the more efficient killing of parasites in R mosquitoes (this point is further discussed below).

Noteworthy, in *TEP1*-deficient mosquitoes we detected a small number of dead ookinetes that were not labelled with TEP1 suggesting that *TEP1* is not the only mosquito gene that controls parasite development and that other genes and gene cascades act in concert to keep the parasite load low during infection. In keeping with this, a leucine-rich repeat immune protein, LRIM1,

was recently shown to be implicated in the mosquito antiparasitic response (Osta et al., 2004). Additional actors involved in mosquito-parasite interactions are likely to be identified by functional screens such as the one we have initiated in our laboratory.

In conclusion, we identified a family of thioester-containing proteins in the mosquito *A. gambiae* and provided the functional characterization of one of them, TEP1. We demonstrate that binding of TEP1 to pathogen surfaces promotes two distinct immune reactions: phagocytosis of bacteria and killing of *Plasmodium* parasites. Many questions still have to be answered (Fig. 9): (i) How does TEP1 become activated? (ii) How is TEP1 targeted to pathogen surfaces? (iii) How does TEP1 binding trigger the destruction of pathogens? (iv) Relevance of the polymorphism of *TEP1* (and other *TEPs*) in the establishment of refractoriness. (v) Comparative functional analysis of the TEP family in *A. gambiae.* Understanding the mechanism of TEP1 binding to and killing of parasites and the differences between the two *TEP1* alleles might provide important clues for the design of new strategies to enhance the mosquito antiparasitic response during the first days after infection, when the immune response is most efficient and when the number of parasites is at its minimum.

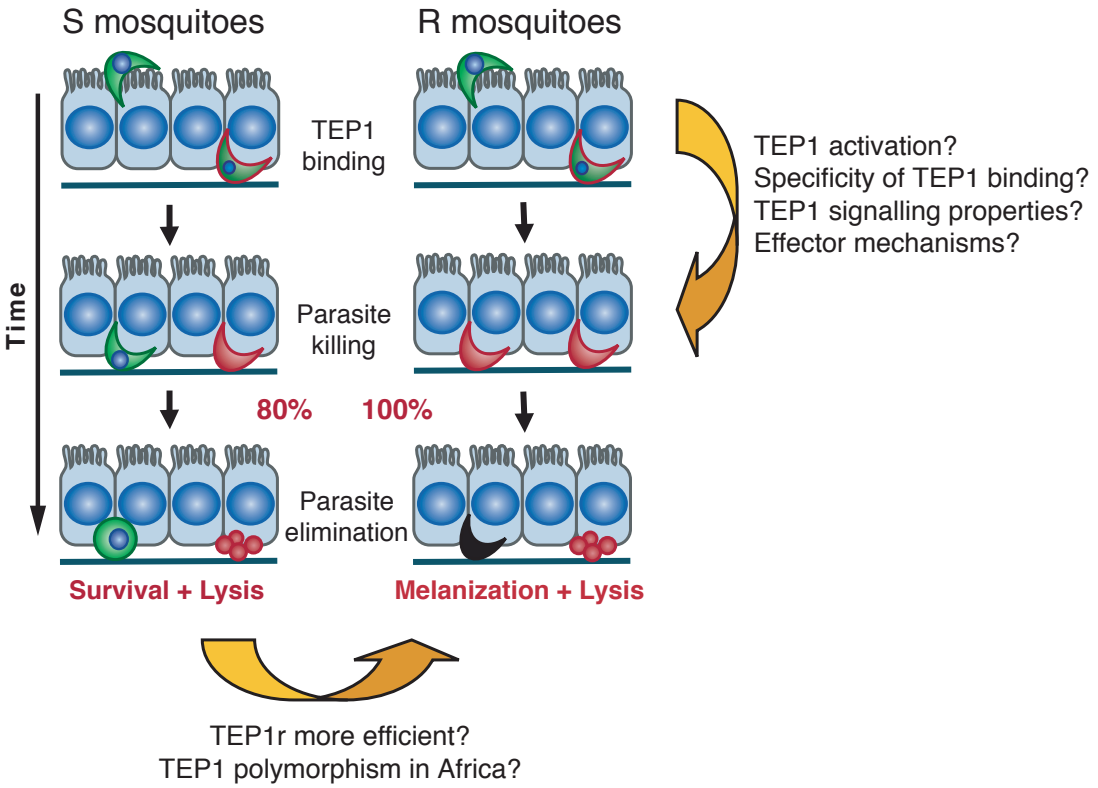


Figure 9. Model for the role of TEP1 in the antiparasitic response of susceptible (S) and refractory (R) mosquitoes. Banana-shaped ookinetes (green) traverse the mosquito midgut epithelium (represented as a cell layer, with microvilli on the lumen side) to reach the basal lamina (thick line) where they become labelled with TEP1 (red line around ookinetes). This labelling targets parasites for killing. Dead parasites (red) loose their nucleus (blue), and are eliminated. Two major differences are observed between susceptible and refractory mosquitoes. First, all parasites are killed in R mosquitoes whereas in S mosquitoes, 20% of parasites survive and transform into oocysts (round and green). Second, in S mosquitoes, dead parasites are disposed by lysis (red balls), whereas in R mosquitoes both lysis and melanization (black) are observed. Large arrows point to new questions that have been raised.

How does TEP1 become activated?

In chapter 1, we reported that proteolytic cleavage of TEP1 is increased in the hemolymph of wounded or bacterium-infected mosquitoes, and that the cleaved C-terminal fragment of TEP1 binds to bacteria in a thioester-dependent manner. At present, it is unclear whether the same proteolytic activation is required for binding of TEP1 to ookinetes and whether the same C-terminal cleavage product is observed on the surface of the parasites. We have developed monoclonal antibodies against the N-terminal fragment of TEP1, aiming to use double staining with antibodies against the N- and C-terminal domains to answer this question. In addition, we do not know which protease(s) cleave TEP1 (at least in the case of TEP1 binding to bacteria): Is it a specific protease complex, as for complement factors? Extensive search of the fruitfly and mosquito genome for genes encoding complement-activating proteases such as MASP, factor B and C2 did not identify homologues, suggesting that insect TEPs are activated by endogenous proteases released upon injury or by proteases of pathogen origin. Finally, we have not examined the possibility that the cleaved N-terminal part might act as a signalling molecule, like complement anaphylatoxins.

We have not explored the role of the thioester in parasite binding. Further mutant analysis of the efficiency of TEP1 binding to the parasites will provide important information for structure-function analysis of this and other thioester-containing proteins.

How is TEP1 targeted to pathogen surfaces?

We reported that TEP1 binds to the surfaces of bacteria and parasites, but we did not detect TEP1 on the surface of mosquito cells by immuno-histochemical analysis. The specific binding of TEP1 to pathogen surfaces could result from:

- the specific recognition of pathogens by TEP1 or associated activating molecules,
- the localized activation of TEP1 by pathogen proteases that are membrane bound or that are secreted by pathogens and thus present in higher concentrations in the vicinity of the pathogen surface,
- specific inhibitory molecules expressed by mosquito cells.

This issue has not been addressed yet.

How does TEP1 binding trigger the destruction of pathogens?

Binding of TEP1 to bacteria promotes their uptake by hemocytes, probably via the interaction with a receptor that remains to be identified. Interestingly, vertebrate α2Ms are recognized by the LDL receptor Related Protein (LRP), a homologue of which has been identified in *A. gambiae* (Luis Moita, unpublished data). Finally, we do not understand the mechanisms that are triggered by TEP1 binding on parasites. It is tempting to speculate that TEP1 may function as a complement-like factor, with the covalent binding of its C-terminal part recruiting the formation of a structure similar to the membrane attack complex in mammals, and directing killing of the parasite.

Transgenic mosquitoes as a tool to dissect TEP1 signalling properties.

