Molecular Strategies for Active Host Cell Invasion by Apicomplexan Parasites

by

Michelle Tonkin
BSc, Abilene Christian University, 2009

A Dissertation Submitted in Partial Fulfillment
of the Requirements for the Degree of

DOCTOR OF PHILOSOPHY

in the Department of Biochemistry and Microbiology

© Michelle Tonkin, 2014
University of Victoria

Supervisory Committee

Molecular Strategies for Active Host Cell Invasion by Apicomplexan Parasites

by

Michelle Tonkin
BSc, Abilene Christian University, 2009

Supervisory Committee

Dr. Martin J. Boulanger (Department of Biochemistry and Microbiology)
Supervisor

Dr. Caroline E. Cameron (Department of Biochemistry and Microbiology)
Departmental Member

Dr. Terry W. Pearson (Department of Biochemistry and Microbiology)
Departmental Member

Dr. Jeremy E. Wulff (Department of Chemistry)
Outside Member

Abstract

Supervisory Committee

Dr. Martin J. Boulanger (Department of Biochemistry and Microbiology)
Supervisor
Dr. Caroline E. Cameron (Department of Biochemistry and Microbiology)
Departmental Member
Dr. Terry W. Pearson (Department of Biochemistry and Microbiology)
Departmental Member
Dr. Jeremy E. Wulff (Department of Chemistry)
Outside Member

Parasites of phylum Apicomplexa cause devastating diseases on a global scale. *Toxoplasma gondii*, the etiological agent of toxoplasmosis, and *Plasmodium falciparum*, the most virulent agent of human malaria, have the most substantial effects on human health and are the most widely studied. The success of these parasites is due in part to a sophisticated molecular arsenal that supports a variety of novel biological processes including a unique form of host cell invasion. Accessing the protective environment of the host cell is paramount to parasite survival and is mediated through an active invasion process: the parasite propels itself through a circumferential ring known as the moving junction (MJ) formed between its apical tip and the host cell membrane. The MJ ring is comprised of a parasite surface protein (AMA1) that engages a protein secreted by the parasite into the host cell and presented on the host cell surface (RON2). Thus, through an intriguing mechanism the parasite provides both receptor and ligand to enable host cell invasion. Prior to the studies described herein, the characterization of the AMA1-RON2 association was limited to low-resolution experiments that provided little insight into the functional and architectural details of this crucial binary complex. Towards elucidating the mechanism of AMA1-RON2 dependent invasion, I first structurally characterized *T. gondii* AMA1 bound to the corresponding binding region of RON2; analysis of the AMA1-RON2 interface along with biophysical data revealed an intimate association likely capable of withstanding the shearing forces generated as the parasite dives through the constricted MJ ring. To investigate the role of the AMA1-RON2 complex across genera, species and life-cycle stages, I next characterized the AMA1-RON2 complex from a distantly related genus within Apicomplexa (*Plasmodium*) and from a divergent pairing within *T. gondii*. By combining structural, biophysical and biological data, I was able to generate a detailed model describing the role of AMA1 and RON2 in MJ

iv

dependent invasion, which is currently supporting efforts to develop novel vaccines and cross-reactive small molecule therapeutics.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

AIDS	Acquired immunodeficiency syndrome
AMA	Apical membrane antigen
Bb	*Babesia bovis*
Bd	*Babesia divergens*
BSA	Buried surface area
CLS	Canadian Light Source
DI	Domain I
DII	Domain II
DIII	Domain III
DMEM	Dulbecco's modified Eagle's medium
DMSO	Dimethyl sulfoxide
EGF	Epidermal growth factor
ELISA	Enzyme-linked immunosorbent assay
Et	*Eimeria tenella*
FBS	Fetal bovine serum
FITC	Fluorescein isothiocyanate
GST	Glutathione-S-transferase
HBS	HEPES-buffered saline
HFF	Human foreskin fibroblasts
His_6	Hexa-histidine tag
HIV	Human immunodeficiency virus
HRP	Horseradish peroxidase
HXGPRT	Hypoxanthine-xanthine-guanine phosphoribosyltransferase
IFA	Immunofluorescence assay
Ig	Immunoglobulin
ITC	Isothermal titration calorimetry
mAb	Monoclonal antibody
MAEBL	Merozoite apical erythrocyte-binding ligand
MJ	Moving junction
MR	Molecular replacement
MWCO	Molecular weight cut-off
Nc	*Neospora caninum*
NMR	Nuclear magnetic resonance
P1, P2, P3	Primary, secondary, and tertiary amplified baculovirus
PAF	Paraformaldehyde picric acid
PAN	Plasminogen/apple/nematode
PBS	Phosphate-buffered saline
PCR	Polymerase chain reaction

PDB	Protein data bank
PEG	Polyethylene glycol
Pf	*Plasmodium falciparum*
Pk	*Plasmodium knowlesi*
Pv	*Plasmodium vivax*
PV	Parasitophorous vacuole
Py	*Plasmodium yoelii*
rmsd	Root mean square deviation
RON	Rhoptry neck protein
RON2sp	RON2 synthetic peptide
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
*Sf*9	*Spodoptera frugiperda 9*
SFM	Serum-free medium
SPR	Surface plasmon resonance
SSRL	Stanford Synchrotron Radiation Lightsource
Tg	*Toxoplasma gondii*
*Tg*Sporo	*T. gondii* sporozoite
TRX	Thioredoxin

Acknowledgements

I'll start by saying a huge thank you to my fantastic supervisor! Dr. Marty Boulanger has been an inspiration throughout my time in the PhD program. His intelligence, support, openness to talk through ideas, and willingness to provide me with opportunities to stretch my abilities have all been invaluable.

A sincere thank you to my committee members – Dr. Terry Pearson, Dr. Caroline Cameron, and Dr. Jeremy Wulff. I truly appreciate the investment of their time and intellect.

Thank you to all my lab mates, both past and present. Special thanks to a few very special people – Jo, Jasleen, Adrienne, Susann, Bianca, Fangni, Sean, and Drew.

Thanks also to my undergraduate supervisor, Dr. Greg Powell, who inspired me to further pursue scientific research.

A big thank you to my wonderful family ☺

Finally, thank you to my amazing fiancé, Matt, for his unconditional encouragement and support - I could not have done this without him (or without our perpetually enthusiastic and cuddly puppy, Zoey).

Dedication

I would like to dedicate my research to the Zambia Mission Fund (ZMF); the extraordinary people working in Zambia with ZMF are immersed in the effects of poverty, AIDS, and apicomplexan disease, particularly malaria, and are devoted to providing assistance on every level. I truly hope that my research will help inform the design and development of new therapeutics that make a difference for the people who need them most.

Chapter 1: Introduction

1.1 Apicomplexan parasites: a global burden

Phylum Apicomplexa contains a diverse group of more than five thousand obligate intracellular parasitic protozoan species, many of which cause devastating diseases on a global scale. The pervasive organisms in this phylum include *Cryptosporidium*, one of the most significant waterborne parasites causing an estimated three quarters of a million annual cases of cryptosporidiosis in the United States alone (Scallan, Hoekstra et al. 2011); *Eimeria*, the etiological agent of poultry coccidiosis presenting a major economic burden to the poultry industry, with a global cost of at least 1.5 billion USD (Ruff 1999); and *Babesia*, one of the most prevalent parasites in free-living cattle (Homer, Aguilar-Delfin et al. 2000, Zintl, Mulcahy et al. 2003). The focus of this dissertation is on the two genera that are the most widely studied and have the most substantial impact on human health, *Toxoplasma* and *Plasmodium*, which are the etiological agents of toxoplasmosis and malaria, respectively. The genus *Toxoplasma* contains just one species, *T. gondii*, which chronically infects up to a third of the world's human population and is especially dangerous for immunocompromised people and congenitally infected neonates (Jackson and Hutchison 1989, Tenter, Heckeroth et al. 2000). In contrast, *Plasmodium* contains more than 200 species, but *P. falciparum* alone causes the most severe human malaria, with more than 250 million new cases every year leading to at least one million deaths (Snow, Guerra et al. 2005, World Health Organization. 2010).

1.2 Apicomplexan classification, ultra-structure, and life cycles

1.2.1 Apicomplexan classification and ultra-structure

With few exceptions, classification into phylum Apicomplexa requires the presence of the apical complex, which is a set of organelles and structures involved in the intricate process of host cell invasion (Dubremetz, Garcia-Réguet et al. 1998); the apicoplast, a chloroplast-like organelle with varied functions; nine singlet centrioles; subcortical microtubules; and a system of flattened alveoli underlying the plasma membrane (Cavalier-Smith 1993, Kim and Weiss 2004). Within phylum Apicomplexa, the two predominant classes are defined

based on the absence or presence of the apical conoid structure: Aconoidasida and Conoidasida (Levine 1988). Class Aconoidasida is populated by organisms such as *Plasmodium* and *Babesia*, while class Conoidasida is populated primarily by the coccidians, which includes genera such as *Toxoplasma*, *Neospora*, *Eimeria*, and *Cryptosporidium*.

As eukaryotes, all apicomplexans maintain a variety of membrane bound organelles, including the nucleus, rough endoplasmic reticulum, Golgi complex, and a single mitochondrion (Figure 1). As noted above, the apicomplexans also have a number of unique structures and organelles. The presence of the apical complex is a predominant distinguishing feature of the apicomplexans, and contains the apical polar ring, micronemes and rhoptries, as well as the conoid structure for members of class Conoidasida; while dense granules are not exclusively apically located they are still considered part of the apical complex (Figure 1) (Dubey, Lindsay et al. 1998, Hu, Johnson et al. 2006). Of particular interest to this study are the micronemes and rhoptries that contain proteins essential for host cell invasion; micronemes are small oval-shaped electron dense organelles condensed at the apical end of the parasite, while rhoptries are larger club-shaped organelles that taper toward the apex and can be further divided into neck and bulb regions although no separating membrane is present (Nichols, Chiappino et al. 1983, Carruthers 1999, Bannister, Hopkins et al. 2000, Boothroyd and Dubremetz 2008). The ultra-structural features depicted in Figure 1 are generally maintained by the major invasive life cycle stages of *T. gondii* (tachyzoites, bradyzoites, and sporozoites) and *P. falciparum* (merozoites and sporozoites) although the size, organization, quantity, and internal composition of the organelles differ significantly between the various stages (Aikawa 1977, Dubey, Lindsay et al. 1998, Bannister, Hopkins et al. 2000).

Apical polar ring
[Conoid]
Micronemes

Rhoptry neck and bulb

Plasma membrane

Inner membrane complex

Mitochondrion

Golgi complex

Dense granules

Apicoplast

Nucleus

Rough endoplasmic reticulum

[Posterior pore]
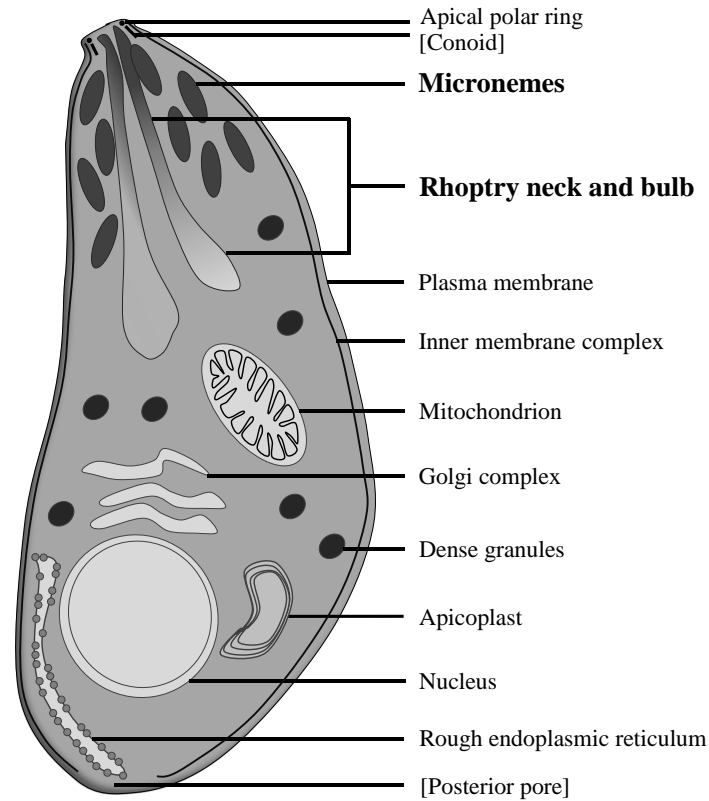
Figure 1: Schematic of the general ultra-structural characteristics of invasive apicomplexan parasite forms, displayed on a *T. gondii* tachyzoite.

Organelles and structures present in nearly all apicomplexan invasive life cycle stages are shown without brackets, while organelles and structures unique to the Conoidasida class are shown in square brackets. *T. gondii* tachyzoites tend to be elongated (~6 to 10 µm long), while *P. falciparum* merozoites are notably smaller (~1.6 µm) and more spherical but still maintain all of the organelles and structures not delineated by square brackets (Dubey, Lindsay et al. 1998, Bannister, Hopkins et al. 2000). In bold are the organelles that will be further discussed in this dissertation.

1.2.2 T. gondii life cycle

T. gondii is an obligate intracellular parasite with a very complex life cycle that consists of three invasive parasite stages (tachyzoite, bradyzoite, and sporozoite) collectively able to infect any nucleated cell within any warm-blooded host, disseminate through the host and even cross the blood-brain barrier, form chronic dormant infections, and be released into

the environment where encysted forms can survive for extended periods even in harsh conditions (Figure 2) (Dubey 2004, Robert-Gangneux and Darde 2012).
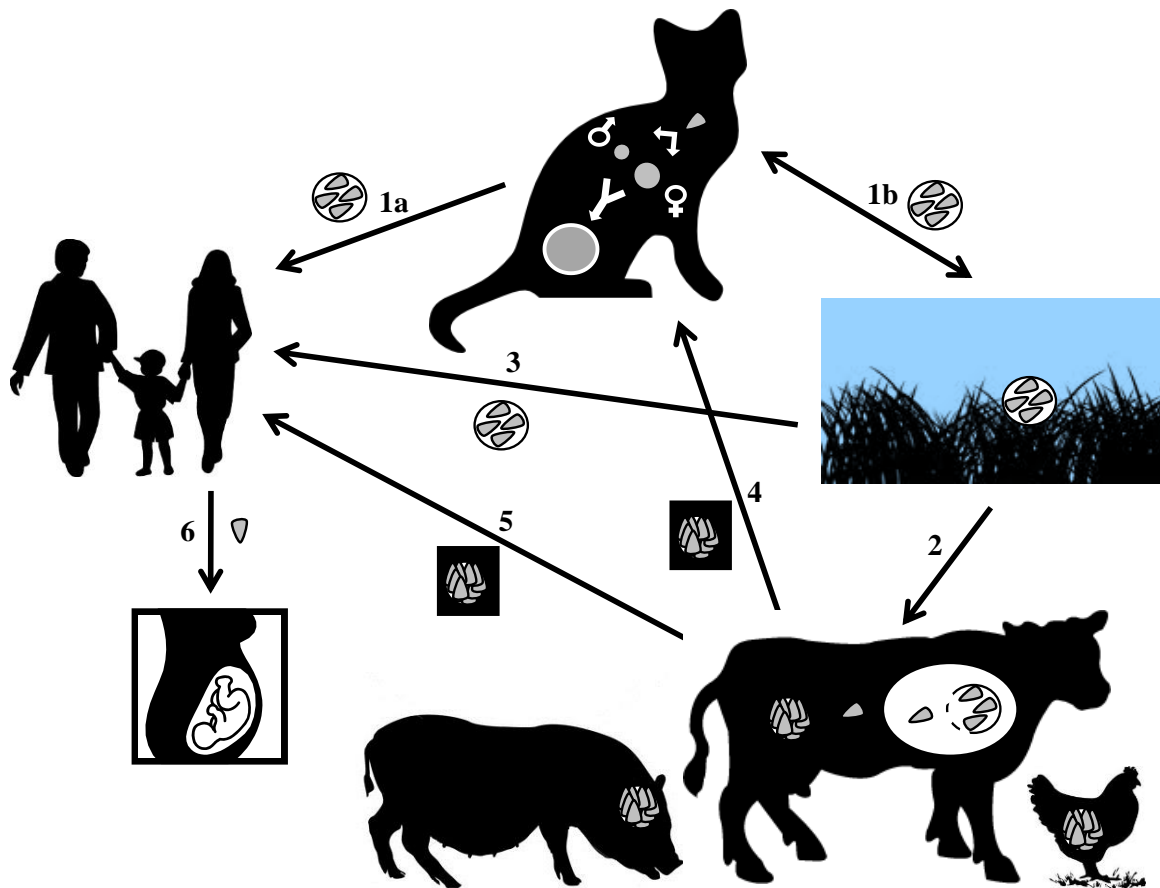
Figure 2: Schematic of the major steps in the life cycle and transmission of *T. gondii*. See main text for description of life cycle stages. For clarity, images are not to scale. Within cat – large grey circle, unsporulated oocyst; smaller grey circles, micro and macrogametes. Encircled grey parasites, encysted sporozoites (sporulated oocysts). Clustered grey parasites, tissue cyst of bradyzoites. Single free parasite, tachyzoite or invasive stage bradyzoite/sporozoite.

Members of the feline family are the definitive host of *T. gondii*; micro and macrogametes develop exclusively in the cat gut, and sexual reproduction leads to the development of unsporulated oocysts that are passed in feces (Figure 2, top) (Dubey and Frenkel 1972, Dubey 1998). A single cat can shed up to 100 million oocysts that will

sporulate in the environment (each sporulated oocyst contains eight sporozoites), ready for transmission and infection (Tenter, Heckeroth et al. 2000). In the case of domesticated cats, these oocysts can be transferred directly to humans (Figure 2 – line 1a), but for feral cats these oocysts are shed into the water or soil and can be ingested by other feral cats (Figure 2 – line 1b) (Tenter, Heckeroth et al. 2000). Oocysts in the environment are very stable and can contaminate food or water destined for animal (Figure 2 – line 2) or human (Figure 2 – line 3) consumption. Once ingested by an intermediate host, invasive sporozoites are released from the sporulated oocysts. Sporozoites will transition to tachyzoites to generate a quickly spreading infection, and some of these tachyzoites will transition to bradyzoites and form latent tissue cysts. The tissue cysts of *T. gondii* bradyzoites are generally found in the brain or muscle, and the number of parasites in a single tissue cyst can vary from just two to a few thousand depending on the age of the cyst (Dubey, Lindsay et al. 1998). If the portion of an animal containing a tissue cyst is ingested by a cat (Figure 2 – line 4) the life cycle is complete. However, these infectious tissue cysts can also be consumed by humans in the form of undercooked meat (Figure 2 – line 5) (Tenter, Heckeroth et al. 2000). Bradyzoites released from tissue cysts by proteolytic digestion of the cyst wall traverse the intestinal barrier and transition to tachyzoites, which disseminate throughout the body (Dubey 1998, Lyons, McLeod et al. 2002). If the infected human is a pregnant woman in her first trimester who has not previously been infected with *T. gondii*, the mother can pass tachyzoites through the placenta to the fetus resulting in congenital toxoplasmosis, which is generally a dead-end for the parasite life cycle (Figure 2 – line 6) (Desmonts, Daffos et al. 1985, Tenter, Heckeroth et al. 2000).

1.2.3 P. falciparum life cycle

Like *T. gondii*, *P. falciparum* is an obligate intracellular parasite, but *P. falciparum* is a vector-borne pathogen exhibiting a complicated life cycle that requires infection of both a human host and a mosquito transmission vector (Figure 3) (Miller, Baruch et al. 2002, Suh, Kain et al. 2004).
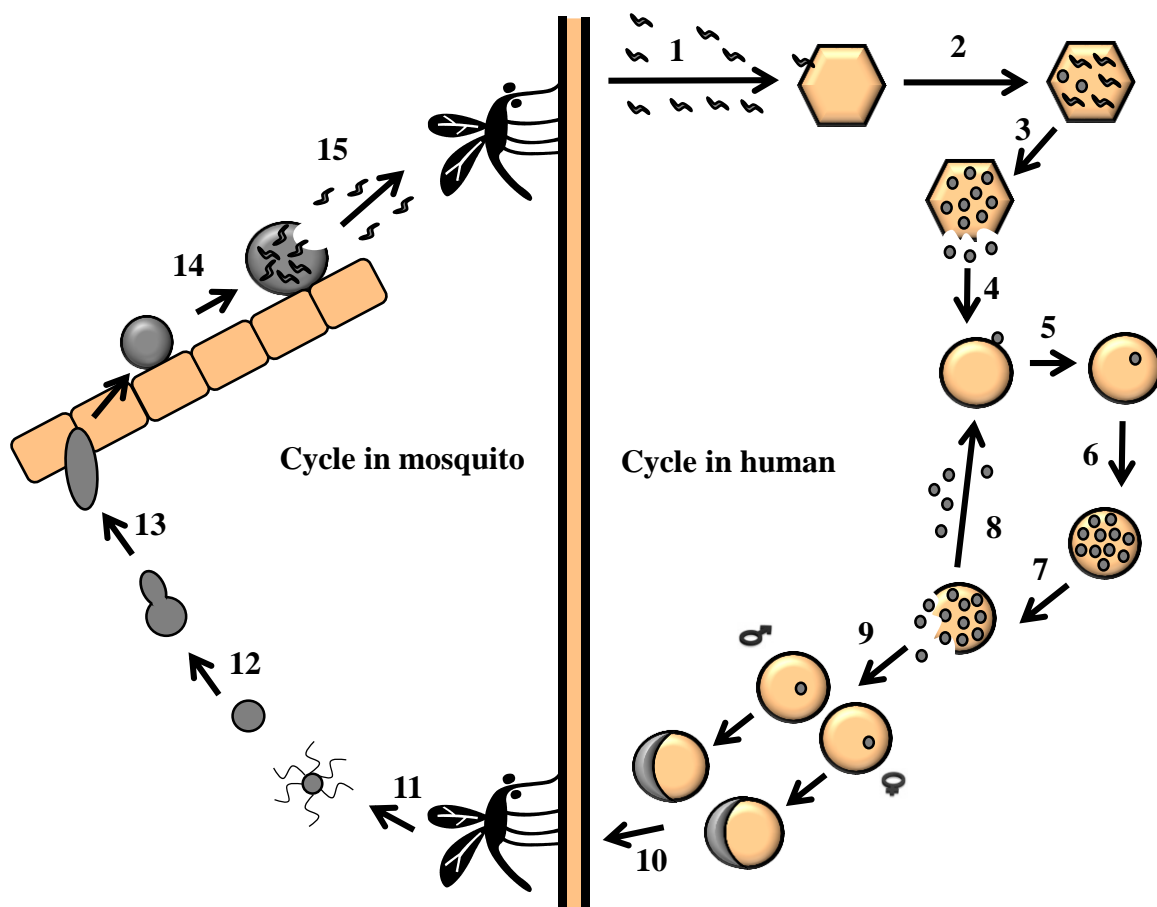
Figure 3: Schematic of the *P. falciparum* life cycle.

See main text for description of life cycle stages. For clarity, images are not to scale. Grey squiggles, sporozoites; beige hexagon, hepatocyte; small grey circles, merozoites; beige circle, erythrocyte; beige circles with moon shape, gametocytes; grey circle with flagella, microgamete; plain grey circle, macrogamete; grey circle with oblong head, zygote; grey oval, ookinete; beige rectangles, mosquito midgut; beveled grey circle, oocyst.

During the *Plasmodium* life cycle, humans become infected with the sporozoite form through the bite of an infected mosquito (Miller, Baruch et al. 2002). Sporozoites migrate from the site of infection to the liver, where they each pass through several hepatocytes before invading a final hepatocyte (Figure 3 – line 1). A single intrahepatocytic sporozoite will replicate and differentiate into tens of thousands of merozoites (Figure 3 – line 2), thereby forming a mature liver schizont (Figure 3 – line 3). In some species of *Plasmodium* (*P. vivax* and *P. ovale*), a few schizonts may remain in the liver in a dormant state and cause a malaria relapse weeks to years later (Suh, Kain et al. 2004). In the case

of *P. falciparum*, the schizonts will burst and the released merozoites subsequently enter the blood stream (Figure 3 – line 4) and invade erythrocytes (Figure 3 – line 5). Multiple rounds of asexual replication over about forty-eight hours lead to an erythrocyte schizont containing approximately twenty parasites (Figure 3 – line 6), which ruptures and releases the merozoites into the bloodstream (Figure 3 – line 7) (Miller, Baruch et al. 2002). Most parasites re-invade in a cycle that leads to acute disease pathology (Figure 3 – line 8), while some intraerythrocytic parasites differentiate into non-pathogenic sexual gametocytes (Figure 3 – line 9). These gametocytes are taken up by a female *Anopheles* mosquito (Figure 3 – line 10), where they mature into micro and macrogametes (Figure 3 – line 11) that undergo fertilization to form a zygote (Figure 3 – line 12). The zygote matures into an ookinete that is capable of invading the mosquito midgut epithelium (Figure 3 – line 13) and subsequently differentiates into a sporozoite-containing oocyst (Figure 3 – line 14). The sporozoites continue to replicate, burst the oocyst, and travel to the salivary glands ready for transmission to the human host, thus completing the life cycle (Figure 3 – line 15) (Miller, Baruch et al. 2002, Suh, Kain et al. 2004).

1.3 Epidemiology and pathology of *T. gondii* and *P. falciparum*

The diverse and highly complex life cycles of the apicomplexans are major contributors to the widespread prevalence of these parasites. Focussing here on the effects of *T. gondii* and *P. falciparum* on the human population, it is evident that just these two members of phylum Apicomplexa cause strikingly devastating diseases.

1.3.1 T. gondii

Epidemiology

Thousands of studies on the prevalence and seroprevalence of *T. gondii* in humans and animals have been published, reporting rates for human prevalence with a wide distribution – from 10 to 80% – that varies with geographic region (Tenter, Heckeroth et al. 2000, Pappas, Roussos et al. 2009). The prevalence of *T. gondii* in both developed and developing nations is likely due to its multiple modes of transmission. Untreated or poorly treated water supplies contaminated with oocysts are likely a major source of infection for poverty-stricken areas, but are also a source of infection for developed municipalities with poor

water treatment protocols such as the 1995 outbreak in Victoria, Canada (Bowie, King et al. 1997, Dubey 2004, Hotez 2008). Foodborne illness, either through ingestion of surface contaminated fruits and vegetables (oocysts) or consumption of undercooked meat (tissue cysts), is a major source of infection particularly noted in Western countries and in Europe (Robert-Gangneux and Darde 2012); a recent study estimated that *T. gondii* causes almost a quarter of all foodborne illness related deaths in the United States (Scallan, Hoekstra et al. 2011). Finally, proximity to infected felines, either wild or domesticated, is a widespread factor of human infections, as an estimated 80% of domesticated felines are seropositive for *T. gondii*; while seropositive cats are not necessarily shedding oocysts, the seroprevalence rate is considered a strong indicator of environmental oocyst burden (Jones and Dubey 2010, Robert-Gangneux and Darde 2012).

Pathology

The high worldwide prevalence of *T. gondii* is a major concern for human health, with implications for both immunocompetent and immunocompromised individuals, and for neonates. Acute infections in the immunocompetent generally present as a brief set of influenza-like symptoms followed by a dormant chronic infection as the parasites retreat into bradyzoite-filled tissue cysts (Lyons, McLeod et al. 2002). However, for those patients who do have symptoms associated with acute infection that are more severe than influenza-like, lymphadenopathy and ocular disease are the most common presentations (Petersen and Liesenfeld 2007). Toxoplasmic retinochoroiditis in particular is quite serious - it is often a recurring disease and is associated with symptoms including eye lesions, blurred vision, glaucoma, and complete or partial vision loss (Petersen and Liesenfeld 2007, Roberts, Kuo et al. 2007). The symptoms of chronic *T. gondii* infection in immunocompetent adults have long been considered to be negligible. However, animal studies showing the ability of *T. gondii* to permanently alter the behavior of its host organism (Webster 2007) have inspired a new branch of research aimed at establishing a link between *T. gondii* brain tissue cysts and numerous neurological conditions and behavioural changes in humans including schizophrenia, increased apprehension, decreased reaction times (suspected to be the link between *T. gondii* infection and an

increased risk for traffic accidents), and depression (Torrey and Yolken 2003, Flegr 2007, Yolken and Torrey 2008, Henriquez, Brett et al. 2009, Flegr 2013).

While both acute and latent infections clearly have some level of risk for immunocompetent individuals, the possibility of contracting new infections or of recrudescence of latent infections in immunosuppressed patients is well-established to be life-threatening. In particular, people with acquired immunodeficiency syndrome (AIDS) and patients undergoing chemotherapy for cancer treatment or organ transplantation are at the highest risk for reactivation (Araujo and Remington 1987). Toxoplasmosis is one of the most prevalent central nervous system opportunistic infections in human immunodeficiency virus (HIV)-positive and AIDS patients (Meira, Vidal et al. 2011, Yan, Huang et al. 2013), and most *T. gondii* infections in immunocompromised patients appear as toxoplasmic encephalitis, which is fatal if left untreated (Contini 2008).

In addition to acquired adult infections, vertical transmission from mother to fetus is possible and can result in life-threatening congenital toxoplasmosis. The severity of *T. gondii* infection in neonates is inversely correlated to the stage of pregnancy, with the most dangerous cases occurring when the mother has an acute infection in the first or second trimester (Desmonts and Couvreur 1974). Acute infections in the first trimester often lead to abortion, while neonates surviving to birth will likely have a significantly reduced quality of life due to devastating symptoms such as mental retardation, seizures, hydrocephaly, and severe eye lesions (Robert-Gangneux and Darde 2012).

Genotypes and virulence

A final important note on the ability of *T. gondii* to cause disease in humans is the relationship between genotypes and virulence. While there is just a single *Toxoplasma* species, *T. gondii* is split into three clonal lineages (Types I, II, and III) defined by a variety of genetic markers. In general, Type I strains are extremely virulent in mice while Type II strains cause the majority of infections in humans (Howe and Sibley 1995). Additionally, in recent years, a number of atypical strains have been identified, some of which can cause severe disease in humans (Carme, Bissuel et al. 2002, De Salvador-Guillouet, Ajzenberg et al. 2006, Khan, Jordan et al. 2006, Miller, Miller et al. 2008). Several characteristics can be identified that play a role in the differing virulence including migration speed,

penetration of the submucosa, growth rate, and interconversion between tachyzoite and bradyzoite stages (Barragan and Sibley 2002, Saeij, Boyle et al. 2005). Additionally, a single family of polymorphic rhoptry kinases and pseudokinases that are injected into the host cell appear to be largely responsible for the varying virulence of the different genotypes in mice (Saeij, Boyle et al. 2006, Taylor, Barragan et al. 2006, Reese, Zeiner et al. 2011). Other studies indicate that virulence is much more complex and depends on a combination of parasite and host characteristics (Robben, Mordue et al. 2004, Carruthers and Suzuki 2007, Maubon, Ajzenberg et al. 2008). A complete picture of the connections between genotypes and infection still remains to be elucidated, but wherever possible the work in this dissertation was done on Type II strains that are widely considered to be the most prevalent and virulent in humans (Ajzenberg, Cogne et al. 2002, Ajzenberg, Yera et al. 2009).

1.3.2 P. falciparum

Epidemiology

Despite the fact that *T. gondii* is so prevalent in the human population, it is not well recognized outside of the scientific community. In contrast, malaria and its etiological agent, *Plasmodium*, are well known in the general populace and are the subject of innumerable clinical, industrial and academic investigations. A major reason for this discrepancy is the death toll associated with malaria − conservative estimates are in the range of 250 to 515 million episodes of clinical malaria each year leading to more than one million deaths, 85% of which are in children under the age of five (Snow, Guerra et al. 2005, World Health Organization. 2010). Over 90% of reported cases of malaria occur in Africa, and worldwide, malaria has the strongest impact on the poverty-stricken, as 60% of deaths occur in the poorest 20% of the population (Suh, Kain et al. 2004).

Pathology

Malaria often begins as asymptomatic, but, if left untreated, can rapidly develop into influenza-like symptoms, fever, splenomegaly, metabolic acidosis, severe anaemia, cerebral malaria with impaired consciousness and seizures, and finally death (Taylor, Borgstein et al. 1993, Marsh, Forster et al. 1995, Suh, Kain et al. 2004). The symptoms of

Plasmodium infection result from parasites propagating within the liver, destroying hepatocytes (Miller, Baruch et al. 2002, Frevert, Usynin et al. 2008), and subsequently invading and destroying erythrocytes (Figure 3), while sequestration of parasitized erythrocytes in the brain and placenta lead to the most serious forms of the disease (MacPherson, Warrell et al. 1985, Achur, Kakizaki et al. 2008). Severe malaria, characterized by acute respiratory distress, cerebral malaria, and renal insufficiency, is almost always associated with *P. falciparum*; while five species of *Plasmodium* are infectious to humans (*P. falciparum*, *P. vivax*, *P. ovale*, *P. malariae*, and *P. knowlesi*) and each one can result in high parasite burdens, only *P. falciparum* has the ability to bind human epithelium, leading to occlusion of the microvasculature and resulting in the symptoms of severe malaria (Miller, Baruch et al. 2002, Suh, Kain et al. 2004).

As with *Toxoplasma* infections, it is more than just the characteristics of the parasite that can affect the clinical outcome of infections with *P. falciparum*. Host factors that can play a role in malaria disease progression include sickle cell disease, glucose-6-phosphate dehydrogenase deficiency, and previous infection resulting in partially protective cellular and humoral immunity (Allison 1954, Ruwende, Khoo et al. 1995, Miller, Baruch et al. 2002, Suh, Kain et al. 2004).

1.4 Toxoplasmosis and malaria: disease prevention, diagnosis and treatment

While there are a number of methods for preventing apicomplexan infections, and for diagnosis and treatment once a person becomes infected, there is still a major push to identify novel targets for vaccines and drugs; no human vaccine is licensed for either *T. gondii* or *P. falciparum*, and current drugs are often too expensive for widespread use, have severe side-effects, or become virtually useless as the parasites develop resistance.

1.4.1 Toxoplasmosis

Prevention

A number of simple, hygiene-based measures can be taken to avoid infection with *T. gondii*. People should practice careful hand-washing after contact with known carriers of *T. gondii* parasites such as raw meat, soil, cats, or cat feces (Robert-Gangneux and Darde

2012). Fruits and vegetables, particularly those that grow in the soil or in close proximity to the soil, should be washed thoroughly before consumption. Domestic cats should be kept inside and fed canned or dried food, and their litter boxes should be changed frequently by a person wearing gloves and a mask. In addition to hygiene measures, destruction of animal tissue cysts by cooking meat to well-done, by freezing it for several days at a temperature below -20 °C, or by choosing meat that has been gamma-irradiated is highly advisable (El-Nawawi, Tawfik et al. 2008). As numerous outbreaks of *T. gondii* infection have been linked to waterborne oocysts, water treatment with appropriate filtration systems is critical (Jones and Dubey 2010). It is incredibly important that pregnant women, particularly those who have not had a previous *T. gondii* infection, follow these guidelines to avoid congenital infection of the fetus. It is also worth noting that infection with one strain of *T. gondii* will not necessarily provide immunity against future infections with other strains, so caution should still be taken even by people who have had a previous infection (Elbez-Rubinstein, Ajzenberg et al. 2009, Robert-Gangneux and Darde 2012). Finally, despite extensive research into the development of a vaccine targeting *T. gondii*, there is no licensed human vaccine currently available (Liu, Singla et al. 2012).

Diagnosis

Diagnosis of *T. gondii* infections depends on the immune status of the individual undergoing testing, but in general detection falls under two distinct categories: direct (parasite visualization or DNA detection) or indirect (antibody-based assay) (Robert-Gangneux and Darde 2012). For immunocompetent individuals, serological tests are most commonly used and are predominately retrospective since most infections quickly resolve into an asymptomatic state. Levels of immunoglobulin (Ig) A (produced during first week of infection, plateau in one month, remain detectable for up to nine months), IgM (produced during first week of infection, plateau in one month, decrease to undetectable levels after one to six months or remain detectable for more than twelve months), IgG (detectable one to three weeks after rise of IgM, plateau at two to three months, drop to residual amounts detectable for the life of the patient), and/or IgE (produced early in infection and quickly disappear) antibodies can be assessed, but there is no established international standard for calibration of results (Robert-Gangneux and Darde 2012). Furthermore, the complex

profiles of production of anti-*T. gondii* antibodies during the course of infection can make it very difficult to accurately date the time of infection, which is of great importance especially when trying to establish the likelihood of congenital infection. A diagnosis of congenital infection during gestation generally requires a positive polymerase chain reaction (PCR) result for the presence of parasite DNA in the amniotic fluid, while parasite detection in the placenta (by PCR and/or inoculation of placental material into mice) can be a useful test after birth. Since immediate detection of acute infection is of necessity for immunocompromised patients, PCR or microscopic examination of tissues and/or fluids is the general course-of-action to quickly identify tachyzoites (Robert-Gangneux and Darde 2012).

Treatment

Due to the varied nature of *T. gondii* infections, treatment regimens are not broadly recommended, but rather designed for each patient and will account for factors such as the location of the infection (eye, brain, etc.), the stage of the infection (acute or latent), the immune status of the individual, any additional infections (HIV-positive and AIDS patients often require specialized treatments), and the pregnancy status if the patient is a woman (Petersen and Liesenfeld 2007). Despite mounting evidence of the effects *T. gondii* tissue cysts can have on the brain (Section 1.3.1), immunocompetent individuals with chronic latent infections are generally not considered to require treatment. The drug combination recommended for most patients with acute infections is pyrimethamine with a sulfonamide or with clindamycin. Pyrimethamine inhibits dihydrofolate reductase, while sulfonamides inhibit dihydrofolic acid synthase, leading to a synergistic therapeutic approach as both drugs target enzymes in folic acid metabolism (Petersen and Liesenfeld 2007). Pyrimethamine in particular leads to bone-marrow suppression and anemia, so folinic acid is routinely given to prevent these effects, while still maintaining efficacy since *T. gondii* cannot uptake folinic acid. For a pregnant woman with a highly suspected or confirmed acute infection, spiramycin (a macrolide antibiotic likely functioning as a protein synthesis inhibitor) treatment is generally prescribed until delivery (Brissonnoel, Trieucuot et al. 1988, Cortina-Borja, Tan et al. 2010). It is important to note that all of these drugs are only active against the tachyzoite stage and not encysted bradyzoites, meaning that treatment of

an acute infection will likely result in the patient finishing the treatment with a latent infection (Petersen and Liesenfeld 2007).

1.4.2 Malaria

Prevention

There are two overlapping approaches to malaria disease prevention – chemoprophylactic treatments for travellers and mosquito bite-prevention for both travellers and people living in endemic areas (Suh, Kain et al. 2004). Widespread chemoprophylaxis is considered unfeasible in malaria-endemic countries primarily due to limited finances. However, for short-term travellers to malaria-endemic areas, prophylactic drugs are very effective although not 100% reliable; people who have travelled to malaria-endemic countries even with the use of prophylactic drugs should seek immediate medical attention if they develop a fever within two months of arriving to their home country (Suh, Kain et al. 2004). A number of chemoprophylatics are currently prescribed including Atovaquone-proguanil (targets the cytochrome *bc*1 complex, inhibiting mitochondrial respiration and indirectly inhibiting pyrimidine biosynthesis), Chloroquine (targets the parasite's ability to detoxify heme), Doxycycline (possibly targets expression of apicoplast genes), Mefloquine (likely targets heme detoxification and solute import into the parasite's digestive vacuole), Primaquine (unknown mechanism of action; only drug known to be effective against latent *P. vivax* infections), and Quinine (similar mechanism of action to Chloroquine is suspected), but few of these drugs are generally well tolerated and multi-drug resistant *Plasmodium* strains have already developed in several regions (Suh, Kain et al. 2004, Petersen, Eastman et al. 2011, Tan, Magill et al. 2011). Both travellers and residents should avoid being outside after dusk, as the *Anopheles* mosquitos are night-biting (Suh, Kain et al. 2004). Skin should be kept covered, preferably with insecticide-impregnated clothing, and an insecticide-impregnated bed net is highly recommended; permethrin is the primary insecticide of choice due to its two modes of action – insect repellent and contact insecticide – but a combination with other insecticides is recommended to reduce resistance (Binka, Kubaje et al. 1996, Goodyer, Croft et al. 2010).

While numerous vaccines have been tested in clinical trials, no effective vaccine against malaria is currently available for human use (Crompton, Pierce et al. 2010). Two

of the major difficulties in malaria vaccine development are the high levels of variability and polymorphism in *Plasmodium* proteins that result in species and strain specificity of developed therapeutics, and the presence of redundant pathways for critical biological mechanisms (Crompton, Pierce et al. 2010, Good and Doolan 2010). In addition to blood-stage and liver-stage vaccines that have so far only achieved a maximum result of partial efficacy (Crompton, Pierce et al. 2010, Olotu, Fegan et al. 2013), transmission blocking vaccines are currently being investigated as a novel method for targeting the spread of *Plasmodium* species through the mosquito vector (Carter, Mendis et al. 2000, Crompton, Pierce et al. 2010, Kappe, Vaughan et al. 2010).

Diagnosis

Historically, the major method for diagnosis of malaria has been peripheral blood smears; this method requires a microscope, staining reagents, a technician with expertise in interpreting blood smears, as well as several blood samples taken over forty-eight hours (Suh, Kain et al. 2004). A set of more user-friendly rapid malaria tests has more recently been developed, with the rapid antigen detection test being the most practical (Suh, Kain et al. 2004, McMorrow, Aidoo et al. 2011). From finger-prick blood samples, rapid antigen detection tests identify the presence of up to three *Plasmodium* proteins and output a positive or negative result within twenty minutes (McMorrow, Aidoo et al. 2011). While cost-effective and easy to use, these tests suffer from a lack of sensitivity and the output cannot be correlated to levels of parasitemia, which is critical for determining an appropriate treatment regimen. Serological tests are not used for diagnosing acute malaria infections in individuals, and while PCR is one of the most specific and sensitive tests, the assay time and required equipment have restricted its use for widespread diagnostics (Snounou, Viriyakosol et al. 1993, Suh, Kain et al. 2004).

Treatment

Similar to toxoplasmosis, there is not a single broadly recommended treatment regimen for malaria. The tactic taken for curing an individual of malaria will depend on the *Plasmodium* species that caused the infection, the likelihood of drug resistance based on geographical location, and the severity of the infection (Suh, Kain et al. 2004, Petersen, Eastman et al.

2011). In general, if characterization of the infecting *Plasmodium* strain cannot be attained, the patient should be treated as if they have acquired a multi-drug resistant strain. Currently, the World Health Organization recommends artemisinin (pro-drug activated by heme-mediated cleavage leading to free radicals that result in cellular damage and death) or its derivatives in combination with one of the classical drugs (see *Prevention* section above), in the form of artemisinin combination therapies (ACT), as the primary approach for treating malaria infections (Klonis, Creek et al. 2013, World Health Organization. 2010); however, it is worth noting that some resistance has already been observed to artemisinin, mostly due to the usage of monotherapies and substandard formulations (Dondorp, Yeung et al. 2010). Treatment often lasts for weeks, and full compliance of the patient is critical to avoiding development of new resistant strains (Petersen, Eastman et al. 2011).

1.5 Apicomplexan active invasion

Despite the broad range of hosts, transmission mechanisms, and disease pathologies, the apicomplexans maintain a broadly conserved obligate intracellular life cycle dependent on a unique mode of active host cell invasion (Figure 4) (Sibley 2010). Intriguingly, active host cell invasion has been suggested to be the "Achilles' heel" of the apicomplexan lifestyle for effective therapeutic intervention (Cowman and Crabb 2006).
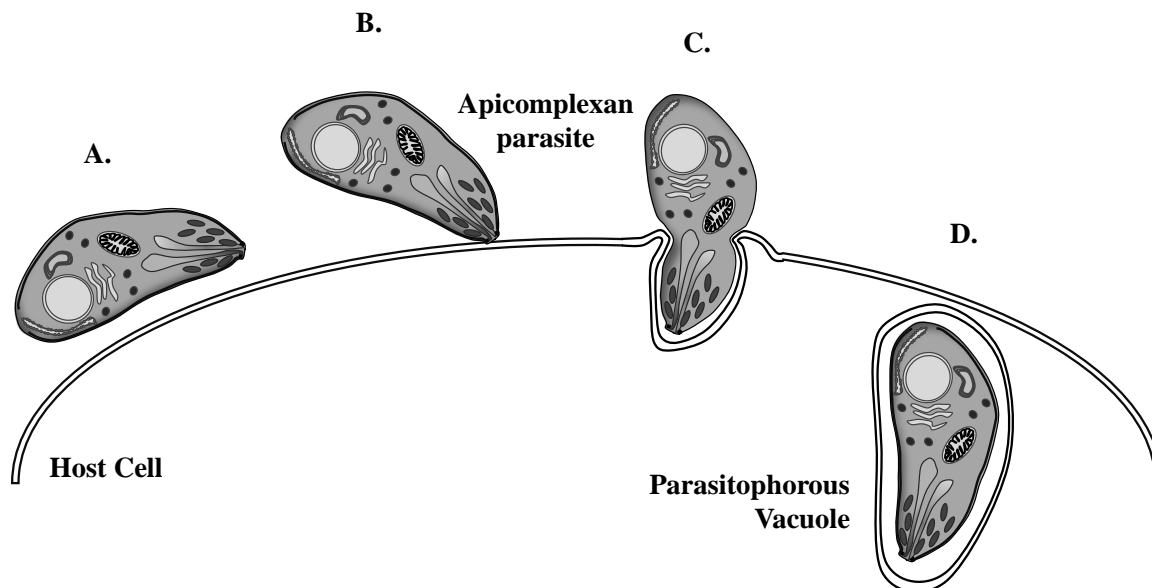
Figure 4: Schematic of the four major steps in the active host cell invasion mechanism employed by apicomplexan parasites.
A. Host cell recognition through reversible attachments. **B.** Apical reorientation and tight, irreversible attachment. **C.** Active host cell invasion through the constricted moving junction ring. **D.** Sealing within the protective, nascent parasitophorous vacuole.

Active invasion by apicomplexan parasites begins with initial recognition and reversible attachment to the host cell (Figure 4A). If a suitable point for invasion is found, the parasite will apically reorient and form an irreversible attachment (Figure 4B). An apicomplexan-unique circumferential ring of adhesion, termed the moving junction (MJ), is then formed between the apical tip of the parasite and the host cell membrane, and the parasite propels itself forward, squeezing through this junction concurrent with the generation of the parasitophorous vacuole (PV) (Figure 4C). Finally, the parasite is sealed within the PV and released into the host cell cytoplasm, where it uses a sophisticated set of proteins and biological processes to harvest nutrients and energy from the host cell and modulate the immune response to enable intracellular survival and replication (Figure 4D) (Cowman and Crabb 2006, Boothroyd and Dubremetz 2008). While clearly an intricate process, the complete invasion sequence can be accomplished in about two minutes, with the step of MJ progression taking less than twenty seconds for both *T. gondii* and *P. falciparum* (Carruthers and Boothroyd 2007, Gilson and Crabb 2009).

Since intracellular parasites are considerably more difficult to target with therapeutics than extracellular parasites, the first three steps of the essential process of active invasion are appealing for the identification of novel therapeutic targets (Macraild, Anders et al. 2011). However, a significant body of work exists on the characterization of the first two steps of active host cell invasion and has revealed that both *T. gondii* and *P. falciparum* employ numerous semi-redundant and highly divergent pathways to accomplish host cell attachment and apical reorientation (Carruthers 1999, Reed, Caruana et al. 2000, Cowman and Crabb 2006, Boulanger, Tonkin et al. 2010, Gaur and Chitnis 2011). More recent studies have suggested a broadly conserved mechanism for MJ formation and progression, making this step the ideal target not only for therapeutic development, but also for learning about an exceptionally unique and interesting biological mechanism that the apicomplexans have evolved to enable their widespread intracellular lifestyle.

1.5.1 Moving junction formation and progression

Identification of the MJ and its importance to successful invasion

As early as 1969, researchers observed that apicomplexan zoites do not penetrate the host cell membrane, but rather enter the host cell through a constriction into an extreme depression of the host cell membrane (Ladda, Aikawa et al. 1969). In 1978, the MJ was characterized by electron microscopy of *Plasmodium knowlesi* parasites invading erythrocytes, and was identified as an electron dense apposition between the parasite and host cell membranes, about 250 nm in length, that forms at the point of contact by the apical tip of the parasite, and moves as a circumferential ring during invasion (Aikawa, Miller et al. 1978). The ability to invade in an MJ dependent fashion is paramount to parasite survival, as evidenced by the fact that parasites phagocytosed by macrophages either escape from the phagosome into a nascent vacuole through an MJ or are rapidly degraded within the phagosome (Morisaki, Heuser et al. 1995). This phenomenon is likely due to the fact that the MJ of apicomplexan parasites not only enables the formation of the PV, but also serves as a selective gate by filtering proteins on the host cell membrane to modify the surface of the nascent vacuole (Ward, Miller et al. 1993, Mordue, Desai et al. 1999, Charron and Sibley 2004). This gate-keeping action by the MJ complex is thought to play

a critical role in the non-fusogenic nature of the resulting vacuole (Mordue, Hakansson et al. 1999).

Molecular composition of the MJ

For decades the MJ was recognized as an important component of the invasion process of apicomplexans, but the molecular composition remained unknown, mainly because it is only assembled for an invasion event that lasts for mere seconds. In 2005, using immunoaffinity purification with an antibody that distinctly stains the MJ of *T. gondii* tachyzoites, one group identified rhoptry neck protein 4 (RON4) at the MJ (Lebrun, Michelin et al. 2005). In the same month a second group used pull-down assays with apical membrane antigen 1 (AMA1), a parasite protein previously shown to be important for invasion, to identify a stable complex between AMA1 and three RONs (2, 4, and 5), and localized both AMA1 and RON4 to the MJ during invasion (Alexander, Mital et al. 2005). AMA1, a type I integral membrane protein with a large ectodomain (~50 kDa) and a small cytoplasmic tail, is localized to the micronemes and translocated to the parasite surface around the time of invasion (Waters, Thomas et al. 1990, Donahue, Carruthers et al. 2000, Healer, Crawford et al. 2002, Bannister, Hopkins et al. 2003, Besteiro, Michelin et al. 2009), while the RONs are very large proteins (up to ~250 kDa) with up to three transmembrane passes and are localized to the neck region of the rhoptry organelles prior to secretion (Alexander, Mital et al. 2005, Bradley, Ward et al. 2005). Based on the surprising findings that micronemal-derived AMA1 associates with a collection of rhoptry neck-derived proteins during invasion, a model was proposed requiring a synchronization between the microneme and rhoptry organelles, resulting in AMA1 and the RONs collaborating on the parasite surface to bind a ligand on the host cell (Alexander, Mital et al. 2005). While these two studies represent the seminal identification of the molecular composition of the MJ, subsequent research has led to significant reorganization and refinement of the original MJ model. In particular, the RON complex was found to be secreted from the rhoptry necks into the host cell cytosol, where RON2 is integrated into the host cell plasma membrane, positioning a portion of RON2 to be exposed on the host cell surface while the remainder of RON2 links to RONs 4 and 5 within the host cytosol. Importantly, more selective pull-down experiments revealed a binary complex formed

between AMA1 and RON2, independent of the other RONs in the complex (Besteiro, Michelin et al. 2009). Further studies reversed the originally suspected orientation for RON2, showing a direct interaction between the AMA1 ectodomain and a short segment of RON2 in the C-terminal third of the protein for both *T. gondii* and *P. falciparum* (Lamarque, Besteiro et al. 2011, Srinivasan, Beatty et al. 2011, Tyler and Boothroyd 2011). Recently, super-resolution immunofluorescence microscopy was used to study *Plasmodium* merozoite invasion, and three-dimensional visualizations clearly showed AMA1 forming a ring structure surrounded by a ring of RON4 (Riglar, Richard et al. 2011). Together with studies identifying sequences encoding AMA1 and RON2/4/5 across the phylum, revealing conservation of the AMA1-RONs interaction in different apicomplexan species, and showing the ability to block the MJ step of invasion with antibodies and peptides that inhibit the AMA1-RONs interaction (Chesne-Seck, Pizarro et al. 2005, Alexander, Arastu-Kapur et al. 2006, Collins, Withers-Martinez et al. 2007, Narum, Nguyen et al. 2008, Cao, Kaneko et al. 2009, Collins, Withers-Martinez et al. 2009, Straub, Cheng et al. 2009, Richard, MacRaild et al. 2010, Besteiro, Dubremetz et al. 2011), this body of work gave rise to the leading model for the primary molecular composition of the MJ core (Besteiro, Dubremetz et al. 2011, Shen and Sibley 2012) (Figure 5). Based on all of these studies, it appears that apicomplexan parasites provide both receptor (AMA1) and ligand (RON2) to enable MJ dependent active invasion, which may explain why the mechanism is so highly conserved despite the enormous diversity of host cell tropisms within phylum Apicomplexa, and further supports the idea of targeting the MJ step of active invasion with novel therapeutics.
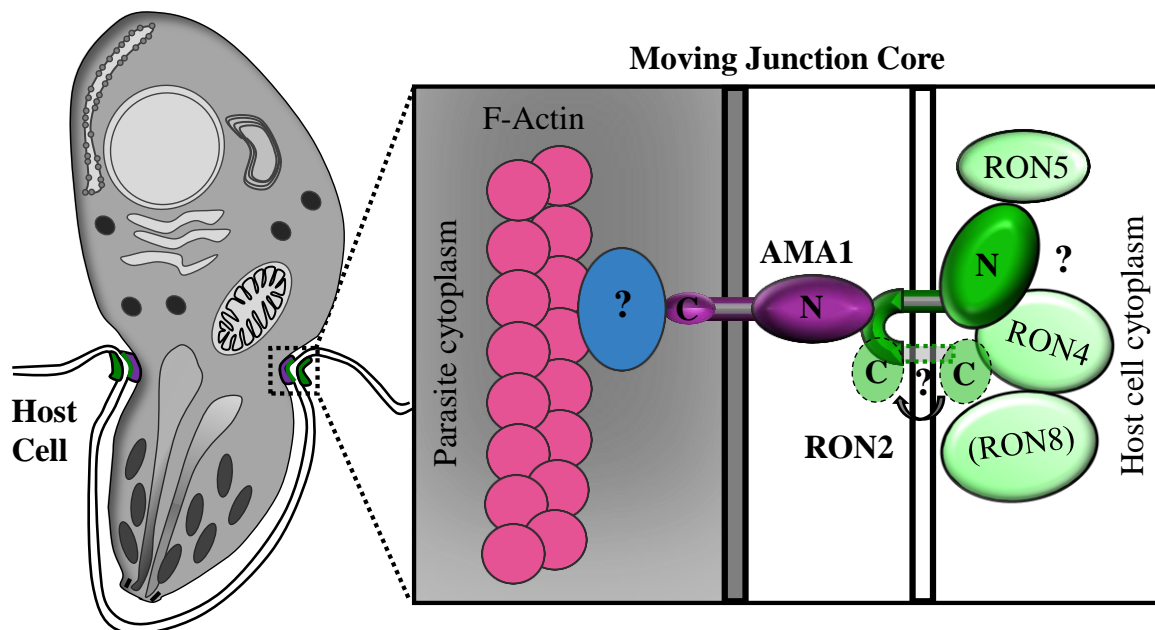
Figure 5: The leading model for the core molecular composition of the apicomplexan MJ complex. Location of the apicomplexan MJ, represented on an invading *T. gondii* tachyzoite. Inset – simplified schematic of the core components of the apicomplexan MJ. Note that F-actin is part of a complicated and unique apicomplexan motor system that drives invasion and is separated from the central cytoplasm by the inner membrane complex (not shown for clarity). The putative connection between the AMA1 cytoplasmic tail and the motor system has been extensively studied, but remains controversial (Bosch, Buscaglia et al. 2007, Treeck, Zacherl et al. 2009, Sheiner, Santos et al. 2010, Srinivasan, Beatty et al. 2011, Shen and Sibley 2014). Furthermore, the positioning of the C-terminal portion of RON2 with respect to the host cell membrane (arrow with question mark), the stoichiometry and intermolecular interactions of the RON complex, and the anchor(s) for the RON complex within the host cell are all unknown. Note that additional RONs, such as RON8 in the coccidians (shown in brackets), may be added to the complex in some species (Besteiro, Michelin et al. 2009, Straub, Cheng et al. 2009).

1.5.2 The extracellular parasite-host cell link proteins: AMA1 and RON2

Diverse biological functions of AMA1

Prior to establishing the molecular composition of the MJ, AMA1 was considered critical for parasite survival, and is now one of the best studied invasion proteins of the apicomplexans. AMA1 was first identified in *P. knowlesi* merozoites in 1982 (Deans,

Alderson et al. 1982), however a comprehensive model of its roles remains enigmatic with many conflicting reports describing its function (reviewed in (Tyler, Treeck et al. 2011)). For example, AMA1 is proposed to have host cell binding capabilities (Urquiza, Suarez et al. 2000, Fraser, Kappe et al. 2001, Kato, Mayer et al. 2005, Valbuena, Rodriguez et al. 2006, Montero, Rodriguez et al. 2009), to properly position the parasite for invasion (Giovannini, Spath et al. 2011), to either play a role in or not be required for initiating close host cell contact and apical reorientation (Thomas, Deans et al. 1984, Mitchell, Thomas et al. 2004, Richard, MacRaild et al. 2010, Srinivasan, Beatty et al. 2011), to be critical for or peripherally involved in invasion (Thomas, Deans et al. 1984, Hehl, Lekutis et al. 2000, Triglia, Healer et al. 2000, Li, Dluzewski et al. 2002, Gaffar, Yatsuda et al. 2004, Silvie, Franetich et al. 2004, Mital, Meissner et al. 2005, Giovannini, Spath et al. 2011), and to participate in signalling through phosphorylation of the cytoplasmic domain (Treeck, Zacherl et al. 2009, Leykauf, Treeck et al. 2010, Treeck, Sanders et al. 2011). Much of the controversy in the field derives from genetic manipulation studies that often select for a minor phenotype and rarely extend to investigating the activation of compensatory or redundant pathways. Despite these conflicting studies, it is clear that AMA1 plays a major role in parasite survival across the phylum.

Targeting AMA1 for therapeutic development: successes and challenges

AMA1 is considered a prime malaria vaccine candidate since anti-AMA1 antibodies can be found in naturally infected patients (Lal, Hughes et al. 1996, Hodder, Crewther et al. 2001, Cortes, Mellombo et al. 2003), and the presence of anti-AMA1 antibodies in humans prior to malaria transmission season or the immunization of animal models with correctly folded AMA1 leads to a significant protective response (Deans, Knight et al. 1988, Anders, Crewther et al. 1998, Narum, Ogun et al. 2000, Stowers, Kennedy et al. 2002, Polley, Mwangi et al. 2004, Remarque, Faber et al. 2008). However, developing a broadly effective vaccine targeting AMA1 has been challenging, due in part to its highly polymorphic nature that imparts species and even strain specificity to the antibodies, with only rare cases of cross-genera reactivity (Hodder, Crewther et al. 2001, Zhang, Compaore et al. 2007). Numerous studies have identified polymorphisms in *P. falciparum* AMA1 (*Pf*AMA1) that are likely due to diversifying selection pressure of the host's protective immune response

(Escalante, Grebert et al. 2001, Hodder, Crewther et al. 2001, Polley and Conway 2001), but recent studies have suggested that including multiple AMA1 alleles in a vaccine may circumvent this issue (Remarque, Faber et al. 2008, Dutta, Dlugosz et al. 2013, Miura, Herrera et al. 2013). AMA1 based vaccines have struggled to induce a protective response in humans despite the fact that antibodies can block invasion *in vitro* (Remarque, Faber et al. 2008, Spring, Cummings et al. 2009, Ouattara, Mu et al. 2010, Thera, Doumbo et al. 2011, Laurens, Thera et al. 2013, Tamminga, Sedegah et al. 2013), highlighting the need for more detailed information regarding AMA1 and its ligands in order to improve vaccine efficacy.

Structural characteristics of the AMA1 ectodomain

In most apicomplexans, AMA1 is synthesized as an approximately 80 kDa protein and localized to the micronemes, following which a pro-domain is cleaved leaving the mature ~65 kDa protein to be translocated to the surface just prior to invasion (Peterson, Marshall et al. 1989, Narum and Thomas 1994, Donahue, Carruthers et al. 2000, Howell, Withers-Martinez et al. 2001, Healer, Crawford et al. 2002, Lovett and Sibley 2003, Singh, Alam et al. 2010). Early predictions based on chemical modification studies suggested that the fully processed ectoplasmic region of AMA1 consists of three disulfide-defined domains (domain I, DI; domain II, DI; domain III, DIII), with a total of sixteen cysteines predicted to form eight disulfide bonds (Hodder, Crewther et al. 1996).

The 1.8 Å resolution crystal structure of the ectoplasmic region of *P. vivax* AMA1 (*Pv*AMA1) revealed a stacked three domain architecture with DI membrane distal, DIII membrane proximal, and DII sandwiched in the center (Pizarro, Vulliez-Le Normand et al. 2005) (Figure 6A). Three DI loops in the predicted apical surface were not observed in the *Pv*AMA1 crystal structure, limiting insight into the surface that would be positioned closest to the host cell and therefore a prime candidate for the ligand binding region. Additionally, one extended loop from DII (termed the "DII loop") was missing density for a stretch of forty residues (amino acids 295 to 334). Despite these disordered surface loops, the core of each domain was well ordered, and revealed that DI and DII each adopt a divergent plasminogen/apple/nematode (PAN) fold, with loops extending from the core consisting of a five-stranded anti-parallel beta-sheet partially curving around a central alpha helix;

PAN domains are found in a diverse array of proteins, many of which adhere to protein or carbohydrate receptors (Tordai, Banyai et al. 1999). In contrast to DI and DII, DIII contains a cystine knot motif (three disulfides with 1-4, 2-5, 3-6 connectivity) and appears to create a stabilizing saddle to support *Pv*AMA1 DI and DII, but it does not conform to any known protein fold.
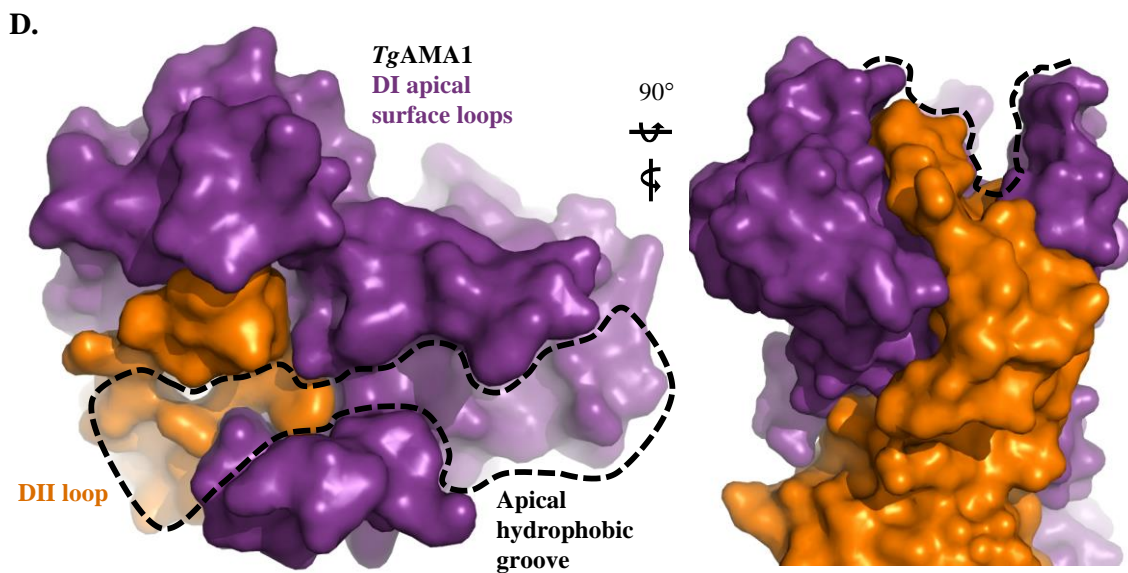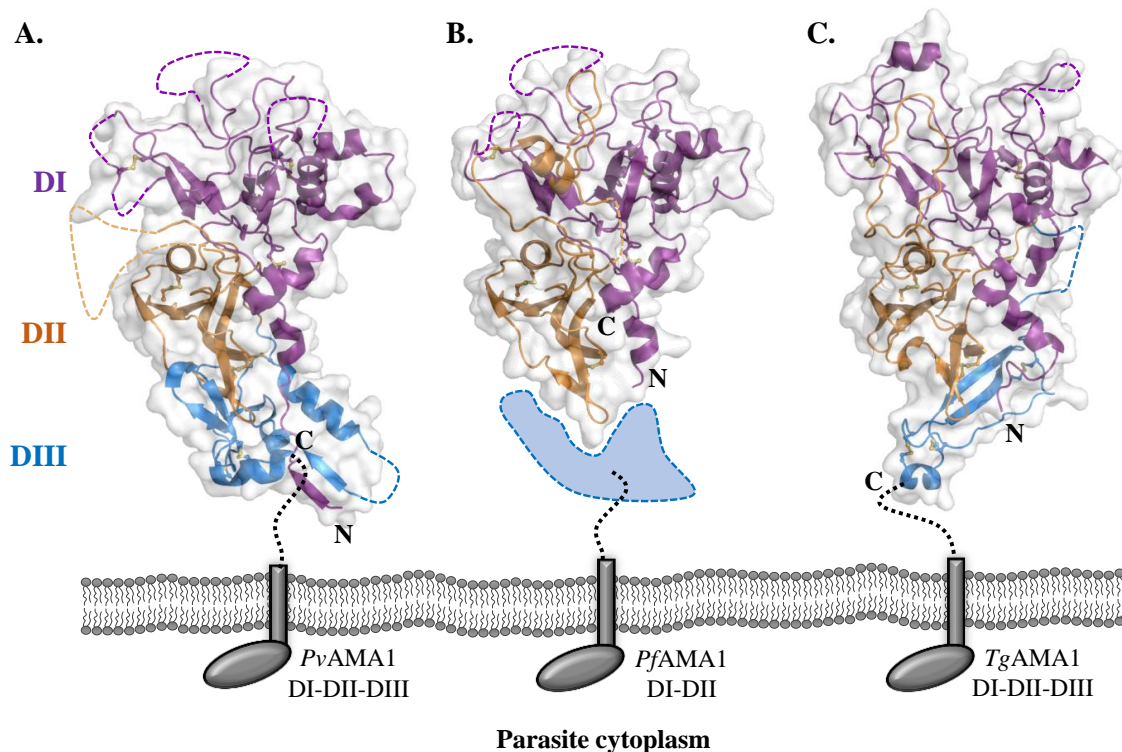
Figure 6: Structural characterization of AMA1 prior to the start of this work.
A. – **C.** Cartoon representations of AMA1 crystal structures with semi-transparent surface, shown as a side view of the predicted orientation to the parasite cell membrane. Dotted lines indicate disordered loops not modeled in the structures. DI, purple; DII, orange; DIII, blue (dotted blue shape represents the missing *Pf*AMA1 DIII); disulfides, yellow ball-and-stick; transmembrane domains, grey rectangles; cytoplasmic domains, grey ovals. *Pv*AMA1 DI-DII-DIII; Protein Data Bank (PDB) ID 1W8K. *Pf*AMA1 DI-DII; PDB ID 1Z40. *Tg*AMA1 DI-DII-DIII; PDB ID 2X2Z. **D.** Left − apical surface view of *Tg*AMA1 DI (purple surface) and DII (orange surface) showing the central groove (dashed shape). Right − end-on view of (left) showing depth of the apical groove (dashed curve).

Structural characterization of *Pf*AMA1 was attempted with both nuclear magnetic resonance (NMR) spectroscopy and X-ray crystallography. Initially, the structures of DII and DIII were determined separately by NMR (Nair, Hinds et al. 2002, Feng, Keizer et al. 2005), but without the context of the other ectoplasmic domains, only limited insight could be gained. In 2005, the high resolution crystal structure of *Pf*AMA1 DI/DII was obtained (Figure 6B) (Bai, Becker et al. 2005, Gupta, Bai et al. 2005). A comparison of the NMR and crystal structures for DII show very little overlap beyond the PAN-associated beta-sheet, highlighting the importance of interdomain interactions to stabilizing the structure of AMA1. Some apical surface loops that were disordered in the *Pv*AMA1 structure were ordered in the structure of *Pf*AMA1, revealing an apical hydrophobic groove nestled in the center of the loops; a significant portion of the DII loop was ordered by packing up the side of DI, imbedding in the apical groove and forming an integral part of the apical surface. Polymorphism mapping showed that the majority of polymorphic *Pf*AMA1 residues cluster on one side of the molecule near the apical groove, suggesting that this region must be protected from the immune system and may be the ligand binding site of AMA1 (Bai, Becker et al. 2005, Dutta, Lee et al. 2007).

No further insight into the structure of AMA1 was provided until the publication of the fully processed *T. gondii* AMA1 (*Tg*AMA1) ectodomain structure (DI, DII, and DIII) (Figure 6C) (Crawford, Tonkin et al. 2010). *Tg*AMA1 has multiple areas of sequence divergence compared to other AMA1s that map to regions of functional importance in the

Plasmodium AMA1 structures. For example, the DI apical surface loops of *Tg*AMA1 are significantly reorganized in structure and chemistry compared to *Plasmodium* AMA1s, prompting the designation of these loops as structural filters restricting access to the central hydrophobic groove, and supporting the theory of the apical groove as the putative ligand binding site (Figure 6D). Furthermore, the DII loop of *Tg*AMA1, which is fourteen residues shorter than *Plasmodium* AMA1s, was ordered in its entirety and anchored into the apical hydrophobic groove by a pair of tryptophan residues burying into a deep pocket bifurcated by a central tyrosine (Figure 6D). The *Tg*AMA1 structure also revealed a divergent and minimalist DIII, which still adopted the structurally ultra-stable cystine knot motif but layered across the bottom of DII (Figure 6C) instead of forming the saddle-like structure observed for *Pv*AMA1 (Figure 6A) (Pizarro, Vulliez-Le Normand et al. 2005).

Evidence for the apical groove as the primary ligand binding site on AMA1

Structural characterization of AMA1 has provided the basis for mapping epitopes of four invasion inhibitory antibodies, which strongly suggest a ligand binding site on the apical surface of AMA1. These studies have focused on *Pf*AMA1, given the global burden of *P. falciparum* malaria. First, monoclonal antibody (mAb) 4G2 binds a portion of the *Pf*AMA1 DII loop that is organized along DI and frames the edge of the apical surface (Pizarro, Vulliez-Le Normand et al. 2005, Collins, Withers-Martinez et al. 2007); importantly, mAb 4G2 was shown to bind apo *Pf*AMA1 but not *Pf*AMA1 in complex with the *Pf*RONs (Collins, Withers-Martinez et al. 2009), suggesting that RON2 likely binds near the AMA1 apical groove and either induces a conformational change or masks the mAb 4G2 epitope. Secondly, mAb 1F9 was shown to bind the *Pf*AMA1 apical groove on the opposite side to mAb 4G2 in a region that encompasses the most polymorphic residue of *Pf*AMA1 (Coley, Parisi et al. 2006, Coley, Gupta et al. 2007). Finally, two very similar invasion inhibitory single-variable-domain antibodies were co-crystallized with *Pf*AMA1, and both bound the apical groove in a region that significantly overlaps with and expands on the mAb 1F9 epitope, contacting residues deep in the hydrophobic groove that are more broadly conserved (Henderson, Streltsov et al. 2007). Altogether, mapping the epitopes of invasion inhibitory antibodies on AMA1 has provided strong evidence for the primary ligand binding site comprising the AMA1 apical groove.

Limited knowledge of RON2, the parasite-derived host cell receptor of AMA1

As previously noted, pull-down studies originally showed AMA1 coordinating a large complex of RONs (Alexander, Mital et al. 2005), and in 2009 a binary complex between AMA1 and RON2 was revealed, with AMA1 positioned on the parasite surface and RON2 integrated into the host cell membrane (Besteiro, Michelin et al. 2009). The topology of RON2 in the host membrane, and thus the region exposed on the host cell surface, has been difficult to define. Recently, one study successfully showed the N-terminal portion to reside within the host cell, with a 65-residue extracellular portion in the C-terminal third of RON2 interacting with AMA1, necessitating at least one transmembrane domain (Lamarque, Besteiro et al. 2011). Additional studies were unable to determine if a second transmembrane domain was present after the primary AMA1-interacting region, and thus orientation of the final C-terminal domain with respect to the host cell membrane remains unknown (Figure 5) (Lamarque, Besteiro et al. 2011, Tyler and Boothroyd 2011). Interestingly, super-resolution imaging of *Pf*RON2 showed a coordinated ring structure during invasion and clear association with a RON4 ring, supporting its localization at the MJ (Zuccala, Gout et al. 2012). Importantly, prior to this work, there was no structural information available for any portion of RON2, limiting insight into its assembly with AMA1.

Outstanding question

While it is clear that a significant body of work supports AMA1 and RON2 forming the critical extracellular link of the MJ complex, the resolution of the information is notably low, with most data on the interaction stemming from immunofluorescent localizations and immunoaffinity enrichments. Although many details about AMA1 and RON2 individually remain to be elucidated, the major outstanding question in the field at the start of the work presented in this dissertation was very clear, "how does AMA1 coordinate RON2 to facilitate MJ dependent invasion?"

1.6 Research objectives

The first two steps of host cell invasion by the apicomplexan parasites are enabled by an arsenal of parasite proteins and corresponding host cell ligands. In contrast, the step of MJ mediated active invasion appears to occur through the interaction of parasite-derived proteins conserved across phylum Apicomplexa. Given the difficulty in targeting these parasites once established within the host cell environment and protected by the PV, blocking the MJ step of invasion is an attractive choice for the development of broadly reactive and highly effective therapeutics. This step of active invasion also holds broader interest on the basic science level as a relatively unique biological event, where two membranes (parasite and host cell) move past each other but do not fuse as they are held in close apposition by protein receptor-ligand pairs. Given the lack of high resolution information detailing the AMA1-RON2 pairing that is critical to apicomplexan parasite host cell invasion and thereby parasite survival, the three main objectives of this dissertation project were as follows:

(1) Define a detailed interaction paradigm of the AMA1-RON2 molecular recognition event through a structural and biochemical investigation of AMA1 with a segment of RON2 that represents the most essential AMA1-binding region.

(2) Broaden the AMA1-RON2 interaction paradigm by assessing the detailed interactions at the interface of divergent AMA1-RON2 pairings in order to identify hot spots or regions likely giving rise to plasticity or specificity.

(3) Assess the relevance of AMA1-RON2 structural data to guiding or supporting the rational development of novel broadly reactive therapeutics.

Chapter 2: Establishing a structural paradigm for the AMA1-RON2 interaction critical to apicomplexan moving junction dependent invasion

Adapted from:

Tonkin, M. L.[*], M. Roques[*], M. H. Lamarque, M. Pugniere, D. Douguet, J. Crawford, M. Lebrun[†] and M. J. Boulanger[†] (2011). "Host cell invasion by apicomplexan parasites: insights from the co-structure of AMA1 with a RON2 peptide." Science **333**(6041): 463-467.

[*] These authors contributed equally

[†] Co-corresponding authors

This manuscript was designated the top Biochemistry manuscript in the associated issue of *Science* and highlighted in a Perspectives piece (Baum and Cowman 2011)**. This manuscript was also recognized as the Editor's Choice for Host-pathogen interactions in *Science Signalling*** (Vinson 2011) **and has already been cited over fifty times.**

Contributions:

JC cloned and generated the viruses for *Tg*AMA1; **MLT** cloned and generated the viruses for *Tg*AMA1 Y230A. **MLT** produced and purified the *Tg*AMA1 proteins. **MLT**, JC, and MJB designed the *Tg*RON2sp peptide. **MLT** crystallized and solved the structures of *Tg*AMA1 Y230A, *Tg*AMA1 Y230A-TgRON2sp, and *Tg*AMA1-TgRON2sp. **MLT** and MJB analyzed the structures. **MLT** and MJB designed the *Tg*AMA1 and *Tg*RON2 mutations. MR, MHL, MP performed the binding, inhibition of invasion, and SPR assays and generated the mutant proteins. DD generated the homology model. **MLT**, MR, ML and MJB wrote the paper.

2.1 Abstract

Apicomplexan parasites such as *Toxoplasma gondii* and *Plasmodium* species actively invade host cells through a moving junction complex assembled at the parasite-host cell interface. Moving junction assembly is initiated by injection of parasite rhoptry neck proteins (RONs) into the host cell, where RON2 spans the membrane and functions as a ligand for apical membrane antigen 1 (AMA1) on the parasite. We have determined the structure of *Tg*AMA1 complexed with a RON2 peptide at 1.95 Å resolution. A stepwise assembly mechanism results in an extensive buried surface area enabling the MJ complex to resist the mechanical forces encountered during host cell invasion. Besides providing insights into host cell invasion by apicomplexan parasites, the structure offers a basis for designing therapeutics targeting these global pathogens.

2.2 Introduction

Apicomplexan parasites cause widespread diseases including malaria (*Plasmodium*) and toxoplasmosis (*Toxoplasma*). The success of these obligate intracellular pathogens is largely due to their mechanism of active invasion (Carruthers and Boothroyd 2007) involving the coordinated secretion of proteins from the microneme and rhoptry secretory organelles (Carruthers and Sibley 1997). In an intriguing process, the parasite provides both ligand and receptor to enable host cell invasion (Boothroyd and Dubremetz 2008).

The molecular basis of active invasion by apicomplexans relies on secretion of a rhoptry neck (RON) complex, consisting of at minimum RONs 2, 4 and 5, into the host cell. RON2 is integrated into the host plasma membrane and RONs 4 and 5 are localized to its cytoplasmic face (Besteiro, Michelin et al. 2009). A surface exposed ectodomain of RON2 serves as the ligand (Lamarque, Besteiro et al. 2011) for the multifunctional, micronemal secreted apical membrane antigen 1 (AMA1) (Waters, Thomas et al. 1990, Alexander, Mital et al. 2005, Kato, Mayer et al. 2005, Mital, Meissner et al. 2005, Treeck, Zacherl et al. 2009, Santos, Ferguson et al. 2011) displayed on the parasite cell surface. The assembled AMA1-RON complex forms a ring-like structure known as the moving junction (MJ), which migrates from the anterior to the posterior of the parasite (Aikawa, Miller et al. 1978), leading to internalization of the parasite into a parasitophorous vacuole (PV). In addition to linking the parasite and host cell membranes, the MJ acts as a diffusion

barrier excluding host transmembrane proteins from the PV (Mordue, Desai et al. 1999), thereby preventing degradation of intracellular parasites (Morisaki, Heuser et al. 1995).

AMA1 and RONs 2, 4 and 5 are conserved across Apicomplexa (Gaffar, Yatsuda et al. 2004, Chesne-Seck, Pizarro et al. 2005, Lebrun, Michelin et al. 2005, Alexander, Arastu-Kapur et al. 2006, Besteiro, Michelin et al. 2009, Cao, Kaneko et al. 2009, Proellocks, Coppel et al. 2010), suggesting a broadly conserved host cell invasion mechanism. In support of this prediction, we have recently shown that the AMA1-RON2 interaction is equally important for the invasive process by *T. gondii* and *P. falciparum* (Lamarque, Besteiro et al. 2011). Structural studies of *Tg*AMA1 (Crawford, Tonkin et al. 2010), *P. vivax* AMA1 (Pizarro, Vulliez-Le Normand et al. 2005) and *Pf*AMA1 (Bai, Becker et al. 2005, Gupta, Bai et al. 2005) revealed a conserved hydrophobic trough in the ectodomain that displays structural hallmarks of a receptor binding surface and thus may be involved in coordinating the binding of RON2. Here we report a structural and functional characterization of *Tg*AMA1 in complex with a complementary *Tg*RON2 peptide.

2.3 Materials and methods

Materials

All basic chemicals for experiments done in the Boulanger Lab were purchased from Sigma Aldrich (Oakville, ON) or Bio Basic Canada (Markham, ON). Enzymes for molecular cloning were purchased from New England Biolabs (Pickering, ON). Gibco Express Five Serum-Free Medium (SFM) for Hi5 cell growth, Gibco *Sf*-900 III SFM for *Spodoptera frugiperda* 9 (*Sf*9) cell growth, Gibco gentamicin reagent solution, Gibco L-glutamine, and Cellfectin insect cell transfection reagent were purchased from Invitrogen (Burlignton, ON). Orbigen Sapphire linearized baculovirus DNA was purchased from Fisher Scientific (Pittsburgh, PA). Greiner bio-one Cellstar six-well sterile tissue culture plates were purchased from VWR (Radnor, PA). Compact 300 crystallization plates were purchased from Emerald Biosystems (Bainbridge Island, WA).

Cloning, protein production, and purification of TgAMA1 and TgAMA1 Y230A

Cloning

Standard protocols were used for DNA cloning, transformation, amplification, and purification. A clone encoding the fully processed ectodomain of *Tg*AMA1 (Ser64 to Thr483, with numbering based on the initiation methionine in the signal sequence; UniProt ID B6KAM0) was amplified from *T. gondii* Type II (strain Me49) cDNA using the following primers (restriction sites are underlined): F, 5'-ACATGA CCATGGGAAGTGCAAGCACGTCGGGGAATC-3' and R, 5'-CTGTCTG GCGGCCGCAGTGGTAGAGCCACATTCATTTTG-3'. The polymerase chain reaction (PCR) (Mullis and Faloona 1987) product was digested with NcoI and NotI and ligated into a modified pAcGP67B (pAcGP67Bmod) Baculovirus transfer vector (Pharmingen; Mississauga, ON) in frame with an N-terminal GP67 insect cell secretion signal and a C-terminal hexa-histidine tag (His$_6$), each separated from the gene by a thrombin cleavage site. Sequencing confirmed that no mutations were introduced during amplification procedures.

Site-directed mutagenesis

The *Tg*AMA1 Y230A construct was generated by site-directed mutagenesis (Winter, Fersht et al. 1982) using the Stratagene QuikChange II XL protocol (Agilent Technologies; Mississauga, ON) with *Tgama1*/pAcGP67Bmod as the template and with the following primers (the mutagenic portion is underlined, TAC to GCC): F, 5'-CGACTGTGCCACATCCTCGCCGTATCGATGCAGCTGATG-3' and R, 5'-CATCAG CTGCATCGATACGGCGAGGATGTGGCACAGTCG-3'.

Baculovirus generation and amplification

Recombinant Type II (Me49) *Tg*AMA1 and *Tg*AMA1 Y230A were produced in insect cells and purified according to the protocols below (Crawford, Tonkin et al. 2010). All transfection and amplification procedures were performed with *Sf*9 cells grown in *Sf*-900 III SFM. The *Tgama1*/pAcGP67Bmod and *Tgama1_y230a*/pAcGP67Bmod constructs were incubated with Sapphire linearized baculovirus DNA and Cellfectin transfection reagent in *Sf*-900 III SFM containing no antibiotics. The mixtures were overlayed on

adherent *Sf*9 insect cells at 60-80% confluence in 6-well tissue culture plates and incubated at 27 °C without shaking for 7 hours. The transfection supernatants were removed and *Sf*-900 III SFM containing 10 µg/mL gentamicin was overlayed on the adherent transfected *Sf*9 cells. The culture plates were wrapped tightly to prevent evaporation and incubated at 27 °C without shaking for 7 days to generate a primary virus (P1). The low titre, small volume P1 virus was used to infect 50 mL of *Sf*9 cells, and the cells were kept below a density of 1.8 x 10^6 cells/mL until less than 0.5 x 10^6 remained viable. The cell debris was removed by centrifugation, with the supernatant containing the higher titre secondary virus (P2). The P2 viruses were then amplified to P3s using the same procedure. The high titre P3 viruses were stored at 4 °C and used for large scale expression of *Tg*AMA1 and *Tg*AMA1 Y230A.

Protein production

For large scale expression, Hi5 cells at 1.8 x 10^6 cells/mL in Express Five SFM containing 10 µg/mL gentamicin and 16.5 mM L-glutamine were infected with an optimized ratio of 3.5 mL of *Tgama1*/pAcGP67Bmod P3 virus per 1 L of cell culture. The baculovirus infected cells were incubated at 27 °C with shaking at 120 rpm for 72 hours to allow for induction of the late stage polyhedron promoter controlling *Tg*AMA1 production. After the 72 hour incubation, the Hi5 cells were at a cell density of 2.2 x 10^6 cells/mL with $10 - 25\%$ nonviable cells. Similar methods were used for *Tg*AMA1 Y230A.

Protein purification

The cell culture was clarified by centrifugation for 15 min at 4 °C and 4500 rpm. The pellets were discarded and the supernatant containing secreted *Tg*AMA1 was filtered sequentially through 1 µm and 0.45 µm filters to remove remaining cellular debris. To reduce the volume of the sample and buffer exchange into a solution suitable for Ni-affinity chromatography (Binding Buffer: 20 mM HEPES pH 8.0, 1 M NaCl, 30 mM imidazole), the sample was applied to a Centramate Tangential Flow Filtration device (Pall Canada; Mississauga, ON) assembled with three stacked 10 kDa molecular weight cut-off (MWCO) membranes. The concentrated sample was applied to a HisTRAP FF 5 mL Ni-affinity column (GE Healthcare Life Sciences; Baie d'Urfe, QC) in Binding Buffer. Bound proteins

were eluted from the Ni column using a gradient of Elution Buffer (20 mM HEPES pH 8.0, 1 M NaCl, 500 mM imidazole), analyzed by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) (Raymond and Weintraub 1959), pooled based on purity, and concentrated using a 10 kDa MWCO Millipore Amicon Centricon spin concentrator (Sigma-Aldrich; Oakville, ON). The sample was buffer exchanged into HEPES-buffered saline (HBS; 20 mM HEPES pH 7.5, 150 mM NaCl) with 2.25 mM calcium chloride and the His_6 tag and signal peptide remnants were cleaved by overnight incubation at 18 °C with Novagen restriction grade thrombin (EMD Millipore through VWR; Radnor, PA). The cleaved *Tg*AMA1 sample was applied to a Superdex 200 HiLoad 16/60 gel filtration column (GE Healthcare Life Sciences; Baie d'Urfe, QC) pre-equilibrated in HBS, and fractions containing monomeric, monodisperse sample were pooled and concentrated to 20 mg/mL in HBS for binding assays and crystallization trials. The purity of *Tg*AMA1 was determined by SDS-PAGE at each stage of the purification and protein concentrations were analyzed by absorbance at 280 nm based on an extinction coefficient of 67810 M^{-1} cm^{-1}. Similar methods were used for *Tg*AMA1 Y230A, and the final yield of both purified TgAMA1 and *Tg*AMA1 Y230A was approximately 2 mg per litre of insect cell culture.