To complete the functional analysis of TEP1 *in vivo*, we have used a complementary gain-of-function approach by overexpression of *TEP1*. Together with John Clayton from our laboratory, we have established two transgenic mosquito lines bearing an over-expression cassette with *TEP1s* under the control of the promoter of the ubiquitously expressed *Drosophila* heat shock protein 70 (*Hsp70*) (Sakai and Miller, 1992). To our great surprise, the phenotype of one of the lines was a knock out of *TEP1* without or after heat shock. Consistently, the phenotype of these transgenic mosquitoes upon parasite infection was similar to that of *dsTEP1*-treated susceptible mosquitoes. The piggyBac-driven insertion is located in the 3' UTR of the ornithine transporter gene (X chromosome), upstream the gene terminator (Fig. 10A). Both the selection cassette (*3xPax3::DsRed*) and expression cassette (*Hsp70::TEP1*) are in opposite orientation compared to that of the ornithine transporter gene. Thus, transcription of the ornithine transporter gene could lead to the production of a large hybrid transcript comprising ornithine gene itself and antisense strains of inserted transposon coding for *TEP1* and *DsRed*. Although expression pattern of the ornithine gene is presently unknown, it is possible that its expression, and hence the production of the hybrid transcript, are not overlapping in time or in place with that of *Pax6* promoter. In contrast, *Hsp70* is a ubiquitous leaky promoter, causing *TEP1* transcription in all mosquito tissues. Should this hypothesis be correct, simultaneous synthesis of sense and antisense transcripts for *TEP1* should induce dsRNA formation and activation of RNAi directed against exogenous and endogenous *TEP1* transcripts (Fig. 10B). To assess this hypothesis, we are planning to perform Northern blotting to analyze RNA extracts of transgenic mosquitoes for the absence of *TEP1* transcript and for the presence of *TEP1*-specific siRNAs. As *TEP1*-depleted mosquitoes are hypersensitive to *P. berghei* infections, we have recently patented the use of these transgenic knockdown mosquitoes for production of high numbers of parasites.

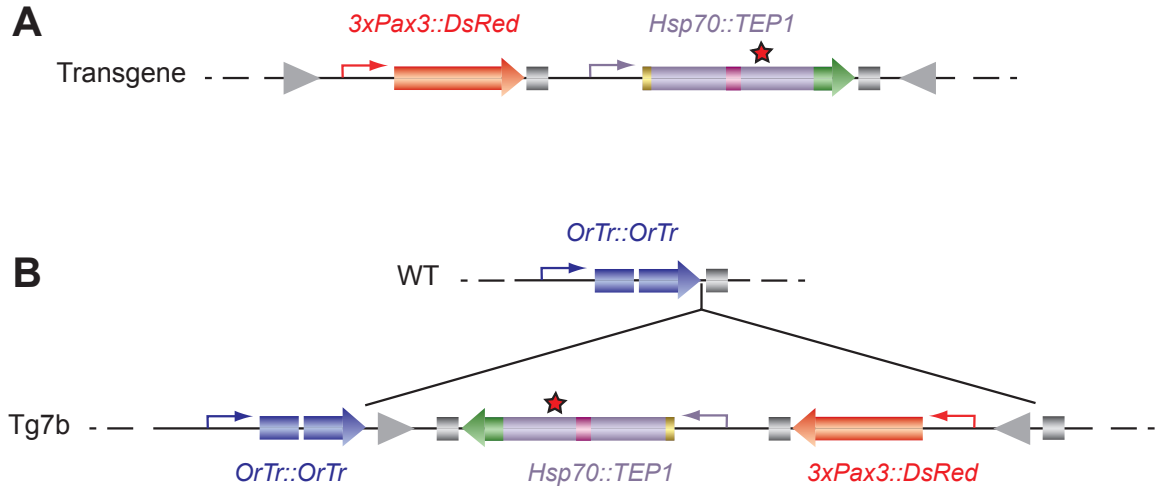


Figure 10. Transgene for *TEP1* overexpression and insertion site in lineTg7b. A. Overexpression cassette based on the Piggy Bac transposon. **B**. Insertion site of the trangene in line Tg7b. Note that the transgene is positioned between the last exon encoded by the Ornithrine transporter gene and the terminator sequence of this gene. Coding sequences are represented with rectangles and the sense direction is indicated with an arrowhead in 3'. Promoters are represented by arrows, terminators by gray rectangles and the inverted repeats of the Piggy Bac transposon by gray triangles.

The phenotype of the second transgenic line was closer to what we expected: Under normal breeding conditions, a slight increase in the TEP1 protein level was detected in the hemolymph of these mosquitoes. We are currently establishing the heat shock conditions to drive the overexpression of *TEP1*. Importantly, the comparison of *TEP1*-overexpressing and *TEP1*-knockout mosquitoes using microarrays should provide important insights into the signalling pathways activated by TEP1 binding to the parasites and bacteria, and the effector mechanisms that are triggered.

Relevance of the polymorphism of TEP1 (and other TEPs) in the establishment of refractoriness.

We have not yet addressed the question of whether a direct correlation exists between the polymorphism detected in *TEP1r* and the more efficient parasite recognition and killing in refractory mosquitoes. Transgenic studies will be necessary to test critically whether the expression of the *TEP1r* allele in the genetic background of S mosquitoes is sufficient to accelerate the kinetics of TEP1 binding to parasites and to augment parasite killing. Should it be the case, further insights into the properties of the TEP1 molecule and the specific importance of different domains could come from a detailed comparison of the properties of the *TEP1s* and *TEP1r* alleles, especially through the analysis of chimeras.

The laboratory model of infection that we have used, *A. gambiae* and *P. berghei,* does not exist in nature. Future studies will focus in part on the role of TEP1 in the immune response of *A. gambiae* to its natural parasites. Our preliminary results, in collaboration with R.E. Sinden (Imperial College, London), indicate that TEP1 recognizes and binds to *P. falciparum* parasites, suggesting that TEP1 binding is not limited to *P. berghei*. Moreover, molecular analysis of natural populations of *A. gambiae*, in collaboration with D. Fontenille (IRD, Montpellier), points to the existence of the *TEP1s* and *TEP1r* alleles in malaria-endemic regions of West Africa. Further studies on the genetics of natural populations will be required to examine whether the distribution of *TEP1* alleles, and potentially of allelic forms of other *TEPs*, correlates with the incidence of mosquito refractoriness in Africa.

Moreover two additional putative allelic forms of *TEP5* and *TEP6*, namely *TEP17* and *TEP18*, are located in the same cluster as *TEP16* (Fig. 8). Our preliminary results indicate that, as for *TEP1* and *TEP16, TEP6* and *TEP18* are specific to the susceptible and refractory strains, respectively. One might speculate that *TEP16, 17* and *18* all belong to the same haplotype that would be specific to R mosquitoes. Finally, a fourth putative allelic pair, *TEP8* - *TEP19*, was detected in the genome.

And the other members of the family?

The annotation of genome revealed the existence of 15 *TEP* genes in *A. gambiae*. The functional analysis of all TEPs by targeted gene disruption *in vivo* will provide new insights in the role of this family in the mosquito immune responses. So far, we have characterized TEP1-4 in some detail at the biochemical (Table 1) and transcriptional levels, and have analyzed their knockout phenotype *in vivo*. Although the expression of both *TEP3* and *TEP4* is induced upon bacterial challenge and parasite infection, analysis of the knockdown phenotypes revealed that TEP3 and 4 do play a role in promoting phagocytosis of bacteria and in the survival of mosquitoes upon bacterial challenge, however, only TEP3 seems to have a role in the antiparasitic response. Yet, most of the comparative analysis remains ahead, and especially that of mosquito-specific TEPs.

	TEP1	**TEP2**	**TEP3**	**TEP4**
Secreted protein	+	-	+	+
Glycosylation	+	- (?)	+	+
Functional TE	+	+	-	+
Proteolytic activation	+	+	+	+
Production site	**hemocytes**	**midgut and salivary glands**	**ND**	**ND**
Binding to pathogens	**bacteria and *Plasmodium***	***Plasmodium***	**ND**	**ND**

Table 1: Biochemical analysis of aTEP1-4.

Interestingly, major structural differences exist between mosquito TEPs. In particular, only 4 mosquito TEPs (TEP1,2,4,15) and maybe 2 more (TEP5 and 7, which sequence is not yet complete), have a conserved thioester motif (for comparison, in *Drosophila*, 9 out of 10 TEPs have a thioester motif). Uncovering the functions and mechanisms of action of the many thioester-less TEPs in mosquitoes might shed some light on the importance of other domains of these proteins.