Protein expression and purification from E. coli

Expression of pGEX-*Tg*RON2-2 was performed using *Escherichia coli* (*E. coli*) and purified as described previously (Lamarque, Besteiro et al. 2011).

TgRON2 synthetic peptide (TgRON2sp) synthesis

A 37 residue region of *Tg*RON2 Type II strain Me49 (*Tg*RON2sp; Asp1297 to Thr1333, with numbering based on the initiation methionine in the signal sequence; UniProt ID B6KV60), was synthesized by Kinexus (Vancouver, BC) with a free amine N-terminus and amidated C-terminus, and was disulfide cyclised. The peptide was confirmed by mass spectrometry to be >95% pure. The lyophilized peptide was solubilized in 100% dimethyl sulfoxide (DMSO) and subsequently diluted in filter sterilized HBS for use in functional and co-crystallization studies. N-terminal fluorescein isothiocyanate (FITC) labelled peptide was similarly synthesized.

Crystallization, data collection and processing

Initial crystals for *Tg*AMA1-*Tg*RON2sp and *Tg*AMA1 Y230A-*Tg*RON2sp were identified after three months in a single condition of the PACT Premier Screen (Molecular Dimensions; Altamonte Springs, FL) - 25% polyethylene glycol (PEG) 1500, 0.1 M MIB buffer pH 4.0. Streak seeding improved the reproducibility and decreased the time to obtain crystals: small crystals were observed after two days and grew to a final size of 0.5 x 2 x 0.2 mm within eighteen days. The final drops consisted of 0.69 µL of *Tg*AMA1 protein (14 mg/mL), 0.11 µL *Tg*RON2sp (13.7mg/mL) (corresponding to an AMA1:RON2 molar ratio of 1:2) with 0.8 µL of reservoir solution equilibrated against 100 µL of reservoir solution. Cryoprotection of the crystals was carried out in reservoir solution supplemented with 12.5% glycerol for 20 seconds and the crystals were flash cooled at 100 K directly in the cryostream. Diffraction data for *Tg*AMA1-*Tg*RON2sp and *Tg*AMA1 Y230A-*Tg*RON2sp were collected on beam line 08ID-1 (CMCF) at the Canadian Light Source (CLS; Saskatoon, SK).

Diffraction quality crystals for *Tg*AMA1 Y230A were grown in sitting drops consisting of 0.8 µL protein (14 mg/mL) and 0.8 µL reservoir solution (21% PEG3350, 0.1 M Tris pH 7.4), and equilibrated against 100 µL of reservoir solution. A single crystal was cryoprotected in reservoir solution supplemented with 12.5% glycerol for 20 seconds and flash cooled at 100 K directly in the cryostream. Diffraction data were collected on beam line 9-2 at the Stanford Synchrotron Radiation Lightsource (SSRL; Menlo Park, CA).

Structure solution and refinement

For all three structures, diffraction data were processed using Imosflm (Battye, Kontogiannis et al. 2011) and Scala (Evans 2006) in the CCP4 suite of programs (Winn, Ballard et al. 2011). The structures of *Tg*AMA1-*Tg*RON2sp, *Tg*AMA1 Y230A-*Tg*RON2sp, and *Tg*AMA1 Y230A were processed to 1.95 Å, 2.55 Å, and 2.35 Å, respectively. Initial phases were obtained by molecular replacement (MR) using PHASER (McCoy, Grosse-Kunstleve et al. 2007) with the unliganded *Tg*AMA1 structure (Protein Data Bank (PDB) ID 2X2Z) as a search model. Adequate electron density was observed in the apical groove of *Tg*AMA1 for both complex structures to trace the majority of the

RON2 peptide *de novo* in COOT (Emsley and Cowtan 2004), with near complete modeling possible upon further cycles of refinements. Solvent molecules were selected using COOT and refinement was carried out using Refmac5 (Murshudov, Vagin et al. 1997). Stereochemical analysis performed with PROCHECK and SFCHECK in CCP4 (Winn, Ballard et al. 2011) showed excellent stereochemistry with more than 94% of the residues in the favored conformations for each structure and no residues modeled in disallowed orientations of the Ramachandran plot. Overall, 5% of the reflections were set aside for calculation of R_{free}. Data collection and refinement statistics are presented in Tables 1 and 3.

Protein Data Bank accession codes

The coordinate and structure factor files have been deposited to the PDB with the following accession codes: *Tg*AMA1-*Tg*RON2sp, 2Y8T; *Tg*AMA1 Y230A, 2Y8R; *Tg*AMA1 Y230A-*Tg*RON2sp, 2Y8S.

Bioinformatics

RON2 sequences used in alignments: *T. gondii* (TGME49_100100), *Neospora caninum* (NCLIV_064620), *P. falciparum* (PF14_0495), *P. vivax* (PVX_117880), and *Babesia bovis* (BBOV_I001630). Sequences were aligned using ClustalW (Thompson, Higgins et al. 1994) and illustrated in ESPript (Gouet, Courcelle et al. 1999).

Binding assays

Techniques for parasite culture, site-directed mutagenesis, fluorescence imaging, as well as analysis of the interaction between *Tg*AMA1 and *Tg*RON2 in enzyme-linked immunosorbent assay (ELISA), mammalian BHK-21 cells and Type I RH strain tachyzoites were conducted as described in previously published protocols (Lamarque, Besteiro et al. 2011). For the binding assay using BHK-21 cells, 0.1 µg/mL to 10 µg/mL of recombinant *Tg*RON2 proteins were used (no binding was observed below 0.1 µg/mL for wild-type *Tg*RON2-2). After five washes in phosphate buffered saline (PBS) to remove unbound protein, cells were fixed in 4% paraformaldehyde picric acid (PAF) and further processed for immunofluorescence using anti-glutathione-S-transferase (GST) antibody

(labelling of recombinant *Tg*RON2-2 mutants) and anti-ectodomain B3.90 antibody (labelling of *Tg*AMA1). For *in situ* binding analysis of *Tg*RON2-2 mutants, fixed intracellular tachyzoites were permeabilized and incubated with *Tg*RON2-2 at various concentrations (0.004 to 10 µg/mL). Binding to *Tg*AMA1 in micronemes was assessed on fixed and permeabilized intracellular parasites by immunofluorescence using anti-GST antibody (labelling of *Tg*RON2-2) and anti-ectodomain B3.90 antibody (labelling of *Tg*AMA1). To test binding of *Tg*RON2sp-FITC, the peptide was used at 0.1 µg/mL for 30 minutes on fixed and permeabilized intracellular tachyzoites. For ELISA, *Tg*AMA1 was coated at 1 µg/mL and *Tg*RON2-2 recombinant proteins used at 0.0009 µg/mL to 10 µg/mL. Results presented are representative of at least three independent experiments. For competitive ELISA assays, *Tg*RON2-2 was used at 1 µg/mL and pre-incubated with increasing concentrations of *Tg*RON2sp for 15 minutes before being added to each well. Surface plasmon resonance (SPR) studies of the *Tg*AMA1-*Tg*RON2-2 assembly were performed on a BIACORE 3000 instrument (GE Healthcare, Biacore AB; Uppsala, Sweden) at 25 °C using running buffer HBSEP (10 mM HEPES pH 7.4, 150 mM NaCl, 3 mM EDTA and 0.005% Biacore TM surfactant), at a flow rate of 50 µL/min. For GST-capture of recombinant *Tg*RON2-2, monoclonal anti-GST (GE Healthcare) was covalently immobilized on a flow cell of a CM5 sensor chip with 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide (EDC)/N-hydroxysuccinimide (NHS) according to the manufacturer's instructions. A control flow cell was prepared with the same chemical treatment, but without antibody. For affinity measurements, each concentration of *Tg*AMA1 or *Tg*AMA1 Y230A was injected following an injection of *Tg*RON2-2 over 2 minutes at 910 nM. Regeneration of the flow cell between each run was achieved with a pulse of 10 mM HCl after a 400 second dissociation step. The binding level of *Tg*AMA1 (520 nM) on *Tg*RON2-2 mutants was measured at 125 seconds, immediately after the end of injection. The absence of binding with GST was controlled in each experiment. For determination of *Tg*RON2sp affinity, *Tg*AMA1 was covalently immobilized on a CM5 sensor chip by amine coupling and *Tg*RON2sp was injected at different concentrations. A pulse of 100 mM HCl was used to regenerate the *Tg*AMA1 coated surface. The kinetic and affinity data were calculated using the BIAevaluation 4.1 software using a Langmuir 1:1 fitting model.

Invasion assays

Inhibition of invasion was performed as described previously (Lamarque, Besteiro et al. 2011) with 200 µg/ml (6 µM) of recombinant *Tg*RON2 proteins. Intracellular parasites were detected by immunofluorescence microscopy using anti-ROP1 antibodies that labelled the nascent PV. When using *Tg*RON2sp, concentrations ranging from 8 nM to 5 µM were used. GST was used to define 100% invasion. Results presented are representative of at least three independent experiments.

Homology modelling

The structural model for the *Pf*RON2 peptide (1201-1239) (UniProt B9A598) was generated based on the structure of *Tg*RON2sp, with which it shares nearly 50% sequence identity, optimized to reflect the insertion of two amino acids in the β-hairpin loop, and filtered through sequence alignments with twelve additional Apicomplexa RON2 sequences. MODELLER 9v8 (Eswar, Webb et al. 2006) was used to perform the multi-template homology modelling of the *Pf*AMA1-*Pf*RON2 peptide complex based on the structures of *Pf*AMA1-IgNAR14l-1 (PDB ID 2Z8V; chain A) and *Tg*AMA1-*Tg*RON2sp (reported here – PDB ID 2Y8T), and sequence alignments of *Pf*AMA1 DI/DII [segments 108-350 and 390-438] (UniProt Q7KQK5, isolate 3D7) and AMA1 from *T. gondii*, *B. bovis*, and *P. vivax*. The N- and C-termini of *Pf*AMA1 were removed to reflect the structurally characterized DI/DII construct, and the DII loop segment (residues 351-389) was also removed as no reliable coordinates are available for this region in *Tg*AMA1-*Tg*RON2sp. The final model was chosen based on its low value of the MODELLER objective function, satisfying the Ramachandran plot (MolProbity) (Chen, Arendall et al. 2010), ERRAT (Colovos and Yeates 1993) and ProQ scores (Wallner and Elofsson 2003), and by visual inspection. No further model optimization was performed.

2.4 Results and discussion

To refine the 65-residue region of *Tg*RON2 previously identified as capable of binding *Tg*AMA1 (*Tg*RON2-2) (Lamarque, Besteiro et al. 2011), we generated deletion constructs targeting both its N- (*Tg*RON2-2 D1) and C-termini (*Tg*RON2-2 D2) (Figure 7A). A

qualitative cellular binding assay showed that both deletion constructs were capable of binding *Tg*AMA1 to the same extent (Figure 7B); however, SPR and invasion studies revealed participation of the N-terminal region of *Tg*RON2-2 while the C-terminal region appeared to be dispensable (Figure 7C). On the basis of these observations, we synthesized a 37 amino acid peptide (*Tg*RON2 synthetic peptide; *Tg*RON2sp) corresponding to residues 1297 to 1333. Because a pair of conserved cysteines present in the central region may endow the peptide with crucial structure, *Tg*RON2sp was disulfide-cyclised. *Tg*RON2sp was active in binding *Tg*AMA1 in cellular binding assays (Figure 7B), host cell invasion inhibition assays at nanomolar concentrations (Figure 7D) [SPR determined K_d of 14.2 nM; Figure 8A], and competition ELISA against *Tg*RON2-2 (Figure 8B).
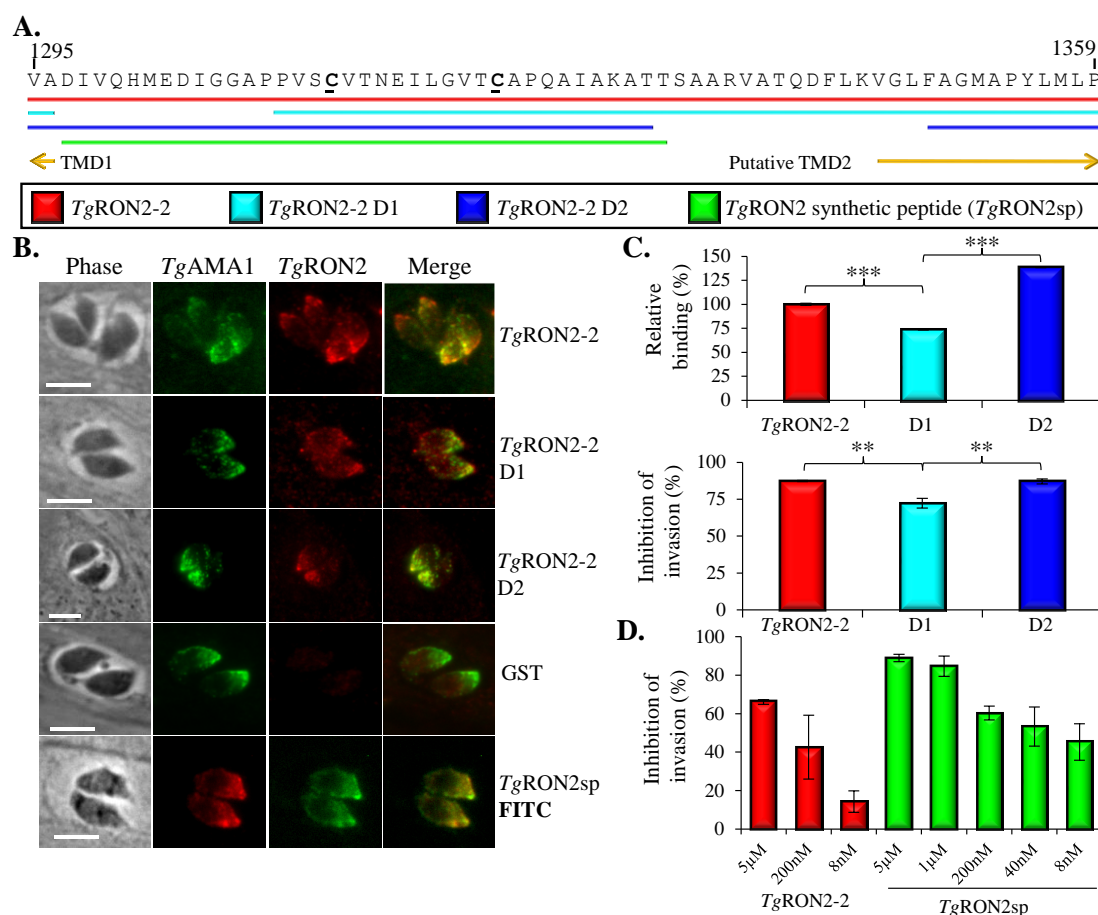
Figure 7: Defining the *Tg*RON2 recognition sequence for *Tg*AMA1.

A. *Tg*RON2-2 deletion constructs; TMD, transmembrane domain. **B.** Immunofluorescence microscopy images of *Tg*RON2 constructs labeling *Tg*AMA1 in micronemes of fixed and permeabilized intracellular parasites using an antibody against GST (labeling of *Tg*RON2-

2) and an antibody against ectodomain B3.90 (labeling of *Tg*AMA1). Scale bars indicate 5 µm. **C.** SPR analysis of *Tg*RON2 constructs binding to *Tg*AMA1 (top) presented as relative percent binding. Immunofluorescence-based measurement showing the ability of *Tg*RON2 constructs to inhibit *T. gondii* invasion of human foreskin fibroblast cells (bottom). Note the statistically significant effect observed relative to *Tg*RON2-2. Student's *t*-test; ** p<0.01, *** p<0.001; means ± standard deviation for *N*=3. **D.** *Tg*RON2sp displays similar invasion inhibition properties to *Tg*RON2-2; means ± standard deviation for *N*=3.

k_a (1/Ms)	k_d (s-1)	K_D (M)
1.8 x10^5	2.55 x10^{-3}	1.42x10^{-8}
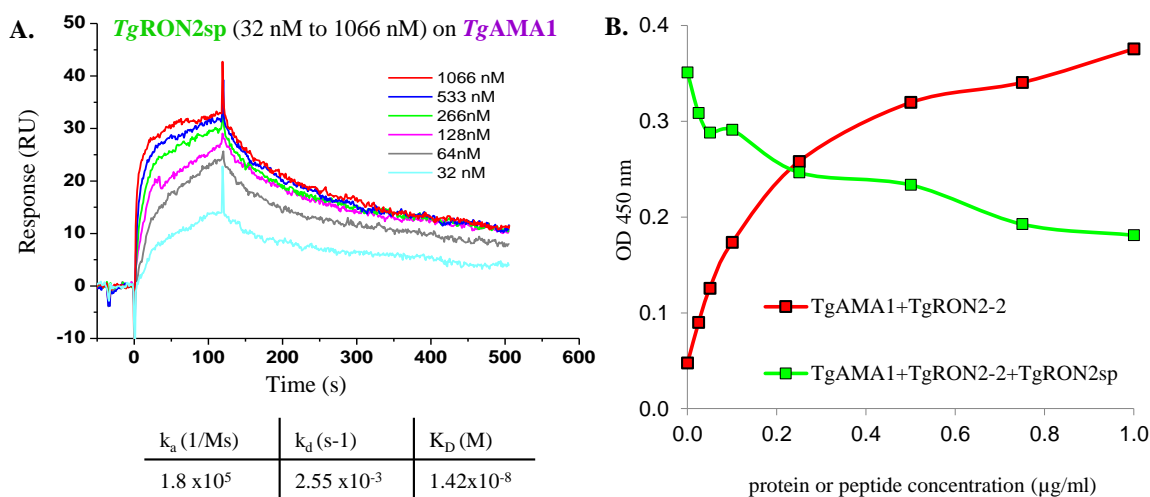
Figure 8: Affinity of *Tg*RON2sp for *Tg*AMA1.

A. SPR based binding kinetics of *Tg*RON2sp on *Tg*AMA1: binding sensorgrams of *Tg*RON2sp at six concentrations (from 32 nM to 1066 nM) to *Tg*AMA1 covalently immobilized (2500-3000 RU) on a CM5 sensor chip. All sensorgrams were obtained by subtracting the signal from the control flowcell. The flow rate was 50 µL/min. **B.** Incubation with *Tg*RON2sp prevents the binding of *Tg*RON2-2 to *Tg*AMA1 in a dose-dependent fashion. Competitive ELISAs were performed by a pre-incubation for 15 minutes of 1 µg/mL *Tg*RON2-2 and increasing concentrations of *Tg*RON2sp (0 to 1 µg/ml) before being added to each well coated with 1 µg/mL of *Tg*AMA1. Values represent means from a representative experiment out of two independent assays.

To structurally characterize the complex, we crystallized *Tg*AMA1 bound to *Tg*RON2sp and refined the structure to a resolution of 1.95 Å (Figure 9A, Table 1). Clear electron density in the apical groove of *Tg*AMA1 enabled accurate tracing of the entire *Tg*RON2sp sequence (Figure 9B). *Tg*RON2sp bound in a U-shaped conformation resulting in a total buried surface area of 3765 Å^2 (Figure 9B, 10, Table 2) with significant shape complementarity (complexation significance score of 1.00 (Krissinel and Henrick 2007)). These features likely endow the MJ with the ability to withstand the shear mechanical force associated with the parasite moving through the constricted MJ ring.
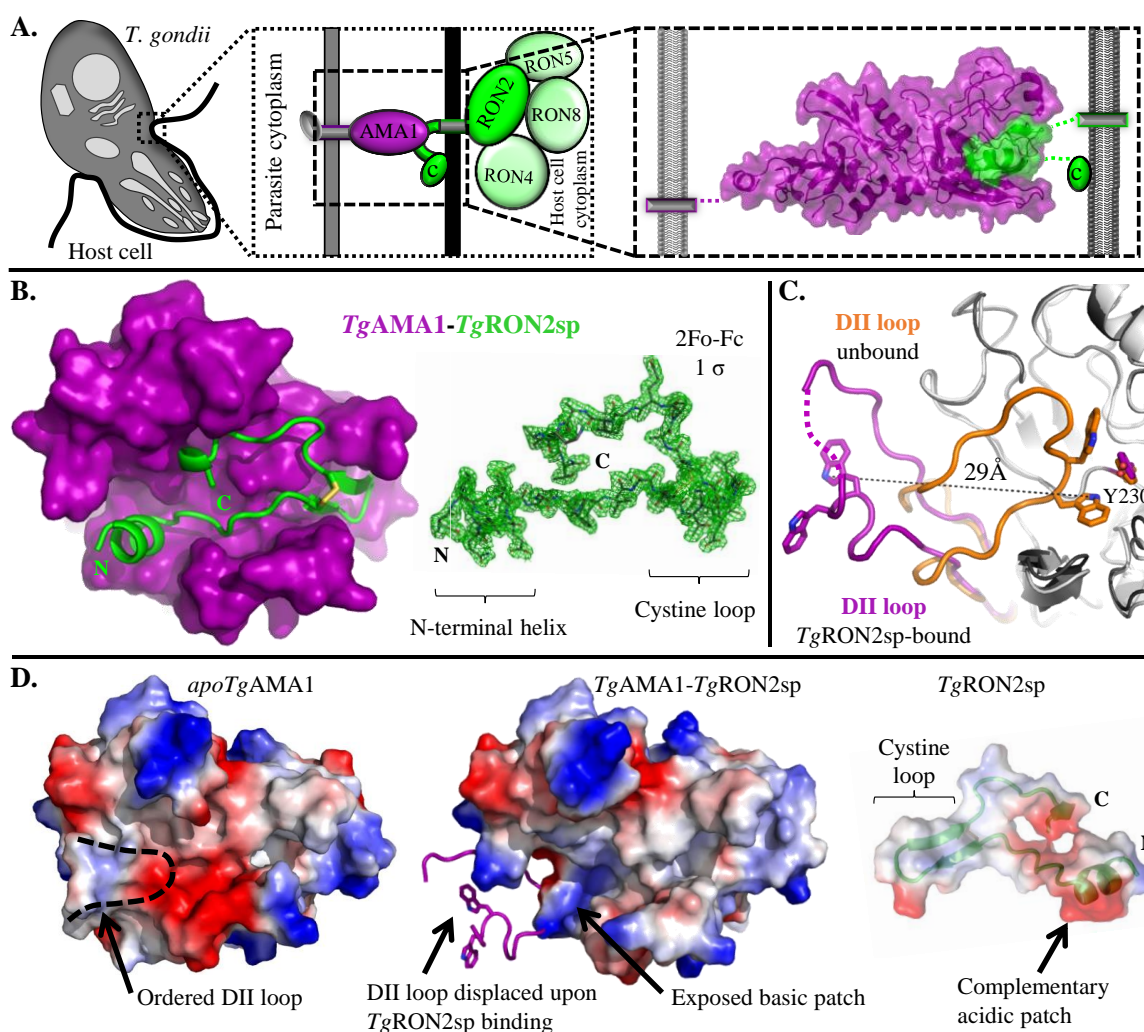
Figure 9: *Tg*RON2sp induces conformational change in *Tg*AMA1.

A. Overview of the essential *Tg*AMA1-*Tg*RON2 complex that links apicomplexan parasites and their target host cells. Proposed extracellular location of *Tg*RON2 C-terminus

denoted by "C" (Tyler and Boothroyd 2011). **B.** Top view of *Tg*RON2sp (green) bound in the *Tg*AMA1 apical groove (purple). Electron density map of *Tg*RON2sp shown in green and contoured at 1.0σ (center) reveals a well-ordered peptide that adopts a U-shaped conformation. **C.** Top view of *Tg*AMA1 showing the 29 Å displacement of the DII loop from its ordered position in the apo form (orange) to its *Tg*RON2sp bound form (purple). To highlight displacement of the DII loop, *Tg*RON2sp is not shown. **D.** Electrostatic renderings of the *Tg*AMA1 apical surface when the DII loop is ordered in the absence of *Tg*RON2sp (left, dotted line highlights position of DII loop) and when it is displaced by *Tg*RON2sp (middle, *Tg*RON2sp not shown in order to visualize the newly exposed basic surface in the *Tg*AMA1 apical groove). *Tg*RON2sp is displayed in an open book view with respect to the apical groove of *Tg*AMA1. The acidic surface on *Tg*RON2sp (right) electrostatically complements the basic patch in the reconfigured *Tg*AMA1 groove (middle).

Table 1: Data collection and refinement statistics for *Tg*AMA1-*Tg*RON2sp.

Data collection statistics	
Spacegroup	P2$_1$
a, b, c (Å)	69.86, 96.38, 78.35
α, β, γ (deg.)	90, 115.64, 90
Wavelength (Å)	0.9795
Resolution range (Å)	45.00 – 1.95 (2.06 – 1.95)
Measured reflections	314359
Unique reflections	68179
Redundancy	4.6 (4.5)
Completeness (%)	100.0 (99.9)
I/σ(I)	10.6 (3.7)
R$_{merge}$[a]	0.105 (0.479)
Refinement Statistics	
Resolution (Å)	39.81 – 1.95
R$_{cryst}$[b]	0.170
R$_{free}$[c]	0.228
No. of atoms	
Protein (A/B/D/E)	3082/260/3131/252
Solvent	604
Boric Acid	8
B-values (Å^2)	
Protein (A/B/D/E)	20.0/21.2/25.1/30.9
Solvent	33.1
Boric Acid	44.1
r.m.s. deviation from ideality	
Bond lengths (Å)	0.020
Bond angles (deg.)	1.94
Ramachandran statistics (%)	
Most favoured	97.5
Allowed	2.5
Disallowed	0.0

Values in parentheses are for the highest resolution shell

[a] R$_{merge}$= $\sum_{hkl} \sum_i |I_{hkl,i} - [I_{hkl}]| / \sum_{hkl} \sum_i I_{hkl,i}$, where $[I_{hkl}]$ is the is the average of symmetry related observations of a unique reflection

[b] R$_{cryst}$=$\sum |F_{obs}-F_{calc}|/\sum F_{obs}$, where F$_{obs}$ and F$_{calc}$ are the observed and the calculated structure factors, respectively

[c] R$_{free}$ is R using 5% of reflections randomly chosen and omitted from refinement

Figure 10: Detailed atomic interactions that define the *Tg*AMA1-*Tg*RON2sp interface.

A. *Tg*AMA1 is shown in purple, *Tg*RON2sp in green and buried solvent molecules as blue spheres. **B.** Top – *Tg*AMA1 Tyr241 forms a pi-pi stacking interaction with His1301 on *Tg*RON2sp. Middle – The base of the *Tg*RON2sp helix is anchored by a combination of five hydrogen bonds incorporating backbone and side chain interactions. Bottom – *Tg*AMA1 Leu99, conserved in sequence and structure with *Pf*AMA1 and *Pv*AMA1, forms the base of the pocket occupied by the N-terminal helix of *Tg*RON2sp. **C.** Shape complementarity is observed as *Tg*RON2sp Pro1309 is caged by a trio of *Tg*AMA1 tyrosines and thus occupies the pocket left vacant by the displacement of Trp354 with the DII loop. **D.** The hydroxyl group of the tyrosine central to the *Tg*AMA1 apical groove (Tyr230) orients the amidated C-terminal threonine of the peptide (Thr1333) with a single hydrogen bond and also the *Tg*RON2sp coil connector with two backbone hydrogen bonds. **E.** Ten of the twenty-three predicted hydrogen bonds are dedicated to ordering the previously disordered Loop 2 of *Tg*AMA1 to form the contiguous antiparallel beta-sheet and anchoring the disulfide bond. **F.** Pro1325 enhances shape complementarity while five hydrogen bonds are spread over five residues of the C-terminal coil of *Tg*RON2sp and likely responsible for ordering this region of the peptide.

Table 2: Interactions at the *Tg*AMA1-*Tg*RON2sp interface.

Hydrogen bonds			Buried Surface Area (BSA) > 35 Å²			
*Tg*RON2sp	*Tg*AMA1	Distance (Å)	*Tg*RON2sp	BSA (Å²)	*Tg*AMA1	BSA (Å²)
Glu1303 [Oε1]	**Arg111 [N]**	3.46	**His1301**	51.04	**Tyr110**	79.48
Glu1303 [Oε2]	Gln361 [Nε2]	2.55	Met1302	45.47	**Arg111**	38.52
Asp1304 [O]	**Arg111 [NH1]**	2.98	Glu1303	86.45	Thr144	48.70
Asp1304 [Oδ1]	Gln361 [Nε2]	2.75	**Asp1304**	78.13	Glu145	76.48
Gly1306 [O]	Met233 [N]	3.65	**Ile1305**	113.42	**Phe163**	107.46
Val1311 [O]	**Tyr230 [OH]**	2.56	Gly1306	60.63	Val164	46.14
Val1311 [N]	**Tyr230 [OH]**	3.54	Ala1308	53.14	Thr165	47.25
Cys1313 [O]	**Met204 [N]**	3.06	**Pro1309**	93.69	**Ile171**	45.01
Thr1315 [O]	Val202 [N]	2.85	Val1311	75.72	Asn184	35.70
Thr1315 [N]	Val202 [O]	2.90	**Cys1313**	39.23	Lys200	62.44
Asn1316 [Nδ2]	**Phe197 [O]**	3.47	Val1314	85.69	**Met204**	89.37
Asn1316 [Nδ2]	Lys200 [O]	3.70	Thr1315	62.38	**Tyr230**	62.87
Asn1316 [Nδ2]	Thr201 [Oγ1]	2.76	Asn1316	67.11	Met233	72.35
Glu1317 [N]	Lys200 [O]	2.99	Glu1317	67.53	**Tyr241**	75.19
Thr1322 [Oγ1]	Thr165 [N]	3.59	**Ile1318**	106.78	Trp253	47.31
Cys1323 [O]	Val164 [N]	2.94	**Leu1319**	106.09	Gln361	47.36
Cys1323 [N]	Val164 [O]	2.92	Thr1322	53.38		
Gln1326 [Nε2]	Thr144 [Oγ1]	3.69	Pro1325	49.86		
Ala1327 [O]	Glu145 [N]	2.87	Gln1326	40.46		
Ala1329 [N]	Pro143 [O]	2.97	Ile1328	79.40		
Lys1330 [N]	**Glu145 [Oε2]**	3.16	Ala1329	68.67		
Ala1331 [O]	Trp253 [Nε1]	3.35	Ala1331	58.61	**Phe197**	29.52
Thr1333 [O]	**Tyr230 [OH]**	2.90	**Thr1332**	98.79	**Tyr213**	20.50
			Thr1333	40.46	**Tyr215**	32.96

XX = *Tg*RON2-2 mutant in this study; **XX** = *Tg*AMA1 mutant in this study

Another feature of the *Tg*AMA1-*Tg*RON2sp co-structure is a substantial conformational change in the DII loop of AMA1 (Figure 9C, 11) - an influential substructure in RON2 coordination (Pizarro, Vulliez-Le Normand et al. 2005, Coley, Gupta et al. 2007, Collins, Withers-Martinez et al. 2009, Crawford, Tonkin et al. 2010). In the apo *Tg*AMA1 structure (Crawford, Tonkin et al. 2010) the DII loop is intimately associated with the base of the hydrophobic groove through numerous interactions, including a pair of tryptophan residues that anchor the loop into hydrophobic pockets in the apical groove (Figure 9C). By inducing the conformational changes leading to displacement of the DII

loop, *Tg*RON2sp is able to expose the fully functional binding surface on *Tg*AMA1 that displays enhanced shape and charge complementarity (Figure 9D). In its reconfigured form, the apical groove presents an extended basic patch that complements an acidic patch on the *Tg*RON2sp N-terminal helix (Figure 9D). These data support a hypothesis of a stepwise binding mechanism where contact is initiated between the end of the AMA1 apical groove opposite the DII loop and the disulfide-tethered β-hairpin loop region of RON2, but elucidation of this hypothesis will require further investigations (Figure 9B).
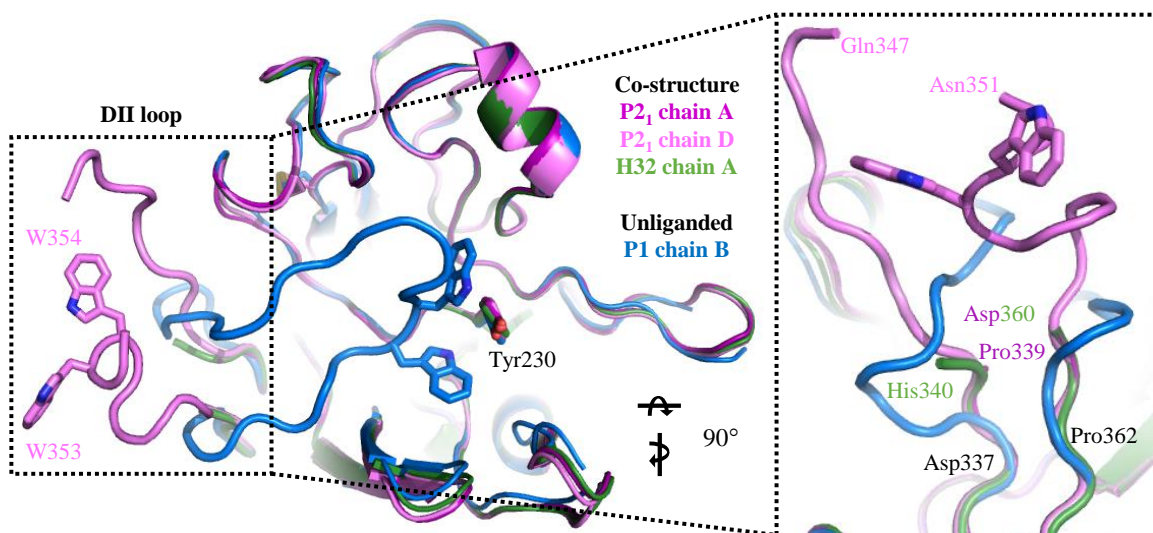
Figure 11: Crystal contacts and an alternate crystal form of *Tg*AMA1-*Tg*RON2sp support the disordered structure of the displaced DII loop.

Overlaying the two *Tg*AMA1 chains of the co-structure in the monoclinic spacegroup (A – purple; D – violet) with the additional single *Tg*AMA1 chain from another crystal form of the complex (hexagonal space group; forest green) and chain B of unliganded *Tg*AMA1 (blue) reveals differences in the ordered structure of the displaced DII loop. In the unliganded structure, the DII loop is modeled in its entirety with the two tryptophans at the tip of the loop (Trp353/354) anchored by the apical groove central tyrosine (Tyr230). Inset – The last spatially conserved residues defining the base of the DII loop between all four structures are Asp337 at the N-terminal end and Pro362 at the C-terminal end. The loop is nearly completely modeled in chain D of the P2₁ crystal form (missing Ala348-Trp350), but the loop in chain A is only modeled to Pro339 and Asp360. A hexagonal crystal form with one complex in the asymmetric unit revealed crystal packing with the DII loops

directed into large solvent channels and subsequently disordered; the extent of ordered structure closely mirrors that seen in chain A of the monoclinic crystal form, with the base of the DII loop modeled to His340 and Asp360.

Guided by our high-resolution *Tg*AMA1-*Tg*RON2sp co-structure and sequence alignments (Figure 12A), we identified key residues on *Tg*RON2-2 and probed their functional relevance. Alanine substitutions in the N-terminal helix (Asp1297, His1301, Asp1304 and Ile1305), the coil connecting the helix and cystine loop (Pro1309), the cystine loop (Cys1313, Ile1318, Leu1319 and Cys1323), and the C-terminal coil (Thr1332) all showed some degree of reduced binding to *Tg*AMA1 as determined by SPR and ELISA (Figure 12B, 13), consistent with an extensive, composite *Tg*AMA1-*Tg*RON2sp interface. Of particular note, the broadly conserved residues Asp1297, Asp1304, Pro1309, Cys1313, and Cys1323, were identified as critical for *Tg*AMA1 binding, as was the less conserved Leu1319 located in the cystine loop. An important role for the cysteine-dependant conformation of RON2 in promoting complex formation is corroborated by the reduced ability of the *Tg*RON2-2 cysteine mutants to bind *Tg*AMA1 in the micronemes of *T. gondii* tachyzoites (Figure 14) and inhibit host cell invasion by *T. gondii* (Figure 12C).
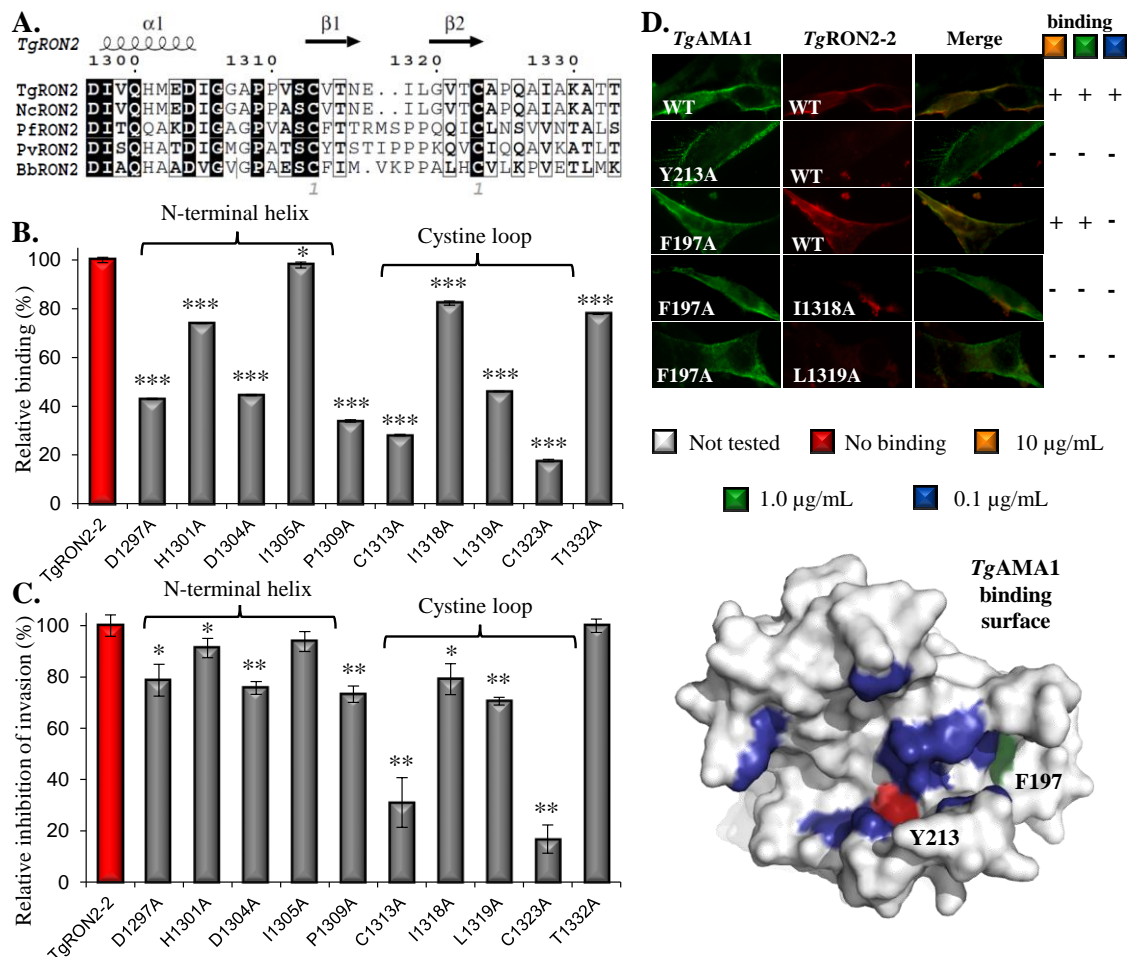
Figure 12: Hot spot residues in *Tg*AMA1-*Tg*RON2sp binding.

A. Alignment of apicomplexan RON2 sequences corresponding to *Tg*RON2sp; secondary structure and numbering are based on *Tg*RON2. **B.** SPR-measured relative binding of *Tg*RON2-2 mutants to *Tg*AMA1. **C.** Effect of *Tg*RON2-2 mutants (6 μM) in inhibiting invasion by *T. gondii*, highlighting the crucial roles of Cys1313 and Cys1323. Student's *t*-test; * $p<0.05$, ** $p<0.01$, *** $p<0.001$; means ± standard deviation for $N = 3$. **D**. Binding consequences of *Tg*AMA1 mutants and synergistic *Tg*RON2-2 mutants in cell-based assays (top) and spatially represented on the surface of *Tg*AMA1 (bottom). WT, wild type.
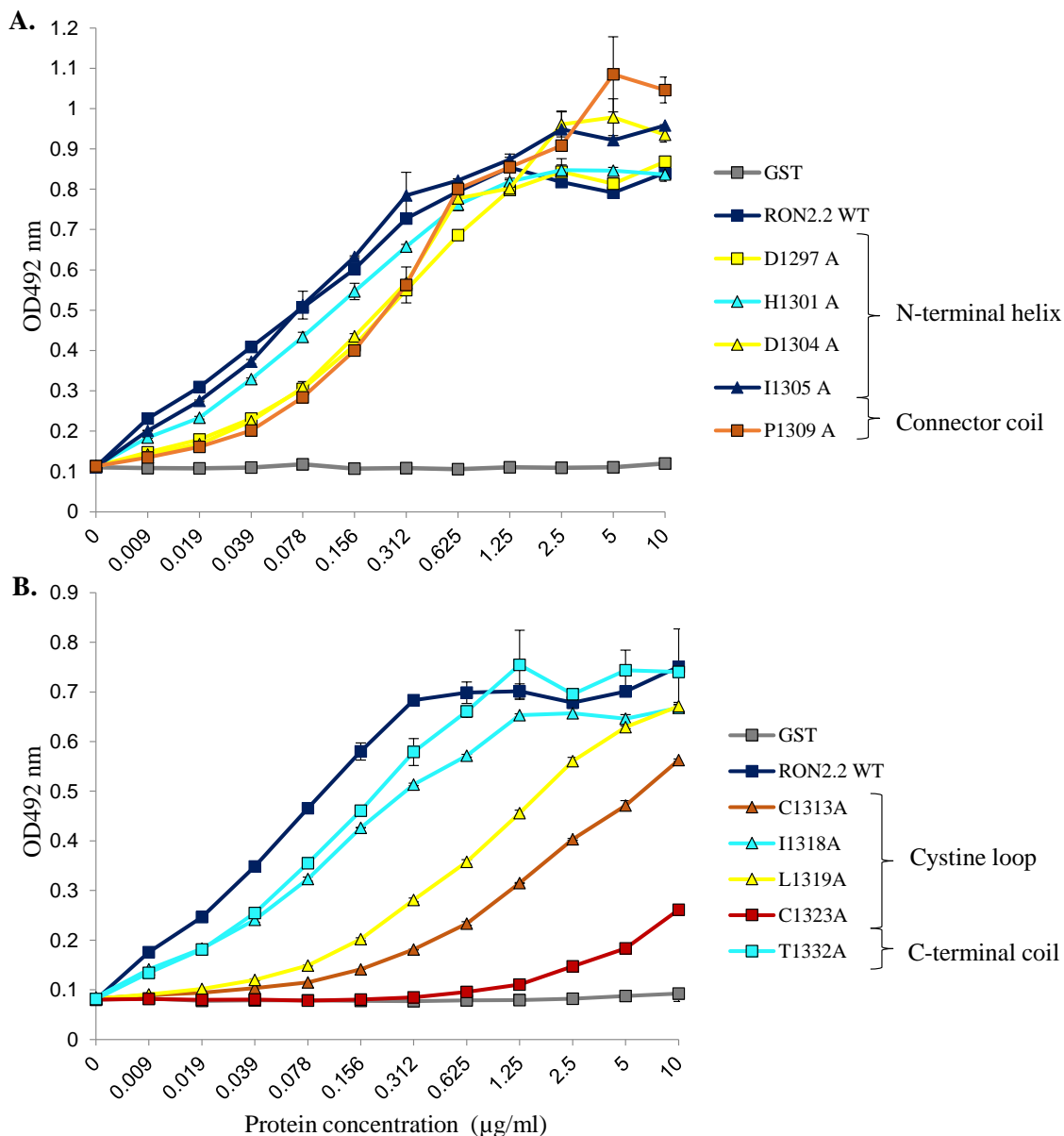
Figure 13: Binding analysis of *Tg*RON2-2 mutants by ELISA.