Finally, although no α2M-like TEP could be identified in insects at the sequence level, some of them might still function as pan-protease inhibitors. This type of activity has not yet been explored for insect TEPs due to technical limitations imposed by the availability of biological material for biochemical studies. Deeper characterization of loss-of-function phenotypes and the development of new tools for further biochemical analysis in insect TEPs will be instrumental for uncovering TEP functions and, eventually for our understanding of this multifaceted family of proteins.

2002, the Anopheline Revolution?

Last years witnessed a real revolution in both tools and information available for research in vector biology. Here I identify the milestones, which changed or are changing our understanding of the principles of malaria transmission:

• **2002, Publication of the genome sequence of *A. gambiae*.** Mining the *A. gambiae* genome will lead to the discovery of genes that are potentially involved in processes like metabolism, behaviour (like the preferential biting habits of *A. gambiae* towards humans). Comparison of genomic information between flies and mosquitoes can give a hint on potential functions for homologues, it can also identify mosquito-specific genes that might have evolved in adaptation to the blood-sucking habit of mosquitoes and to their specific pathogenic environment. A better understanding of mosquito physiology will be instrumental for the identification of genes involved in insecticide resistance and of potential targets for new insecticides or for chemicals to manipulate mosquito immune reactions, behaviour, etc. (reviewed in Craig et al., 2003).

• **2002, Large-scale transcriptional analysis by microarrays.** The analysis of global changes in development, growth and response to parasite infection and to external factors will assist the search for effector molecules and pathways.

• **2002, Reverse genetics available in *A. gambiae*.** This method will be essential to confirm the function of genes potentially implicated in a given biological process, and to conduct large-scale reverse genetic functional screens for the identification of new effector genes.

• **2000-2001, Stable transformation of anopheline mosquitoes.** The genetic engineering of mosquitoes represents a powerful laboratory research tool. In addition, although quite controversial, the release of transgenic mosquitoes refractory to parasite infection could help to control disease transmission.

The results that are presented in this thesis could not have been obtained without these important developments. However it is clear that we are still missing important insights into biology of mosquito-parasite interactions, which will ultimately lead to the design of new strategies to alleviate the unacceptably high load of suffering caused by malaria and other vector-borne diseases.

Thèse en Français

Étude du Rôle des Protéines à Thioester dans la Réponse Immunitaire du Moustique *Anopheles gambiae*

Rôle des Protéines à Thioester dans la Réponse immunitaire du Moustique *Anopheles gambiae*

Introduction

Le paludisme est de loin la maladie transmise par les insectes qui affecte le plus d'humains. Elle est provoquée par un parasite eucaryote unicellulaire, *Plasmodium*, et est transmise à l'homme par un moustique vecteur, *Anopheles*. Cette maladie, actuellement endémique en Afrique, en Amérique du sud et en Asie du sud-est, est particulièrement meurtrière en Afrique sub-saharienne, où elle fait au moins 90% de ses victimes. Chaque année, entre 300 et 500 millions de personnes contractent le paludisme et plus 2 millions en meurent, principalement des enfants de moins de 5 ans. Parmi les adultes, ce sont les femmes enceintes et les touristes non-immunisés qui sont les plus vulnérables. Outre les conséquences humaines, cette maladie est aussi un frein au développement économique des pays endémiques : la grande majorité de ces pays sont aussi ceux qui ont le plus faible produit national brut par habitant et les taux de croissance économique les plus bas (Sachs and Malaney, 2002).

Contrôle du paludisme : vers de nouvelles stratégies?

Plusieurs stratégies ont été mises en œuvre pour tenter de limiter la transmission du paludisme (Greenwood and Mutabingwa, 2002). Historiquement, les plus grands succès ont été obtenus grâce au contrôle des populations de moustiques, notamment à l'aide d'insecticides comme le DTT. La campagne d'éradication globale du paludisme des années 50-60 a permis d'éradiquer ou au moins de contrôler le paludisme dans certains pays, en particuliers ceux où le climat est tempéré et où la transmission était modeste. Cependant ces efforts n'ont pas été soutenus du fait, entre autres, de leur coût élevé à long terme et de l'émergence de résistances à cet insecticide parmi les populations de moustiques. Par ailleurs, d'autres facteurs sont venus aggraver la situation du paludisme dans le monde. De loin le plus important d'entre eux est le développement par *Plasmodium falciparum* de résistance à des médicaments peu chers, et dans certaines parties d'Asie du Sud-est, à quasiment tous les médicaments antipaludéens connus. D'autres tentatives se sont concentrées sur la prévention de la transmission de la maladie aux humains grâce à la vaccination. Cependant aucun vaccin efficace n'a été mis au point jusqu'à présent, principalement à cause du fait que le parasite change rapidement ses structures antigéniques, rendant ainsi les vaccins rapidement inefficaces (Richie and Saul, 2002). Le début des années 1990 a vu renaître

une prise de conscience internationale de l'ampleur inquiétante de cette maladie et, en 1992, le contrôle du paludisme est redevenu une priorité de l'Organisation Mondiale de la Santé (OMS). En 1998, l'OMS a lancé l'association « Roll Back Malaria » qui rassemble les principaux groupes œuvrant pour le contrôle de cette maladie et dont le but est de réduire de moitié le fardeau paludéen d'ici à l'an 2010. Une large partie des fonds est consacrée à rendre accessibles les moyens d'intervention déjà disponibles mais sous utilisés dans les pays endémiques, tels que les moustiquaires traitées avec des insecticides, le traitement intermittent préventif ou thérapeutique des femmes enceintes et des enfants, etc. De plus, un soutien financier est apporté à la recherche fondamentale et appliquée, afin de promouvoir le développement de nouveaux outils (nouveaux médicaments, insecticides), mais aussi de nouvelles stratégies pour contrôler la capacité vectorielle des moustiques. La mise au point récente de techniques permettant la modification génétique de l'anophèle rend possible la création de lignées de moustiques réfractaires aux parasites. En effet, le moustique peut être transformé pour exprimer un composé exogène inhibant le développement du parasite (Ito et al., 2002), pour exprimer de façon erronée une molécule endogène nécessaire au parasite, ou pour optimiser des mécanismes de défense déjà existants chez le moustique. Bien que les conséquences écologiques de la mise en liberté de moustiques transgéniques, même stérilisés, soient très difficiles à évaluer, cette méthode reste une option attractive pour pouvoir contrôler leur capacité vectorielle. Dans tous les cas, la mise en œuvre de nouvelles stratégies et la mesure de leur efficacité et des risques associés requièrent une meilleure compréhension des interactions moustique-parasite au niveau moléculaire.

Le cycle de vie de Plasmodium

Les parasites responsables du paludisme sont transmis à l'homme par un moustique femelle infecté lorsque celle-ci prend un repas de sang. Afin de pouvoir infecter l'hôte suivant, le parasite doit suivre un cycle de développement complexe dans le moustique (Fig. 11). À l'intérieur de l'intestin, les gamétocytes males et femelles se développent en gamètes ; la fécondation a alors lieu, quelques heures seulement après le repas de sang. Les zygotes ainsi formés se transforment en oocinètes mobiles qui envahissent et traversent l'épithélium intestinal 24 à 48h après l'infection. Une fois qu'ils ont atteint le côté basal, ils forment des kystes, appelés ookystes, entre l'épithélium intestinal et la membrane basale. Pendant les 10 jours qui suivent, à l'intérieur de chaque ookyste, une méiose suivie de nombreux cycles mitotiques, produisent des milliers de sporozoïtes haploïdes. Quatorze à 16 jours après infection, les sporozoites sont libérés dans l'hemocoele (équivalent du système circulatoire des vertébrés) du moustique, migrent et envahissent les glandes salivaires du moustique. Le cycle du parasite dans le moustique est achevé quand le moustique injecte ces sporozoïtes infectieux dans l'hôte suivant.