ELISAs were carried out with increasing concentrations of recombinant *Tg*RON2-2 mutants on *Tg*AMA1 coated at 1 µg/mL. GST alone was used as a control. Bound *Tg*RON2-2 protein was detected using rat anti-GST antibodies (1:1000) followed by an anti-rat Immunoglobulin (Ig) G horseradish peroxidase conjugate (1:2500) (ZyMAX, Invitrogen; Burlington, ON). Data are represented as mean ± standard deviation of three independent experiments. **A.** Mutants in the N-terminal helix and the coil connecting the helix and cystine loop; colored based on a thermal scale relative to wild-type *Tg*RON2-2. **B.** Mutants in the cystine loop and the C-terminal coil; coloured as in **A**.

µg/ml	10	1	0.5	0.1	0.02	0.04	
GST	-	-	-	-	-	-	
TgRON2-2	+	+	+	+	+	-	
TgRON2-2 D1297A	+	+	+	+	-	-	N-terminal helix
TgRON2-2 H1301A	+	+	+	+	-	-	
TgRON2-2 D1304A	+	+	+	+	-	-	
TgRON2-2 I1305A	+	+	+	+	-	-	
TgRON2-2 P1309A	+	+	+	-	-	-	Connector coil
TgRON2-2 C1313A	-	-	-	-	-	-	Cystine loop
TgRON2-2 I1318A	+	+	+	-	-	-	
TgRON2-2 L1319A	+	+	+	-	-	-	
TgRON2-2 C1323A	-	-	-	-	-	-	
TgRON2-2 T1332A	+	+	+	+	+	-	C-terminal coil

Figure 14: *In vitro* binding analysis of *Tg*RON2-2 mutants.

Fixed intracellular tachyzoites were permeabilized and incubated with *Tg*RON2-2 at various concentrations (0.04 to 10 µg/mL). Binding of *Tg*RON2-2 to *Tg*AMA1 in micronemes was assessed by immunofluorescence using anti-GST antibody (labelling of *Tg*RON2-2) and anti-ectodomain B3.90 antibody (labelling of *Tg*AMA1). Scale bar, 5µm.

To probe the role of specific AMA1 residues in promoting complex formation, we selected a set of eleven *Tg*AMA1 residues for mutagenesis studies (Figure 15 and Table 2). Tyr230 was initially targeted as the analogous residue in *Pf*AMA1 (Tyr251) was previously implicated in complex formation (Collins, Withers-Martinez et al. 2009). The Tyr230 to Ala230 (Y230A) substitution showed no detectable effect in protein conformation (Figure 16A), or cell-based (Figure 16B) and ELISA (Figure 16C) assays, and only an order of magnitude weaker binding by SPR (K_D of 10.2 nM versus K_D of 1.0 nM with the GST fusion of *Tg*RON2-2) (Figure 16D, E). A minimal effect for Tyr230 in promoting or stabilizing the *Tg*AMA1-*Tg*RON2sp complex was confirmed by the TgAMA1 Y230A structure and the *Tg*AMA1 Y230A-*Tg*RON2sp co-structure (Table 3,

Figure 17). The remaining *Tg*AMA1 mutants were expressed on the surface of BHK-21 cells and tested for their ability to bind *Tg*RON2-2 (Figure 12D, 18). Substitutions of Phe197 (base of cystine loop) or Tyr213 (base of tyrosine cage surrounding Pro1309) to alanine impaired binding. In fact, Y213A showed no binding at any of the concentrations tested, suggesting that this highly conserved residue is a critical anchor point during complex formation. The F197A mutant showed a more moderate effect with binding disrupted only at the lowest concentration. Thus, the F197A variant provided an opportunity to test the potential of synergistic *Tg*AMA1-*Tg*RON2 mutations. Total abrogation of complex formation was observed between *Tg*AMA1 F197A and either *Tg*RON2-2 I1318A or L1319A (tip of the cystine loop). These synergistic effects prompted us to generate three additional *Tg*RON2-2 constructs comprising (i) the N-terminal helix to cystine loop, (ii) the coil connector to C-terminus, and (iii) the N-terminal helix to coil connector (Figure 19). Analysis of these constructs by ELISA demonstrated clear binding for the first two, but no interaction with the third, effectively isolating the most critical region of *Tg*RON2-2 to the cystine loop. These results enable the refinement of our proposed mechanism of binding where the cystine loop in combination with the Tyr213-Pro1309 bridge act as a primary structural brace to anchor the peptide in the apical groove and promote displacement of the DII loop.
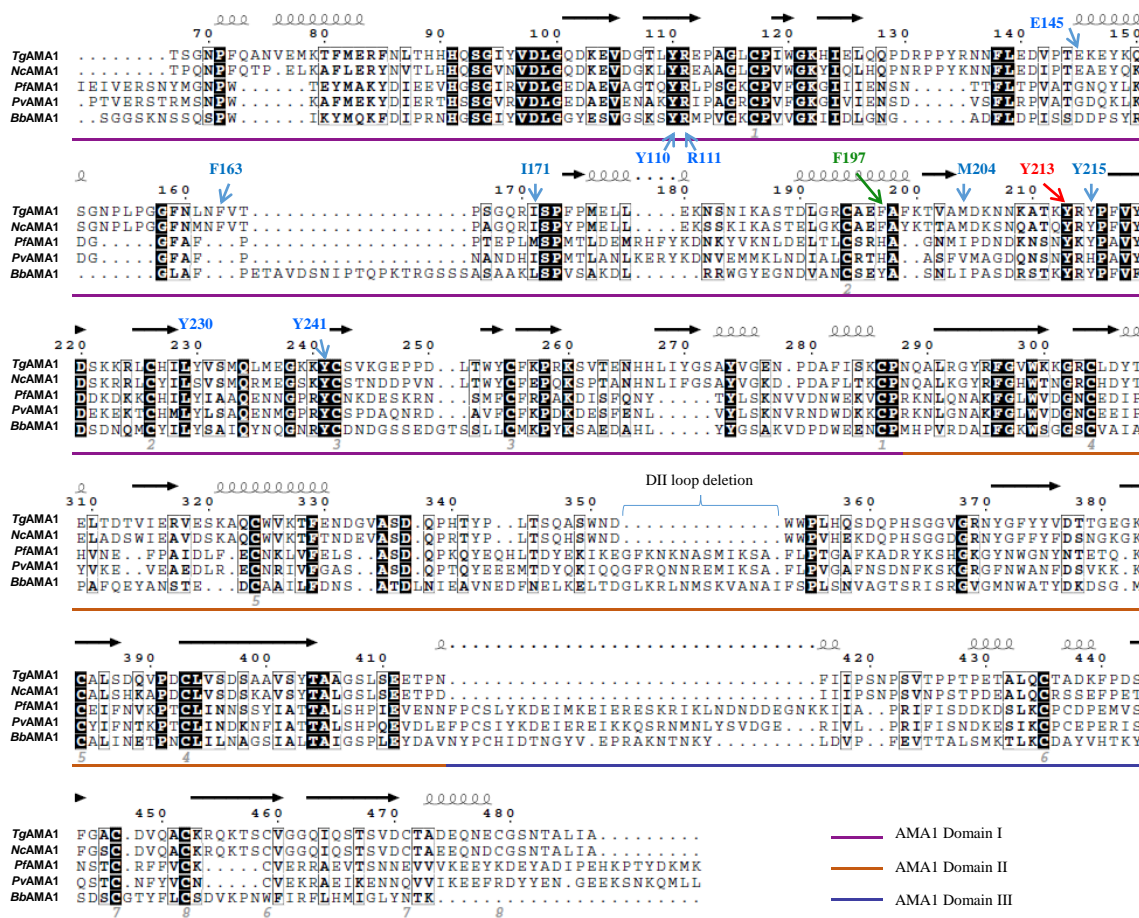
Figure 15: *Tg*AMA1 mutations used in this study.

AMA1 sequence alignment of the fully processed three-domain extracellular region using Kalign and ESPript. Accession numbers for aligned AMA1 sequences are as follows: *T. gondii* (ME49_055260), *Neospora caninum* (BAF45372), *P. falciparum* (XP_001348015.1), *P. vivax* (XP_001615447), and *Babesia bovis* (AAS58045.1), with the *P. vivax* sequence modified to represent that crystallized by Pizarro et al 2005 (Pizarro, Vulliez-Le Normand et al. 2005). Secondary structure and numbering is based on *Tg*AMA1. Mutants in this study are indicated, and coloured as in Figure 12D.

A. Absorbance (mAU)

Elution volume (mL)

kDa
58
46
MW *Tg*AMA1 *Tg*AMA1 Y230A

I: 75kDa
II: 43kDa
III: 29kDa

B.

DIC   *Tg*AMA1 Y230A

*Tg*RON2-2   Merge

C. OD492 nm

Protein concentration (ng/ml)

*Tg*AMA1
*Tg*AMA1 Y230A

D. *Tg*AMA1 (2.5 nM_350 nM) on *Tg*RON2-2

Response (RU)

Time (s)

k_a (1/Ms)	k_d (s-1)	K_D (M)
4.3×10^5	4.26×10^{-4}	9.89×10^{-10}

*Tg*AMA1 Y230A

Response (RU)

348 nM
174 nM
87 nM
43 nM
21 nM
10 nM
5 nM
2.5 nM

Time (s)

k_a (1/Ms)	k_d (s-1)	K_D (M)
$3,53 \times 10^5$	3.61×10^{-3}	1.02×10^{-8}

E. Response (RU)

*Tg*AMA1

*Tg*AMA1 Y230A

*Pf*AMA1
*Pv*AMA1

Time (s)

Figure 16: Preliminary analysis of *Tg*AMA1 Y230A binding to *Tg*RON2-2.

A. Elution of Ni-affinity purified native and mutant *Tg*AMA1 off the size exclusion column revealed similar elution profiles suggesting the structural core of the mutant was conserved. Inset – SDS-PAGE showed the two proteins migrating in a similar pattern at the expected molecular weight of 48 kDa. **B.** *Tg*RON2-2 bound to *Tg*AMA1 Y230A expressed on the surface of BHK-21 cells. **C.** ELISA was carried out with increasing concentrations of *Tg*RON2-2 and recombinant *Tg*AMA1 (orange) or *Tg*AMA1 Y230A (yellow) proteins coated at 1 μg/mL. Bound *Tg*RON2-2 protein was detected using rat anti-GST antibody (1:1000) followed by an anti-rat IgG horseradish peroxidase conjugate (1:2500) (ZyMAX,

Invitrogen) Binding was visualized using SIGMAFAST-o-phenylenediamine tablet and absorbance read at 492 nm. Data are represented as mean ± standard deviation of three independent experiments. **D.** Multi concentration SPR binding study of *Tg*AMA1 and *Tg*AMA1 Y230A on *Tg*RON2-2: binding sensorgrams of *Tg*AMA1 (top) and *Tg*AMA1 Y230A (bottom) at 6 concentrations (from 2.5 nM to 350 nM) to *Tg*RON2-2 (GST fusion) captured at 612 RU on immobilized anti-GST antibody. All sensorgrams were obtained by subtracting the signal from the control flowcell. The flow rate was 50 µL/min. **E.** Simplified SPR sensorgram of *Tg*RON2-2 on *Tg*AMA1 (350 nM) highlights the higher dissociation rate of *Tg*AMA1 Y230A relative to *Tg*AMA1. Note that *Pf*AMA1 and *Pv*AMA1 have no cross-reactivity with *Tg*RON2-2.

Table 3: Data collection and refinement statistics for *Tg*AMA1 Y230A and *Tg*AMA1 Y230A-*Tg*RON2sp.

	*Tg*AMA1 Y230A	*Tg*AMA1 Y230A-*Tg*RON2sp
Data collection statistics		
Spacegroup	P1	P2$_1$
a, b, c (Å)	67.52, 76.09, 88.78	70.52, 96.83, 78.33
α, β, γ (deg.)	71.89, 73.37, 73.50	90, 116.75, 90
Wavelength (Å)	0.9795	0.9795
Resolution range (Å)	38.59 – 2.35 (2.48 – 2.35)	62.97 – 2.55 (2.69 – 2.55)
Measured reflections	152470	173635
Unique reflections	63415	30766
Redundancy	2.4 (2.4)	5.6 (5.6)
Completeness (%)	96.9 (96.1)	100.0 (100.0)
I/σ(I)	7.2 (1.9)	11.9 (4.1)
R$_{merge}$[a]	0.100 (0.456)	0.108 (0.415)
Refinement statistics		
Resolution (Å)	38.59 – 2.45	56.70 – 2.55
R$_{cryst}$[b]	0.212	0.185
R$_{free}$[c]	0.305	0.263
No. of atoms		
Protein (A/B/D/E)	3158/3178/3211/3122	3037/252/3091/245
Solvent	533	212
Boric Acid	N/A	8
Glycerol	N/A	18
B-values (Å^2)		
Protein (A/B/D/E)	25.5/29.1/28.8/32.2	28.4/31.4/36.3/50.9
Solvent	27.91	33.51
Boric Acid	N/A	37.80
Glycerol	N/A	51.27
r.m.s. deviation from ideality		
Bond lengths (Å)	0.014	0.016
Bond angles (deg.)	1.62	1.70
Ramachandran statistics (%)		
Most favoured	94.6	95.3
Allowed	5.4	4.7
Disallowed	0.0	0.0

Values in parentheses are for the highest resolution shell

[a] R$_{merge}$= $\sum_{hkl} \sum_i |I_{hkl,i} - [I_{hkl}]| / \sum_{hkl} \sum_i I_{hkl,i}$, where $[I_{hkl}]$ is the average of symmetry related observations of a unique reflection

[b] R$_{cryst}$=$\sum|F_{obs}-F_{calc}|/\sum F_{obs}$, where F$_{obs}$ and F$_{calc}$ are the observed and the calculated structure factors, respectively

[c] R$_{free}$ is R using 5% of reflections randomly chosen and omitted from refinement

Figure 17: Structural analysis of the role of *Tg*AMA1 Y230 in DII loop stabilization and RON2 coordination.

Left – Apical view of apo *Tg*AMA1 Y230A DI surface with secondary structure DII loop (yellow). Inset - overlay with apo *Tg*AMA1 (orange) shows reorganization of the DII loop. Right – Apical view of *Tg*AMA1 Y230A in complex with *Tg*RON2sp (dark grey). Inset - overlay with *Tg*AMA1-*Tg*RON2sp (purple-green) highlights a limited role for Tyr230 in forming a bifurcated hydrogen bond.
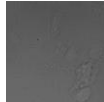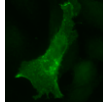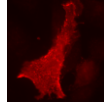
	TgRON2-2			TgRON2-2 I1318A			TgRON2-2 L1319A			DIC	TgAMA1	TgRON2
µg/ml	10	1	0.1	10	1	0.1	10	1	0.1			
TgAMA1	+	+	+	+	+	+/-	+	+	+/-			
TgAMA1 R111A	+	+	+									
TgAMA1 Y241A	+	+	+									
TgAMA1 Y110A	+	+	+									
TgAMA1 Y213A	-	-	-									
TgAMA1 Y215A	+	+	+									
TgAMA1 Y230A	+	+	+									
TgAMA1 F163A	+	+	+									
TgAMA1 I171A	+	+	+									
TgAMA1 F197A	+	+	-	-	-	-	-	-	-			
TgAMA1 M204A	+	+	+									
TgAMA1 E145A	+	+	+									

N-terminal helix — Connector coil — Cystine loop — C-terminal coil

Figure 18: *In vitro* binding analysis of *Tg*AMA1 mutants.

Transiently transfected BHK-21 cells expressing *Tg*AMA1 or a mutated version of *Tg*AMA1 at their surface were incubated with *Tg*RON2-2 or a mutated version of *Tg*RON2-2 at concentrations ranging from 0.1 µg/mL to 10 µg/mL, for one hour. After five washes in PBS to remove unbound proteins, cells were fixed in 4% PAF and further processed for immunofluorescence. Binding was considered as negative when cells expressing a strong signal of *Tg*AMA1 at their surface (using anti-ectodomain B3.90 antibody) were absolutely negative for the detection of *Tg*RON2-2 using anti-GST antibody.
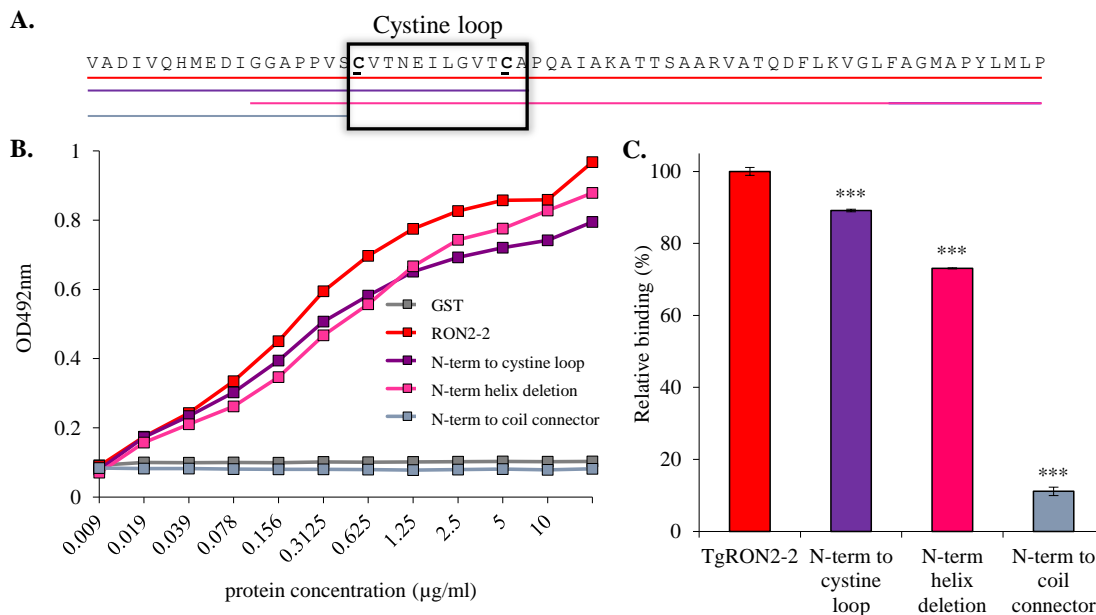
Figure 19: Demonstration of the cystine loop as the critical substructure for *Tg*AMA1-*Tg*RON2sp complex formation.

A. *Tg*RON2-2 (red) with three additional truncation constructs honing in on the role of individual substructures; region after the cystine loop (purple), N-terminal helix (pink), and cystine loop (blue). **B.** An ELISA assay and **C.** a SPR analysis shows a limited role for the complete section of *Tg*RON2-2 from the end of the cystine loop to the C-terminus (purple), and a construct missing just the N-terminal helix is still competent to bind *Tg*AMA1. However, once the cystine loop is removed (blue) there is no detectable binding, supporting the insufficiency of the N-terminal helix and coil connector alone to displace the DII loop. Together, these constructs highlight a key role for the cystine loop in mediating complex formation and support our proposed binding mechanism. SPR-measured binding of *Tg*RON2 segments to *Tg*AMA1 in **C**. is expressed in % of binding relative to *Tg*RON2-2. Data are represented as mean ± standard deviation of three independent experiments; *t*-test, *** p<0.001.

Having analyzed the AMA1-RON2 interaction in *T. gondii*, we set out to evaluate its molecular conservation across Apicomplexa. We recently showed that the AMA1-RON2 binding event is conserved in *Plasmodium* species (Lamarque, Besteiro et al. 2011). Sequence analysis revealed a two amino acid insertion in the *Pf/Pv*RON2 cystine loop (Figure 12A) and, correspondingly, the *Pf*AMA1 apo structure (Bai, Becker et al. 2005)

displayed an extended apical groove relative to *Tg*AMA1 (Crawford, Tonkin et al. 2010) that appeared ideally suited to complement the predicted *Pf*RON2 cystine loop extension (Figure 20A). Thus, although the cystine loop region is important for the initial interaction, it may also serve to regulate cross genera specificity.

A. *Pf*AMA1-*Pf*RON2sp homology model *Tg*RON2sp

*Tg*RON2sp

*Tg*AMA1, extended *Pf*AMA1 platform

Extended *Pf*RON2 loop

B. *Pf*AMA1 - mAb 4G2   *Pf*AMA1 - mAb 1F9

*Pf*AMA1 - IgNAR14l-1   *Pf*AMA1 - R1 peptide
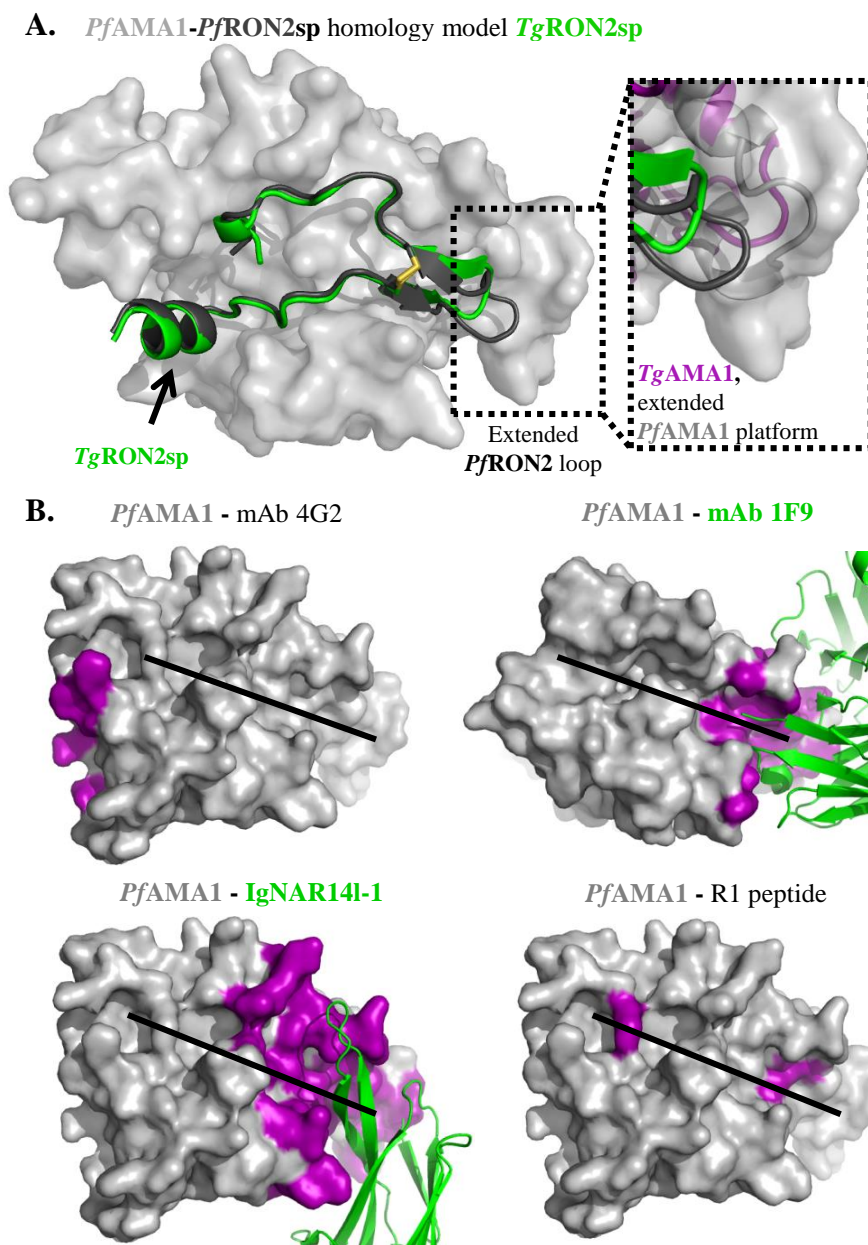
Figure 20: A model for *Plasmodium* invasion.

A. Homology model of *Pf*AMA1-*Pf*RON2sp complex reveals an extended platform in the apical groove that accommodates the extended *Pf*RON2sp cystine loop (black) relative to the overlaid *Tg*RON2sp (green). **B.** Apical surface views of *Pf*AMA1 (grey) bound to

invasion inhibitory molecules (green) with *Pf*AMA1 residues implicated in binding shown in purple. PDB ID for IgNAR14l-1, mAb 4G2 and R1 is 2Z8V, and for mAb 1F9 is 2Q8A. Straight black lines denote the ligand binding groove.

The *Tg*AMA1-*Tg*RON2sp structure provides a basis for inhibitory mechanisms derived from recent antibody (monoclonal antibody (mAb) 4G2, mAb 1F9, IgNAR14l-1) and peptide (R1) studies (Figure 20B). Mutagenesis studies suggested that the cross-reactive mAb 4G2 binds to the stem of the *Pf*AMA1 DII loop (Pizarro, Vulliez-Le Normand et al. 2005, Collins, Withers-Martinez et al. 2009). It is now clear that mAb-mediated anchoring of the DII loop would impair the necessary conformational changes induced by RON2 binding. Opposite the DII loop, mAb 1F9 binds the *Pf*AMA1 DI apical surface (Coley, Parisi et al. 2006, Coley, Gupta et al. 2007), whereas IgNAR14l-1 forms an extensive interface with the same region (Henderson, Streltsov et al. 2007), both effectively occluding the RON2 cystine loop binding surface. Nuclear magnetic resonance (NMR) analysis showed that the R1 peptide contacts residues throughout the groove (Richard, MacRaild et al. 2010), and interpreted in the context of our data, may displace the DII loop in a similar fashion to *Tg*RON2sp. Collectively, these data provide a valuable resource for rational design of inhibitory small molecules, peptide mimetics and engineered vaccines. Indeed, the ability of *Tg*RON2sp to inhibit invasion at nanomolar concentrations is a promising step towards these goals.

Chapter 3: Investigating the structural basis for cross-genera and cross-species specificity in the moving junction proteins AMA1 and RON2

Adapted from:

Vulliez-Le Normand, B.[*], **M. L. Tonkin**[*], M. H. Lamarque[*], S. Langer, S. Hoos, M. Roques, F. A. Saul, B. W. Faber, G. A. Bentley[†], M. J. Boulanger[†] and M. Lebrun[†] (2012). "Structural and functional insights into the malaria parasite moving junction complex." PLoS Pathogens **8**(6): e1002755.

[*] These authors contributed equally

[†] Co-corresponding authors

Contributions:

MLT cloned the *Pf*AMA1 gene. SL generated the *Pf*AMA1 virus. **MLT** produced and purified the protein. **MLT** designed the *Pf*RON2sp1 peptide, crystallized the complex with *Pf*AMA1, and solved and refined the *Pf*AMA1-*Pf*RON2sp1 co-structure. BVN and GAB solved the *Pf*AMA1-*Pf*RON2sp2 and *Pf*AMA1-R1 co-structures. **MLT**, GAB and MJB analyzed the three co-structures. MHL, MR and ML performed the SPR, binding and invasion experiments. BVN and GAB performed the ITC experiment. BVN, **MLT**, MHL, GAB, MJB and ML wrote the paper.

3.1 Abstract

Members of phylum Apicomplexa, which include the malaria parasite *Plasmodium*, share many features in their invasion mechanism in spite of their diverse host cell specificities and life cycle characteristics. The formation of a moving junction between the membranes of the invading apicomplexan parasite and the host cell is common to these intracellular pathogens. The moving junction contains two key parasite components: the surface protein apical membrane antigen 1 (AMA1) and its receptor, the rhoptry neck protein (RON) complex, which is targeted to the host cell membrane during invasion. In particular, RON2, a transmembrane component of the RON complex, interacts directly with AMA1. Here, we report the crystal structure of AMA1 from *Plasmodium falciparum* in complex with a

peptide derived from the extracellular region of *Pf*RON2, highlighting clear specificities of the *P. falciparum* RON2-AMA1 interaction. The receptor-binding site of *Pf*AMA1 comprises the hydrophobic groove and a region that becomes exposed by displacement of the flexible Domain II loop. Mutations of key contact residues of *Pf*RON2 and *Pf*AMA1 abrogate binding between the recombinant proteins. Although *Pf*RON2 contacts some polymorphic residues, binding studies with *Pf*AMA1 from different strains show that these have little effect on affinity. Moreover, we demonstrate that the *Pf*RON2 peptide inhibits erythrocyte invasion by *P. falciparum* merozoites and that this strong inhibitory potency is not affected by AMA1 polymorphisms. In parallel, we have determined the crystal structure of *Pf*AMA1 in complex with the invasion-inhibitory peptide R1 derived by phage display, revealing an unexpected structural mimicry of the *Pf*RON2 peptide. These results identify the key residues governing the interactions between AMA1 and RON2 in *P. falciparum* and suggest novel approaches to anti-malarial therapeutics.

3.2 Author summary

Malaria arises from infection of erythrocytes by single-cell parasites belonging to the genus *Plasmodium*, with the species *P. falciparum* causing the most severe forms of the disease. The formation of a moving junction between the membranes of the parasite and its host cell is essential for invasion. Two important components of the moving junction are apical membrane antigen 1 (AMA1) on the parasite surface and the *Plasmodium* rhoptry neck (RON) protein complex that is translocated to the erythrocyte membrane during invasion. The extracellular region of RON2, a component of this complex, interacts with AMA1, providing a bridge between the parasite and its host cell that is crucial for successful invasion. The parasite thus provides its own ligand for AMA1 and accordingly this critical interaction is not subject to evasive adaptations by the host. We present atomic details of the interaction of *Pf*AMA1 with the carboxy-terminal region of RON2 and shed light on structural adaptations by each apicomplexan parasite to maintain an interaction so crucial for invasion. The structure of the RON2 ligand bound to AMA1 thus provides an ideal basis for drug design as such molecules may be refractory to the development of drug resistance in *P. falciparum*.

3.3 Introduction

Plasmodium species, and *P. falciparum* in particular, are devastating global pathogens that place nearly half the human population at risk to malaria, leading to more than 250 million cases yearly and over one million deaths (World Health Organization. 2010). The success of the malaria parasite can be attributed to its intracellular lifestyle, invading host cells both in liver and blood stages. Invasion of red blood cells is an active process involving a moving junction (MJ), which is formed by intimate contact between erythrocyte and parasite membranes and is thought to be coupled to the parasite's actin-myosin motor (Aikawa, Miller et al. 1978, Baum, Papenfuss et al. 2006). A number of merozoite antigens, either exposed on the surface or stored in secretory organelles, play a role in the invasion process (Cowman and Crabb 2006). One of these is apical membrane antigen 1 (AMA1), a type I transmembrane protein secreted from the micronemes to the merozoite surface and present at the MJ (Waters, Thomas et al. 1990, Riglar, Richard et al. 2011). AMA1 is highly conserved in the *Plasmodium* genus (Waters, Thomas et al. 1990) and, moreover, in the Apicomplexa phylum to which *Plasmodium* belongs (Hehl, Lekutis et al. 2000, Gaffar, Yatsuda et al. 2004), suggesting a common functional role in diverse host cell invasion scenarios. In the apicomplexan organism *Toxoplasma gondii*, the ligand for AMA1 was shown to be rhoptry neck protein 2 (RON2), a component of the parasite-derived RON protein complex that is secreted into the host cell during invasion and integrated into the host cell membrane (Besteiro, Michelin et al. 2009, Tyler and Boothroyd 2011). This interaction was subsequently confirmed in *P. falciparum* as well (Lamarque, Besteiro et al. 2011, Srinivasan, Beatty et al. 2011). Apicomplexans thus provide both receptor and ligand to drive active invasion.

In many malaria-endemic regions, *P. falciparum* has become resistant to classic drugs, such as chloroquine, and has begun to show some resistance to recently introduced drugs. Since both AMA1 and RON2 are specific to Apicomplexa and essential for invasion, interruption of the AMA1-RON2 interaction presents an ideal new target for the design and development of inhibitors. This is supported by the recent observation that the invasion-inhibitory peptide R1 (Harris, Casey et al. 2005, Harris, Casey et al. 2009) blocks interaction between AMA1 and the RON complex in *P. falciparum* (Richard, MacRaild et al. 2010), but due to the polymorphism of AMA1, the effectiveness of this peptide inhibitor

is limited to a subset of parasite isolates. Interestingly, R1 does not prevent apical contact but no formation of a functional MJ ensues from this event (Richard, MacRaild et al. 2010).

Crystal structures of *Pf*AMA1 in complex with invasion-inhibitory antibodies (Coley, Gupta et al. 2007, Henderson, Streltsov et al. 2007) have implicated a hydrophobic groove on Domain I (DI) of *Pf*AMA1 as being critical for function. The topological nature of the *Pf*AMA1 groove (Bai, Becker et al. 2005) is conserved in *P. vivax* AMA1 (Pizarro, Vulliez-Le Normand et al. 2005) and *T. gondii* AMA1 (Crawford, Tonkin et al. 2010), and contains a number of residues that are conserved or semi-conserved across *Plasmodium* species, as well as other members of Apicomplexa (Chesne-Seck, Pizarro et al. 2005), suggesting that it contributes to the ligand-binding site of AMA1. This was recently confirmed by the crystal structure of *Tg*AMA1 in complex with a synthetic peptide, *Tg*RON2sp, which inserts in the groove of *Tg*AMA1 (Tonkin, Roques et al. 2011).

Here, we report the crystal structure of the complex formed between *Pf*AMA1 and peptide segments of *Pf*RON2, which, together with our previous structural results on the *Tg*AMA1-*Tg*RON2sp co-structure (Tonkin, Roques et al. 2011), highlights a conserved, crucial interaction in apicomplexan host cell invasion. Functional characterization of hot-spot residues driving *Pf*AMA1-*Pf*RON2 complex formation leads to a deeper understanding of key interactions occurring at the MJ of *P. falciparum* and reveals the molecular basis of cross-strain reactivity while preserving specificity for the species. We also describe the crystal structure of *Pf*AMA1 in complex with the invasion-inhibitory peptide R1 (Harris, Casey et al. 2005), and show that this peptide presents an intriguing structural mimicry of *Pf*RON2. Collectively, our results provide an important structural basis for designing cross-strain reactive molecules that inhibit invasion by *P. falciparum*.

3.4 Materials and methods

Materials

All materials were purchased as described in section 2.3, with the following exception and addition. An improved transfection reagent, Cellfectin II, was purchased from Invitrogen (Burlington, ON). Tunicamycin was purchased from Calbiochem (EMD Millipore; Billerica, MA).

Cloning of the Pfama1 gene

Due to the extensive AT codon bias of *Plasmodium* genes (Gardner, Hall et al. 2002), a clone encoding Domains I and II of the ectoplasmic region of *Pf*AMA1 strain 3D7 (Gupta, Bai et al. 2005) (Asn104 to Glu438, with numbering based on the initiation methionine in the signal sequence; UniProt ID Q7KQK5) was codon optimized and synthesized by GenScript (Piscataway, NJ), and subsequently sub-cloned into a pAcGP67B vector modified as described in section 2.3 for expression in insect cells.

Protein Production and purification

Baculovirus generation and amplification, protein production and purification

The *Pfama1*/pAcGP67Bmod primary virus (P1) was prepared as described in section 2.3, with the exception that Cellfectin II was used as the lipid transfection reagent reducing the required incubation of the transfection solution with the adherent *Spodoptera frugiperda 9* (*Sf*9) cells to five hours. The P2 was generated and stored as described in section 2.3 and shown to be of a high enough titre to not require amplification to a P3.

For large scale expression, 5.4 L of Hi5 cells at 2.0×10^6 cells/mL in Express Five serum-free media (SFM) containing 10 µg/mL gentamicin and 16.5 mM L-glutamine were infected with an optimized ratio of 5 mL of *Pfama1*/pAcGP67Bmod P2 virus per 1 L of cell culture and tunicamycin was added to a final concentration of 0.3 µg/mL. The baculovirus infected cells were incubated at 27 °C with shaking at 120 rpm for 65 hours. After the 65 hour incubation, the Hi5 cells were at a cell density of 1.6×10^6 cells/mL with 45% nonviable cells.

The cell culture was clarified and the *Pf*AMA1 protein purified from the supernatant as described for *Tg*AMA1 in section 2.3, with the following exceptions. A 2 mL bed volume of Ni-chelated sepharose beads (GE Healthcare Life Sciences; Baie d'Urfe, QC) pre-equilibrated in Binding Buffer was added to the concentrated, buffer exchanged sample and allowed to batch bind at 4 °C for 1 hour. The beads were separated from the solution using a 60 mL Chemglass Buchner funnel with a coarse frit (VWR; Radnor, PA) under vacuum. The sample was cleaved with thrombin as previously described, and further purified by gel filtration chromatography. The final sample was concentrated to 18 mg/mL in HEPES-buffered saline (HBS) for co-crystallization trials. The concentration of

*Pf*AMA1 throughout purification was analyzed by absorbance at 280 nm based on an extinction coefficient of 52425 M^{-1} cm^{-1}. The final yield of recombinant protein was approximately 3 mg per litre of insect cell culture

AMA1 from P. pastoris

Synthetic genes were optimized for *Pf*AMA1 coding residues 97-442, from strains 3D7 (Genbank accession number U33274), CAMP (accession number M34552) and HB3 (accession number U33277). Potential N-glycosylation sites were mutated and genes were cloned EcoRI-KpnI in the pPicZalpha A vector (Invitrogen), resulting in an 11-residue sequence extension followed by myc-epitope and hexa-His tags at the C-terminus, expressed in *P. pastoris*, and purified as described (Kocken, Withers-Martinez et al. 2002). Yield after purification was approximately 20 mg per litre of culture. *Pf*AMA1 FVO (residues 25-545, no tags, accession number AJ277646) was produced as described before (Faber, Remarque et al. 2008). The DiCo3 protein was modified compared to the published protein (Remarque, Faber et al. 2008); it includes the *Pf*AMA1 FVO prodomain (amino acids 25-96) and one additional mutation to minimize proteolytic cleavage Lys376Arg (B. Faber, unpublished results). The *Pf*AMA1 3D7mut (Ile225Asn, residues 25-545, no tags) mutant was generated by site-directed mutagenesis (GenScript) and produced in *P. pastoris* in a similar fashion to the native protein (Faber, Remarque et al. 2008).

Peptide synthesis

A 39 amino acid region of *Pf*RON2 strain 3D7 (PfRON2sp1; Asp2021 to Ser2059, numbering based on the initiation methionine in the signal sequence; UniProt ID Q8IKV6) was synthesized by Kinexus (Vancouver, BC) with a free amine N-terminus and amidated C-terminus, and was disulfide cyclised. The peptide was confirmed by mass spectrometry to be >95% pure. Lyophilized *Pf*RON2sp1 was solubilised in 100% dimethyl sulfoxide (DMSO) and subsequently diluted in filter sterilized HBS for use in co-crystallization studies.

Peptides *Pf*RON2sp2 (residues Lys2027 to Asn2054) and R1 were synthesized by PolyPeptide (Strasbourg, France) and solubilized in 3.5% DMSO for subsequent use.

Crystallization, data collection and processing

Initial crystals of *Pf*AMA1-*Pf*RON2sp were identified in Crystal Screen I/II (Hampton Research; Aliso Viejo, CA), and diffraction quality crystals were obtained after seven days in 30% polyethylene glycol (PEG) 400, 100 mM Tris-HCl pH 8.5, 200 mM tri-sodium citrate dihydrate. The final 2.0 µL drops consisted of equal volumes protein (*Pf*AMA1 3D7 (5 mg/mL final concentration) incubated with *Pf*RON2sp1 (1:2 molar excess)) and reservoir solution equilibrated against 100 µL of reservoir solution. Cryoprotection was carried out in reservoir solution supplemented with 6.25% glycerol for 20 s and the crystal was flash cooled at 100K directly in the cryostream. Diffraction data were collected on beam line 9-2 at SSRL (Menlo Park, CA).

Crystals of *Pf*AMA1 CAMP *Pf*RON2sp2 were obtained in 20% PEG 4000, 100 mM Tris-HCl pH 8.6, 100 mM sodium acetate and 20% isopropanol and the protein (6.4 mg/mL final concentration) incubated with *Pf*RON2sp2 (1:5 molar excess). Diffraction data were collected from a crystal in cryoprotectant buffer at 100K on beam line ID29 at European Synchrotron Radiation Facility (Grenoble, France). Crystals of *Pf*AMA1 3D7 R1 were obtained in 15% PEG 4000, 100 mM Tris-HCl pH 8.5, 100 mM sodium acetate and 10% isopropanol and the protein (5.4 mg/mL final concentration) incubated with R1 (1:6 molar excess). Diffraction data were collected at 100K on beam line PROXIMA 1 at SOLEIL (St. Aubin, France).

Structure solution and refinement

Diffraction data were processed using Imosflm (Battye, Kontogiannis et al. 2011) or XDS (Kabsch 2010), and Scala (Evans 2006) in the CCP4 suite of programs (Winn, Ballard et al. 2011). Initial phases were obtained by MR using PHASER (McCoy, Grosse-Kunstleve et al. 2007) or AMoRe (Navaza 2001) with the unliganded *Pf*AMA1 structure (Protein Data Bank (PDB) ID 1Z40). Tracing of the *Pf*RON2 and R1 peptides, and addition of solvent molecules, was performed manually in COOT (Emsley and Cowtan 2004) and refinement in Refmac5 (Murshudov, Vagin et al. 1997) or autoBUSTER (Global Phasing Ltd, Cambridge, UK). Stereochemical analysis was performed with PROCHECK and SFCHECK in CCP4 (Winn, Ballard et al. 2011). Data collection and refinement statistics are presented in Tables 6 and 7.

Protein Data Bank accession codes

The coordinate and structure factor files have been deposited to the PDB with the following accession codes: *Pf*AMA1-*Pf*RON2sp1, 3ZWZ; *Pf*AMA1-*Pf*RON2sp2, 3SRI; *Pf*AMA1-R1, 3SRJ.

Binding studies by SPR

Surface plasmon resonance (SPR) measurements were made with a Biacore 2000 instrument (Biacore AB). AMA1 proteins diluted in 10 mM sodium acetate pH 4.5 for 3D7, CAMP, HB3 and FVO strains, or pH 4.0 for 3D7mut and Dico3, were covalently immobilized by an amine-coupling procedure on CM5 sensor chips (GE Healthcare). The reference flow cell was prepared by the same procedure in absence of protein. Binding assays were performed at 25°C in phosphate-buffered saline (PBS) and 0.005% Tween 20 by injecting a series of peptide (*Pf*RON2sp1 and *Pf*RON2sp2 on 3D7, CAMP, HB3 and FVO, and R1 on 3D7mut and Dico3) concentrations at a constant flow rate of 5 µL/min. A heterologous peptide was used to verify the absence of non-specific binding. Peptide dissociation was realized by injecting the running buffer, and the surface was regenerated by injecting glycine/HCl pH 1.5 followed by 0.05% SDS. Control flow cell sensorgrams were subtracted from the ligand flow cell sensorgrams and averaged buffer injections were subtracted from analyte sensorgrams. For peptide R1, steady-state signals (Req) were obtained directly from the plateau region of the sensorgrams, while for *Pf*RON2 peptides, estimated values of Req were obtained by extrapolation from the experimental curves since the association phase did not reach a final equilibrium state. All calculations were made using the BIAevaluation 4.2 software (BIAcore AB). The saturation curves obtained by plotting Req versus the peptide concentration were fitted with a steady-state model to obtain the Rmax and the apparent equilibrium dissociation constants, K_D. To normalize the response for the different ligands, these curves were reported as the percentage of bound sites (ratio Req/Rmax) versus the analyte concentration.

Isothermal titration calorimetry

Isothermal titration calorimetry (ITC) measurements were made using an ITC200 calorimeter (MicroCal). *Pf*AMA1 3D7 (*P. pastoris*) and peptide R1 were diluted in PBS to

final concentrations of 0.6 µM and 55 µM, respectively. *Pf*AMA1 3D7 (initial volume 200 µL) was titrated at 25°C by consecutive injections of the peptide R1 (2 µL aliquots at 3 min intervals). Raw data were normalized and corrected for the heat of dilution of R1 in PBS. Binding stoichiometry was determined by fitting the final data to a 1:1 interaction model using the Origin7 software (OriginLab).

P. falciparum cultures and invasion assays

The *P. falciparum* cell cultures and the invasion assays were performed as described previously (Lamarque, Besteiro et al. 2011). Briefly, highly synchronized *P. falciparum* 3D7 and HB3 schizonts (1.5% hematocrit, 1.5% parasitemia) were incubated with R1 or *Pf*RON2sp1 peptides. Blood smears were collected 16 hours post-invasion and used for counting ring-stage parasites. The results presented are representative of three independent experiments, each performed in triplicate.

Transient transfection experiments and cell binding assays

Cell binding assays using *Pf*AMA1-expressing BHK-21 cells and recombinant glutathione-S-transferase (GST)-*Pf*RON2-5 fusion proteins were performed as previously described (Lamarque, Besteiro et al. 2011). Although not quantitative, this cell-binding assay truly reflects the interaction between AMA1 and RON2 as we carefully checked all the experimental steps as well as the image recording as described below. Transfections were carried out using Lipofectamine Reagent (Invitrogen) as instructed by the manufacturer with 3 x 10^5 BHK-21 cells grown on coverslips for 24 hours in 6 well plates. Cells were grown for an additional 24 hours post-transfection before subsequent analysis. Expression and correct folding of *Pf*AMA1 (and the mutants) at the host cell surface was verified by immunofluorescence assay (IFA) performed with or without permeabilization, using antibodies either specific to the cytoplasmic tail (anti-myc tag) or specific to the extracellular ectodomain of *Pf*AMA1 (mouse monoclonal antibody (mAb) F8.12.19 (Igonet, Vulliez-Le Normand et al. 2007)). For binding assays, coverslips from a same transfection experiment were washed in HBSS (Invitrogen) before addition of recombinant *Pf*RON2-5 wild type or mutants diluted in HBSS at 10, 1 or 0.1 µg/ml. Coverslips incubated with GST were systematically used as a control. After five washes in PBS to

remove unbound protein, cells were fixed in 4% paraformaldehyde picric acid (PAF) and further processed for IFA as described above (Lamarque, Besteiro et al. 2011). The binding characteristics of RON2 (anti-GST labelling) on the *Pf*AMA1 mutant were only considered valid when its signal was identical to that of wild type *Pf*AMA1. All other micrographs were obtained with a Zeiss Axiophot microscope equipped for epifluorescence. Adobe photoshop (Adobe Systems, Mountain View, CA) was used for image processing. Matching pairs of images were recorded with the same exposure time and processed identically. The *Pf*AMA1 and GST-*Pf*RON2 mutated constructs were generated by site-directed mutagenesis using QuikChange II XL protocol (Stratagene).

3.5 Results

PfRON2sp specifically binds to PfAMA1

From the 67-residue construct, *Pf*RON2-5, that we previously showed to have affinity for *Pf*AMA1 (Lamarque, Besteiro et al. 2011), and guided by the *Tg*AMA1-*Tg*RON2sp structure (Tonkin, Roques et al. 2011), we synthesized two analogous *Pf*RON2 peptides: *Pf*RON2sp1 (residues 2021-2059; numbering from the initiation methionine in PF14_0495), and *Pf*RON2sp2 (residues 2027-2055). Significantly, there is no polymorphism in this sequence among *P. falciparum* isolates. Both constructs incorporate a disulfide-bound β-hairpin loop proposed to be critical in complex formation (Tonkin, Roques et al. 2011) while *Pf*RON2sp2 is truncated at both the N- and C-termini (Figure 21A). Since the extracellular region of *Pf*RON2 is non-polymorphic, we determined the affinity of both peptides for *Pf*AMA1 by SPR measurements using the 3D7, CAMP, FVO and HB3 proteins to explore the possible effects of AMA1 polymorphisms. The affinity of *Pf*AMA1 from 3D7 for *Pf*RON2sp1 is 25-fold higher than for *Pf*RON2sp2 (Figure 21B to E, Table 4), highlighting a moderate, yet influential, role for the N- and C-terminal tails. Interestingly, K_D values for *Pf*RON2sp1 showed no significant variation in binding to *Pf*AMA1 from the four strains.
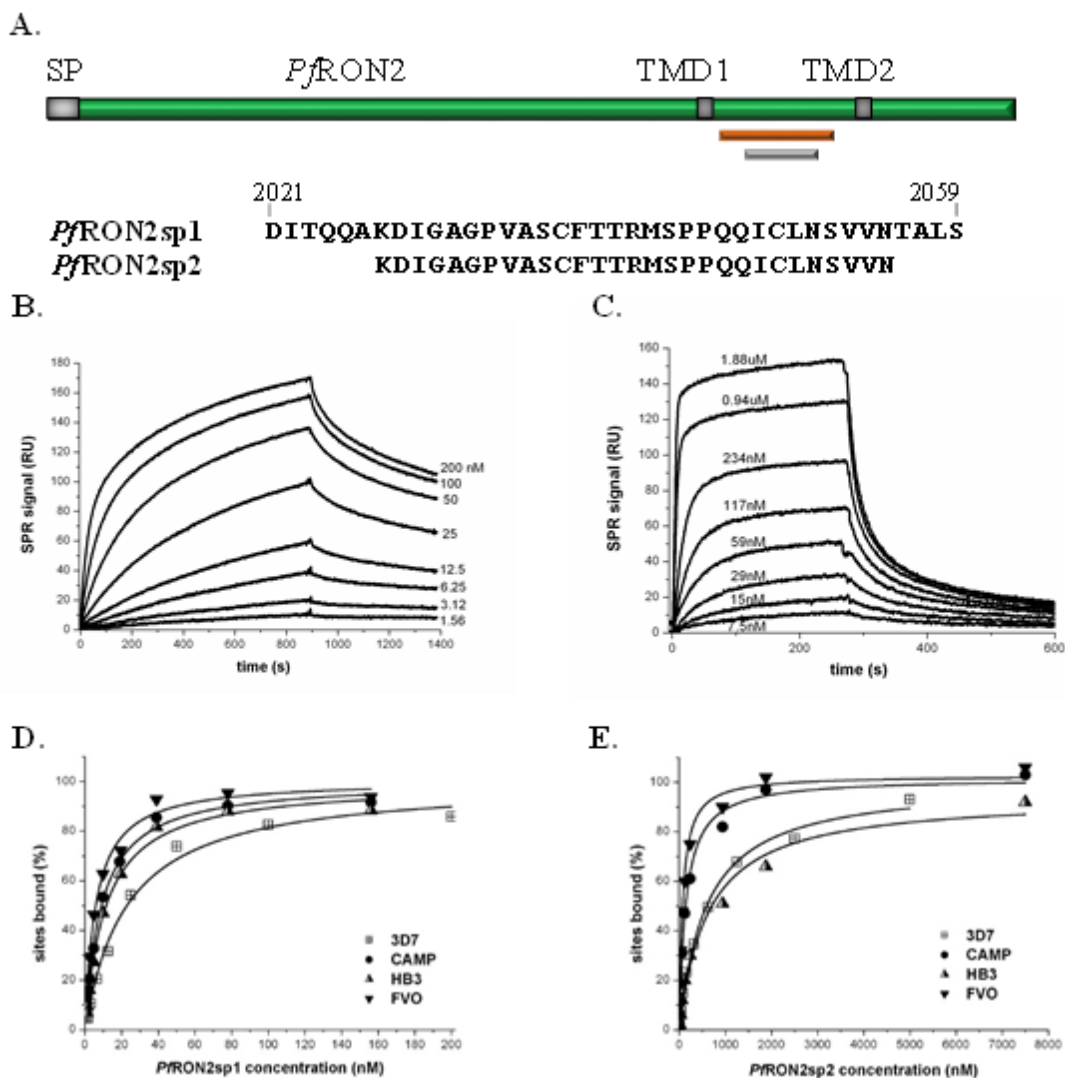
Figure 21: SPR studies of peptides *Pf*RON2sp1 and *Pf*RON2sp2 binding to recombinant *Pf*AMA1 from multiple strains reveal that *Pf*RON2sp1 has a consistently higher affinity.

A. *Pf*RON2sp1 (orange) and *Pf*RON2sp2 (grey) represent peptides of *Pf*RON2 (green). SP, signal peptide. TMD, putative transmembrane domain. **B.** Sensorgrams showing *Pf*RON2sp1 (analyte) binding to *Pf*AMA1 3D7 (immobilized). The *Pf*RON2sp1 concentrations are indicated for each curve (nM). **C.** Sensorgrams showing *Pf*RON2sp2 (analyte) binding to *Pf*AMA1 CAMP (immobilized), with *Pf*RON2sp2 concentrations indicated. **D, E.** Variation percentage of bound sites (deduced from the steady-state response) with respect to analyte concentration (**D**, *Pf*RON2sp1; **E**, *Pf*RON2sp2) obtained from binding to immobilized recombinant *Pf*AMA1 from strains 3D7 (shown in **B**), CAMP

(shown in **C**), FVO and HB3. The derived apparent equilibrium dissociation constants KD are given in Table 4.

Table 4: Apparent equilibrium dissociation constants K_D (nM) for the binding of peptides _Pf_RON2sp1 and _Pf_RON2sp2 to AMA1 from different strains of _P. falciparum_.

Strain	_Pf_RON2sp1	_Pf_RON2sp2
3D7	20.3 ± 6.3	520 ± 74
CAMP	14.6 ± 3.8	165 ± 42
FVO	9.2 ± 3.0	80 ± 15
HB3	18.3 ± 4.6	680 ± 180

Independent experiments were performed at least three times and the values represent the mean $\pm$ SD.

_Pf_RON2sp1 and _Pf_RON2sp2 were co-crystallized with the first two ectoplasmic domains (DI, DII) of recombinant _Pf_AMA1 3D7 or CAMP strains, respectively. The co-structure of _Pf_AMA1 3D7 with _Pf_RON2sp1 (PDB ID 3ZWZ) was refined to 2.2 Å resolution, while _Pf_AMA1 CAMP with _Pf_RON2sp2 (PDB ID 3SRI) was refined to 1.6 Å resolution (Table 5). The two co-structures overlay with a root mean square deviation (rmsd) of 0.81 Å over 304 Cα positions, and the two peptides alone overlay with a rmsd of 0.34 Å over the complete length of the modeled _Pf_RON2sp2 (25 Cα) (Figure 22A). These data confirm that the reduced affinity of _Pf_RON2sp2 is due to the truncated N- and C-termini. Since _Pf_RON2sp1 is more biologically relevant than its truncated counterpart, it was used for the following analyses unless otherwise noted.

Table 5: Data collection and refinement statistics for *Pf*AMA1 3D7-*Pf*RON2sp1 and *Pf*AMA1 CAMP-*Pf*RON2sp2.

	PfAMA1 3D7-PfRON2sp1	*PfAMA1 CAMP-PfRON2sp2*
Data collection statistics		
Spacegroup	P2$_1$	P2$_1$
a, b, c (Å)	70.15, 38.26, 70.75	70.72, 38.14, 72.08
α, β, γ (deg.)	90, 99.73, 90	90, 97.72, 90
Wavelength (Å)	0.9795	0.9537
Resolution range (Å)	45.41 – 2.10 (2.21 – 2.10)	46.97 – 1.60 (1.69 – 1.60)
Measured reflections	109520	153050
Unique reflections	22041	48207
Redundancy	5.0 (5.0)	3.2 (3.2)
Completeness (%)	100.0 (100.0)	94.9 (92.8)
I/σ(I)	8.7 (3.2)	12.7 (1.7)
R$_{merge}$[a]	0.148 (0.470)	0.056 (0.618)
Refinement statistics		
Resolution (Å)	34.87 – 2.10	35.04 – 1.60
R$_{cryst}$[b]	0.164	0.176
R$_{free}$[c]	0.201	0.195
No. of atoms		
Protein (A/B)	2377/259	2309/190
Solvent	226	265
Glycerol	30	N/A
B-values (Å^2)		
Protein (A/B)	17.3/29.9	27.5/48.3
Solvent	28.7	37.6
Glycerol	39.1	N/A
r.m.s. deviation from ideality		
Bond lengths (Å)	0.02	0.01
Bond angles (deg.)	1.52	1.05
Ramachandran statistics (%)		
Most favoured	97.6	96.7
Allowed	2.4	3.3
Disallowed	0.0	0.0

Values in parentheses are for the highest resolution shell

[a] R$_{merge}$= $\sum_{hkl} \sum_i |I_{hkl,i}$ - $[I_{hkl}]| / \sum_{hkl} \sum_i I_{hkl,i}$, where $[I_{hkl}]$ is the average of symmetry related observations of a unique reflection

[b] R$_{cryst}$=$\sum|F_{obs}$-$F_{calc}|/\sum F_{obs}$, where F$_{obs}$ and F$_{calc}$ are the observed and the calculated structure factors, respectively

[c] R$_{free}$ is R using 5% of reflections randomly chosen and omitted from refinement

Figure 22: Structure of *Pf*AMA1 complexed with *Pf*RON2-derived peptides.

A. Top - Co-crystal structures of *Pf*AMA1 (blue surface) with *Pf*RON2sp1 (orange) and *Pf*RON2sp2 (grey), show a disulfide-anchored U-shaped conformation in the apical groove of *Pf*AMA1. Bottom – 2Fo-Fc electron density map (orange) for *Pf*RON2sp1 contoured at 1.0 σ, highlighting well-ordered density from the N-terminal helix, through the cystine loop, to the C-terminal coil. **B.** Notable changes in the structure of *Pf*AMA1 between the apo structure (green; PDB ID 1Z40) and the *Pf*AMA1-*Pf*RON2sp1 co-structure (blue-orange) as observed from a side view. Box 1 - The DII loop of apo *Pf*AMA1 is ejected

from the apical groove during binding to *Pf*RON2sp1, leaving room for the *Pf*RON2sp1 N-terminal helix to occupy the space vacated by the DII loop helix. Box 2 - The β-strands of the *Pf*RON2sp1 cystine loop order a *Pf*AMA1 surface loop, generating a contiguous three-stranded β-sheet. **C.** In the region of the *Pf*RON2sp1 N-terminal helix, there is notable structural mimicry to the *Pf*AMA1 apo DII loop, including several conserved residues, and a conserved hydrogen bonding network incorporating three buried water molecules. **D.** Arg2041, specific to *P. falciparum*, fits snugly into a deep pocket in the surface of *Pf*AMA1 and is stabilized through a complex network of seven hydrogen bonds.

*Pf*RON2sp1, traced from Thr2023 to Leu2058, includes a disulfide bridge between Cys2037 and Cys2049 and makes several direct contacts with *Pf*AMA1 (Figure 23), resulting in a total buried surface area of 3154 Å^2 (1441 Å^2 for *Pf*AMA1 and 1713 Å^2 for *Pf*RON2sp1). Overall, the binding paradigm established by *Tg*AMA1-*Tg*RON2sp (Tonkin, Roques et al. 2011) is maintained, with an N-terminal helix seated at one end of the AMA1 ligand-binding groove and extended through an ordered coil to a disulfide-closed β-hairpin loop, generating a U-shaped conformation (Figure 22A). Similarly, exposing a functional ligand-binding groove on AMA1 requires displacement of the extended non-polymorphic DII loop, which adopts a disordered state (not modeled between Lys351 to Ala387); this region is stabilized by DI in apo *Pf*AMA1 (Figure 22B). Intriguingly, the backbone of the N-terminal helix and additional coil of *Pf*RON2sp1 (2024-QQAKDIGAG-2032) overlays remarkably well with a section of the apo *Pf*AMA1 DII loop (360-YEKIKEGFK-368) (rmsd < 0.4 Å), which also includes a helical region (Figure 22B - box 1). Three water molecules buried by the DII loop in the apo form are retained in the ligand-bound state and facilitate a network of hydrogen bonds that bridge *Pf*AMA1 DI to either the DII loop or *Pf*RON2sp in apo *Pf*AMA1 or the complex, respectively (Fig. 22C). The majority of intermolecular contacts are formed by the segment Lys2027-Met2042 of *Pf*RON2sp1. An influential residue on *Pf*RON2 appears to be Arg2041, a residue specific to the *P. falciparum* species, located at the tip of the β-hairpin with its guanidyl group fitting snugly into a preformed pocket of *Pf*AMA1 (Figure 22D).

A.

| | Polar interactions | | Hydrophobic interactions |
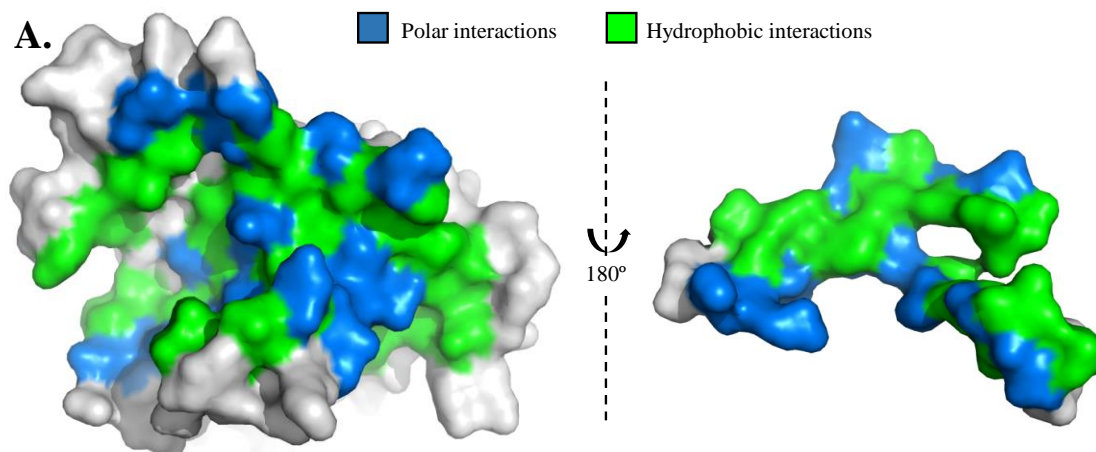
B.

Hydrogen bonds			Buried Surface Area (BSA) of interfacing residues		
*Pf*AMA1	*Pf*RON2sp	Distance (Å)	*Pf*AMA1	BSA (Å²) >20	*Pf*RON2sp	BSA (Å²) >5
Arg143 [NH1]	Asp2028 [O]	2.83	Tyr142	67.63	Gln2025	67.72
Tyr390 [OH]	Asp2028 [Oδ2]	3.53	Arg143	47.53	Lys2027	66.93
Ala254 [N]	Gly2030 [O]	3.49	Ala170	26.16	Asp2028	70.00
Tyr251 [OH]	Pro2033 [O]	3.81	Thr171	45.92	Ile2029	122.00
Tyr251 [OH]	Ala2035 [N]	3.42	Gly172	35.01	Gly2030	61.36
Asp227 [Oδ1]	Ser2036 [N]	3.63	Leu176	43.26	Ala2031	16.28
Asp227 [Oδ2]	Ser2036 [N]	3.46	Pro185	38.95	Gly2032	21.16
Asp227 [N]	Cys2037 [O]	3.06	Thr186	42.23	Pro2033	117.32
Asp227 [Oδ1]	Cys2037 [N]	2.69	Glu187	38.90	Val2034	22.08
Ile225 [N]	Thr2039 [O]	2.78	Pro188	43.58	Ala2035	31.92
Ile225 [O]	Thr2039 [N]	3.10	Phe201	26.87	Ser2036	10.10
Asn223 [O]	Arg2041 [N]	2.81	Asn205	34.39	Cys2037	29.71
Asn223 [Oδ1]	Arg2041 [N]	2.96	Asn223	75.29	Phe2038	96.80
Gly222 [O]	Arg2041 [NH1]	2.87	Met224	44.27	Thr2039	40.02
Met224 [O]	Arg2041 [NH1]	3.34	Ile225	82.05	Arg2041	179.42
Tyr234 [O]	Arg2041 [NH1]	3.33	Pro226	33.78	Met2042	116.83
Tyr234 [O]	Arg2041 [NH2]	3.08	Asp227	71.82	Gln2046	42.04
Ser232 [O]	Arg2041 [NH2]	2.70	Tyr234	29.46	Gln2047	6.14
Asn223 [Oδ1]	Met2042 [N]	2.84	Lys235	17.49	Ile2048	81.38
Glu187 [N]	Cys2049 [O]	2.97	Tyr251	88.03	Cys2049	24.15
Glu187 [O]	Cys2049 [N]	3.25	Ala254	38.63	Leu2050	41.23
Glu187 [N]	Asn2051 [Oδ1]	3.53	Tyr262	46.63	Ser2052	39.62
Pro185 [O]	Asn2051 [N]	3.19	Asn271	24.39	Val2053	17.59
Gly172 [N]	Val2053 [O]	2.83	Met273	35.67	Val2054	57.91
Ala170 [O]	Asn2055 [N]	2.83	Pro350	46.09	Asn2055	85.00
Gln174 [O]	Asn2055 [Nδ2]	3.27	Tyr390	33.58	Thr2056	0.00
Asn271 [Oδ1]	Asn2055 [Nδ2]	2.94			Ala2057	47.96
					Leu2058	94.78

Figure 23: Detailed analysis of interactions at the *Pf*AMA1-*Pf*RON2sp1 interface.

A. Open-book surface representation of *Pf*AMA1 (left) and *Pf*RON2sp1 (right) showing the extensive involvement of residues from both molecules in forming a complex interface. Residues involved in hydrogen bonding are colored blue, while residues contributing significant buried surface area (BSA > 20 Å² for *Pf*AMA1, > 5 Å² for *Pf*RON2sp1) are

colored green. **B.** Table of residues involved in hydrogen bonding at the *Pf*AMA1-*Pf*RON2sp1 interface (left) and residues contributing significant BSA (right), as calculated by PISA (Krissinel and Henrick 2007).

R1 occupies the PfRON2sp-binding site on PfAMA1

The invasion-inhibitory peptide R1, comprising twenty residues (VFAEFLPLFSKFGSRMHILK) (Harris, Casey et al. 2005), has been shown by nuclear magnetic resonance (NMR) to bind to the *Pf*AMA1 hydrophobic groove, but this study gave little structural detail of the interaction (Richard, MacRaild et al. 2010). We therefore crystallized *Pf*AMA1 3D7 (DI and II) with R1 to compare with the *Pf*RON2 complex (Table 6). Surprisingly, two molecules of R1 are bound to *Pf*AMA1, which we denote respectively as the major peptide (R1-major), lying deeply in the binding groove, and the minor peptide (R1-minor), lying above R1-major and making fewer contacts with *Pf*AMA1 (Figure 24 and Table 7). Several solvent molecules bridge directly between *Pf*AMA1 and R1-major. As in the *Pf*AMA1-*Pf*RON2sp complexes, the N-terminus of R1-major binds to a region of *Pf*AMA1 that becomes exposed after displacement of the DII loop.

Table 6: Data collection and refinement statistics for *Pf*AMA1 3D7-R1.

Data collection statistics	
Spacegroup	$P2_12_12_1$
a, b, c (Å)	38.32, 144.32, 145.64
α, β, γ (deg.)	90, 90, 90
Wavelength (Å)	0.9791
Resolution range (Å)	40.28 – 2.15 (2.25 – 2.15)
Measured reflections	156625
Unique reflections	42798
Redundancy	3.7 (2.5)
Completeness (%)	95.3 (75.7)
I/σ(I)	13.3 (2.2)
R_{merge}[a]	0.075 (0.485)
Refinement statistics	
Resolution (Å)	37.06 – 2.15
R_{cryst}[b]	0.171
R_{free}[c]	0.214
No. of atoms	
Protein (A/B/C/D/E/F)	2375/2385/157/60/135/77
Solvent	450
B-values (Å^2)	
Protein (A/B/C/D/E/F)	36.4/40.5/50.6/77.8/61.5/92.6
Solvent	46.1
r.m.s. deviation from ideality	
Bond lengths (Å)	0.01
Bond angles (deg.)	1.10
Ramachandran statistics (%)	
Most favoured	96.3
Allowed	3.7
Disallowed	0.0

Values in parentheses are for the highest resolution shell

[a] $R_{merge} = \sum_{hkl} \sum_i |I_{hkl,i} - [I_{hkl}]| / \sum_{hkl} \sum_i I_{hkl,i}$, where $[I_{hkl}]$ is the average of symmetry related observations of a unique reflection

[b] $R_{cryst} = \sum |F_{obs} - F_{calc}| / \sum F_{obs}$, where F_{obs} and F_{calc} are the observed and the calculated structure factors, respectively

[c] R_{free} is R using 5% of reflections randomly chosen and omitted from refinement

Figure 24: Structure of *Pf*AMA1 complexed with R1 peptide.

A. The co-crystal structure of *Pf*AMA1 (blue surface) with R1 reveals two bound peptides, R1-major (yellow) and R1-minor (purple). **B.** Detailed analysis of interactions at the *Pf*AMA1–R1-major, *Pf*AMA1–R1-minor, and R1-major–R1-minor interfaces. Surface representation of *Pf*AMA1 (blue), with R1-major (yellow) and R1-minor (purple) shown as cartoons. Box 1 – R1-major anchors its N-terminus to *Pf*AMA1 through three backbone hydrogen bonds. Box 2 – the central region of the *Pf*AMA1 apical groove is occupied by R1-major through both hydrophobic and polar interactions. Box 3 – R1-minor forms most of its anchor points to *Pf*AMA1 through the apical loops and does not contact the base of the groove, which is occupied by R1-major. Panel 4 – Backbone hydrogen bonds between R1-minor and R1-major generate a β-sheet, while R1-major is further pinned to the *Pf*AMA1 groove through three hydrogen bonds. Panel 5 – R1-major integrates into

*Pf*AMA1 with the use of an arginine knob-in-hole interaction stabilized by six hydrogen bonds, which is also exploited by *Pf*RON2sp.

Table 7: Polar interactions and buried surface areas in the *Pf*AMA1-R1 crystal structure.

Top – polar contacts between *Pf*AMA1 3D7 and R1-major (column 1), and buried surface areas of individual residues of *Pf*AMA1 3D7 (column 2) and R1-major (column 3). Salt bridges are indicated in bold. Middle – polar contacts between *Pf*AMA1 3D7 and R1-minor (column 1), and buried surface areas of individual residues of *Pf*AMA1 3D7 (column 2) and R1-minor (column 3). Bottom – polar contacts between R1-major and R1-minor (column 1), and buried surface areas of individual residues of R1-major (column 2) and R1-minor (column 3).

Hydrogen bonds			Buried Surface Area (BSA) of interfacing residues			
*Pf*AMA1	R1-major	Distance (Å)	*Pf*AMA1	BSA (Å^2) >10	R1-major	BSA (Å^2)
Gln141 [Oε1]	Val-P1 [N]	3.63	Asp134	13.78	Val-P1	54.34
Gln141 [N]	Val-P1 [O]	2.90	Val137	20.45	Phe-P2	110.47
Gln141 [O]	Ala-P3 [N]	3.01	Thr140	33.80	Ala-P3	50.43
Tyr175 [OH]	Leu-P6 [O]	2.64	Gln141	73.43	Glu-P4	16.60
Tyr251 [OH]	Pro-P7 [O]	2.61	Tyr142	87.11	Phe-P5	125.26
Tyr251 [OH]	Phe-P9 [N]	2.83	Arg143	18.79	Leu-P6	138.29
Asn228 [N]	Lys-P11 [O]	3.65	Thr171	11.91	Pro-P7	22.10
Asp227 [Oδ2]	**Lys-P11 [Nζ]**	**2.94**	Tyr175	53.31	Leu-P8	122.82
Ile225 [O]	Gly-P13 [N]	2.98	Leu176	29.27	Phe-P9	70.36
Ile225 [N]	Gly-P13 [O]	2.92	Phe183	19.69	Ser-P10	1.60
Asn223 [O]	Arg-P15 [N]	2.82	Pro184	15.55	Lys-P11	56.81
Met224 [O]	Arg-P15 [NH1]	3.04	Thr186	14.81	Phe-P12	96.18
Gly222 [O]	Arg-P15 [NH1]	2.86	Phe201	34.84	Gly-P13	60.80
Ser232 [O]	Arg-P15 [NH2]	2.81	Tyr202	17.00	Ser-P14	25.79
Tyr234 [O]	Arg-P15 [NH2]	2.87	Asn205	15.30	Arg-P15	176.50
Asn223 [Oδ1]	Met-P16 [N]	3.13	Asn223	61.30	Met-P16	111.19
			Met224	16.97	His-P17	5.16
			Ile225	69.29	Ile-P18	0.00
			Pro226	14.06		
			Asp227	91.34		
			Asn228	46.98		
			Tyr234	20.22		
			Lys235	18.99		
			Tyr251	73.04		
			Ala254	24.11		
			Asn257	13.16		
			Met273	30.45		

PfAMA1	R1-minor	Distance (Å)	PfAMA1	BSA (Å^2) >10	R1-minor	BSA (Å^2)
Thr171 [Oγ1]	Glu-P4 [Oε1]	2.94	Thr171	45.67	Glu-P4	53.48
Pro185 [O]	Leu-P6 [N]	3.72	Pro184	22.02	Phe-P5	71.78
Glu187 [N]	Leu-P6 [O]	2.91	Pro185	55.63	Leu-P6	75.53
Glu187 [O]	Leu-P8 [N]	3.55	Thr186	40.35	Pro-P7	27.11
Lys230 [Nζ]	**Ser-P10 [O]**	**3.10**	Glu187	74.90	Leu-P8	103.31
			Pro188	13.18	Phe-P9	39.80
			Met190	16.34	Ser-P10	34.03
			Met224	13.75		
			Asn228	52.41		
			Lys230	25.76		
R1-major	R1-minor	Distance (Å)	R1-major	BSA (Å^2) >0	R1-minor	BSA (Å^2)
Phe-P12 [O]	Phe-P9 [N]	2.84	Leu-P8	25.72	Glu-P4	39.27
Phe-P12 [N]	Pro-P7 [O]	3.31	Phe-P9	7.71	Phe-P5	1.11
Ser-P14 [N]	Phe-P9 [O]	3.41	Ser-P10	43.07	Leu-P6	57.16
			Lys-P11	44.63	Pro-P7	31.17
			Phe-P12	69.58	Leu-P8	54.38
			Gly-P13	18.84	Phe-P9	71.79
			Ser-P14	24.70	Ser-P10	38.45
			His-P17	35.68		

R1-major makes several direct contacts with PfAMA1 (113 interatomic distances < 3.8 Å), including nineteen hydrogen bonds and a salt bridge between the amino group of Lys-P11 (R1 peptide residue numbers are prefixed by P) and the Asp227 carboxylate group of PfAMA1 (Table 7A). Contacts made by R1-minor to PfAMA1 are fewer (26 contacts < 3.8 Å) and include only five hydrogen bonds (Table 7B). Interactions between R1-major and R1-minor are maintained by a total of twenty-four interatomic contacts, including three hydrogen bonds (Table 7C). In total, 3025 Å^2 of molecular surface is buried between PfAMA1 and the two peptides, with R1-major contributing about 75% to this area. The buried surface between R1-major and R1-minor is 563 Å^2, reflecting the smaller number of close interatomic contacts between these two components.

Since the structure of the PfAMA1 3D7-R1 complex revealed two bound peptide molecules, binding measurements of R1 to PfAMA1 3D7 were made by ITC to examine the stoichiometry (Figure 25). The measured K_D of 145 nM is comparable with previous measurements by SPR (Harris, Casey et al. 2009) and the deduced stoichiometry was 1:1 over the peptide concentrations used. This implies that the second binding site in the crystal structure (R1-minor) has an affinity that could not be determined under the experimental

conditions used for ITC but can be estimated to be at least ten-fold weaker than the major site.
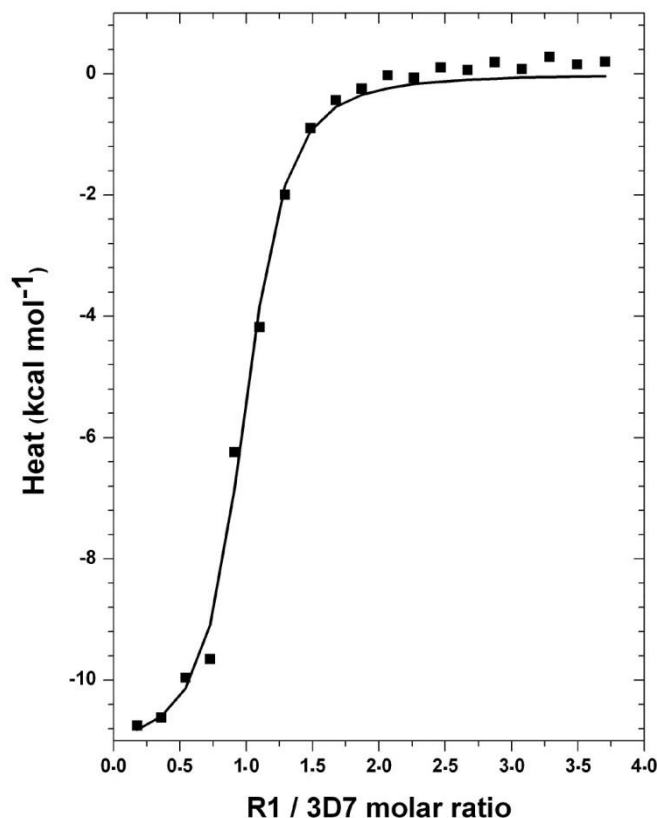
Figure 25: ITC of peptide R1 binding to *Pf*AMA1 3D7.

R1 mimicry of PfRON2

While R1-major follows the general contour of the receptor-binding groove, it does so in a linear fashion rather than the U-shaped conformation adopted by *Pf*RON2sp1 (Figure 26A). R1-minor occupies a similar region in space as the second strand of the *Pf*RON2sp β-hairpin, contacting the same DI loop of *Pf*AMA1 but running in the opposite direction to form a parallel two-stranded β-sheet with the major peptide (Figure 26A). Portions of R1-major exhibit structural similarity to *Pf*RON2, displaying a 1.2 Å rmsd over twelve Cα positions (*Pf*RON2sp1, Ala2031 to Met2042; R1-major, Phe-P5 to Met-P16) (Figure 26A). Moreover, sequence alignment based on the structural superposition reveals a remarkable similarity between the central regions of the two ligands; the segments Ala2031-Met2042

of *Pf*RON2 and Phe-P5–Met-P16 of R1 have five identical amino acids and two conservative differences (Figure 26B). R1-major residue Arg-P15 contributes the most contacts to *Pf*AMA1 and is positioned within the same pocket of *Pf*AMA1 as *Pf*RON2 Arg2041 (Figure 26A - box 3) where it maintains six of the seven hydrogen bonds observed for *Pf*AMA1-*Pf*RON2sp. Interestingly, while *Pf*RON2 mimicry is observed in the cystine loop-binding region (Phe2038/Phe-P12 to Arg2041/Arg-P15) (Figure 26A - box 2), R1-major establishes clear anchor points in the hydrophobic groove different from *Pf*RON2; Phe-P2 and Phe-P5 brace the peptide N-terminus in the region exposed by displacement of the DII loop, with Phe-P5 occupying the pocket left vacant by Phe367 of *Pf*AMA1 (Figure 26A - box 1).
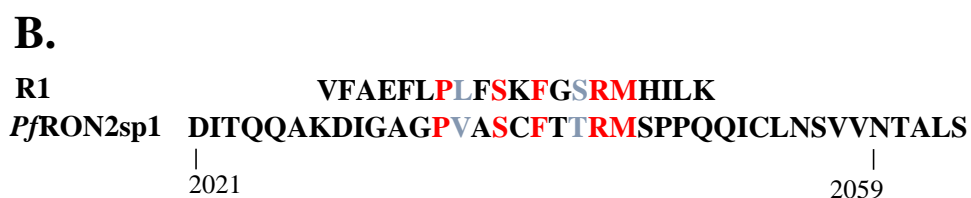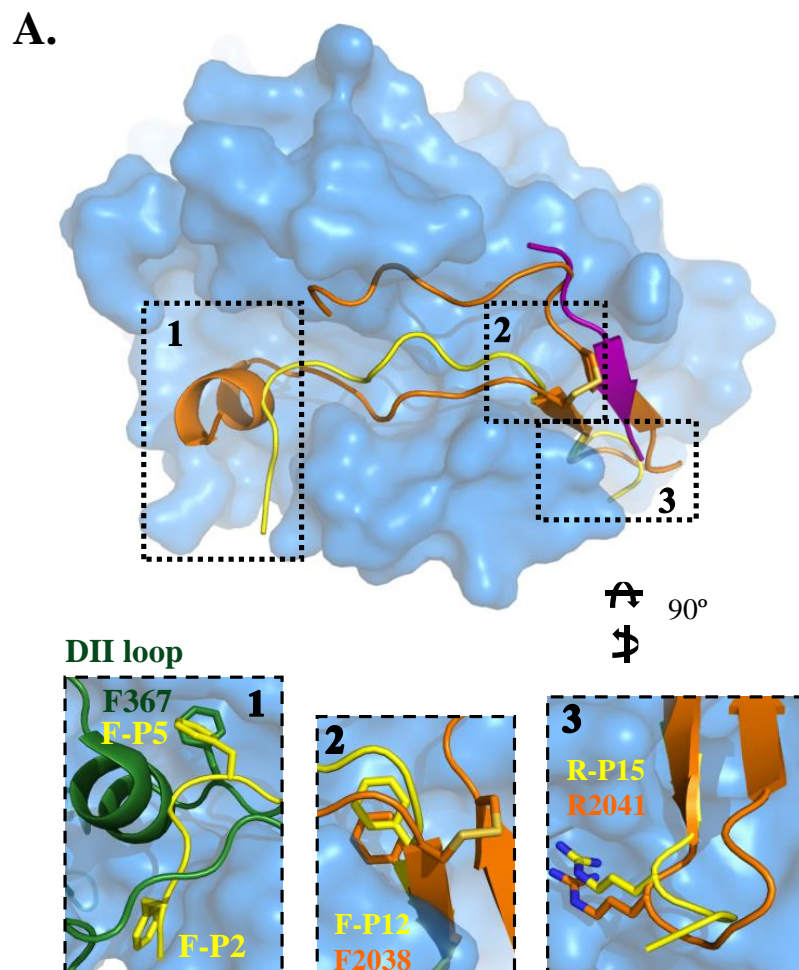
Figure 26: Structural mimicry of *Pf*RON2 by peptide R1 in binding to *Pf*AMA1.
A. Top (left) and end-on (right) views of *Pf*AMA1-*Pf*RON2sp1 (orange cartoon) overlayed on *Pf*AMA1-R1-major (yellow)/R1-minor (purple), show that the *Pf*AMA1 groove is capable of accepting only *Pf*RON2sp1 or the two R1 peptides at one time. Box 1 shows that Phe-P5 of R1 mimics Phe367 of the DII loop, while boxes 2 and 3 highlight spatial conservation of a phenylalanine anchor at the center of the groove, and a knob-in-hole interaction incorporating the peptide Arg-P15. R1-major is shown in yellow, *Pf*RON2sp1 in orange and apo *Pf*AMA1 in green. **B.** Comparison of the R1 and *Pf*RON2sp1 sequences reveals five identical (red) and two similar (blue) residues.

PfAMA1 polymorphisms at positions 175 and 225 are determinant for the 3D7 specificity of R1

R1 is strain specific, binding to *Pf*AMA1 from the 3D7 (cognate antigen) and D10 strains, but with much reduced affinity to the HB3 or W2mef proteins, as determined by enzyme-linked immunosorbent assay (ELISA) (Harris, Casey et al. 2005) or SPR (Harris, Casey et al. 2009) measurements (recapitulated in Table 8). In contrast, *Pf*RON2sp1 bound to all the *Pf*AMA1 proteins tested (Table 4) with a higher affinity than for R1 peptide. Consistent with these values, *Pf*RON2sp1 displayed a higher capacity to inhibit red blood cell invasion by *P. falciparum* 3D7 than the R1 peptide (Figure 27). Moreover, *Pf*RON2sp1 shows cross-strain inhibition of invasion as expected from its biological function (Table 4), contrasting with the more restricted strain specificity of R1 (Figure 27, Table 8) (Harris, Casey et al. 2005).
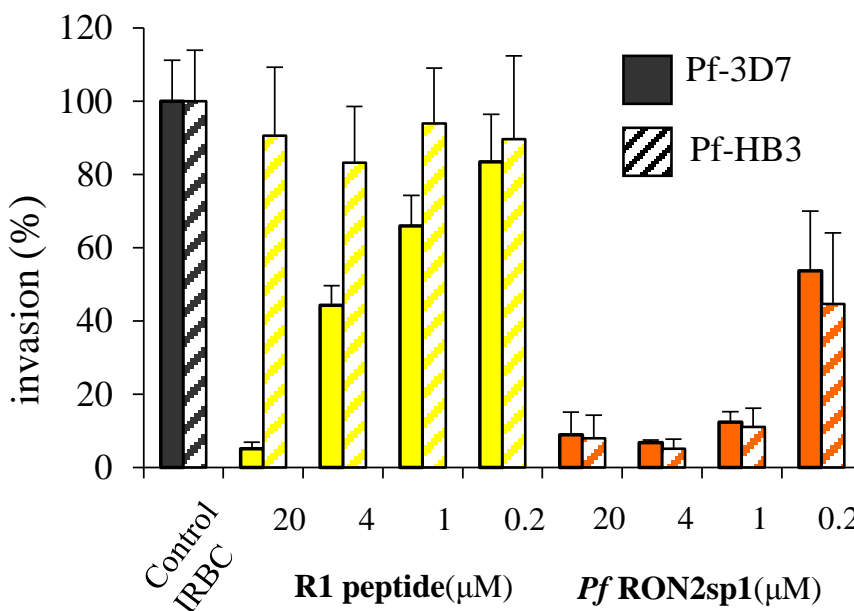
Figure 27: Highly potent cross-strain inhibition of red blood cell invasion by *Pf*RON2sp1.