Des moustiques susceptibles et des moustiques réfractaires

À cause de la complexité des interactions entre le parasite et le moustique, la plupart des parasites ne peuvent se développer que dans un nombre très restreint d'espèces de moustiques et ainsi, seul un nombre limité de combinaisons *Plasmodium* - *Anopheles* peuvent transmettre le paludisme

dans un groupe particulier de vertébrés. Outre ces variations interspécifiques, la capacité de chaque moustique à supporter le développement du parasite (ou capacité vectorielle) varie énormément entre individus d'une même espèce. Aussi, les interactions moustique – parasite constituent un aspect essentiel de la transmission de la maladie et représentent des cibles potentielles pour les stratégies visant à contrôler le paludisme. Les facteurs qui font qu'un moustique n'est pas capable de transmettre le parasite sont de deux ordres : le moustique est dépourvu d'un élément requis pour le développement du parasite, comme par exemple l'acide xanthurenic (Billker et al., 1998) qui est nécessaire à la gamétogénèse et/ou la réponse immunitaire du moustique est efficace pour supprimer le développement des parasites.

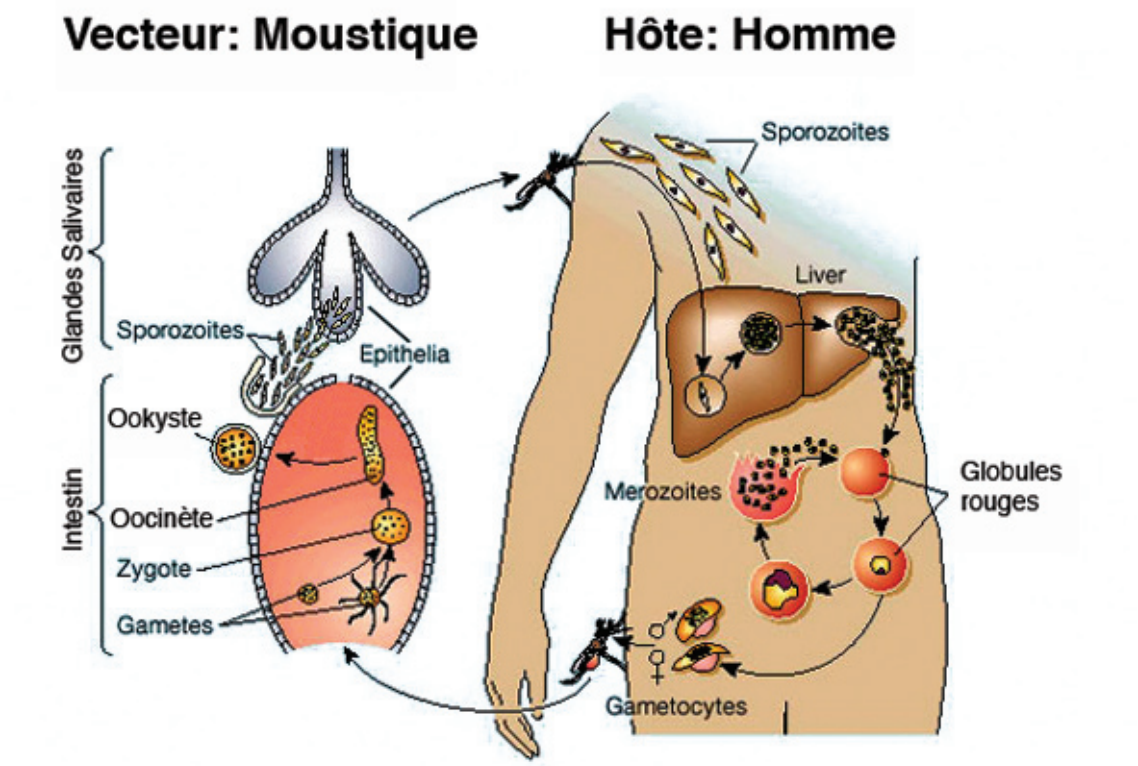


Figure 11. Cycle de vie de *Plasmodium*. Le parasite passe par de nombreux stades développementaux qui sont adaptés à l'invasion et à multiplication dans des types de cellules différents chez l'hôte humain et chez le moustique vecteur.

La réponse immunitaire du moustique

Plusieurs faits semblent en effet indiquer que les moustiques se sont dotés d'une réponse antiparasitaire efficace. Au cours de son développement chez le moustique, *Plasmodium* subit de nombreuses pertes et traverse plusieurs goulets d'étranglement, indiquant que le moustique est capable de limiter le développement du parasite. Ceci est confirmé par le cas extrême de deux souches de moustiques réfractaires sélectionnées au laboratoire, où le développement du parasite est bloqué par son encapsulation mélanotique (Collins et al., 1986) ou par sa lyse (Vernick et al., 1995) au niveau du tube digestif. Cependant les mécanismes moléculaires qui sous-tendent ces

résistances, et plus généralement les défenses immunitaires du moustique restent à découvrir. L'étude du système immunitaire du moustique a très largement bénéficié de la connaissance plus précise que nous avions des systèmes immunitaires des vertébrés et d'autres Arthropodes, et plus particulièrement de la mouche *Drosophila melanogaster*, qui, comme le moustique, est un insecte diptère.

De nombreux gènes potentiellement impliqués dans la réponse immunitaire du moustique ont été identifiés par homologie avec des gènes de l'immunité chez d'autres insectes, ou bien parce que leur expression est modulée lors d'une infection. Cependant, leur rôle dans la réponse immunitaire du moustique, et en particulier dans la défense antiparasitaire, n'avait pas été confirmé par une analyse fonctionnelle. Au moment où nous avons commencé ce projet, plusieurs laboratoires avaient déjà utilisé la méthode d'interférence à ARN pour supprimer spécifiquement l'expression des gènes dans différents organismes, et notamment, dans des cultures de cellules de drosophile. Aussi, nous avons décidé d'adapter cette méthode chez l'anophèle afin de pouvoir étudier la fonction des gènes.

La mise en place de nouvelles techniques telles que l'interférence à ARN (Blandin et al., 2002), la transformation génétique (Catteruccia et al., 2000b; Grossman et al., 2001), l'analyse de l'expression des gènes à grande échelle grâce aux puces à ADN (Dimopoulos et al., 2002), couplée avec l'achèvement du séquençage du génome d'*Anopheles gambiae* (Holt et al., 2002) a amorcé une véritable révolution dans le domaine de l'immunité du moustique. Dans une revue que nous avons récemment publié (Blandin and Levashina, 2004a), nous faisons état de nos connaissances antérieures et voyons en quoi ces progrès techniques ont changé notre vision du système immunitaire de l'anophèle.

La famille des protéines à thioester

Les bases moléculaires de la reconnaissance des pathogènes chez le moustique, comme chez la drosophile d'ailleurs, sont longtemps restées inconnues. Afin d'aborder ce problème, nous avons choisi d'étudier les protéines à thioester (TEPs pour "Thioester-containing proteins") en tant que molécules potentiellement impliquées dans la reconnaissance des pathogènes et dans l'activation des effecteurs de la réponse immunitaire. En effet, chez les vertébrés, les membres de cette famille sont d'importants acteurs de la réponse immunitaire, on y trouve :

- les α2-macroglobulines (α2Ms), qui sont des inhibiteurs universels de protéases, probablement impliqués dans l'élimination des protéases qui sont libérées lors d'une blessure ou qui sont produites par des pathogènes,
- les facteurs C3, C4 et C5 du complément, qui se lient à la surface des pathogènes et favorisent ainsi leur élimination par phagocytose ou par lyse cellulaire.