Comparison of *Pf*RON2sp1 and R1 peptides (concentrations 0.2 to 20 µM) in inhibiting red blood cell invasion by *P. falciparum* 3D7 or HB3 highlights the higher inhibitory efficiency and cross-strain reactivity of *Pf*RON2sp1. Parasitemia of control infected red

blood cells (IRBC) sixteen hours post-invasion was used as the 100% invasion reference. Means (± standard deviation for N = 3) are shown.

Table 8: Polymorphic residues of *Pf*AMA1 contacting peptide R1.
A. Polymorphic residues contacting R1-major showing the sequence for strains analyzed using ELISA (*) (Harris, Casey et al. 2005), SPR ([+]) (Harris, Casey et al. 2009) and in this study using SPR (°). **B**. Polymorphic residues contacting R1-minor, showing the sequence for strains as presented in **A**. **C**. Binding to *Pf*AMA1, classified as strong (S) or weak (W) for the studies presented in **A** and **B**.

A.

Residue no.	3D7[+*]	D10[+]	HB3[+*]	W2mef[+*]	3D7mut°	Dico3°
175	Tyr	Tyr	Asp	Tyr	Tyr	Asp
224	Met	Met	Met	Met	Met	Met
225	Ile	Ile	Asn	Asn	Asn	Ile

B.

Residue no.	3D7*[+]	D10[+]	HB3[+*]	W2mef[+*]	3D7mut°	Dico3°
187	Glu	Glu	Glu	Lys	Glu	Glu
230	Lys	Lys	Lys	Lys	Lys	Lys

C.

	3D7[+*]	D10[+]	HB3[+*]	W2mef[+*]	3D7mut°	Dico3°
K_D (µM)	0.080	-	68	16	15	22
Binding	S	S	W	W	W	W

The *Pf*AMA1 3D7-R1 crystal structure shows that three polymorphic residues (175, 224 and 225) contact R1-major (Table 8). The 224 polymorphism, Met/Leu (strains tested only contained Met, but Leu can be found at position 224 in other strains), is conservative and since contacts are formed by the main chain only, this should not affect R1 specificity. The 3D7 and D10 antigens both carry Tyr175 and Ile225; for the W2mef and HB3 antigens,

residue 175 is Tyr and Asp, respectively, and residue 225 is Asn in both. Thus, polymorphisms at positions 225 and possibly 175 appear to be determinant for the 3D7 specificity of R1 at the major peptide-binding site (Table 8A). R1-minor contacts polymorphic residue 230, which is Lys in all strains studied (Table 8B). As our data suggest a weak affinity for this binding site, however, it is unlikely that this polymorphism has a significant effect on the specificity for R1. We examined these polymorphisms further using the mutant PfAMA1 Dico3 (Remarque, Faber et al. 2008), which differs only at residue 175 for the R1-contacting residues (Table 8A), and a 3D7 mutant with the substitution Ile225Asp, which we call 3D7mut. The equilibrium K_D, determined from the SPR steady-state responses to R1 binding, was 15.2±1.9 μM for 3D7mut and 22.3±3.3 μM for Dico3, showing a reduction in affinity of over 200-fold with respect to the native 3D7 antigen (Figure 28, Table 8C). This affinity is comparable to that observed for HB3 and W2mef (Harris, Casey et al. 2009) (recapitulated in Table 8), and confirms that both Tyr175 and Ile225 are important for the strain-specific recognition of R1. Tyr175, located at the tip of a flexible DI loop that is solvent-exposed in the apo antigen (Bai, Becker et al. 2005), becomes buried by R1-major and forms a hydrogen bond to this ligand *via* the phenol group. Ile225 is also buried by R1-major, forming a pair of hydrogen bonds *via* its main chain to the R1-major main chain.
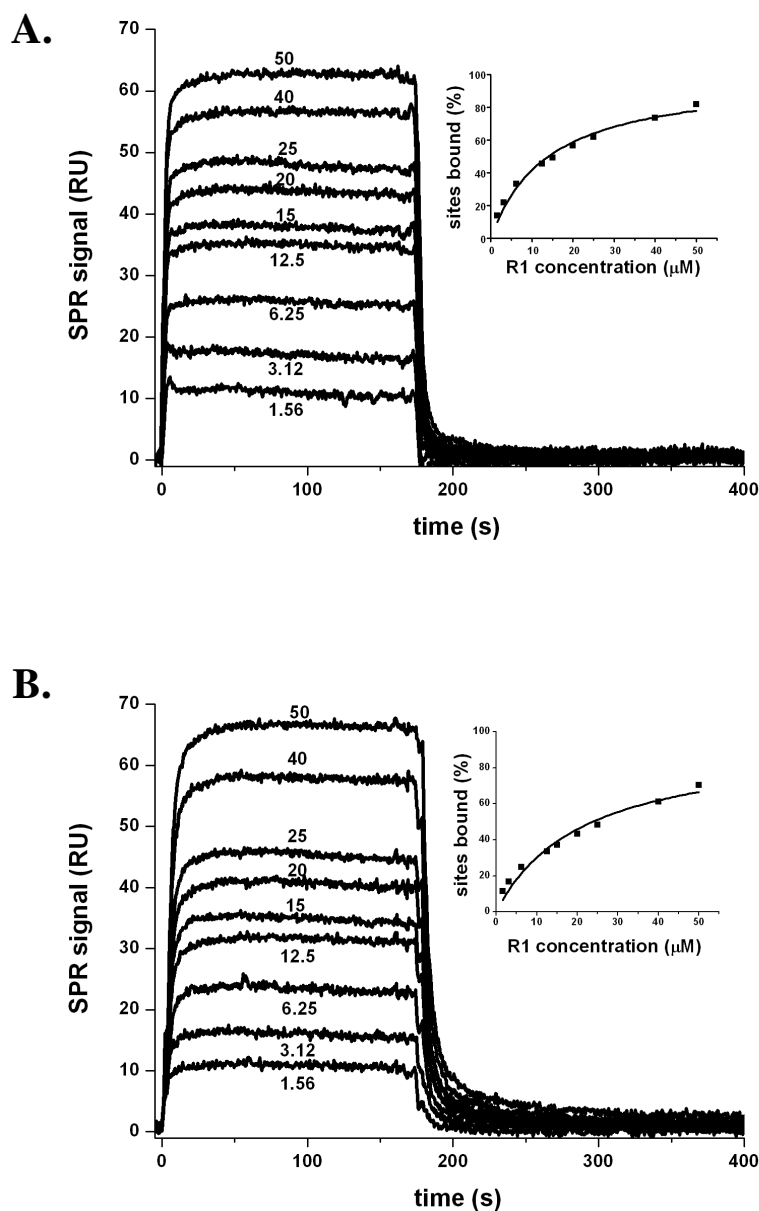
Figure 28: SPR studies of peptide R1 binding to *Pf*AMA1 mutants 3D7mut and Dico3.
A. Left - sensorgrams, showing R1 (analyte) binding to *Pf*AMA1 3D7mut (immobilized). R1 concentrations are indicated for each curve (µM). Right - the variation in percentage of bound sites (deduced from the steady-state response) with respect to analyte concentration. **B.** Left - sensorgrams, showing R1 (analyte) binding to Dico3 (immobilized), with R1 concentrations indicated. Right - the variation in percentage of bound sites (deduced from the steady-state response) with respect to analyte concentration. The equilibrium

dissociation constant K_D derived from the steady state binding curves is 15.2 µM for 3D7mut and 22.3 µM for Dico3.

Hot spots driving specific PfAMA1-PfRON2 complex formation

Guided by the similarities between the *Pf*RON2sp and R1 co-structures, and the conservation of key contact residues (Figure 29A), we probed the functional importance of a subset of *Pf*RON2 residues by testing the binding to BHK-21 cells expressing *Pf*AMA1 of GST-*Pf*RON2-5 fusion proteins carrying single alanine mutations at: Pro2033 (aligns structurally with Pro7 of peptide R1, which was shown to be critical for binding (Lee, Yao et al. 2011)), Phe2038 (interacts with invariant residue Phe183 in the hydrophobic groove and aligns structurally with Phe12 of R1), Arg2041 (extensive contacts with *Pf*AMA1 and structurally equivalent to Arg-P15 of R1) and Pro2044 (the peptide bond Ser2043-Pro2044 is *cis* and is thus important for the β-hairpin conformation). Consistent with the structure, mutation of Arg2041 to Ala abrogated binding to *Pf*AMA1 (Fig. 29B). Similar effects were observed with Pro2044, Phe2038 and Pro2033 mutations, and the latter was also shown to be a key residue in the *Tg*AMA1-*Tg*RON2 interaction (Tonkin, Roques et al. 2011).
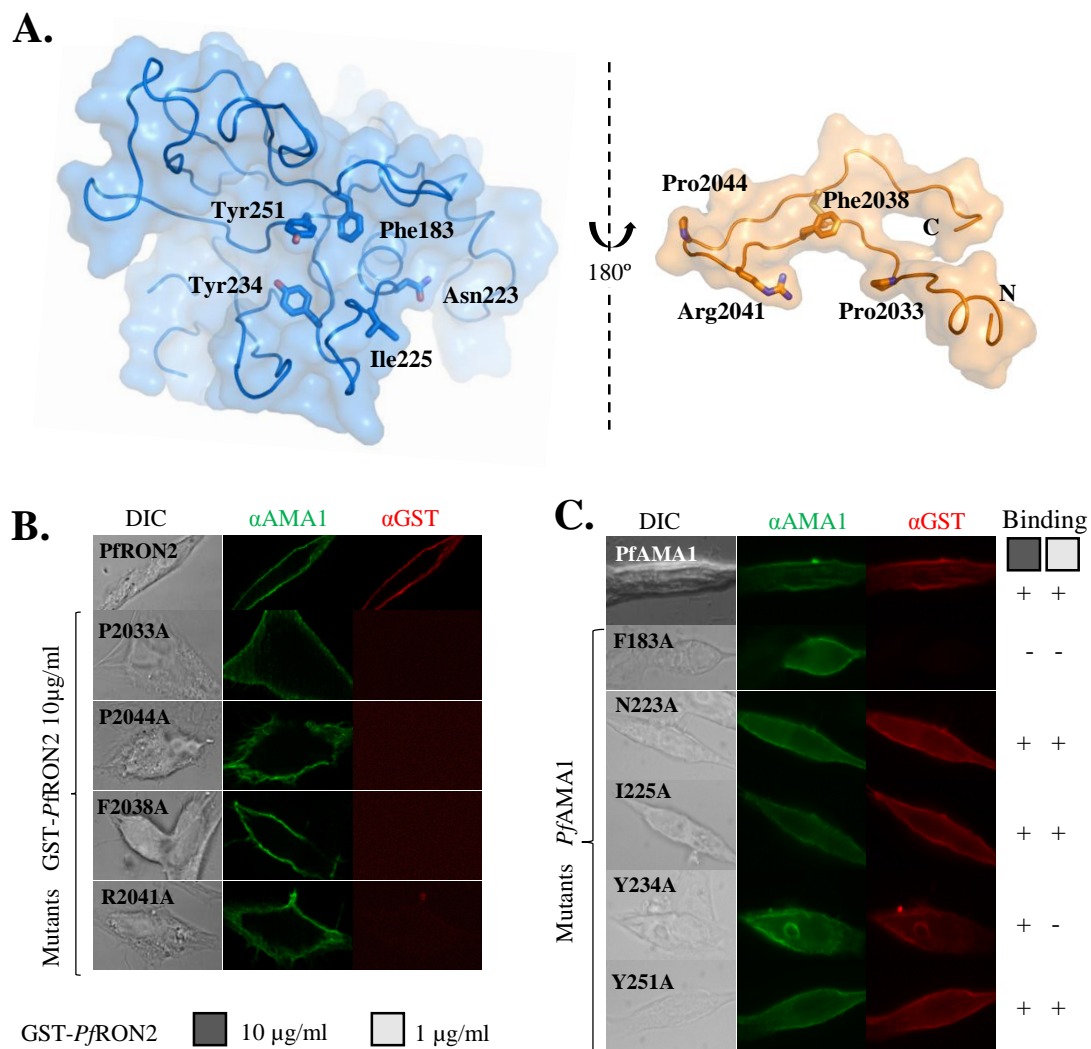
Figure 29: Mutations of *Pf*AMA1 and *Pf*RON2-5 reveal residues critical for high affinity interaction.

A. Interface between *Pf*AMA1 and *Pf*RON2sp1 shown in open-book presentation. Residues of both components that were mutated are labeled. **B.** Binding characteristics of recombinant GST-*Pf*RON2-5 mutants to dissect hot-spot residues in *Pf*RON2. *Pf*AMA1-expressing BHK-21 cells were incubated with 10 µg/ml of *Pf*RON2 or mutated proteins (GST-fusion proteins), washed and the binding of recombinant *Pf*RON2 fragment was revealed with anti-GST antibody. *Pf*AMA1 was detected with mAb F8.12.19, which recognizes extracellular Domain III. **C.** Binding consequences of *Pf*AMA1 mutations. Mutated versions of *Pf*AMA1 were expressed on the surface of BHK-21 cells and incubated with wild-type *Pf*RON2 recombinant proteins at 10 and 1 µg/ml.

Similarly, a subset of key *Pf*AMA1 residues was also chosen for mutation: Phe183 (an invariant residue that contributes to the hydrophobic groove and that interacts with Phe2038 of *Pf*RON2 *via* aromatic interactions), Asn223 (which makes important polar interactions with *Pf*RON2), residue 225 (a polymorphic residue that contributes many contacts to *Pf*RON2 in the structure of both the CAMP (Asn225) and 3D7 (Ile225) complexes), Tyr234 (which makes polar contacts to Arg2041 of *Pf*RON2) and Tyr251 (which has been suggested by previous studies to be important (Collins, Withers-Martinez et al. 2009, Srinivasan, Beatty et al. 2011)). A clear role for Phe183 in *Pf*AMA1-*Pf*RON2 complex formation was evident when expressed on the surface of BHK-21 cells and tested for the ability to bind GST-*Pf*RON2-5 fusion protein (Figure 29C). A less pronounced role of Tyr234 was observed and none for the remaining residues, including Tyr251. Although these conclusions differ from those of others (Collins, Withers-Martinez et al. 2009, Srinivasan, Beatty et al. 2011), these results are consistent with the limited contacts shown by this residue in the structures and with our earlier findings on the *Tg*AMA1-*Tg*RON2 interaction, where the equivalent *Tg*AMA1 residue, Tyr230, had a minimal effect on the binding (Tonkin, Roques et al. 2011).

3.6 Discussion

The structure of *Pf*AMA1 in complex with the extracellular region of its ligand *Pf*RON2 and the accompanying functional analyses reveal atomic details of the interaction between two key partners at the MJ. The binding site on *Pf*AMA1 includes the hydrophobic groove and a region that becomes exposed by displacement of the flexible DII loop from its apo conformation. Comparison of residues from both components at the *Pf*AMA1-*Pf*RON2 interface with those of other apicomplexan homologs underscores the separate co-evolution of the receptor-ligand pair in members of the phylum.

The DII loop displays a strong propensity for mobility in *P. falciparum* (Bai, Becker et al. 2005, Coley, Gupta et al. 2007) and *P. vivax* AMA1 structures (Pizarro, Vulliez-Le Normand et al. 2005), particularly at its N- and C-terminal extremities (weak or absent electron density); the central region of the DII loop is more structured and stabilized by contacts with DI, and is better defined in some of these AMA1 structures. Here, we show that the DII loop is displaced by *Pf*RON2sp, as well as by the R1 peptide. In *T. gondii*, the

DII loop is fourteen residues shorter than in the *Plasmodium* orthologues and appears less mobile (Crawford, Tonkin et al. 2010) but nonetheless is readily displaced by *Tg*RON2sp (Tonkin, Roques et al. 2011). Flexibility may therefore have an important functional role: it protects a significant portion of the binding site in apo AMA1 against the host's immune response but can be readily displaced to extend the hydrophobic groove for effective binding to RON2. The anti-*Pf*AMA1 invasion-inhibitory mAb 4G2, which binds to the N- and C-termini of the DII loop (Pizarro, Vulliez-Le Normand et al. 2005), probably prevents its displacement for effective binding to *Pf*RON2. The absence of polymorphisms in the DII loop in spite of immune targeting of this region underlines its important functional role (Chesne-Seck, Pizarro et al. 2005).

We have previously demonstrated an evolutionary constraint on the AMA1–RON2 interaction within apicomplexan parasites (Lamarque, Besteiro et al. 2011). Our functional analysis of the *Tg*AMA1-*Tg*RON2sp co-structure led to our hypothesis that the cystine loop initially anchors the receptor to the hydrophobic groove, causing expulsion of the DII loop to promote interaction throughout the entire binding site (Tonkin, Roques et al. 2011). Comparison of the *Tg*AMA1-*Tg*RON2sp and *Pf*AMA1-*Pf*RON2sp co-structures reveals that the cystine loop, while conserved across the two genera, is the most divergent region within this portion of RON2 (Figure 30). The separate co-evolution of the AMA1-RON2 pair in Apicomplexa is clearly illustrated by the difference between the cystine loop conformations of *Pf*RON2sp and *Tg*RON2sp. In particular, this allows Arg2041 to access the specific *Pf*AMA1 pocket (Figure 22D), where it participates in an intricate network of polar interactions. From mutagenesis, we have demonstrated a crucial role of Arg2041 in complex formation (Figure 29B). Moreover, this region of the cystine loop also appears to play an influential role in species selectivity as superposition of the *Pv*AMA1 structure (Pizarro, Vulliez-Le Normand et al. 2005) onto *Pf*AMA1-*Pf*RON2sp shows that Arg2041 would be sterically hindered at the interface but Thr, the equivalent residue in RON2 from *P. vivax*, can be accommodated (Figure 31A). This accounts for our prior observation that the original 67-residue segment of *Pf*RON2 does not bind to *Pv*AMA1 (Lamarque, Besteiro et al. 2011).

A.

*Tg*RON2	DIVQHMEDIGGAPPVSCVTNEI--LGVTCAPQAIAKATT
*Nc*RON2	DIVQHMEDIGGAPPASCVTNEI--LGVTCAPQAIAKATT
*Pf*RON2	DITQQAKDIGAGPVASCFTTRMSPPQQICLNSVVNTALS
*Pv*RON2	DISQHATDIGMGPATSCYTSTIPPPKQVCIQQAVKATLT
*Pk*RON2	DITQHASDIGMGPVTSCYTSTIPPPKQVCIQQAVKVTLT
*Py*RON2	DITQHATDIGMGPSTSCYTSLLPPPKSICIQQTVKTVLT
*Bb*RON2	DIAQHAADVGVGPAESCFI-MVKPPALHCVLKPVETLMK

Cystine loop

B.

*Pf*AMA1-*Pf*RON2sp-*Tg*RON2sp

C.

*Pf*AMA1--*Pf*RON2sp   *Tg*AMA1--*Tg*RON2sp

90°

Figure 30: The RON2 cystine loop governs specificity.

A. Alignment of RON2 sequences truncated to correlate *Pf*RON2sp1 with RON2 sequences from the following accession numbers: *Tg*RON2 - TGME49_100100, *Neospora caninum* RON2 (*Nc*RON2) - NCLIV_064620, *Pf*RON2 - PF14_0495, *Pv*RON2 - PVX_117880, *P. yoelii* RON2 (*Py*RON2) – PY_06813, *Babesia bovis* RON2 (*Bb*RON2)

(BBOV_I001630). **B.** Overlay of *Tg*RON2sp (green; PDB ID 2Y8T) onto *Pf*AMA1-*Pf*RON2sp (blue-orange) shows that both peptides adopt a helix/coil/cystine loop/coil architecture in the AMA1 groove, with the highest divergence localized to the cystine loop (black arrow). **C.** Electrostatic surface renderings of *Pf*AMA1 (left) and *Tg*AMA1 (right), with the secondary structure of the RON2 binding partner and residues defining the base of the cystine loop shown, illustrates that both interactions are complementary, but highly genus specific.
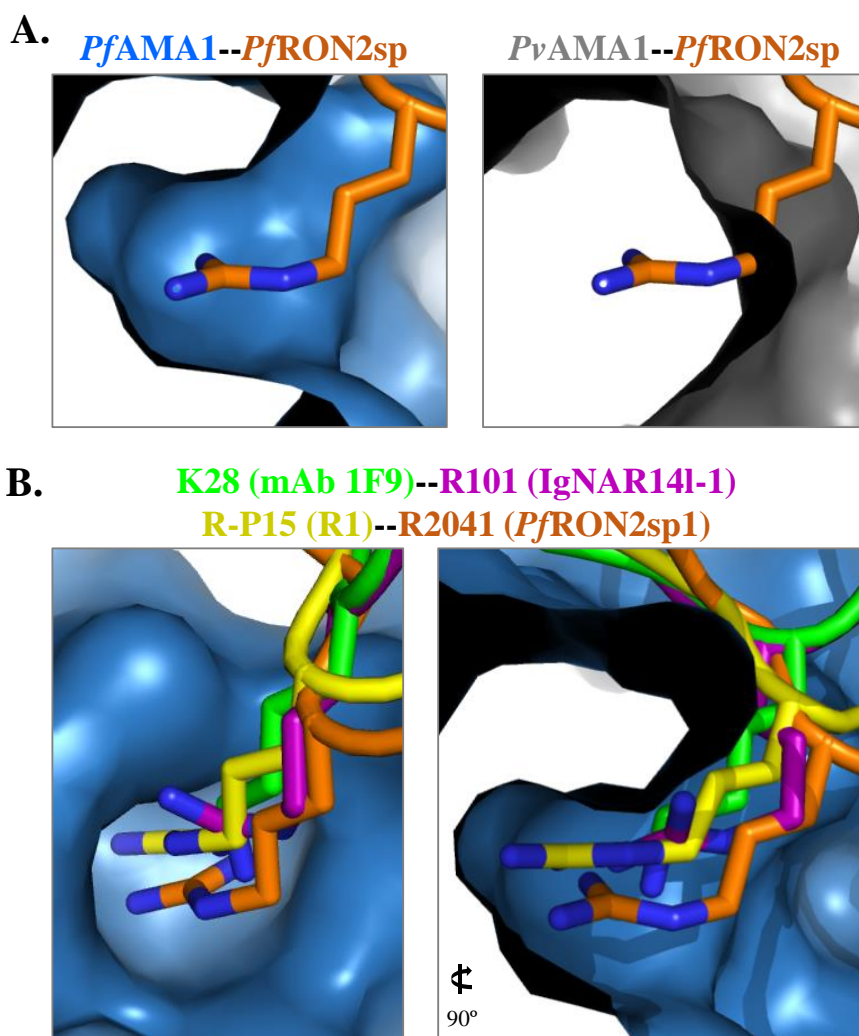
Figure 31: The Arg knob-in-hole interaction is critical for species selectivity and interaction with invasion inhibitory antibodies and peptides.

A. Left - A cut-away surface of *Pf*AMA1 (blue), reveals that Arg2041 of *Pf*RON2sp1 (orange) integrates deeply into a well-defined pocket. Right - However, no analogous

pocket is observed in *Pv*AMA1 (grey; PDB ID 1W8K). **B.** Peptides and antibodies known to be invasion inhibitory for *P. falciparum* occupy the key Arg binding site, as shown by orthogonal views of the *Pf*AMA1-*Pf*RON2sp1 co-structure (blue-orange) overlayed with the mAb 1F9 co-structure (1F9, green; PDB ID 2Q8B), IgNAR14l-1 co-structure (IgNAR, purple; PDB ID 2Z8V), and R1 co-structure (R1, yellow; reported here).

An additional feature of the *Pf*RON2sp cystine loop region is the presence of a *cis* peptide bond between Ser2043 and Pro2044; the Ser-Pro-Pro segment contributes negligible buried surface area but is important for maintaining the β-hairpin conformation for efficient complex formation. Sequence alignment reveals that the Pro duo (Pro2044-Pro2045) is preserved in all analyzed *Plasmodium* species (Fig. 30A) and is thus likely important for specific recognition of *Plasmodium* AMA1s. We propose that it provides necessary internal structure at the tip of the cystine loop and places the disulfide bond in the proper orientation to brace the AMA1-RON2 interaction. The influential role of Pro2044 is confirmed by mutagenesis where substitution with Ala, which would disfavor the *cis* peptide bond, abrogates *Pf*AMA1-*Pf*RON2 binding (Figure 29B). While *T. gondii* does not share the conserved proline pair, its cystine loop is two residues shorter (Figure 30A), which mirrors the narrower groove of *Tg*AMA1. Altogether, the overall U-shape architecture of RON2 in complex with AMA1 appears to be remarkably well maintained within apicomplexan parasites but specific features are clearly visible in the cystine loop of *Pf*RON2 and *Tg*RON2, highlighting how a receptor-ligand complex has evolved to maintain a common and crucial event in the biology of these parasites.

Although the *Pf*AMA1-*Pf*RON2 interface is highly conserved, five polymorphic residues of *Pf*AMA1 contact the non-polymorphic *Pf*RON2sp (Cao, Kaneko et al. 2009). Of these, however, only residue 225 (Asn/Ile) varies significantly. The remaining polymorphisms should not affect binding as they involve main chain contacts only (residues 172, 174, 187 and 224). Our study allows a detailed structural assessment of polymorphism at residue 225 since complexes with *Pf*AMA1 from the 3D7 (Ile225) and CAMP (Asn225) strains were determined. The 3D7 and CAMP orthologues both maintain two hydrogen bonds between the main chain of residue 225 and *Pf*RON2 Thr2039. However, Ile225 presents a deep pocket to Arg2041 with apolar contacts formed between

the aliphatic regions of these two side chains, while Asn225 presents a shallower pocket to Arg2041 with the Asn225 amide group stacking against the guanidyl group. Nonetheless, our binding studies by SPR show no significant difference in the affinity of these two *Pf*AMA1 homologs for *Pf*RON2sp2. Sequence variations at *Pf*RON2-interacting positions, 172 (Glu/Gly), 187 (Glu/Asn) and 225 (Ile/Asn) are represented by the strains 3D7, CAMP, FVO and HB3 that we have analyzed by SPR; the very similar K_D constants, ranging from approximately 10 to 20 nM, confirm that these exert little effect on the strength of the interaction.

Peptide R1 shows a more restricted specificity as it binds strongly to the cognate 3D7 and closely related D10 antigens but only weakly to orthologues that do not carry the same polymorphic amino acids at position 175 or 225 (Table 8). Tyr175 in *Pf*AMA1 3D7 makes a hydrogen bond to the main chain of R1-major but, as this residue is located in a flexible loop with some freedom to adapt to the *Pf*AMA1-R1 interface, it is unclear why the Asp175 polymorphism leads to reduced affinity. In the case of Ile225 of *Pf*AMA1 3D7, the main chain forms two hydrogen bonds to the main chain of R1-major but the preference of R1 for the Ile225 polymorphism remains unexplained as it contrasts with *Pf*RON2sp where main chain hydrogen bonds are also formed by both Ile225 (3D7) and Asn225 (CAMP) to the main chain of *Pf*RON2. This emphasizes that specificity differences may present subtleties that are difficult to decipher. Here, the crystal structure of R1 in complex with the 3D7mut (Ile225Asn) and Dico3 (Tyr175Asp) mutants of *Pf*AMA1 would provide invaluable insights into this question. Taken together, these results highlight that unlike the natural ligand *Pf*RON2, R1, which was selected by phage display, is highly susceptible to polymorphisms.

R1 exhibits a close structural similarity to *Pf*RON2, with the major/minor peptide pair displaying a similar boomerang form as *Pf*RON2, binding to the same region of *Pf*AMA1 and following the same general contour of the binding-site groove. Our structural data show that binding of R1-minor is dependent upon prior binding of R1-major as it lies above the latter in the binding groove and makes fewer contacts to *Pf*AMA1. This, indeed, is consistent with the ITC measurements that show a stoichiometry of 1:1, indicating a weaker affinity for the minor peptide-binding site. R1-major is thus favored as the principle

inhibitor of the interaction with *Pf*RON2, but this does not preclude a contribution by the minor peptide-binding site at high peptide concentrations.

Therapeutic strategies aimed at inhibiting the interaction between *Pf*AMA1 and *Pf*RON2 should be very effective in treating malaria as they address a critical phase in the life cycle of the parasite and, importantly, should not be compromised by polymorphism since the *Pf*AMA1-*Pf*RON2 interface is highly conserved. Our results provide a structural basis for designing inhibitors against the most virulent malaria parasite. The *Pf*RON2sp1 peptide used in this study has a very high affinity to *Pf*AMA1 and is very efficient at inhibiting invasion. Moreover, in contrast to the less strongly binding peptide R1, *Pf*RON2sp1 is not strain specific. Structural details of the *Pf*AMA1-*Pf*RON2 interaction offer the possibility to design molecules with the desired specific inhibitory properties by *in silico* screening and structural validation. The binding of *Pf*RON2 Arg2041 to a specific pocket on *Pf*AMA1 could be a critical target region. Indeed, the important role played by Arg-P15 at the *Pf*AMA1-R1 interface closely mirrors the equivalent interaction in the *Pf*AMA1-*Pf*RON2sp complexes and, interestingly, the same pocket is occupied by Arg and Lys in *Pf*AMA1 complexes with the invasion inhibitory antibodies IgNAR (Henderson, Streltsov et al. 2007) and mAb 1F9 (Coley, Gupta et al. 2007), respectively (Figure 31B). Phe2038 (corresponding to Phe-P12 in R1) is also a key residue, as its substitution by Ala affected binding. The importance of this sub-site is further highlighted by the concomitant loss in affinity when Phe183 (with which it interacts) was mutated in *Pf*AMA1. Collectively, these data provide a firm basis for designing molecules with optimal inhibitory properties to treat malarial infection.

Chapter 4: Revealing the structural basis for intra-species specificity in the moving junction proteins AMA1 and RON2

Adapted from:

Poukchanski, A., H. M. Fritz[*], **M. L. Tonkin**[*], M. Treeck, M. J. Boulanger and J. C. Boothroyd (2013). "*Toxoplasma gondii* sporozoites invade host cells using two novel paralogues of RON2 and AMA1." <u>PLoS One</u> **8**(8): e70637.

[*] These authors contributed equally

Contributions:

MLT designed and cloned the sporoAMA1full and sporoAMA1c constructs. **MLT** produced, purified, crystallized, and solved the structure for sporoAMA1c. **MLT** designed and cloned the sporoRON2-D3 construct. **MLT** produced, co-purified, co-crystallized, and solved the co-structure for sporoAMA1c-sporoRON2-D3. **MLT** and MJB analyzed the structural data. AP and HMF did the bioinformatic analyses, pull-down assays, localization studies, and invasion assays. MT sequenced the sporoAMA1 gene and created the sporoAMA1-HA plasmid. AP, HMF, **MLT**, MJB, and JCB wrote the manuscript.

4.1 Abstract

Toxoplasma gondii is an obligate intracellular parasite of the phylum Apicomplexa. The interaction of two well-studied proteins, apical membrane antigen 1 (AMA1) and rhoptry neck protein 2 (RON2), has been shown to be critical for invasion by the asexual tachyzoite stage. Recently, two paralogues of these proteins, dubbed sporoAMA1 and sporoRON2 (or RON2$_{L2}$), respectively, have been identified but not further characterized in proteomic and transcriptomic analyses of *Toxoplasma* sporozoites. Here, we show that sporoAMA1 and sporoRON2 localize to the apical region of sporozoites and that, *in vitro*, they interact specifically and exclusively, with no detectable interaction of sporoAMA1 with generic RON2 or sporoRON2 with generic AMA1. Structural studies of the interacting domains of sporoRON2 and sporoAMA1 indicate a novel pairing that is similar in overall form but distinct in detail from the previously described interaction of the generic pairing. Most

notably, binding of sporoRON2 domain 3 to domains I/II of sporoAMA1 results in major alterations in the latter protein at the site of binding and allosterically in the membrane-proximal domain III of sporoAMA1 suggesting a possible role in signaling. Lastly, pretreatment of sporozoites with domain 3 of sporoRON2 substantially impedes their invasion into host cells while having no effect on tachyzoites, and vice versa for domain 3 of generic RON2 (which inhibits tachyzoite but not sporozoite invasion). These data indicate that sporozoites and tachyzoites each use a distinct pair of paralogous AMA1 and RON2 proteins for invasion into host cells, possibly due to the very different environment in which they each must function.

4.2 Introduction

Studies on host cell invasion by apicomplexan parasites, including *Toxoplasma gondii,* have generally focused on the asexual stages. It is a complex process, whereby the parasite invades the host cell in an active manner involving largely the use of the parasite's own machinery (Dobrowolski and Sibley 1996, Keeley and Soldati 2004) and a coordinated secretion of multiple proteins stored in at least two different organelles, the micronemes and the rhoptries (Alexander, Mital et al. 2005, Lebrun, Michelin et al. 2005, Alexander, Arastu-Kapur et al. 2006, Besteiro, Michelin et al. 2009, Collins, Withers-Martinez et al. 2009, Straub, Cheng et al. 2009, Richard, MacRaild et al. 2010). Invasion begins with a tight attachment, reorientation (or high-affinity apical attachment) and the onset of gliding motility to help the parasite propel its way into the host cell. This latter step involves the formation of an intimate ring of attachment between the plasma membranes of the host cell and parasite (Aikawa, Miller et al. 1978, Michel, Schupp et al. 1980) that migrates down the length of the parasite as it invades. This transient structure is referred to as the moving junction (MJ; also sometimes referred to as the tight junction) and has multiple roles, including generating the parasitophorous vacuole (PV) (Suss-Toby, Zimmerberg et al. 1996) as the parasite pushes into the host cell.

In *T. gondii*, the MJ has been characterized extensively for tachyzoites, the rapidly-dividing, asexual form. The tachyzoite MJ is a multimeric protein complex known to be composed of rhoptry neck proteins (RON) 2, 4, 5, and 8 (Alexander, Mital et al. 2005, Lebrun, Michelin et al. 2005) and apical membrane antigen 1 (AMA1). RON2 has been

predicted to span the host plasma membrane as it interacts with RONs 4, 5, and 8 on the host cytosolic side, and AMA1 on the parasite surface (Tonkin, Roques et al. 2011, Tyler, Treeck et al. 2011). AMA1, a type I transmembrane protein, is conserved across all apicomplexans (Hehl, Lekutis et al. 2000) and its knockdown has been shown to markedly reduce invasion (Mital, Meissner et al. 2005). Furthermore, blocking the AMA1 ectodomain either by antibodies or small peptides can inhibit invasion of host cells by both *T. gondii* and *Plasmodium* species asexual stages (Collins, Withers-Martinez et al. 2009, Treeck, Tamborrini et al. 2009, Lamarque, Besteiro et al. 2011, Tyler and Boothroyd 2011). The intimate, high-affinity interaction of domain 3 of RON2 and the ectodomain of AMA1 is crucial for efficient invasion (Tonkin, Roques et al. 2011, Tyler and Boothroyd 2011) and structural analyses of the association for both *T. gondii* and *P. falciparum* asexual stages have shown an extensive, buried region of interaction between the two proteins (Tonkin, Roques et al. 2011, Vulliez-Le Normand, Tonkin et al. 2012).

Contrary to the well-characterized tachyzoite invasion, very little is known about the mechanism of how *Toxoplasma* sporozoites invade. Sporozoites develop over the course of several days inside the oocysts that are shed by felids into the external environment. Upon ingestion by an intermediate host, sporozoites excyst and invade the host's distal small intestine. At some point soon after the initial invasion, sporozoites convert into tachyzoites, which then disseminate throughout the host (Dubey, Speer et al. 1997). The exact mechanism of host cell invasion by sporozoites has not been studied but they have been reported to use a two-step process whereby invasion first produces a distended, primary vacuole from which the parasite then proceeds to elaborate a tighter, secondary vacuole in which it then grows (Tilley, Fichera et al. 1997). The machinery used in these various steps has not been identified or investigated.

Recently, *T. gondii* sporozoites were subjected to detailed transcriptomic and proteomic analyses (Fritz, Bowyer et al. 2012, Fritz, Buchholz et al. 2012). It was found that, in addition to the well-characterized "generic" AMA1 and RON2, sporozoites also express two paralogues dubbed sporoAMA1 and sporoRON2, respectively, which are not expressed at detectable levels in tachyzoites or bradyzoites (Fritz, Buchholz et al. 2012). The identification of these paralogues drove the question as to the precise interactions and roles played by the two sets of AMA1/RON2 paralogues during sporozoite invasion. Here,

we show that the generic and sporozoite-specific paralogues interact in a mutually exclusive manner. We also perform invasion inhibition assays with sporozoites and show that the sporoAMA1-sporoRON2 complex formation is critical for sporozoite invasion of the host cell, while the "generic" AMA1/RON2 interaction is dispensable for invasion of this life cycle stage. Structural studies reveal the molecular basis for these observations.

4.3 Materials and methods

Materials

All materials were purchased as described in section 2.3, with the following exceptions. An improved transfection reagent, Promega FuGENE HD, was purchased from Fisher Scientific (Pittsburgh, PA). FlashBAC ULTRA, a more highly engineered linearized baculovirus DNA, was purchased from Oxford Expression Technologies (Oxford, UK). BioWhittaker Lonza Insect-Xpress cell growth media for Hi5 and *Spodoptera frugiperda 9* (*Sf*9) cells was purchased from VWR (Radnor, PA).

Host cell culture and parasites

Human foreskin fibroblasts (HFFs) were cultured in complete Dulbecco's modified Eagle's medium (DMEM; Invitrogen, Carlsbad, CA) supplemented with 10% heat-inactivated fetal bovine serum (Hyclone, Logan, UT), 2 mM L-glutamine, 100 U ml^{-1} penicillin and 100 μg ml^{-1} streptomycin. *T. gondii* (Type I RH and Type II ME49 strains) lacking a functional hypoxanthine-xanthine-guanine phosphoribosyltransferase (*HXGPRT*) gene (designated RHΔ*hxgprt* or ME49Δ*hxgprt,* respectively) (Donald, Carter et al. 1996) were cultured by serial passage on confluent monolayers of HFFs in complete DMEM at 37°C with 5% CO2.

Generation of GST-RON2 (gD3 and sD3) fusion proteins

To generate RON2 fusion proteins with an N-terminal glutathione-S-transferase (GST) tag, genericRON2 and sporoRON2 sequences were aligned to determine the location of domain 3 (D3) of sporoRON2. Once determined, both generic RON2D3 and sporoRON2D3 were polymerase chain reaction (PCR)-amplified from *T. gondii* ME49 genomic DNA and introduced into pGEX-6P1 (Agilent) using the BamHI and EcoRI sites. GST-gD3 was

generated as described previously (Tyler and Boothroyd 2011). To generate GST-sD3, the coding sequence for *Tg*SporoRON2 amino acids 995 to 1048 was PCR amplified using primers (GGATCCGACATCGCTCAGTTCCTCACCGAC and CTCGAGTCACTTGAAGACATCCGACAGCGCAG). Production and purification of the GST proteins from *E. coli* strain Rosetta (Novagen) were done essentially as described previously (Brymora, Valova et al. 2004). Concentrated, purified proteins were stored at -80°C in buffer containing 10 mM Tris-HCl pH 8.0, 150 mM NaCl, and 10% glycerol.

Generation of sporoAMA1-HA plasmid

RNA was extracted from sporozoites using Trizol reagent as previously described (Fritz, Buchholz et al. 2012), cDNA was generated using random hexamer primer and the sporoAMA1 gene was amplified using the following primers: sporoAMA1-FW (ATGCCTACAGAATCTCGAAGTAT) and sporoAMA1-REV (GAACTCTGC GTCGACGGCCCT). The resulting product was then subcloned into TOPO for sequencing. After verifying the correct sequence the gene was cloned using the cold-fusion kit (Systems Biosciences, Mountain View, USA), into pSAG1-CDPK3::HA (Garrison, Treeck et al. 2012) using the primers: sporoAMA1_cold-FW (CGAGTATG catgccATGCCTACAGAATCTCGAAGTATCTTGGCTAGGGCGGAAGAGACC) and sporoAMA1_cold-REV: (CAACGGTGAttaATTAATCAGAACTCTGCGTCGACGG CCCTGGAACCCAGAAGCGACT) to generate pSAG1-sporoAMA1::HA.

Transfection of RHΔhxgprt parasites

2×10^7 RHΔ*hxgprt* tachyzoites were washed once with 1 x phosphate-buffered saline (PBS). After one wash in Cytomix (10 mM KPO$_4$, 120 mM KCl, 0.15 mM CaCl$_2$, 5 mM MgCl$_2$, 25 mM Hepes, 2 mM EDTA), the parasites were transfected with 50 μg of pSAG1-sporoAMA1::HA as previously published (Soldati and Boothroyd 1993). Parasites were plated on a fully confluent HFF monolayer, allowed to invade for 24 hours before the media was changed to DMEM to remove residual cytomix.

GST pull-down experiments

Approximately 4×10^8 extracellular RHΔ*hxgprt* or RHΔ*hxgprt*-sporoAMA1-HA parasites were washed three times in 1 x PBS and then lysed on ice in 1 ml of lysis buffer (10 mM

Tris-HCl pH 8.0, 150 mM NaCl, 1 mM EDTA, 0.1% NP-40), supplemented with Complete EDTA-free Protease Inhibitors (Roche). The cleared, NP-40-solublized lysate was divided equally into three tubes and each fraction was supplemented with 4B Glutathione-sepharose beads (GE Healthsciences) that were prebound with 0.5 μM of GST, GST-gD3, or GST-sD3. The lysate suspensions were rotated at room temperature for approximately two hours and the bound beads were then centrifuged at 70 x g for 2 minutes and the supernatant ("flow-through") collected. The pelleted beads were then washed three times in lysis buffer, followed by elution of the GST fusion proteins and any co-purified parasite proteins by boiling for ~5 minutes in 2 x SDS sample buffer (125 mM Tris-HCl pH 7.0, 4% SDS, 20% glycerol, 0.005% bromophenol blue) supplemented with 10% β-mercaptoethanol ("pull-down").

Western blot analyses

Samples were separated on 4-12% gradient Bis-Tris gels (Invitrogen) and analyzed by Western Blot using the following antibodies. Generic AMA1 was detected with mouse monoclonal B3.90 (Donahue, Carruthers et al. 2000). SporoAMA1-HA was detected using the anti-HA rat monoclonal antibody 3F10 conjugated to horseradish peroxidase (HRP) (Roche). SAG1 was detected using rabbit polyclonal sera (a gift from M. Grigg, NIH). Goat anti-mouse secondary antibodies were HRP-conjugated (Biorad).