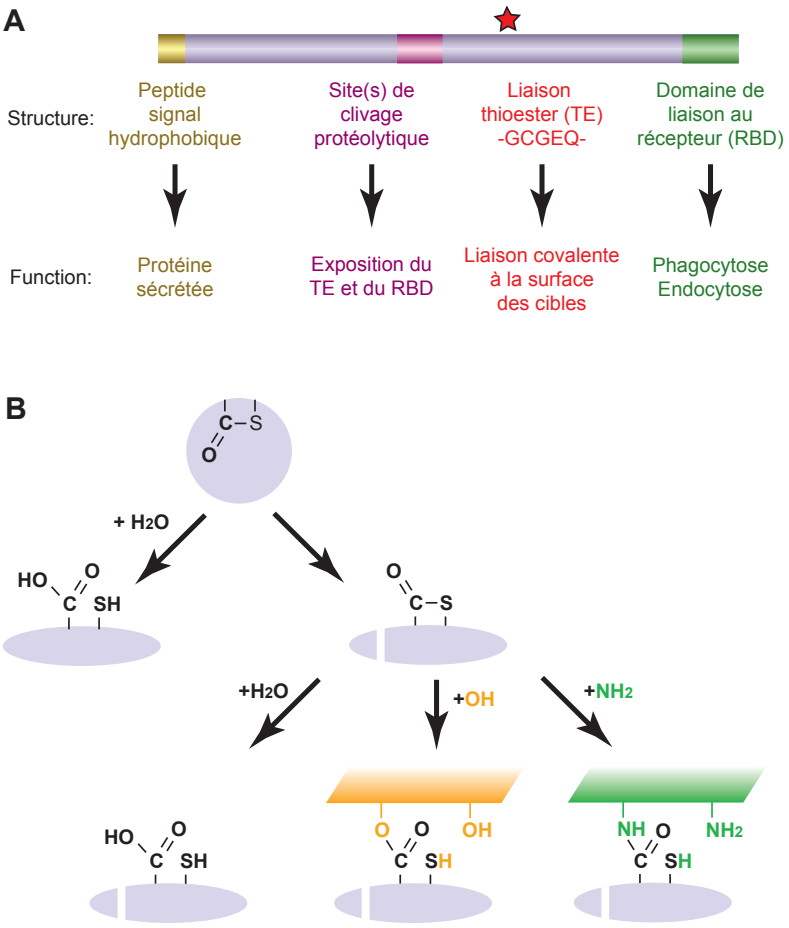


Figure 12. Caractéristiques des TEPs et propriétés de leur liaison thioester. A. Prototype d'une protéine à thioester. **B**. Propriétés de la liaison thioester. Cette liaison hautement réactive est enfouie dans la protéine inactive où elle ne peut réagir qu'avec de petits nucléophiles (méthylamine, H_2O). Le clivage protéolytic de la TEP induit un changement conformationnel (symbolisé par le passage du cercle à l'ovale) et expose la liaison thioester qui peut alors réagir avec des groupements amines ou hydroxyles proches. Localement, la plus grande concentration d'hydroxyles provenient des molécules d'H_2O. Ainsi, toute liaison thioester qui n'a pas réagit rapidement est inactivée par hydrolyse, limitant la dispersion de ces molécules très réactives.

Les TEPs ont en commun plusieurs caractéristiques (Fig. 12): ces protéines de grande taille (~1500 acides aminés) sont (i) sécrétées sous forme inactive. (ii) Pour devenir fonctionnelles, elles doivent être activées par clivage protéolytique. (iii) Les TEPs se lient de façon covalente à la surface de leurs cibles (des pathogènes dans le cas des facteurs du complément, et des protéases dans le cas des α2Ms) via leur liaison thioester (iv) hautement réactive et très conservée parmi les TEPs. L'activation protéolytique ainsi que la liaison de la TEP à sa cible exposent son domaine C-terminal, permettant ainsi l'élimination du complex par phagocytose ou endocytose médiées par un récepteur. Dans le cas du complément, le clivage génère deux fragments: le fragment C-terminal qui contient le thioester et qui est lié à la cible et le fragment N-terminal (anaphylatoxine) qui déclenche l'inflammation au niveau du site d'infection. Toutes ces caractéristiques fonctionnelles imposent des contraintes particulières sur l'organisation structurale de ces protéines.

Objectifs

Aussi, le but principal de mon projet a été l'identification et la caractérisation fonctionnelle des protéines à thioester chez l'un des vecteurs les plus importants du paludisme, le moustique *Anopheles gambiae*. Pour ceci, j'ai dû répondre aux trois questions suivantes :

Est-ce que les moustiques ont des protéines à thioester ?
Nous avions de bonnes raisons de penser que la réponse à cette question serait affirmative. En effet, des molécules de la même famille avaient été caractérisées chez des protostomes, et notamment chez plusieurs Arthropodes (une limule et un tick). De plus, M. Lagueux (IBMC, Strasbourg) avait déjà isolé 6 gènes encodant des protéines à thioester dans le génome de la drosophile.

Comment pouvons-nous analyser la fonction de ces protéines ?
Nombreuses sont les limitations inhérentes à l'élevage des moustiques qui font obstacle à l'utilisation des méthodes classiques de génétique, telles que les cribles de mutants à grande échelle, ou l'analyse de lignées génétiquement modifiées. Aussi, aucun outil pour l'analyse fonctionnelle des gènes n'était disponible chez le moustique. Nous avons donc décidé d'adapter la technique d'interférence à ARN afin d'inhiber spécifiquement l'expression des gènes et de pouvoir en déduire leur fonction. Là encore, nous avions bon espoir de réussir cette entreprise, étant donné que l'interférence à ARN avait déjà été utilisée avec succès dans des cultures de cellules de drosophile (Hammond et al., 2000).

Les protéines à thioester sont-elles impliquées dans la réponse immunitaire du moustique, et en particulier dans la réponse antiparasitaire ?
La réponse à cette question était sûrement la moins évidente… et la plus intéressante ! Au moment où nous avons commencé ce projet, toutes les protéines à thioester caractérisées chez les protostomes avaient été isolées pour leurs propriétés d'inhibiteurs de protéases (Levashina et al., 2003). Ces molécules étaient aussi plus proches des α2Ms que des facteurs du complément au niveau de la séquence et leur fonction biologique était inconnue. De plus, aucune fonction n'avait encore été mise en évidence pour les six *Teps* identifiées dans le génome de la drosophile. Seules deux indications indirectes tirées de la littérature suggéraient que la reconnaissance des parasites par le moustique pouvaient reposer sur des réactions de type complément : (i) le fait que l'activation de la mélanisation soit limitée à la surface des parasites, et (ii) le travail de M. Gorman *et al* démontrant l'importance des caractéristiques de la surface des corps étrangers dans leur reconnaissance par le système immunitaire du moustique.

Discussion

Identification des gènes TEPs *de l'anophèle et comparaison phylogénétique*

Nous avons participé à un effort collectif pour annoter le génome d'*Anopheles gambiae* (Christophides et al., 2002). Ceci nous a permis d'identifier 15 gènes *TEPs* et 4 séquences alléliques. En outre, au moins 9 des 15 gènes sont dépourvus du motif thioester caractéristique de cette famille. L'analyse des fonctions et des mécanismes d'action de ces protéines devrait nous renseigner sur l'importance des autres domaines des TEPs.

L'analyse phylogénétique de l'ensemble des TEPs connues dans le monde animal a mis en évidence plusieurs points importants (Blandin and Levashina, 2004b):

- la ségrégation des TEPs en trois groupes : les facteurs du complément, les α2Ms et les TEPs des invertébrés (les TEPs de l'anophèle et de la drosophile se trouvant dans ce dernier groupe),
- le nombre très limité de TEPs orthologues entre la drosophile et l'anophèle et la présence de nombreuses séquences spécifiques à chaque espèce,
- le fait que la très grande majorité des TEPs qui ont jusqu'ici été caractérisées comme inductibles lors d'une infection figurent parmi les TEPs spécifiques à chaque espèce,

Il semble donc que les grandes familles de TEPs chez l'anophèle et chez la drosophile, soient issues de duplications successives, indépendantes dans chacune de ces espèces, et de leur diversification en réponse à des environnements pathogéniques différents chez la mouche et chez le moustique. Aussi, l'analyse fonctionnelle des TEPs de l'anophèle, et en particulier de celles qui lui sont spécifiques, devrait-elle nous permettre de mieux comprendre la réponse antiparasitaire de ce moustique.

Développement d'outils pour l'analyse fonctionnelle des gènes…

… dans la réponse immunitaire de l'anophèle.