Cloning of the TgSporoama1 and TgSpororon2d3 genes

TgSporoAMA1

The sequence encoding the complete ectodomain of *Tg*SporoAMA1 Type II strain Me49 was synthesized by GenScript (Piscataway, NJ) and codon optimized for insect cell expression. A construct encompassing the conserved portions of the three ectoplasmic domains was sub-cloned out of the synthesized gene (Gln98 to Glu481, with numbering based on the initiation methionine in the signal sequence; sequence determined by the John C. Boothroyd lab (Stanford, CA); *Tg*SporoAMA1_conserved, *Tg*SporoAMA1c) using the following primers: F, 5'-ACATGACCATGGGACAAAACCCC TGGGCTACC-3' and R, 5'-CTGTCTGGCGGCCGCTTCAGGCTCTGTACAGCA AGC-3'. The construct was NcoI – NotI digested and pasted into a pAcGP67B vector modified as described in section

2.3. Sequencing confirmed that no mutations were incorporated during amplification procedures.

TgSporoRON2D3-TRX

The sequence encoding the region of *Tg*SporoRON2 Type II strain Me49 corresponding to *Tg*RON2sp (Asp452 to Ser487, with numbering based on the initiation methionine in the signal sequence; UniProt ID B6KLP1; *Tg*SporoRON2_D3) was synthesized by GenScript (Piscataway, NJ) and codon optimized for expression in *E. coli*. *Tgspororon2d3* was cloned NheI – NotI into a modified pET32a vector (Novagen through EMD Millipore; Billerica, MA) containing N-terminal thioredoxin (TRX) and hexa-histidine (His$_6$) tags separated from the gene of interest by a thrombin site.

Protein Production and co-purification

Baculovirus generation and amplification

The *Tg*SporoAMA1c P1 and P2 viruses were prepared as described in section 2.3, with the exception that FlashBAC ULTRA linearized baculovirus DNA and FuGENE HD lipid transfection reagent were used. In addition, all cells were grown in Lonza media, and the P2 virus was high enough titre to not require amplification to a P3.

Protein production – TgSporoAMA1c

For large scale expression, 2 L of Hi5 cells at 2.0 x 10^6 cells/mL in Lonza media containing 10 µg/mL gentamicin were infected with an optimized ratio of 4 mL of *Tgsporoama1c*/pAcGP67Bmod P2 virus per 1 L of cell culture. The baculovirus infected cells were incubated at 27 °C with shaking at 120 rpm for 65 hours. After the 65 hour incubation, the Hi5 cells were large and granulated and at a cell density of 1.8 x 10^6 cells/mL with 25% nonviable cells.

Protein production – TgSporoRON2D3

For large scale expression, the *Tgspororon2d3*/pET32amod2 clone was transformed into BL21 codon plus (DE3) cells (Novagen through EMD Millipore; Billerica, MA), plated on Luria-Bertani (LB) agar plates containing 50 µg/mL ampicillin and 35 µg/mL

chloramphenicol, and incubated overnight at 37 °C. A single colony was picked and inoculated into 50 mL 2xYT media containing 50 µg/mL ampicillin, 35 µg/mL chloramphenicol, and 0.5 % glucose. The starter culture was incubated with shaking at 200 rpm for 12 hours at 37 °C. After 12 hours, the cell suspension was centrifuged at 2200 rpm for 7 min at 4 °C, the supernatant was discarded, and the pellet was resuspended in Novagen Overnight Express Instant TB autoinduction media (EMD Millipore; Billerica, MA) and used to inoculate 1 L autoinduction media. The large scale culture was grown at 37 °C for 4 hours followed by 12 hours at 30 °C, shaking at 180 rpm. After 16 hours, the cell pellet was harvested by centrifugation at 6300 rpm for 12 min at 4 °C. The supernatant was discarded, and the pellet was resuspended in Binding Buffer and frozen at -80 °C.

Protein co-purification

*Tg*SporoAMA1c was purified from the insect cell culture supernatant, as described for *Pf*AMA1 in section 3.4. The Ni-purified sample was thrombin cleaved overnight and subsequently purified by size exclusion chromatography to remove both the cleaved His_6 tag and thrombin.

 *Tg*SporoRON2D3-TRX was purified from the *E. coli* cell lysate. The frozen pellet was partially thawed overnight by incubation at -20 °C, and subsequently thawed in ice cold water. Sixty µL DNase (2.5 mg/mL) was added to the sample and mixed by end-over-end rotation, and was passed through a French Press twice. The sample was diluted with Binding Buffer and insoluble material was removed by centrifugation at 13000 rpm for 30 min at 4 °C.

 Purified, cleaved *Tg*SporoAMA1c was added directly to the *Tg*SporoRON2D3-TRX clarified cell lysate and allowed to incubate at 4 °C for 30 min with stirring. *Tg*SporoRON2D3-TRX and any bound *Tg*SporoAMA1c was purified by Ni batch binds as described for *Pf*AMA1 in section 3.4. The TRX tag was removed by overnight thrombin cleavage using an optimized ratio of 1 µL of thrombin per 10 mg protein.

 The *Tg*SporoAMA1c-*Tg*SporoRON2D3 complex was purified from thrombin, thioredoxin tag, excess *Tg*SporoRON2D3 and other contaminating proteins by size exclusion chromatography. The purified sample was concentrated to 30 mg/mL in HBS for crystallization trials. The purity of the complex was determined by sodium dodecyl sulfate

polyacrylamide gel electrophoresis (SDS-PAGE) at each stage of the purification and protein concentrations were analyzed by absorbance at 280 nm based on an extinction coefficient of 67350 M^{-1} cm^{-1} for *Tg*SporoAMA1c and 5625 M^{-1} cm^{-1} for *Tg*SporoRON2D3 (19730 M^{-1} cm^{-1} for *Tg*SporoRON2D3-TRX).

Crystallization, data collection and processing

TgSporoAMA1c

Initial crystals of *Tg*SporoAMA1c were identified in MCSG-1 (Microlytic; Burlington, MA) after twelve days. The optimized crystals grew to their final size within eight weeks in a condition of 1.0 M succinic acid pH 7.0, 100 mM HEPES pH 7.0, 1% polyethylene glycol (PEG) monomethyl ether 2000, and 100 mM glycine. The final 2.4 µL drops consisted of equal volumes *Tg*SporoAMA1c (30 mg/mL) and reservoir solution equilibrated against 100 µL of reservoir solution. Cryoprotection was carried out in reservoir solution supplemented with 25% glycerol for 20 seconds and the crystal was flash cooled at 100K directly in the cryostream. Diffraction data were collected on beam line 9-2 at SSRL (Menlo Park, CA).

TgSporoAMA1c-TgSporoRON2D3

Crystal trials for *Tg*SporoAMA1c-*Tg*SporoRON2D3 were set using a Crystal Gryphon (Art Robbins Instruments; Sunnyvale, CA) crystallization robot using several different commercial crystallization screens and three different protein to reservoir drop ratios. Initial crystals of *Tg*SporoAMA1c-*Tg*SporoRON2D3 were identified in Index (Hampton Research), and diffraction quality crystals were obtained after two days in 200 mM magnesium chloride hexahydrate, 100 mM HEPES pH 7.5, 25% PEG3350. The final 0.6 µL drops consisted of 0.2 µL protein (*Tg*SporoAMA1c and *Tg*SporoRON2D3 (15 mg/mL combined)) and 0.4 µL reservoir solution equilibrated against 50 µL of reservoir solution. Cryoprotection was carried out in reservoir solution supplemented with 25% glycerol for 20 seconds and the crystal was flash cooled at 100K directly in the cryostream. Diffraction data were collected on the microfocus beam line 12-2 at SSRL (Menlo Park, CA).

Structure solution and refinement

TgSporoAMA1c

Diffraction data were processed to 2.35 Å resolution using Imosflm (Battye, Kontogiannis et al. 2011) and Scala (Evans 2006) in the CCP4 suite of programs (Winn, Ballard et al. 2011). Initial phases were obtained by MR using PHASER (McCoy, Grosse-Kunstleve et al. 2007) with the DI and DII domains of the unliganded *Tg*AMA1 structure (Protein Data Bank (PDB) ID 2X2Z; 37% sequence identity) trimmed with CHAINSAW (Schwarzenbacher, Godzik et al. 2004) to better reflect the *Tg*SporoAMA1 sequence. No MR solution was found for DIII, which was manually traced into the electron density after several rounds of refinement of the core structure. Addition of solvent molecules was performed manually in COOT (Emsley and Cowtan 2004) and refinement in Refmac5 (Murshudov, Vagin et al. 1997). Stereochemical analysis was performed with PROCHECK and SFCHECK in CCP4 (Winn, Ballard et al. 2011). Data collection and refinement statistics are presented in Table 9.

TgSporoAMA1c-TgSporoRON2D3

Diffraction data were processed to 3.10 Å resolution using Imosflm (Battye, Kontogiannis et al. 2011) and Scala (Evans 2006) in the CCP4 suite of programs (Winn, Ballard et al. 2011). Initial phases were obtained by MR using PHASER (McCoy, Grosse-Kunstleve et al. 2007) with the DI and DII domains of the unliganded *Tg*SporoAMA1c structure with the DII loop removed. No MR solution was found for DIII, so DIII was manually traced into the electron density after multiple rounds of refinement. Tracing of the *Tg*SporoRON2 peptide, and addition of solvent molecules, was performed manually in COOT (Emsley and Cowtan 2004) and refinement in Refmac5 (Murshudov, Vagin et al. 1997). Stereochemical analysis was performed with PROCHECK and SFCHECK in CCP4 (Winn, Ballard et al. 2011). Data collection and refinement statistics are presented in Table 10.

Protein Data Bank accession code

The atomic coordinates and structure factors have been deposited in the PDB under the following codes: *Tg*SporoAMA1c, 3ZLE; *Tg*SporoAMA1c-*Tg*SporoRON2D3, 3ZLD.

Generation of anti-TgSporoAMA1 ectodomain antibodies

Antibodies were raised against the *Tg*SporoAMA1 ectodomain that was purified as stated above. The ectodomain was dissolved in PBS and 100 µg was injected into BALB/c mice in RIBI adjuvant (Corixa). Identical boosts were given at 21-day intervals. Polyclonal antisera was collected after the second boost and screened for reactivity by immunofluorescence assay (IFA) against tachyzoites transfected with *Tg*SporoAMA1-HA. Antibodies used in this study showed colocalization with anti-HA parasites expressing *Tg*SporoAMA1-HA but no reactivity with untransfected parasites indicating their specificity for *Tg*SporoAMA1.

Generation of anti-TgSporoRON2 Domain 4 Antibodies

Antibodies were raised to GST fusions of Domain 4 of *Tg*SporoRON2 (amino acids 1069 to 1167). Proteins were purified as described above (GST purification). Proteins (500 µg) were injected subcutaneously into rabbits (New Zealand White, Harlan laboratories). A 1:1 mix with Freund's complete adjuvant was used for the first injection only. Freund's incomplete adjuvant was used for all subsequent boosts. Boosts were done every two weeks until a sufficiently high antibody titer was obtained. Polyclonal antisera were collected after the third injection and at each subsequent boost. Bleeds were screened for reactivity to sporozoites (by immunofluorescence), and showed no reactivity with tachyzoites (as determined by Western blot and immunofluorescence).

Sporozoite excystation

Oocysts were produced in kittens and harvested from feces as previously described (Fritz, Buchholz et al. 2012). Sporulated M4 oocysts that had been stored in 2% sulfuric acid at 4 ºC were washed three times in 1 x PBS to remove sulfuric acid. The final washed pellet was suspended in 10% Clorox® bleach (diluted in 1 x PBS) and incubated on ice for 30 minutes. Oocysts were then washed two times with 1 x PBS and a third time in DMEM media (without serum) to remove bleach. The final washed oocyst pellet was suspended in DMEM and transferred to a 1.5 ml screw-top microcentrifuge tube containing 350 mg acid-washed glass beads (200-400 mm, Invitrogen) and vortexed at maximum speed in three 30-second intervals (90 seconds total) at which time approximately 90% of the oocysts

were broken open with free sporocysts. Broken oocysts/sporocysts were collected and pelleted by centrifugation. The resulting pellet was suspended in 5% sodium taurodeoxycholate hydrate (Sigma) in DMEM and incubated at 37 ℃ for 10 minutes. Sporozoites were then washed two times in cold DMEM. A third wash was performed in DMEM supplemented with 2% fetal bovine serum (FBS). The washed sporozoites were suspended in DMEM supplemented with 2% FBS for invasion assay.

Immunofluorescence assays

To visualize *Tg*SporoAMA1 in extracellular sporozoites, sporozoites were excysted and placed on a 12-well glass side (Tekdon, Inc.). The sporozoites were allowed to dry on the slide before being fixed in 1 x PBS with 2.5% formaldehyde (EM Biosciences) and permeabilized using 0.2% Triton X-100 in PBS and 3% Bovine Serum Albumin (Sigma). To visualize *Tg*SporoAMA1 in invaded sporozoites, HFF monolayers infected with sporozoites were fixed in 1 x PBS containing 2.5% formaldehyde (EM Biosciences), 2-3 hours post-infection followed by permeabilization as for extracellular parasites. The monolayers were stained with polyclonal mouse-anti-*Tg*SporoAMA1 serum, and either rabbit anti-MIC10 or rabbit-anti-MIC5 polyclonal sera. The primary staining was followed by AlexaFluor5940-labeled goat-anti-mouse antibody and AlexaFluor488-labeled goat-anti-rabbit antibody, respectively. All the secondary AlexaFluor-conjugated antibodies were obtained from Molecular Probes. Coverslips were mounted onto glass slides using Vectashield (Vector Laboratories) and then examined using a 100x oil-immersion objective on an Olympic BX60 upright fluorescence microscope.

To visualize *Tg*SporoRON2 in apical compartment of sporozoites, HFF monolayers infected with sporozoites were fixed in 100% ice-cold methanol for 3 minutes at room temperature. These fixed monolayers were then blocked in PBS and 3% bovine serum albumin (Sigma). The monolayers were stained with polyclonal rabbit-anti-*Tg*SporoRON2 serum and either mouse-anti-RON4, mouse-anti-ROP2/3/4 or mouse-anti-*Tg*SporoAMA1. The primary staining was followed by AlexaFluor594-labeled goat-anti-mouse antibody and AlexaFluor488-labeled goat-anti-rabbit antibody, respectively (Invitrogen). Coverslips were mounted onto glass slides using Vectashield (Vector Laboratories) and then examined using a 100x oil-immersion objective on an Olympus BX60 upright

fluorescence microscope. All digital images were obtained using Image-Pro Plus and the same exposure parameters were used for all comparison sets.

Invasion assay

ME49 tachyzoites were released from infected HFFs by scraping and passage through a 27 gauge needle. Released tachyzoites and M4 excysted sporozoites (as described above) were washed in DMEM supplemented with 2% FBS. Parasites were then incubated in DMEM with 2% FBS supplemented with: 5 μM GST, 5 μM GST-gD3, 5 μM GST-sD3, or 2.5 μM of both GST-gD3 and GST-sD3 at 37C for 3 minutes (all proteins were in 10 mM Tris-HCl pH 8.0, 150 mM NaCl, 10% glycerol). Parasites were added to pre-chilled HFF monolayers grown on glass coverslips. Temperature synchronization of invasion was accomplished by allowing parasites to settle onto the HFFs in an ice water bath for 10 minutes prior to invasion. To initiate invasion, the plates were then transferred to a 37°C water bath for 45 minutes. Infected monolayers were washed twice in 1 x PBS and then fixed in 1 x PBS containing 2.5% formaldehyde. To stain only the extracellular parasites, fixed monolayers were stained with polyclonal rabbit-anti-*Toxoplasma* serum followed by AlexaFluor488-labeled goat-anti-rabbit. To stain all parasites, the infected monolayers were then permeabilized with 1x PBS containing 0.2% triton X-100 followed by staining with the anti-SAG1 mouse monoclonal antibody DG52 (Burg, Perelman et al. 1988) and AlexaFluor594-labeled goat-anti-mouse. The numbers of green (intracellular and extracellular) and red (extracellular) tachyzoites were counted in ten randomly selected fields on each of three separately mounted coverslips for each condition and visualization was performed using a 20x objective on a Nikon Eclipse TE300 microscope. The same exact process was carried out for sporozoites. All image acquisition and analysis was performed blinded. All digital images were obtained using Image-Pro Plus and parasites were quantified using ImageJ (Schneider, Rasband et al. 2012).

4.4 Results

TgSporoRON2 and TgSporoAMA1 are distinct from their "generic" paralogues.

The existence of sporozoite-specific versions of "generic" RON2 and "generic" AMA1 in *T. gondii* sporozoites begs the question of their role in sporozoite invasion. To address this,

we first asked how prevalent are these proteins in related parasites with similar life cycle stages? This was done by creating a rooted phylogenetic tree using ClustalW algorithms on the full amino acid sequences of the RON2 homologues present in these related species. As seen in Figure 32A, *Neospora* and *Eimeria* have orthologues of *Toxoplasma* sporoRON2 that segregate in a distinct and separate clade from the orthologues of the original *Toxoplasma* generic RON2 in these species. Only a single RON2 homologue is present in the representative *Plasmodium*, *Babesia*, and *Theileria* species examined and the clade that includes these latter RON2 sequences is distinct from both the generic RON2 and sporoRON2 clades seen with the Eimeriorina (*Toxoplasma*, *Eimeria* and *Neospora*). This suggests that the duplication that led to the two RON2 clades in Eimeriorina occurred after its split from the Haemosporina (including *Plasmodium*, *Babesia*, and *Theileria*). This is consistent with the fact that the two RON2 versions in the Eimeriorina appear equally closely related to the single RON2 in the Haemosporina. To further elucidate the differences between these proteins, we aligned the critical domain 3 of representative generic RON2 and sporoRON2 orthologues (Figure 32B). This domain of generic RON2 has been previously shown to interact with generic AMA1 in a critical step for invasion (Tyler and Boothroyd 2011). As seen in Figure 32B, the clustering of the RON2-like sequences in the other species apparent at the whole protein level plays out similarly for the crucial domain 3. In fact, for species like *T. gondii* that have generic and sporozoite-specific paralogues, there is very little conservation between the two at the individual amino acid level, with a few notable exceptions including a pair of cysteines that are known for generic RON2 to form an intramolecular disulfide bond (Tonkin, Roques et al. 2011).

A.

B.

C.
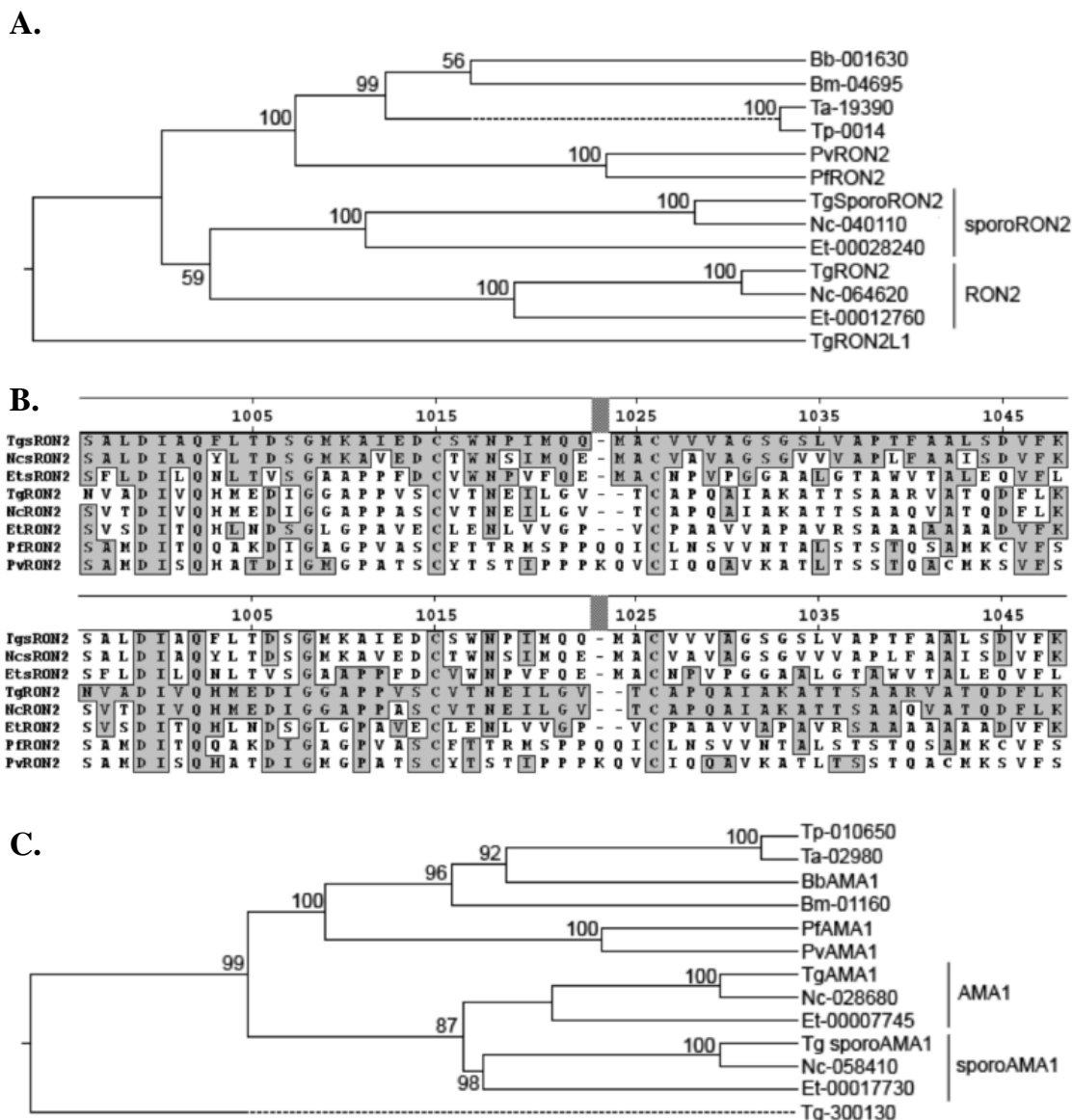
Figure 32: *Tg*SporoRON2 and *Tg*SporoAMA1 are conserved in other apicomplexans and are distinct from generic RON2 and generic AMA1.

A. The *T. gondii* RON2 polypeptide sequences were aligned with their respective homologues in *Eimeria tenella, Neospora caninum, Plasmodium* species (*P. falciparum* and *P. vivax*), *Babesia* species (*B. bovis* and *B. microti*), and *Theileria* species (*T. annulata* and *T. parva*) using ClutalW (as part of MegAlign software (Lasergene)) and an anchored phylogenetic tree was generated using the standard algorithm. The clusters including generic and sporozoite-specific versions of each protein are so-labeled. Bootstrapping analysis was performed to determine confidence intervals (1000 trials). **B.** Alignment of

domain 3 (D3) of the indicated RON2 homologues was performed in ClustalW. Residues identical to that of *Tg*SporoRON2 are indicated with shading on the upper panel, while residues identical to those of *T. gondii* generic RON2 are boxed on the lower panel. Numbers indicate amino acid position of *Tg*SporoRON2 from the starting methionine. **C.** As for **A** except using the AMA1 polypeptide sequences.

Similar analyses were performed for the AMA1 homologues in these species. Construction of an anchored tree revealed a similar distribution of clades, with one set of AMA1 homologues in *Toxoplasma*, *Eimeria* and *Neospora* forming a distinct clade that includes *Tg*SporoAMA1 and another that includes the generic AMA1 of *Toxoplasma* (Figure 32C). The *Plasmodium*, *Babesia* and *Theileria* species analyzed have only a single AMA1 homologue each and these form a separate clade that is equally closely related to the generic and *Tg*SporoAMA1 clades of the Eimeriorina. The existence, however, of an orthologue of *Tg*SporoAMA1 in the genera having orthologues of *Tg*SporoRON2 supported the hypothesis that these sporozoite-specific versions of a well-described generic pairing might themselves be specifically interacting.

TgSporoRON2D3 forms a specific interaction with TgSporoAMA1

It has previously been shown that there is an extensive and tight interaction between generic AMA1 and domain 3 of generic RON2 (Richard, MacRaild et al. 2010, Tonkin, Roques et al. 2011, Tyler and Boothroyd 2011). As the sequences of domain 3 of generic RON2 (gD3) and sporoRON2 (sD3) differ substantially (Figure 32B), we predicted that sD3 would show little if any ability to interact with generic AMA1. To test this prediction, we generated a fusion of sD3 (amino acids 995 to 1048) with GST. In parallel, we prepared a GST fusion of gD3 (amino acids 1293-1345), as previously described (Tyler and Boothroyd 2011). Tachyzoite lysates were mixed with these and glutathione-coupled beads were used to capture parasite proteins that bound to each. The results showed that, as previously reported, generic AMA1 was specifically enriched in the eluate fraction incubated with GST-gD3 but no such binding was observed with GST-sD3 or GST (Figure 33A). These data indicate that, as predicted from its sequence divergence, GST-sD3 has little if any ability to specifically interact with generic AMA1.
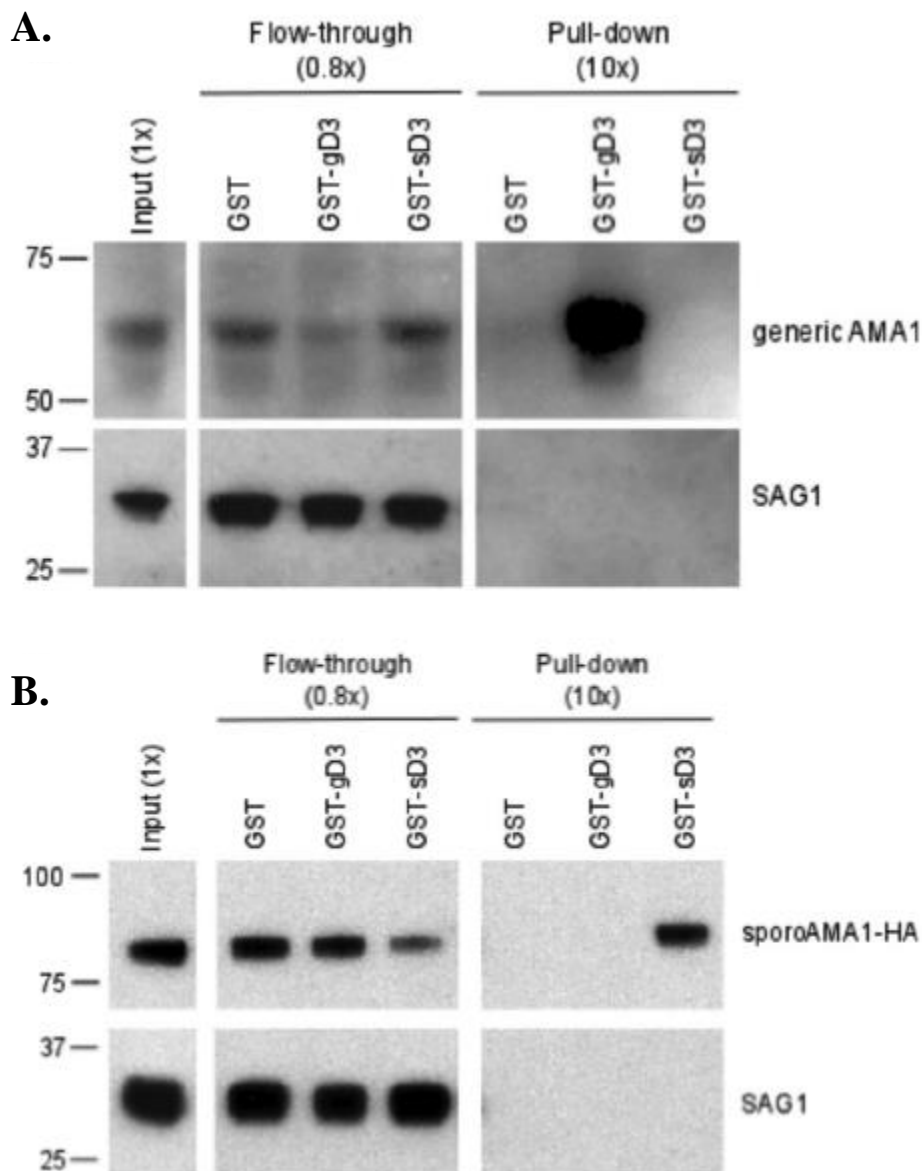
Figure 33: The sporo- and generic versions of RON2 domain 3 interact only with their respective sporo- and generic AMA1 partners.

A. Molar equivalents of GST, GST-gD3, or GST-sD3 were added to lysates of RHΔ*hxgprt* and after NP-40 solubilization the material that did not bind to the GST fusions ("flow-through") or that was pelleted with the fusions ("pull-down") was resolved by polyacrylamide gel electrophoresis and analyzed by immunoblotting with antibodies to generic AMA1 or SAG1 as a control for loading and nonspecific pelleting. Parentheses indicate parasite equivalents of the different fractions relative to input (1x). Size markers indicated in kDa. **B.** GST-pull-down experiments were performed as described in **A** except

using RHΔ*hxgprt* that were transiently expressing sporoAMA1-HA and the sporoAMA1 was detected using antibodies to the HA-epitope tag.

The existence of a sporozoite-specific version of AMA1 suggested this might be the binding partner of *Tg*SporoRON2. To test this, we repeated our GST pull-down experiment. This time, however, tachyzoites were transiently transfected with a plasmid containing full-length *Tg*SporoAMA1 with a C-terminal HA tag. After allowing the transfectants to grow for 48 hours, the parasites were lysed and GST, GST-sD3 and GST-gD3 used for affinity purification. The input, flow-through and eluate were analyzed by immunoblotting for the presence of the HA-tagged sporoAMA1 (Figure 33B). The results show *Tg*SporoAMA1 was specifically enriched in the eluate fraction of GST-sD3, and not in the eluates of GST or GST-gD3 indicating that *Tg*SporoRON2 and *Tg*SporoAMA1 do indeed form a specific interaction, independent of their generic paralogues.

TgSporoAMA1 presents an extensively guarded apical surface

The structure of generic AMA1, co-crystallized with a synthetic peptide representing generic RON2 domain 3 (*Tg*RON2 synthetic peptide (sp)) has previously been reported (Tonkin, Roques et al. 2011). To better understand the mutually exclusive nature of the interaction between the sporozoite and generic AMA1/RON2 pairings, we set out to solve the crystal structure of apo *Tg*SporoAMA1 and of *Tg*SporoAMA1 bound to a portion of sporoRON2D3 (sD3). The structure of the conserved three-domain ectoplasmic region of apo *Tg*SporoAMA1 was solved to 2.35 Å (Table 9). Overall, the structure of apo *Tg*SporoAMA1 conforms to the expected three-domain architecture observed for other AMA1 structures (Fig. 34A) (Bai, Becker et al. 2005, Pizarro, Vulliez-Le Normand et al. 2005, Crawford, Tonkin et al. 2010, Tonkin, Crawford et al. 2013), but reveals several divergent substructures including a network of apical surface loops that provide an unexpected degree of integration with the apical groove (Figure 34B). In particular, apical loops 1 and 2 of domain I (DI) and the extended loop of domain II (DII loop) cover the region occupied by RON2 in previously determined AMA1/RON2 co-structures from both *T. gondii* and *P. falciparum* (Figure 34B) (Tonkin, Roques et al. 2011, Vulliez-Le Normand, Tonkin et al. 2012). In addition, *Tg*SporoAMA1 loop 2 is slightly longer than

the analogous loop in generic AMA1 and has an N-linked glycosylation at Asn230 that is positioned atop the groove (Figure 34B; note that while generic AMA1 has been reported to be naturally N-glycosylated in tachyzoites (Fauquenoy, Morelle et al. 2008), no information exists on *Tg*SporoAMA1 in this regard but it does have a consensus N-linked site which is efficiently N-glycosylated in the insect cells used to generate the recombinant *Tg*SporoAMA1 studied here). Loop 2 folds in over the top of the groove, and positions Pro227 directly over the tip of the DII loop, which appears to compensate for the lack of the central anchoring tyrosine observed in generic AMA1 (*Tg*SporoAMA1 Ser252; *T. gondii* generic AMA1 Tyr230) (Crawford, Tonkin et al. 2010), with Pro227 effectively pinning DII loop Phe376 and Trp377 into the base of the apical groove (Figure 34B inset). Based on the degree to which the apical groove of *Tg*SporoAMA1 is guarded, numerous specific interactions between *Tg*SporoAMA1 and *Tg*SporoRON2, particularly in the predicted initial binding site at the cystine loop, are likely to be required to overcome the energy barriers of DI apical loop rearrangement and displacement of the DII loop.

Table 9: Data collection and refinement statistics for apo *Tg*SporoAMA1c.

Data collection statistics	
Spacegroup	P2$_1$
a, b, c (Å)	179.19, 155.52, 180.59
α, β, γ (deg.)	90, 92.31, 90
Wavelength (Å)	0.9795
Resolution range (Å)	79.30 – 2.35 (2.48 – 2.35)
Measured reflections	1391080
Unique reflections	391285
Redundancy	3.6 (2.4)
Completeness (%)	95.4 (75.8)
I/σ(I)	10.0 (2.1)
R$_{merge}$[a]	0.081 (0.341)
Refinement statistics	
Resolution (Å)	78.04 – 2.35
R$_{cryst}$[b]	0.209
R$_{free}$[c]	0.244
No. of atoms	
Protein (A/B/C/D/E/F/	2927/2899/2899/2897/2927/2889/
G/H/I/J/K/L)	2927/2913/2865/2891/2921/2921
Solvent	1695
Glycerol	108
B-values (Å^2)	
Protein (A/B/C/D/E/F/	37.5/35.7/36.3/35.6/36.0/38.2/
G/H/I/J/K/L)	38.2/38.7/46.1/48.3/49.9/50.0
Solvent	43.7
Glycerol	43.6
r.m.s. deviation from ideality	
Bond lengths (Å)	0.013
Bond angles (deg.)	1.31
Ramachandran statistics (%)	
Most favoured	96.1
Allowed	3.9
Disallowed	0.0

Values in parentheses are for the highest resolution shell

[a] R$_{merge}$= $\sum_{hkl} \sum_i |I_{hkl,i} - [I_{hkl}]| / \sum_{hkl} \sum_i I_{hkl,i}$, where $[I_{hkl}]$ is the average of symmetry related observations of a unique reflection

[b] R$_{cryst}$=$\sum|F_{obs}-F_{calc}|/\sum F_{obs}$, where F$_{obs}$ and F$_{calc}$ are the observed and the calculated structure factors, respectively

[c] R$_{free}$ is R using 5% of reflections randomly chosen and omitted from refinement

Figure 34: SporoAMA1 presents a highly guarded apical groove.

A. Stacked three domain architecture of *Tg*SporoAMA1 shown in the predicted organization to the *T. gondii* sporozoite plasma membrane with the three ectodomains indicated as DI in burgundy, DII in green and DIII in blue. Disulfides are shown as yellow sticks. Dotted line indicates extended Pro/Glu rich region between the conserved portion of DIII and the transmembrane domain (grey rectangle) that leads through to the C-terminal domain (grey oval). **B.** Apical view of apo sporoAMA1 structure, with core structure shown as grey surface and DI surface loops that guard the apical groove shown as burgundy cartoon and semi-transparent surfaces, and the DII loop as a green cartoon and semi-transparent surface. N-linked glycosylation on Asn230 shown as sticks. Numbers indicate

surface loops that frame the apical groove. Inset: sporoAMA1 DII loop residues Phe376 and Trp377 (green) are pinned into the apical groove by Pro227 at the tip of loop 2 (burgundy).

The apical groove of TgSporoAMA1 is specific for TgSporoRON2D3.

The co-structure of *Tg*SporoAMA1 with *Tg*SporoRON2D3, solved to 3.1 Å resolution (Table 10), reveals significant complementarity, with a buried surface area of 3274 Å^2, 23 intermolecular hydrogen bonds (Table 11), and a complexation significance score of 1.00 indicating physiological relevance (Krissinel and Henrick 2007). *Tg*SporoRON2D3 is integrated into the apical groove of *Tg*SporoAMA1 through an N-terminal α-helix seated in the area exposed by displacement of the domain II loop and connecting coil ordered through the center of the groove to the disulfide-bound beta hairpin loop at the opposite end of the apical surface (Figure 35A left). This organization mimics the overall binding paradigm previously observed for generic AMA1/RON2 complexes from *T. gondii* tachyzoites (Figure 35A right) and *P. falciparum* (Tonkin, Roques et al. 2011, Vulliez-Le Normand, Tonkin et al. 2012).

Table 10: Data collection and refinement statistics for *Tg*SporoAMA1c-*Tg*SporoRON2D3.

Data collection statistics	
Spacegroup	$C222_1$
a, b, c (Å)	49.37, 124.18, 171.91
α, β, γ (deg.)	90, 90, 90
Wavelength (Å)	0.9795
Resolution range (Å)	85.96 – 3.10 (3.27 – 3.10)
Measured reflections	39279
Unique reflections	9716
Redundancy	4.0 (3.7)
Completeness (%)	97.8 (95.8)
I/σ(I)	8.2 (3.5)
R_{merge}[a]	0.114 (0.311)
Refinement statistics	
Resolution (Å)	85.96 – 3.10
R_{cryst}[b]	0.205
R_{free}[c]	0.263
No. of atoms	
Protein (A/B)	2745/229
Solvent	7
B-values ($Å^2$)	
Protein (A/B)	49.5/58.6
Solvent	31.7
r.m.s. deviation from ideality	
Bond lengths (Å)	0.011
Bond angles (deg.)	1.22
Ramachandran statistics (%)	
Most favoured	95.9
Allowed	4.1
Disallowed	0.0

Values in parentheses are for the highest resolution shell

[a] $R_{merge} = \sum_{hkl} \sum_i |I_{hkl,i} - [I_{hkl}]| / \sum_{hkl} \sum_i I_{hkl,i}$, where $[I_{hkl}]$ is the average of symmetry related observations of a unique reflection

[b] $R_{cryst} = \sum |F_{obs} - F_{calc}| / \sum F_{obs}$, where F_{obs} and F_{calc} are the observed and the calculated structure factors, respectively

[c] R_{free} is R using 10% of reflections randomly chosen and omitted from refinement

Table 11: Hydrogen bond interactions observed in the *Tg*SporoAMA1c-*Tg*SporoRON2D3 (reported here) and generic AMA1-RON2sp (PDB 2Y8T; chains A and B, respectively) co-structures, aligned based on RON2 sequences with similar interactions bolded and the cystine loop region shaded light grey.

sporoRON2D3	sporoAMA1	Distance (Å)	genRON2sp	genAMA1	Distance (Å)
			Glu1303 [Oε1]	Gln361 [Nε2]	2.55
			Glu1303 [Oε2]	Arg111 [N]	3.46
Asp1006 [O]	**Arg135 [NH1]**	**3.54**	**Asp1304 [O]**	**Arg111 [NH1]**	**2.98**
Asp1006 [Oδ1]	Arg135 [NH1]	3.86	Asp1304 [Oδ1]	Gln361 [Nε2]	2.75
Asp1006 [Oδ2]	Ser383 [N]	3.78			
Gly1008 [O]	**Ala255[N]**	**3.51**	**Gly1306 [O]**	**Met233 [N]**	**3.65**
			Val1311 [N]	Tyr230 [OH]	3.54
Glu1013 [O]	Gln183 [Nε2]	3.04	Val1311 [O]	Tyr230 [OH]	2.56
Asp1014 [O]	Gln183 [Nε2]	3.23			
Cys1015 [N]	Val225 [O]	3.35			
Cys1015 [O]	**Val225 [N]**	**3.06**	**Cys1313 [O]**	**Met204 [N]**	**3.06**
Ser1016 [N]	Gln183 [Oε1]	3.44			
Ser1016 [Oγ]	Gln183 [Oε1]	2.32			
Trp1017 [N]	**Ile223 [O]**	**2.75**	**Thr1315 [N]**	**Val202 [O]**	**2.90**
Trp1017 [O]	**Ile223 [N]**	**2.96**	**Thr1315 [O]**	**Val202 [N]**	**2.85**
Asn1018 [Nδ2]	Tyr185 [OH]	3.46	Asn1316 [Nδ2]	Phe197 [O]	3.47
Asn1018 [Oδ1]	Thr222 [N]	3.52	Asn1316 [Nδ2]	Lys200 [O]	3.70
Asn1018 [Oδ1]	Thr222 [Oγ1]	3.59	Asn1316 [Nδ2]	Thr201 [Oγ1]	2.76
			Glu1317 [N]	Lys200 [O]	2.99
Met1021 [O]	Arg202 [NH2]	3.47			
Gln1023 [Nε2]	Ser187 [Oγ]	3.69			
Gln1023 [Oε1]	Thr186 [N]	3.49			
Gln1023 [Oε1]	Ser187 [N]	3.08			
Met1024 [O]	Thr186 [N]	3.28			
Met1024 [O]	Thr186 [Oγ1]	3.65			
			Thr1322 [Oγ1]	Thr165 [N]	3.59
Cys1026 [N]	**Val184 [O]**	**2.69**	**Cys1323 [N]**	**Val164 [O]**	**2.92**
Cys1026 [O]	**Val184 [N]**	**2.82**	**Cys1323 [O]**	**Val164 [N]**	**2.94**
			Gln1326 [Nε2]	Thr144 [Oγ1]	3.69
			Ala1327 [O]	Glu145 [N]	2.87
			Ala1329 [N]	Pro143 [O]	2.97
			Lys1330 [N]	Glu145 [Oε1]	3.16
			Ala1331 [O]	Trp253 [Nε1]	3.35
			Thr1333 [O]	Tyr230 [OH]	2.90
			Thr1333 [Oγ1]	Tyr230 [OH]	3.70
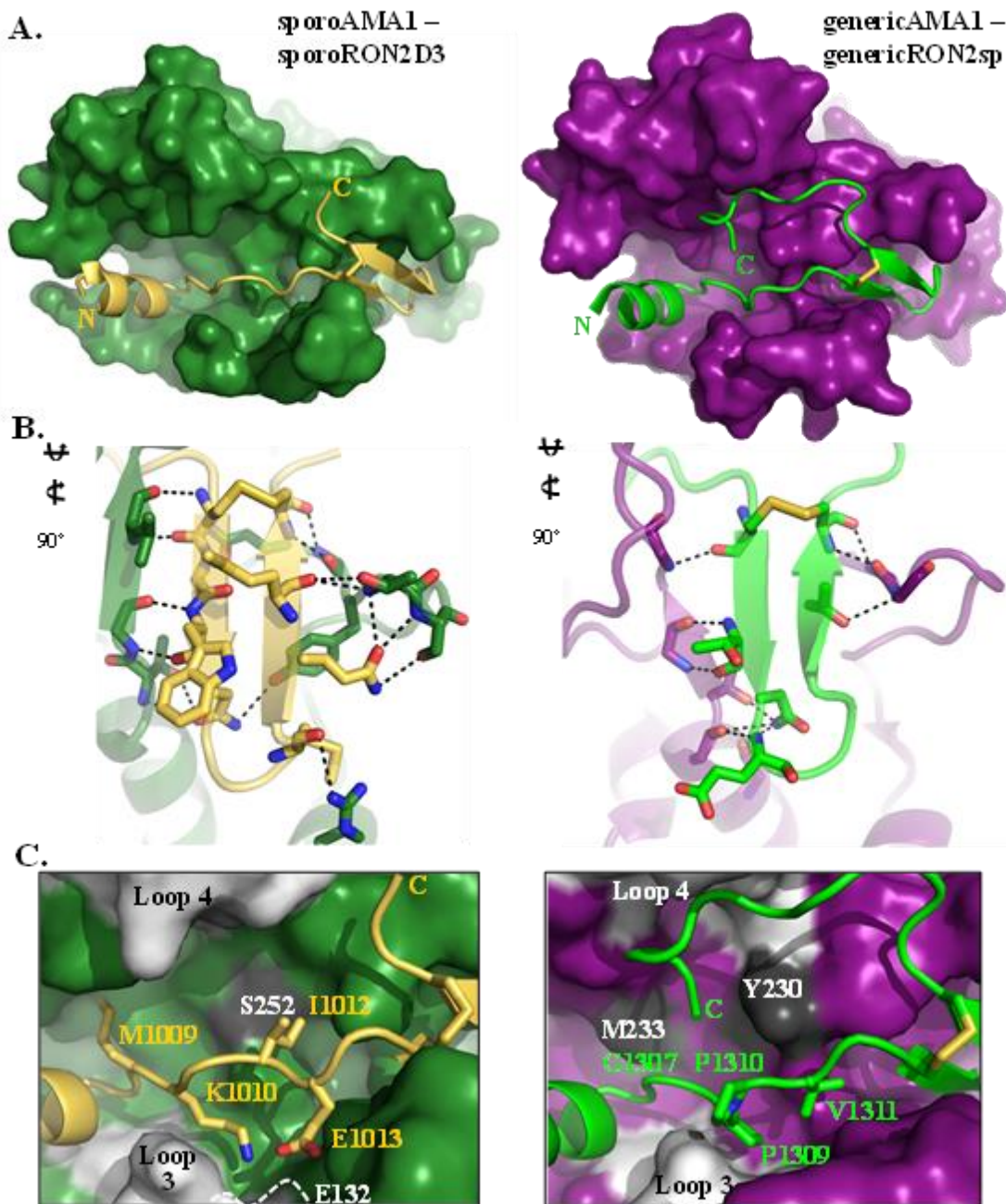
Figure 35: Specificity between *Tg*SporoAMA1 and *Tg*SporoRON2D3 is achieved through interactions within both the cystine loop and the connecting coil.