Nous avons développé des méthodes simples et rapides pour l'analyse génétique indirecte chez *A. gambiae*. Dans un premier temps, nous avons établi les conditions pour la suppression de l'expression des gènes dans une culture de cellules de moustique par transfection avec de l'ARN double brin (ARNdb) spécifique (Levashina et al., 2001). Le knock-down qui en résulte peut alors être analysé *in vitro* pour son phénotype dans des processus tels que la phagocytose. Cependant, cette technique présente des limitations pour l'analyse d'autres types de réponses immunitaires moins facilement modélisables *in vitro* et, c'est en particulier le cas de la réponse antiparasitaire. C'est pourquoi nous avons adapté cette méthode pour les moustiques adultes, où l'expression des gènes peut être efficacement et spécifiquement inhibée par simple injection d'ARNdb (Blandin et al., 2002). Afin d'analyser l'effet de ces knock-downs sur divers aspects de l'immunité du moustique, nous avons mis en place un ensemble de tests phénotypiques qui mesurent:

- la survie du moustique après infection bactérienne,

• la phagocytose des bactéries *in vivo* (test qui a été établi par L. Moita et R. Wang, résultats non publiés)
• le développement du parasite dans les moustiques susceptibles
• la mélanisation du parasite dans les moustiques réfractaires.

Ces méthodes sont actuellement utilisées dans le laboratoire afin de réaliser le crible fonctionnel de plus de 100 gènes dont l'expression est induite lors d'une infection, mais dont la fonction est encore inconnue.

… chez la drosophile.

En outre, nous avons élargi l'application de cette méthode de knock-down par injection d'ARNdb à la drosophile adulte (Goto et al., 2003). Il est important de noter que l'efficacité du knock-out dépend de l'âge de la mouche, un phénomème que nous n'avons pas observé chez le moustique. Le knock-out est efficace lorsque des mouches de 4 jours ou plus sont injectées, et il dure plus d'une semaine. Des études épistatiques peuvent aussi être menées via l'injection d'un mélange de deux ARNdb. Dans un modèle comme celui de la drosophile, où de nombreux outils d'analyse génétique sont déjà disponibles, nous pensons que le knock-out de gènes par injection d'ARNdb est une méthode importante de génétique indirecte pour tester rapidement les fonctions des gènes, en particulier de ceux pour lesquels les mutations sont létales lors du développement.

Rôle de Defensin 1 dans la réponse immunitaire du moustique

La mise en place de ces outils nous a tout d'abord permis de caractériser fonctionnellement le peptide antimicrobien Defensin 1 (Blandin et al., 2002). L'analyse du knock-down a révélé que ce peptide joue un rôle essentiel dans la réponse immunitaire du moustique contre les bactéries à Gram-positif, mais pas dans la réponse contre les bactéries à Gram-négatif, et que Defensin1 n'est pas un facteur important de la réponse antiparasitaire.

Rôle de TEP1 dans la réponse immunitaire du moustique

Nous avons pu montrer que l'opsonisation des bactéries par TEP1 est nécessaire pour activer leur phagocytose par les hémocytes (Levashina et al., 2001, et données non publiées).

D'autre part, nous avons démontré que TEP1 est un facteur essentiel de la réponse antiparasitaire et qu'il est donc, de ce fait, impliqué dans l'établissement de la capacité vectorielle chez *A. gambiae* (Blandin et al., 2004c). En effet, TEP1 se fixe sur les oocinètes et déclenche leur élimination. Chez les moustiques susceptibles, le knockout de *TEP1* provoque une augmentation par 5 du nombre d'ookystes se développant sur le tube digestif, tandis qu'il transforme les moustiques réfractaires en moustiques susceptibles.

Les résultats de notre étude nous ont conduits à proposer le modèle suivant pour expliquer le rôle de TEP1 dans la réponse antiparasitaire de l'anophèle:

- Dans un premier temps, après avoir traversé l'épithélium intestinal, les parasites entrent en contact avec les composés solubles de l'hémolymphe, mais pas avec les hémocytes qui se situent de l'autre côté de la membrane basale. Un de ces composés, TEP1, reconnaît et de fixe à la surface des oocinètes, provoquant leur mort par un mécanisme inconnu.
- Dans un deuxième temps, les parasites morts sont éliminés soit par lyse dans les moustiques susceptibles, soit par lyse et mélanisation dans le cas de la souche réfractaire que nous avons étudiée. Le modèle prédit que ces mécanismes d'élimination ne dépendent pas de TEP1.

La fixation de TEP1 sur les parasites et leur élimination se produisent aussi bien chez les moustiques susceptibles que chez les moustiques réfractaires, mais ces phénomènes sont plus rapides chez les réfractaires et d'autre part, ils entraînent l'élimination de tous les parasites chez les moustiques réfractaires, alors qu'au moins 20% d'entre eux survivent chez les moustiques susceptibles. Lors de l'analyse du génome, nous avions remarqué que l'une des 4 séquences alléliques putatives, *TEP16*, était fortement similaire à *TEP1*. Nous avons démontré que *TEP1*, désormais *TEP1s*, était spécifiquement associée à la souche susceptible, tandis que *TEP16*, désormais *TEP1r,* était exclusivement détectée dans le génome des moustiques réfractaires. La comparaison des deux allèles a montré que la plupart des substitutions étaient concentrées dans la région de type C3d, qui contient le site thioester, et qui, dans les facteurs du complément, se lie directement au substrat. Le polymorphisme entre TEP1r et TEP1s pourrait, au moins partiellement, expliquer pourquoi le marquage des parasites par TEP1 et leur élimination sont plus rapides dans la souche réfractaire que dans la souche susceptible.

En conclusion, nous avons identifié une famille de protéines à thioester dans le moustique *A. gambiae* et caractérisé l'une d'entre elles, TEP1, au niveau fonctionnel. Nous avons démontré que la fixation de TEP1 sur la surface des pathogènes déclenche deux types de réactions immunitaires : la phagocytose des bactéries et la destruction des parasites. De nombreuses nouvelles questions ont été soulevées par ce travail : (i) Comment TEP1 est-elle activée ? (ii) Quels sont les mécanismes qui sont responsables de sa fixation préférentielle sur les surfaces des pathogènes ? (iii) Comment TEP1 déclenche t-elle la destruction des pathogènes sur lesquels elle est fixée ? (iv) Quel est la part du polymorphisme de *TEP1* dans l'établissement du phénotype réfractaire ? (v) Quels sont les rôles des autres TEPs de l'anophèle ? L'étude des mécanismes par lesquels TEP1 se fixe sur les parasites et provoque leur destruction, ainsi que des différences entre les deux allèles de *TEP1* pourrait être essentielle à la mise au point de nouvelles stratégies ayant pour but d'améliorer la réponse antiparasitaire du moustique pendant les premiers jours post infection, au moment où la réponse immunitaire est la plus efficace et où le nombre de parasites est à son minimum.

REFERENCES

References

Adams, M. D., and Sekelsky, J. J. (2002). From sequence to phenotype: reverse genetics in *Drosophila melanogaster*. Nat Rev Genet *3*, 189-198.

Ambros, V. (2003). MicroRNA pathways in flies and worms: growth, death, fat, stress, and timing. Cell *113*, 673-676.

Avila, A., Silverman, N., Diaz-Meco, M. T., and Moscat, J. (2002). The *Drosophila* atypical protein kinase C-ref(2)p complex constitutes a conserved module for signaling in the toll pathway. Mol Cell Biol *22*, 8787-8795.

Billker, O., Lindo, V., Panico, M., Etienne, A. E., Paxton, T., Dell, A., Rogers, M., Sinden, R. E., and Morris, H. R. (1998). Identification of xanthurenic acid as the putative inducer of malaria development in the mosquito. Nature *392*, 289-292.

Blandin, S., Moita, L. F., Kocher, T., Wilm, M., Kafatos, F. C., and Levashina, E. A. (2002). Reverse genetics in the mosquito *Anopheles gambiae*: targeted disruption of the *Defensin* gene. EMBO Rep *3*, 852-856.

Blandin, S., and Levashina, E. A. (2004a). Mosquito immune responses against malaria parasites. Curr Opin Immunol *16*, 16-20.

Blandin, S., and Levashina, E. A. (2004b). Thioester-containing proteins and insect immunity. Mol Immunol *40*, 903-908.

Blandin, S., Shiao, S.-H., Moita, L. F., Janse, C., Waters, A. P., Kafatos, F. C., and Levashina, E. A. (2004c). Complement-like protein TEP1 is a determinant of vectorial capacity in the malaria vector *Anopheles gambiae*. Cell, accepted.

Bulet, P., Hetru, C., Dimarcq, J. L., and Hoffmann, D. (1999). Antimicrobial peptides in insects; structure and function. Dev Comp Immunol *23*, 329-344.

Catteruccia, F., Nolan, T., Blass, C., Muller, H. M., Crisanti, A., Kafatos, F. C., and Loukeris, T. G. (2000a). Toward *Anopheles* transformation: Minos element activity in anopheline cells and embryos. Proc Natl Acad Sci U S A *97*, 2157-2162.

Catteruccia, F., Nolan, T., Loukeris, T. G., Blass, C., Savakis, C., Kafatos, F. C., and Crisanti, A. (2000b). Stable germline transformation of the malaria mosquito *Anopheles stephensi*. Nature *405*, 959-962.

Choe, K. M., Werner, T., Stoven, S., Hultmark, D., and Anderson, K. V. (2002). Requirement for a peptidoglycan recognition protein (PGRP) in Relish activation and antibacterial immune responses in *Drosophila*. Science *296*, 359-362.

Christophides, G. K., Zdobnov, E., Barillas-Mury, C., Birney, E., Blandin, S., Blass, C., Brey, P. T., Collins, F. H., Danielli, A., Dimopoulos, G., et al. (2002). Immunity-related genes and gene families in *Anopheles gambiae*. Science *298*, 159-165.

Collins, F. H., Sakai, R. K., Vernick, K. D., Paskewitz, S., Seeley, D. C., Miller, L. H., Collins, W. E., Campbell, C. C., and Gwadz, R. W. (1986). Genetic selection of a *Plasmodium*-refractory strain of the malaria vector *Anopheles gambiae*. Science *234*, 607-610.

Craig, A., Kyes, S., Ranson, H., and Hemingway, J. (2003). Malaria parasite and vector genomes: partners in crime. Trends Parasitol *19*, 356-362.

Denli, A. M., and Hannon, G. J. (2003). RNAi: an ever-growing puzzle. Trends Biochem Sci *28*, 196-201.

Dimopoulos, G., Casavant, T. L., Chang, S., Scheetz, T., Roberts, C., Donohue, M., Schultz, J., Benes, V., Bork, P., Ansorge, W., 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*, 6619-6624.

Dimopoulos, G., Christophides, G. K., Meister, S., Schultz, J., White, K. P., Barillas-Mury, C., and Kafatos, F. C. (2002). Genome expression analysis of *Anopheles gambiae*: responses to injury, bacterial challenge, and malaria infection. Proc Natl Acad Sci U S A *99*, 8814-8819.

Dimopoulos, G., Muller, H. M., Levashina, E. A., and Kafatos, F. C. (2001). Innate immune defense against malaria infection in the mosquito. Curr Opin Immunol *13*, 79-88.

Dimopoulos, G., Seeley, D., Wolf, A., and Kafatos, F. C. (1998). Malaria infection of the mosquito *Anopheles gambiae* activates immune-responsive genes during critical transition stages of the parasite life cycle. Embo J *17*, 6115-6123.

Dodds, A. W., and Law, S. K. (1998). The phylogeny and evolution of the thioester bond-containing proteins C3, C4 and alpha 2-macroglobulin. Immunol Rev *166,* 15-26.

Dzitoyeva, S., Dimitrijevic, N., and Manev, H. (2001). Intra-abdominal injection of double-stranded RNA into anesthetized adult *Drosophila* triggers RNA interference in the central nervous system. Mol Psychiatry *6*, 665-670.

Ferrandon, D., Jung, A. C., Criqui, M., Lemaitre, B., Uttenweiler-Joseph, S., Michaut, L., Reichhart, J., and Hoffmann, J. A. (1998). A d*rosomycin-GFP* reporter transgene reveals a local immune response in *Drosophila* that is not dependent on the Toll pathway. Embo J *17*, 1217-1227.

Fire, A., Xu, S., Montgomery, M. K., Kostas, S. A., Driver, S. E., and Mello, C. C. (1998). Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. Nature *39*1, 806-811.

Franke-Fayard, B., Khan, S. M., Ramesar, J., Mendoza, J., van der Keur, M., van der Linden, R., Sinden, R. E., Waters, A. P., and Janse, C. J. (2004). A *Plasmodium berghei* reference line that constitutively expresses *GFP* at a high level throughout the complete life cycle. submitted.

Gay, P., and Contamine, D. (1993). Study of the *ref(2)P* locus of *Drosophila melanogaster*. II. Genetic studies of the 37DF region. Mol Gen Genet *239*, 361-370.

Gobert, V., Gottar, M., Matskevich, A. A., Rutschmann, S., Royet, J., Belvin, M., Hoffmann, J. A., and Ferrandon, D. (2003). Dual activation of the *Drosophila* toll pathway by two pattern recognition receptors. Science *302*, 2126-2130.

Gorman, M. J., Schwartz, A. M., and Paskewitz, S. (1998). The role of surface characteristics in eliciting humoral encapsulation of foreign bodies in *Plasmodium*-refractory and -susceptible strains of *Anopheles gambiae*. Journal of Insect Physiology *44*, 947-954.

Goto, A., Blandin, S., Royet, J., Reichhart, J. M., and Levashina, E. A. (2003). Silencing of Toll pathway components by direct injection of double-stranded RNA into *Drosophila* adult flies. Nucleic Acids Res *31*, 6619-6623.

Gottar, M., Gobert, V., Michel, T., Belvin, M., Duyk, G., Hoffmann, J. A., Ferrandon, D., and Royet, J. (2002). The *Drosophila* immune response against Gram-negative bacteria is mediated by a peptidoglycan recognition protein. Nature *416*, 640-644.

Greenwood, B., and Mutabingwa, T. (2002). Malaria in 2002. Nature *415*, 670-672.

Grossman, G. L., Rafferty, C. S., Clayton, J. R., Stevens, T. K., Mukabayire, O., and Benedict, M. Q. (2001). Germline transformation of the malaria vector, *Anopheles gambiae*, with the piggyBac transposable element. Insect Mol Biol *10*, 597-604.

Hammond, S. M., Bernstein, E., Beach, D., and Hannon, G. J. (2000). An RNA-directed nuclease mediates post-transcriptional gene silencing in *Drosophila* cells. Nature *404*, 293-296.

Hoffmann, J. A. (2003). The immune response of *Drosophila*. Nature *426*, 33-38.

Holt, R. A., Subramanian, G. M., Halpern, A., Sutton, G. G., Charlab, R., Nusskern, D. R., Wincker, P., Clark, A. G., Ribeiro, J. M., Wides, R., et al. (2002). The genome sequence of the malaria mosquito *Anopheles gambiae*. Science *298*, 129-149.

Ito, J., Ghosh, A., Moreira, L. A., Wimmer, E. A., and Jacobs-Lorena, M. (2002). Transgenic anopheline mosquitoes impaired in transmission of a malaria parasite. Nature *417*, 452-455.

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

Janeway, C. A., Travers, P., Walport, M. J., and Schlomchik, M. J. (2001). Innate Immunity. In Immunobiology: the Immune system in Health and Disease (New York, Garland Publishing).

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

Kaminishi, H., Miyaguchi, H., Tamaki, T., Suenaga, N., Hisamatsu, M., Mihashi, I., Matsumoto, H., Maeda, H., and Hagihara, Y. (1995). Degradation of humoral host defense by *Candida albicans* proteinase. Infect Immun *63*, 984-988.

Lagueux, M., Perrodou, E., Levashina, E. A., Capovilla, M., and Hoffmann, J. A. (2000). Constitutive expression of a complement-like protein in *toll* and *JAK* gain-of-function mutants of *Drosophila*. Proc Natl Acad Sci U S A *97*, 11427-11432.

Lemaitre, B., Kromer-Metzger, E., Michaut, L., Nicolas, E., Meister, M., Georgel, P., Reichhart, J. M., and Hoffmann, J. A. (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*, 9465-9469.

Lemaitre, B., Nicolas, E., Michaut, L., Reichhart, J. M., and Hoffmann, J. A. (1996). The dorsoventral regulatory gene cassette spatzle/Toll/cactus controls the potent antifungal response in *Drosophila* adults. Cell *86*, 973-983.

Lemaitre, B., Reichhart, J. M., and Hoffmann, J. A. (1997). *Drosophila* host defense: differential induction of antimicrobial peptide genes after infection by various classes of microorganisms. Proc Natl Acad Sci U S A *94*, 14614-14619.

Levashina, E. A., Moita, L. F., Blandin, S., Vriend, G., Lagueux, M., and Kafatos, F. C. (2001). Conserved role of a complement-like protein in phagocytosis revealed by dsRNA knockout in cultured cells of the mosquito, *Anopheles gambiae*. Cell *104*, 709-718.

Levashina, E. A., Blandin, S., Moita, L. F., Vriend, G., Lagueux, M., and Kafatos, F. C. (2003). Thioester-containing proteins of protostomes. In: Ezekowitz, R. A. B., Hoffmann, J. A. (Eds) Infectious disease: Innate Immunity. *Humana Press Inc., Totowa, NJ*, pp155-173.

Matzke, M., and Matzke, A. J. (2003). RNAi extends its reach. Science *301*, 1060-1061.

Meister, M. (2004). Blood cells of *Drosophila*: cell lineages and rol in host defence. Curr Opin Immunol *16*, 1-6.

Mette, M. F., Aufsatz, W., van der Winden, J., Matzke, M. A., and Matzke, A. J. (2000). Transcriptional silencing and promoter methylation triggered by double-stranded RNA. Embo J *19*, 5194-5201.

Michel, T., Reichhart, J. M., Hoffmann, J. A., and Royet, J. (2001). *Drosophila* Toll is activated by Gram-positive bacteria through a circulating peptidoglycan recognition protein. Nature *414*, 756-759.

Mukabayire, O., Cornel, A. J., Dotson, E. M., Collins, F. H., and Besansky, N. J. (1996). The Tryptophan oxygenase gene of *Anopheles gambiae.* Insect Biochem Mol Biol *26*, 525-528.

Naitza, S., and Ligoxygakis, P. (2004). Antimicrobial defences in *Drosophila*: the story so far. Mol Immunol *40*, 887-896.

Nonaka, M., and Yoshizaki, F. (2004). Evolution of the complement system. Mol Immunol *40*, 897-902.

Nusslein-Volhard, C., and Wieschaus, E. (1980). Mutations affecting segment number and polarity in *Drosophila*. Nature *287*, 795-801.

Oduol, F., Xu, J., Niare, O., Natarajan, R., and Vernick, K. D. (2000). Genes identified by an expression screen of the vector mosquito *Anopheles gambiae* display differential molecular immune response to malaria parasites and bacteria. Proc Natl Acad Sci U S A *97*, 11397-11402.

Osta, M. A., Christophides, G. K., and Kafatos, F. C. (2004). Positive and negative effects of mosquito genes on *Plasmodium* development. Science *In press*.

Ramet, M., Manfruelli, P., Pearson, A., Mathey-Prevot, B., and Ezekowitz, R. A. (2002). Functional genomic analysis of phagocytosis and identification of a *Drosophila* receptor for *E. coli*. Nature *416*, 644-648.

Richie, T. L., and Saul, A. (2002). Progress and challenges for malaria vaccines. Nature *415*, 694-701.

Richman, A. M., Bulet, P., Hetru, C., Barillas-Mury, C., Hoffmann, J. A., and Kafalos, F. C. (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*, 203-210.

Richman, A. M., Dimopoulos, G., Seeley, D., and Kafatos, F. C. (1997). *Plasmodium* activates the innate immune response of *Anopheles gambiae* mosquitoes. Embo J *16*, 6114-6119.

Sachs, J., and Malaney, P. (2002). The economic and social burden of malaria. Nature *415*, 680-685.

Sakai, R. K., and Miller, L. H. (1992). Effects of heat shock on the survival of transgenic *Anopheles gambiae* (Diptera: Culicidae) under antibiotic selection. J Med Entomol *29*, 374-375.

Schramke, V., and Allshire, R. (2003). Hairpin RNAs and retrotransposon LTRs effect RNAi and chromatin-based gene silencing. Science *301*, 1069-1074.

Shahabuddin, M., Fields, I., Bulet, P., Hoffmann, J. A., and Miller, L. H. (1998). *Plasmodium gallinaceum*: differential killing of some mosquito stages of the parasite by insect defensin. Exp Parasitol *89*, 103-112.

Sinden, R. E. (1999). *Plasmodium* differentiation in the mosquito. Parassitologia *41*, 139-148.

Sinden, R. E. (2002). Molecular interactions between *Plasmodium* and its insect vectors. Cell Microbiol *4*, 713-724.

Song, X. M., Perez-Casal, J., Bolton, A., and Potter, A. A. (2001). Surface-expressed mig protein protects *Streptococcus dysgalactiae* against phagocytosis by bovine neutrophils. Infect Immun *69*, 6030-6037.

St Johnston, D. (2002). The art and design of genetic screens: *Drosophila melanogaster*. Nat Rev Genet *3*, 176-188.

Tzou, P., Ohresser, S., Ferrandon, D., Capovilla, M., Reichhart, J. M., Lemaitre, B., Hoffmann, J. A., and Imler, J. L. (2000). Tissue-specific inducible expression of antimicrobial peptide genes in *Drosophila* surface epithelia. Immunity *13*, 737-748.

van der Krol, A. R., Mur, L. A., Beld, M., Mol, J. N., and Stuitje, A. R. (1990). Flavonoid genes in petunia: addition of a limited number of gene copies may lead to a suppression of gene expression. Plant Cell *2*, 291-299.

Vargas, L. E. (1949). Culine and aedine mosquitoes and the malaria infections of lower animals. In Malariology, M. F. Boyd, ed. (Philadelphia & London, WB Saunders Company), pp. 526-538.

Vernick, K. D., Fujioka, H., Seeley, D. C., Tandler, B., Aikawa, M., and Miller, L. H. (1995). *Plasmodium gallinaceum*: a refractory mechanism of ookinete killing in the mosquito, *Anopheles gambiae*. Exp Parasitol *80*, 583-595.

Vizioli, J., Bulet, P., Charlet, M., Lowenberger, C., Blass, C., Muller, H. M., Dimopoulos, G., Hoffmann, J., Kafatos, F. C., and Richman, A. (2000). Cloning and analysis of a cecropin gene from the malaria vector mosquito, *Anopheles gambiae*. Insect Mol Biol *9*, 75-84.

Vizioli, J., Bulet, P., Hoffmann, J. A., Kafatos, F. C., Muller, H. M., and Dimopoulos, G. (2001a). Gambicin: a novel immune responsive antimicrobial peptide from the malaria vector *Anopheles gambiae*. Proc Natl Acad Sci U S A *98*, 12630-12635.

Vizioli, J., Richman, A. M., Uttenweiler-Joseph, S., Blass, C., and Bulet, P. (2001b). The defensin peptide of the malaria vector mosquito *Anopheles gambiae*: antimicrobial activities and expression in adult mosquitoes. Insect Biochem Mol Biol *31*, 241-248.

Wodarz, A., Ramrath, A., Grimm, A., and Knust, E. (2000). *Drosophila* atypical protein kinase C associates with Bazooka and controls polarity of epithelia and neuroblasts. J Cell Biol *150*, 1361-1374.

Woo, P. T. (2003). *Cryptobia* (Trypanoplasma) *salmositica* and salmonid cryptobiosis. J Fish Dis *26*, 627-646.

Zasloff, M. (2002). Antimicrobial peptides of multicellular organisms. Nature *415*, 389-395.

Zheng, L., Cornel, A. J., Wang, R., Erfle, H., Voss, H., Ansorge, W., Kafatos, F. C., and Collins, F. H. (1997). Quantitative trait loci for refractoriness of *Anopheles gambiae* to *Plasmodium cynomolgi* B. Science *276*, 425-428.

Zheng, L., Wang, S., Romans, P., Zhao, H., Luna, C., and Benedict, M. Q. (2003). Quantitative trait loci in *Anopheles gambiae* controlling the encapsulation response against *Plasmodium cynomolgi* Ceylon. BMC Genet *4*, 16.