A. Apical view of *Tg*SporoAMA1 (green surface) bound to *Tg*SporoRON2D3 (gold cartoon) (left), showing conservation of the overall AMA1/RON2 binding paradigm with generic AMA1 (purple surface) – generic RON2 synthetic peptide (sp; green cartoon) (right; PDB 2Y8T). Note that the extreme C-terminal portion of the sporoRON2 peptide is disordered as a result of its relatively "early" exit from the stabilizing environment of the

hydrophobic groove and therefore is not resolved in this structure. **B.** Cystine loop interactions clearly differ between *Tg*SporoAMA1-*Tg*SporoRON2D3 (left) and generic AMA1-generic RON2sp (right). Hydrogen bonds shown as dotted black lines. Colored as in **A**. **C.** Additional specificity is gained through interactions with *Tg*SporoAMA1 (green surface) or generic AMA1 (purple surface) and the RON2 coil that connects the N-term helix to the cystine loop. Central groove residue (*Tg*SporoAMA1 Ser252 or generic AMA1 Tyr230) and generic AMA1 groove residue Met233 colored dark grey. Beta-hairpin loop 3 and variable loop 4 colored light grey. Side chains of *Tg*SporoRON2D3 (gold cartoon) and generic RON2sp (green cartoon) involved in specificity shown as sticks.

Despite the similar overall binding paradigm, a detailed analysis of the interactions between each paralogue pair revealed the basis for the biochemically observed specificity. While cross-genera specificity of the AMA1/RON2 interaction has been mainly attributed to interactions within the cystine loop (Tonkin, Roques et al. 2011, Vulliez-Le Normand, Tonkin et al. 2012), between the *T. gondii* paralogue pairs there appear to be major contributions to specificity from both the cystine loop and the coil connecting the N-terminal helix to the cystine loop. Nearly three quarters of the hydrogen bonds formed between sporoAMA1 and sporoRON2-D3 are found within the cystine loop region, and only the strictly backbone interactions are observed at the generic AMA1-generic RON2sp interface, which contains approximately half the number of hydrogen bonds observed for generic AMA1 and the generic RON2 cystine loop (Table 11; Figure 35B). Moving outward from the cystine loop region, the coil connecting the N-terminal helix to the cystine loop has three notable shifts in the intermolecular interface between the paralogue pairings (Figure 35C). Firstly, *Tg*SporoRON2 Met1009 buries into a deep cleft in the side of *Tg*SporoAMA1, while the corresponding residue in generic RON2 is Gly1307 and there is no cleft in this region of generic AMA1 due to the presence of Met233 (Figure 35C). Secondly, beta-hairpin loop 3 is four residues shorter in *Tg*SporoAMA1 than generic AMA1, which together with the introduction of a glutamic acid residue in this loop (*Tg*SporoAMA1 Glu132), presents an enlarged pocket with sufficient shape and charge complementarity to accommodate an extended lysine residue of *Tg*SporoRON2 (*Tg*SporoRON2 Lys1010) (Figure 35C left). In contrast, other AMA1/RON2 co-structures

present a shorter AMA1 pocket accommodating a compact RON2 proline (generic *Tg*RON2 Pro1309; *Pf*RON2 Pro2033) (Figure 35C right). Finally, the flattened surface caused by the presence of groove-central Ser252 in *Tg*SporoAMA1 readily accommodates a bowed out conformation of *Tg*SporoRON2D3 consisting of Ile1012 and Glu1013, whereas the corresponding central residue in generic AMA1, Tyr230, limits the accessible area in this region and can only accommodate a valine (Val1311) flipped down towards the base of the groove and held in position by Pro1310 (Figure 35C). All three of these interactions clearly show receptor-ligand co-evolution that provides specificity for the generic and sporozoite-specific AMA1/RON2 pairs of *T. gondii*.

The C-terminal region of TgSporoRON2D3 is not required for coordination

While some differences between the paralogue pairings contribute to specificity, others are likely the result of capturing different biologically relevant conformations. The C-terminus of *Tg*SporoRON2D3, for example, does not extend the length of the groove as observed in the generic AMA1-RON2sp co-structure, but rather exits the groove shortly after the disulfide of the beta hairpin loop. This conformation of *Tg*SporoRON2D3 likely results from *Tg*SporoAMA1 loop 4 being four residues shorter than the corresponding helical backstop of generic AMA1, which results in an approximately 5 Å encroachment of the loop towards the center of the groove (Figure 35C). *Tg*SporoAMA1 loop 4 shows clear mobility in the crystal structure, and combined with the flexibility inherent to the C-terminal sequence of *Tg*SporoRON2D3 (CVVVAGSGS) suggests that two conformations of the complex likely exist in solution: first, AMA1 loop 4 is directed towards the groove center and the RON2D3 C-terminus exits the groove, and second with AMA1 loop 4 displaced from the groove center and the RON2D3 C-terminal sequence threading back through the groove, as seen for generic RON2sp. This rationale is supported by a comparison with the two previously determined structures of *Pf*AMA1 in complex with *Pf*RON2 peptides of different lengths; in the *Pf*AMA1-*Pf*RON2sp1 co-structure, *Pf*AMA1 loop 4 is displaced from the groove and the *Pf*RON2sp1 C-terminal region threads back through the groove similar to generic RON2sp, while in the *Pf*AMA1-*Pf*RON2sp2 co-structure, *Pf*AMA1 loop 4 is shifted about 4 Å towards the groove center and the two post-cystine loop residues modeled appear to exit the groove in the same fashion as

*Tg*SporoRON2D3 (Vulliez-Le Normand, Tonkin et al. 2012). In addition, the way that *Tg*SporoRON2D3 forms contacts on both edges of the groove (Met1009 and Lys1010; Figure 35C) may explain why it does not require coordination of the post-cystine loop region for binding, and provides an interesting comparison with the recently determined *Pf*AMA1-R1 structure, where trans-groove contacts participate in coordination of the linear R1 peptide (Vulliez-Le Normand, Tonkin et al. 2012).

Localization of TgSporoAMA1 and TgSporoRON2 in sporozoites.

In tachyzoites, it has been shown that generic AMA1 resides in the micronemes (Alexander, Mital et al. 2005) and generic RON2 is localized in the rhoptry necks (Bradley, Ward et al. 2005, Tyler and Boothroyd 2011). To determine the approximate localization of *Tg*SporoAMA1 in the sporozoite stage, we raised antibodies to the *Tg*SporoAMA1 ectodomain (amino acids 79-569) in mice. Extracellular sporozoites were placed on glass slides, fixed, permeabilized and stained with antibodies to MIC10 and *Tg*SporoAMA1 (Figure 36A). In extracellular sporozoites, *Tg*SporoAMA1 localizes to the apical end of the parasite, similar to the pattern seen with MIC10 and similar to what has been observed in tachyzoites (Brydges, Sherman et al. 2000, Hoff, Cook et al. 2001). Attempts to stain unpermeabilized sporozoites gave highly variable patters that were not considered reliable (data not shown). We then examined *Tg*SporoAMA1 localization in intracellular sporozoites. Freshly invaded sporozoites were fixed and permeabilized followed by incubation with anti-sporoAMA1 polyclonal sera and with anti-MIC10 and anti-MIC5, another protein that has been localized previously to the micronemes of tachyzoites (Brydges, Sherman et al. 2000, Hoff, Cook et al. 2001), although like MIC10, it has not been studied in sporozoites (Figure 36B-D). Although all three markers showed consistent apical staining, no significant co-localization was observed between *Tg*SporoAMA1 and either of the other two markers, with *Tg*SporoAMA1 localizing distally to the extreme apical location of the two MIC markers.
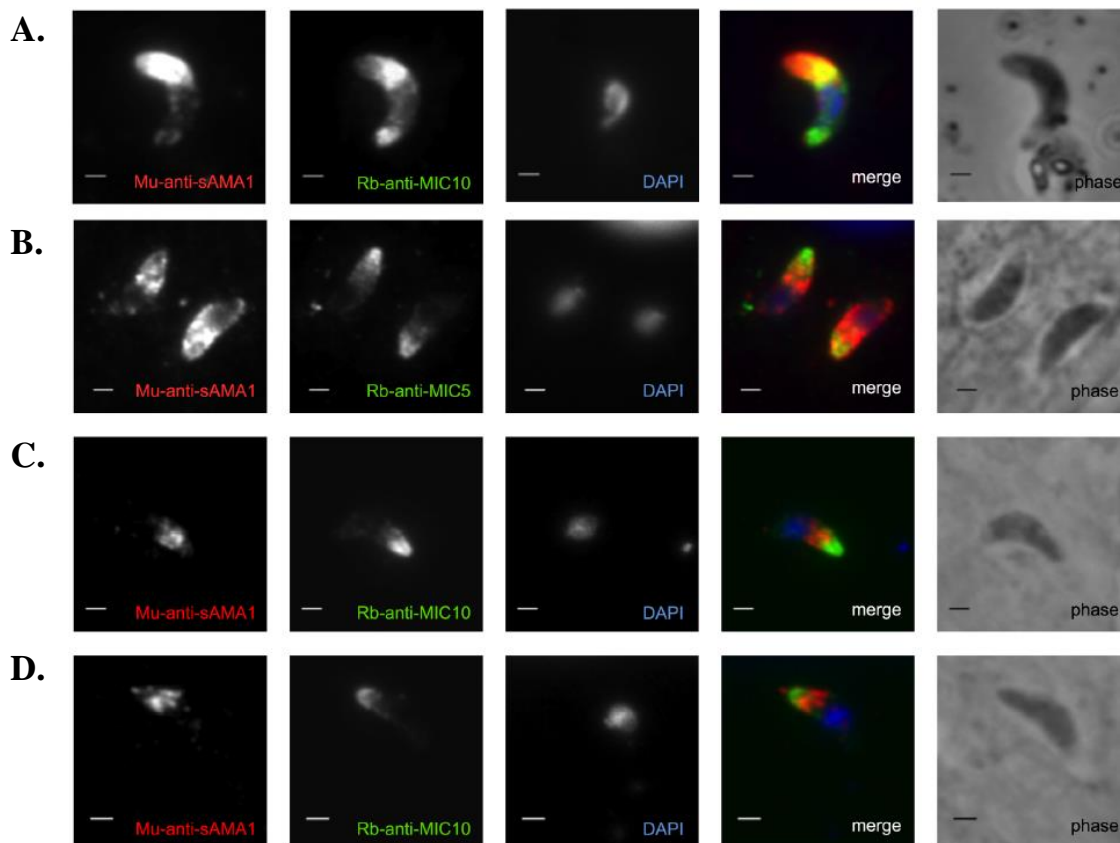
Figure 36: *Tg*SporoAMA1 localizes apically in sporozoites.

Extracellular sporozoites (**A**) or HFF monoloayers infected with sporozoites for 2-3 hours (**B-D**), were formaldehyde-fixed, permeabilized with triton X-100, and stained with mouse anti-*Tg*SporoAMA1 (Mu-anti-sAMA1) and rabbit (Rb) anti-MIC5 (**B**) or anti-MIC10 (**A**, **C** and **D**). Images shown in **C** and **D** are for adjacent parasites in the same field that were too far apart to capture in one image; both are shown to convey the reproducibility of the pattern observed. Scale bars represent 2 μm.

To determine localization of *Tg*SporoRON2, we raised rabbit antibodies to domain 4 (amino acids 1069-1167) of *Tg*SporoRON2 that was fused to GST. Freshly invaded sporozoites were fixed in methanol and stained with anti-*Tg*SporoRON2 polyclonal sera. The results were compared with the localization of either RON4 (a rhoptry neck protein detected in tachyzoites and the proteome of sporozoites) (Figure 37A) or ROP2/3/4 (rhoptry bulb proteins also abundantly detected in proteomes of both stages) (Figure 37B) (Bradley, Ward et al. 2005). Surprisingly, *Tg*SporoRON2 exhibited no colocalization with

RON4 but partial colocalization with ROP2/3/4. No colocalization was found between *Tg*SporoAMA1 and *Tg*SporoRON2 (Figure 37C). Although we cannot exclude the possibility that the antibody is detecting an unrelated sporozoite protein, they were raised to a relatively short portion of *Tg*SporoRON2 (~100 amino acids) that has no significant similarity to any other predicted protein in *T. gondii* and the antibodies were shown to not react to generic RON2 (or any other protein) in tachyzoites (data not shown). Unfortunately, we were not able to obtain enough sporozoites for studies by immune-electron microscopy, and there are no well-studied markers for sporozoite organelles. Hence, we can draw no definitive conclusions about the exact localization of any of these proteins but the results seen suggest substantial differences from the paradigms developed with generic AMA1 and RON2 in tachyzoites. Note that despite several attempts, we were unable to identify sporozoites that were unambiguously in the process of invasion and so cannot comment on which, if either, of the AMA1/RON2 pairs described here is at the MJ formed by invading sporozoites.
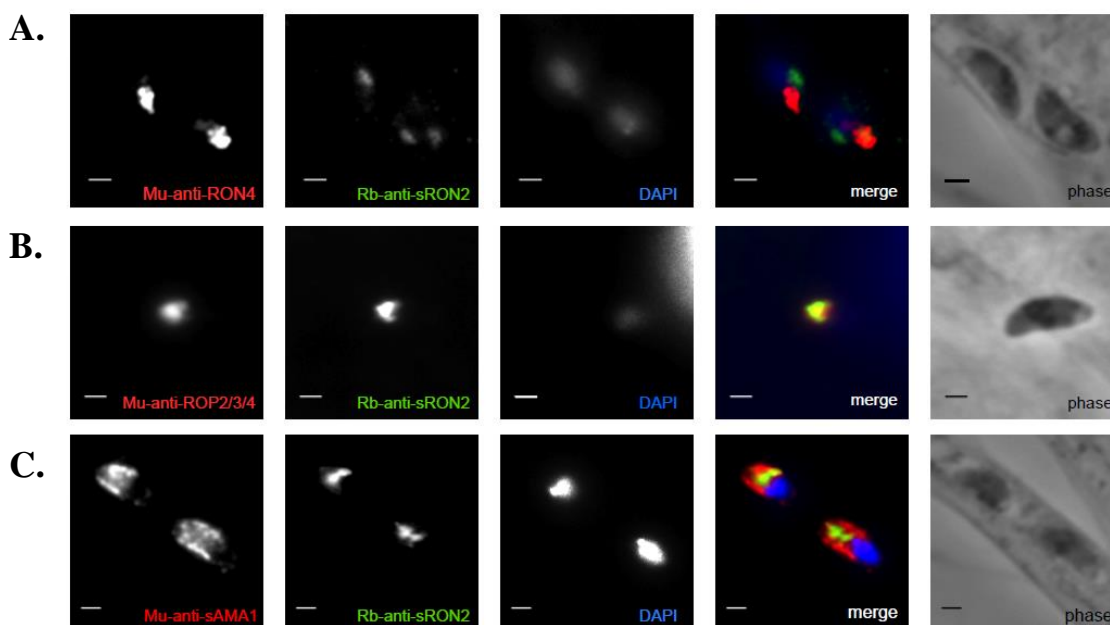

Figure 37: *Tg*SporoRON2 shows partial colocalization with ROP2/3/4 but little if any with RON4.

Infected HFF monolayers were infected with M4 sporozoites for 2-3 hours, and then were methanol-fixed, and stained with rabbit anti-*Tg*SporoRON2 (Rb-anti-sRON2) and either

mouse (Mu) anti-RON4 (**A**), anti-ROP2/3/4 (**B**), or anti-*Tg*SporoAMA1 (**C**). Images where two parasites were present in the same field are shown except for (B) where no such fields were found. The image shown in (B), however, is representative of the pattern consistently observed with these two antibodies. Scale bars represent 2 μm.

Treatment of sporozoites with TgSporoRON2D3 impedes sporozoites invasion.

Blocking the generic AMA1/RON2 interaction by incubating *T. gondii* tachyzoites or *P. falciparum* merozoites with generic RON2 domain 3 inhibits invasion (Lamarque, Besteiro et al. 2011, Srinivasan, Beatty et al. 2011, Tyler and Boothroyd 2011). To test whether the generic AMA1/RON2 complex and/or the sporozoite AMA1/RON2 complex play a role in sporozoite invasion, we performed invasion-inhibition assays of sporozoites and tachyzoites in the presence of equimolar amounts of either GST-gD3, GST-sD3, both or GST alone. Extracellular vs. intracellular parasites were identified by sequential staining with polyclonal rabbit anti-*T. gondii* antisera before detergent permeabilization of host cells, and with mouse anti-SAG1 sera after permeabilization. This allowed us to discriminate successful from blocked invasion events. The results (Figure 38A) showed that treatment of tachyzoites with GST-gD3 significantly decreases the efficiency of invasion, as previously reported (Tyler and Boothroyd 2011), while GST-sD3 has no such effect, as predicted from our biochemical and structural studies.
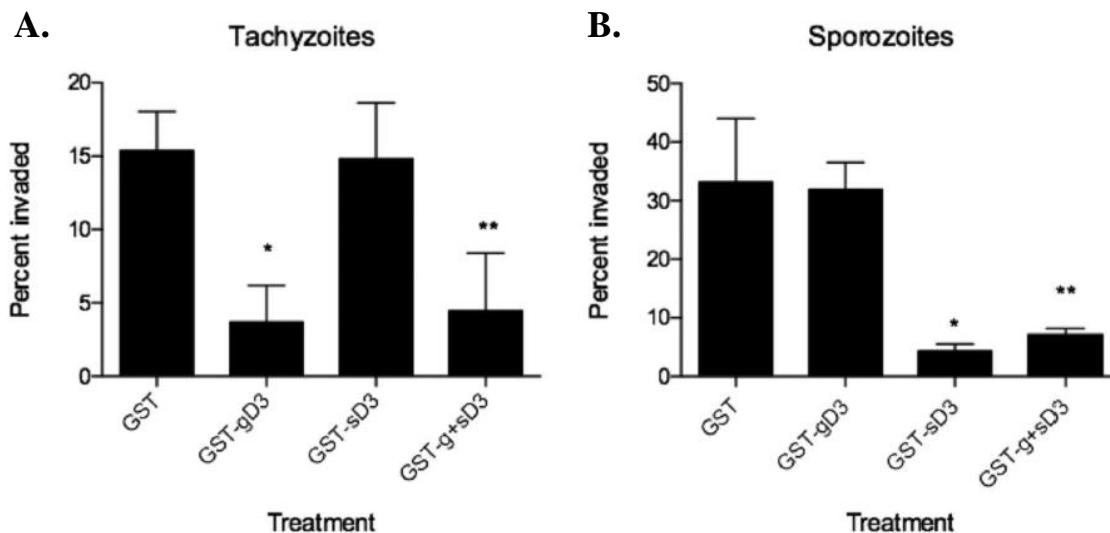
Figure 38: Preincubation of parasites with GST-*Tg*SporoRON2-D3 specifically impedes sporozoite but not tachyzoite invasion.

Tachyzoites (**A**) and sporozoites (**B**) were pretreated with molar equivalents of GST, GST-gD3, GST-sD3, and a mixture of GST-gD3/GST-sD3, and then permitted to invade a monolayer of HFFs for 45 minutes, following temperature synchronization. The number of intracellular parasites was determined by differential staining before and after permeabilization. The percent of invaded (intracellular) parasites relative to the total number was determined by counting parasites in ten randomly selected fields from each of three coverslips for each condition. The counting and analysis were done blinded. A single asterisk indicates $p < 0.05$ for the difference relative to the GST control; double asterisks indicate $p < 0.01$ relative to this control.

Interestingly, even though the generic AMA1/RON2 pair is expressed in sporozoites, treatment with GST-gD3 had no effect on sporozoite invasion. Treatment with GST-sD3, however, resulted in a significant decrease (~88%) in the invasion rate (Figure 38B). Sporozoites treated with both GST-gD3 and GST-sD3 were also markedly decreased in their invasion efficiency although to a slightly lesser degree (~70%), likely due to the fact that there was only half as much GST-sD3 in that fraction compared to the GST-sD3 alone (to keep total protein added equal). These results indicate that the interaction of *Tg*SporoAMA1 and *Tg*SporoRON2D3 is important for successful invasion of host cells by

sporozoites, while the interaction of generic AMA1/RON2 appears dispensable, at least in these conditions.

4.5 Discussion

The results presented here demonstrate that invasion by *T. gondii* sporozoites depends on the interaction of two previously uncharacterized proteins, *Tg*SporoRON2 and *Tg*SporoAMA1. These proteins form a specific pairing, similar to the previously described interaction of generic AMA1 and generic RON2 (Tonkin, Roques et al. 2011, Tyler and Boothroyd 2011). This finding begs the question as to why the sporozoites evolved to use a novel pairing rather than the generic counterparts. Homologues of *Tg*SporoRON2/*Tg*SporoAMA1 are found in *Neospora* and *Eimeria*, but are missing in *Theileria* and *Plasmodium* species. Unlike their Haemospororina relatives, which are insect-transmitted, sporozoites in the Eimeriorina suborder have to cross the intestinal epithelium of their intermediate hosts to initiate a new infection. SporoRON2 and sporoAMA1, therefore, may represent an effective solution to that unique challenge. The mammalian gut presents a formidable barrier to invading microorganisms with defenses ranging from the unique morphology of enterocytes (extensive microvilli and a robust cytoskeleton) to the presence of abundant mucus and a somewhat alkaline pH (Camilleri, Madsen et al. 2012, Hansson 2012). It has been previously reported that *T. gondii* sporozoites may cross the intestinal epithelial cells before replicating in the cells of lamina propria (Barragan and Sibley 2002, Barragan, Brossier et al. 2005), and it may be that the sporoAMA1/sporoRON2 pairing allows the sporozoites of Eimeriorina species to cross this barrier more efficiently than the generic AMA1/RON2 pair. If crossing the intestinal epithelium is the major explanation for the presence of sporoAMA1/sporoRON2, one might expect *Toxoplasma* bradyzoites, which are also orally infectious, to deploy this pairing. Based on transcriptomic analyses, however, bradyzoites do not appear to express either of these genes (Fritz, Buchholz et al. 2012).

Our data do not address why generic versions of AMA1/RON2 are also present in sporozoites. It has been previously reported that, upon invasion, sporozoites form an unusually large primary vacuole, in which the parasite remains for the first 12-18 hours (Speer, Tilley et al. 1995, Tilley, Fichera et al. 1997). Following this, the parasite creates

another MJ and forms a secondary vacuole, in which it then replicates (note that while we did not observe any particularly spacious (i.e. primary) vacuoles, we did not follow invasion by video microscopy and so may have missed this two-step process if under our conditions primary vacuoles are not readily distinguishable from secondary vacuoles). Such a two-step process could be another way in which the two AMA1/RON2 pairings operate: e.g., the sporozoite-specific pairing might play a role in the first step (as argued by our invasion inhibition studies) while the generic pair operates in the more tachyzoite-like formation of the secondary vacuole. Finally, it might also be that the generic and sporozoite-specific AMA1/RON2 pairs are secreted under different invasion conditions; this possibility cannot be excluded as we have looked only *in vitro* and only with fibroblasts as the target host cell.

Previous studies have also postulated that the generic AMA1/RON2 interaction is involved in a signaling event that provides information to the intracellular environment upon formation of the MJ complex (Treeck, Zacherl et al. 2009, Tonkin, Roques et al. 2011, Tonkin, Crawford et al. 2013). While AMA1 DI and DII are both involved in coordinating RON2, and a significant conformational change occurs with the displacement of the DII loop, this signal would still have to be passed through DIII in order to reach the cell interior. A possible mechanism for this signaling is observed when the DIII positioning of apo and bound *Tg*SporoAMA1 is compared (Figure 39). In the apo structure, the highly compact cystine-knot containing DIII of sporoAMA1 extends across the base of DI and DII (Figure 39 left), similar to the conformations observed for DIII in the generic AMA1 of *Toxoplasma* and of *Neospora* and *Babesia* (Crawford, Tonkin et al. 2010, Tonkin, Crawford et al. 2013). However, in the *Tg*SporoAMA1-*Tg*SporoRON2D3 co-structure reported here, *Tg*SporoAMA1 DIII is rotated away from the DI/DII core by about 90°, displacing the C-terminus by more than 40 Å (Fig. 39 right). It is noteworthy that only the compact three-domain organization is observed in the structural studies of generic AMA1. It is tempting to speculate that this conformational change at the base of *Tg*SporoAMA1 is tied to the role of DIII as the conduit for a signal moving from the ectoplasmic region of AMA1 through to the intracellular domain. While a signaling role mediated by an articulating DIII is speculative, our structural data clearly demonstrates the inherent ability of DIII to adopt structurally distinct conformations. It is also worth noting that DIII of

sporoAMA1 is the least conserved portion of the protein relative to its generic paralogue with an extended Pro-Glu linker between the cystine knot and the transmembrane domain (displayed as a dashed line in Figure 39). Since this linker is not included in the structurally characterized *Tg*SporoAMA1, it is possible that the complete ectodomain tethered to the parasite cell surface might preferentially stabilize one of the two DIII orientations observed, independent of *Tg*SporoRON2D3.
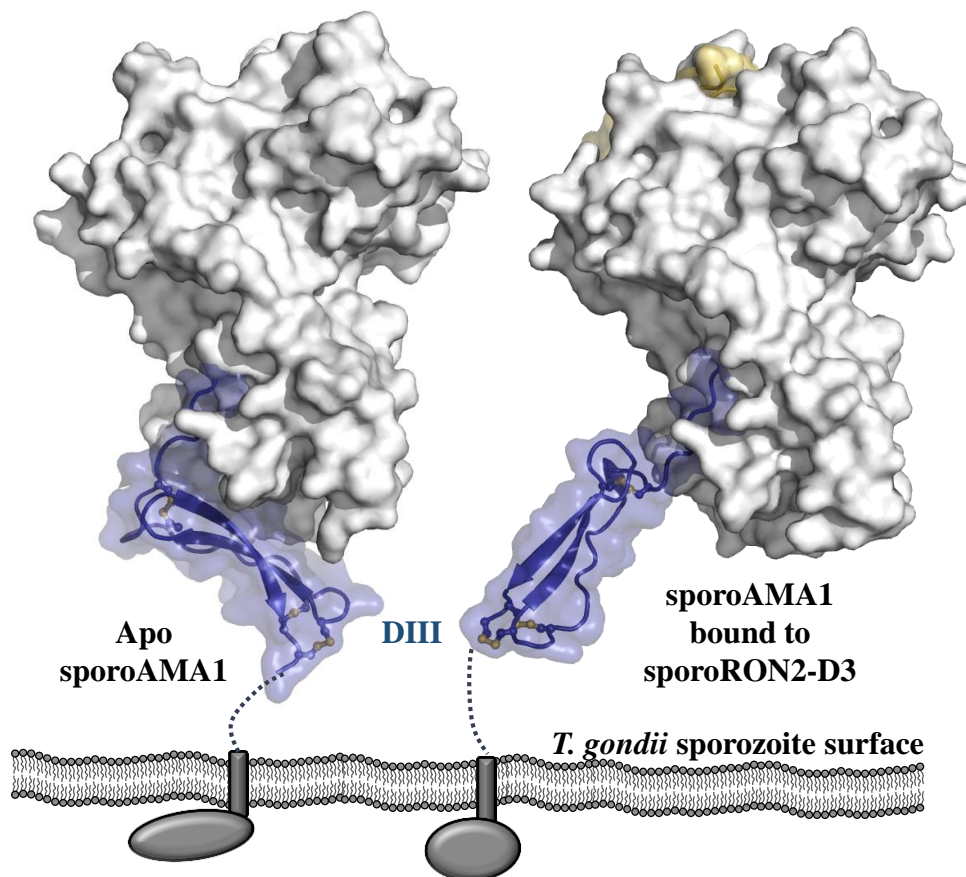
Figure 39: *Tg*SporoAMA1 DIII reorganization upon ligand binding provides possible insight into signal transduction mechanisms.

*Tg*SporoAMA1 shown in predicted organization to the *T. gondii* sporozoite plasma membrane, with DI and DII shown as a grey surface and DIII shown as blue cartoon with a semi-transparent blue surface, *Tg*SporoRON2D3 shown as a gold surface, and disulfides shown as yellow sticks. Dotted lines indicate extended Pro/Glu rich region between the conserved portion of DIII and the transmembrane domain (grey rectangle) that leads through to the C-terminal domain (grey oval/sphere). Left – Apo *Tg*SporoAMA1. Right – *Tg*SporoAMA1 bound to *Tg*SporoRON2D3.

Our attempts to localize *Tg*SporoAMA1 and *Tg*SporoRON2 in sporozoites showed the expected apical staining but surprising differences with other micronemal and rhoptry markers. Sporozoites have not been well studied because of the difficulty of working with them and so no well described, "gold standards" exist for the various organelles. It is known, however, that sporozoites have far more micronemes than tachyzoites and it could be that there are different subpopulations of micronemes with distinct cargoes in the sporozoites. The lack of colocalization between *Tg*SporoRON2 and RON4 is more surprising and not easily explained. There is no *a priori* reason why a paralogue of RON2, which in tachyzoites colocalizes with RON4 in the rhoptry necks, must also go to this compartment in the sporozoites. The localization data suggest that *Tg*SporoRON2 might function independently of RON4 and, again, there is no *a priori* reason why this should not be the case. As mentioned above, the presumptive signaling that occurs when *Tg*SporoRON2 binds *Tg*SporoAMA1 could well operate independent of RON4 whose interaction with RON2 and overall role in the MJ complex has not been elucidated.

Taken together, our data provide new insight into the AMA1/RON2 interaction and show that evolution has favored a structurally distinct pairing in the form of sporoRON2/sporoAMA1 in *Toxoplasma* sporozoites and, we presume, sporozoites of the other Eimeriorina where orthologues for these genes are observed. This provides further evidence that this interaction is critical for apicomplexan invasion of host cells although the enigma remains of why the generic AMA1/RON2 pair is expressed in sporozoites if not necessary for their invasion.

Chapter 5: Discussion

5.1 Summary of research objectives

The organisms in phylum Apicomplexa cause devastating diseases on a global scale, leading to over a million human deaths every year. To enable their widespread success, these obligate intracellular parasites have evolved a number of phylum-specific proteins that facilitate novel biological processes. One important example is the intricate multi-step process of host cell invasion. Across the phylum, numerous different strategies have evolved to enable recognition of and attachment to host cells (Cowman and Crabb 2006). In contrast, a refined mechanism of active host cell invasion is broadly conserved: a circumferential ring of adhesion, termed the moving junction (MJ), is assembled at the apical tip of the parasite where it contacts the host cell, and the parasite dives through this ring into the nascent parasitophorous vacuole. The only identified extracellular bridging proteins at the MJ are apical membrane antigen 1 (AMA1) and rhoptry neck protein 2 (RON2). Prior to this dissertation, only low resolution insight into the assembly of AMA1 and RON2 was available, representing a clear lack in the knowledge required to elucidate how AMA1 coordinates RON2 to structure the essential MJ complex at the parasite-host cell interface. In the last five years a significant amount of work has resulted in a blossoming of our understanding of the detailed molecular mechanisms of apicomplexan parasite invasion, a small portion of which has been presented in this dissertation. The work presented in the preceding three chapters targeted the first two research objectives by developing and refining a high-resolution binding paradigm for the molecular recognition event between AMA1 and RON2. Detailed insight into the assembly of AMA1 and RON2 from *P. falciparum* merozoites and *T. gondii* sporozoites broadened this paradigm to define regions at the interface that are hot spots for molecular organization as well as the regions governing cross-genera, cross-species, and intra-species specificity. The aim of this discussion chapter is to investigate how detailed knowledge of the AMA1-RON2 interaction can enrich our understanding of apicomplexan biology, and to address the final research objective by assessing the relevance of AMA1-RON2 structural information to guiding rational small molecule drug design or enhancing vaccine efficiency.

5.2 Enhancing our understanding of apicomplexan biology through high resolution analyses of AMA1-RON2 complexes

The highly orchestrated process of apicomplexan active host cell invasion is fascinating on a number of levels, with MJ formation and progression being one of the most intriguing steps. The parasite and host cell membranes move past each other but do not fuse, as they are held in close apposition by a dense assembly of receptors (AMA1) on the parasite surface and cognate ligands (RON2) presented on the host cell surface; this AMA1-RON2 structured MJ ring must withstand the shearing forces of invasion as the parasite literally squeezes through the circumferential constriction. Furthermore, given the essential nature of active invasion to their obligate intracellular lifestyle, the apicomplexan parasites must maintain the AMA1-RON2 interaction in different invasion environments and when under attack by the host immune system. High resolution insights into the mechanism of AMA1-RON2 assembly and further research inspired by these insights increase our understanding of how this novel pairing facilitates the unique biological process of apicomplexan active host cell invasion.

AMA1-RON2 assembly is an intimate molecular recognition event

A structural perspective on the AMA1-RON2 complex in the context of other protein-protein mediated pathogen invasion mechanisms clearly displays the extent to which AMA1 has been optimized for RON2 binding. Among the diverse mechanisms employed by viruses, pathogenic bacteria, and protozoan parasites to enter host cells, one commonality is the utilization of protein complexes capable of promoting a sufficiently intimate interaction to enable access of the pathogen to the host cell cytoplasm (Figure 40). Within the last several years there has been a dramatic increase in our understanding of viral-based structural solutions to mediate host cell adhesion and subsequent membrane fusion; the collection of co-structures reveals a variety of molecular recognition events, with two of the largest interface areas represented by Nipah virus attachment glycoprotein complexed with human Ephrin B2 (Buried Surface Area (BSA): 2786 Å^2) (Bowden, Aricescu et al. 2008) and Human immunodeficiency virus (HIV) gp120 complexed with human CD4 (BSA: 2051 Å^2) (Diskin, Marcovecchio et al. 2010). In terms of bacterial

mechanisms of protein-protein mediated host cell entry, although Enteropathogenic *E. coli* (EPEC) do not enter the host cell, they bind tightly to the surface and induce cytoskeletal changes using an elegant mechanism whereby the EPEC translocated intimin receptor (Tir) is injected into the host cell and links to intimin on the bacterial cell surface (BSA: 1294 $Å^2$) (Luo, Frey et al. 2000). In contrast, *Listeria monocytogenes*, an intracellular pathogen, utilizes Internalin A to envelop E-cadherin on the human host cell surface (BSA: 2630 $Å^2$) (Schubert, Urbanke et al. 2002) and initiate its own phagocytosis. The characterization of *Tg*AMA1-*Tg*RON2sp represents the first atomic level characterization of a host-pathogen protein bridge from a protozoan parasite, and the formation of this complex results in the most extensive buried surface area (3765 $Å^2$) of reported host-pathogen binary protein complexes. In the context of other host-pathogen adhesion and invasion complexes, it is clear that AMA1 and RON2 have co-evolved for a biological function that requires their intimate interaction, which clarifies the ability of this complex to structure the constricted MJ ring.

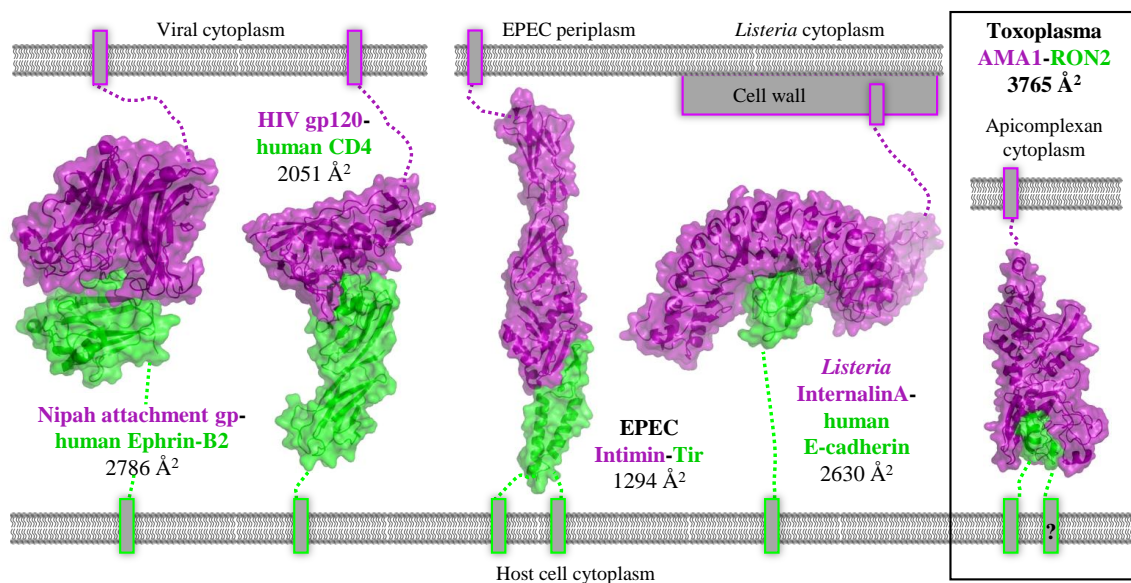**Figure 40: A novel host-pathogen protein ligand-receptor complex.**
Pathogen surface proteins (left – viral, center – bacterial, and right - protozoan) are shown in purple and proteins presented on the host cell shown in green. Buried surface areas are shown below the labels and indicate the combined contribution of surface area from both the pathogen and host cell surface molecules. Note that EPEC Tir is likely a homodimer,

with each Tir individually binding one intimin receptor. Protein Data Bank (PDB) IDs left to right: 2VSM, 3LQA, 1F02, 1O6S, 2Y8T.

AMA1-RON2 biological importance is supported by versatile assemblies

Two recent studies on the biological roles of AMA1 and the RONs during invasion using new genetic manipulation techniques led the authors to conclude that these proteins have separate roles during invasion and that AMA1 is completely dispensable for active host cell invasion; in particular, when AMA1 is knocked down, 10% of parasites are still able to invade in a RON2/RON4 MJ ring dependent mechanism (Giovannini, Spath et al. 2011, Bargieri, Andenmatten et al. 2013). These studies raise the question of why some data provides evidence for uncoupled roles of AMA1 and the RONs during invasion and for a non-essential role of AMA1, while a significant amount of data, including the structural data presented in this dissertation, supports an essential and combined role. In order to answer this question, it is important to take into account that apicomplexans are sophisticated parasites that may have evolved redundant mechanisms for invasion. By investigating at a deeper level how the parasites behave when AMA1 is knocked down, we have recently identified the AMA family of proteins in *T. gondii*, consisting of AMA1, AMA2, AMA3 (sporoAMA1), and AMA4 (Figure 41) (Lamarque, Roques et al. 2014). Despite just 32% sequence identity between *Tg*AMA1 and *Tg*AMA2, in *T. gondii* tachyzoites *Tg*AMA2 is able to serve as a functional analogue for *Tg*AMA1 by binding to *Tg*RON2 (generic RON2) and enabling invasion in the absence of *Tg*AMA1. While we have only shown the functionality of *Tg*AMA3 (*Tg*SporoAMA1) binding to *Tg*RON2$_{L2}$D3 (*Tg*SporoRON2) in sporozoites (Poukchanski, Fritz et al. 2013), it is possible that this complex could be upregulated in the absence of *Tg*AMA1-*Tg*RON2 and/or *Tg*AMA2-*Tg*RON2. *Tg*AMA4 is only 20% identical to *Tg*AMA1 and the ectodomain is predicted to be more than twice the size of the *Tg*AMA1 ectodomain, with a significant portion of this extra sequence centered on the DI-like domain and epidermal growth factor (EGF) domains between the end of DII and the transmembrane domain (Figures 41). We have shown that *Tg*AMA4 is upregulated in the absence of *Tg*AMA1 and capable of binding a third *T. gondii* RON2 paralogue (*Tg*RON2$_{L1}$) (Lamarque, Roques et al. 2014). Interestingly, *Tg*AMA4 is a homologue of the merozoite apical erythrocyte-binding ligand (MAEBL)

proteins in *Plasmodium* species that contain a tandem pair of AMA1 DI-like and DII domains in the ectoplasmic region (Chesne-Seck, Pizarro et al. 2005). While *T. gondii* and other coccidian parasites appear to have more AMA and RON2 family members than *Plasmodium* species (Besteiro, Dubremetz et al. 2011, Oakes, Kurian et al. 2013), if MAEBL is capable of binding RON2, this would provide evidence for why *Plasmodium* AMA1 can be compensated for during MJ-mediated invasion of host cells (Giovannini, Spath et al. 2011). In *T. gondii*, *Tg*AMA1, *Tg*AMA3, *Tg*RON2, *Tg*RON2$_{L1}$, and *Tg*RON2$_{L2}$ have all been shown to be expressed in the sporozoite stage suggesting that the parasites maintain multiple AMA-RON2 options to ensure structuring a successful MJ (Fritz, Buchholz et al. 2012, Possenti, Fratini et al. 2013). Intriguingly, AMAs most commonly expressed only in the sporozoite stage of both *Plasmodium* and *Toxoplasma* have long extensions between the end of DII and the transmembrane domain (between 100 and 700 amino acids) (Figure 41). These extensions may impact the width and flexibility of the established MJ, and may also play a role in establishing the MJ in unique host cell conditions such as the mosquito salivary glands (*Pf*MAEBL (Kariu, Yuda et al. 2002)) or the intestinal tract (*Tg*AMA3 (Poukchanski, Fritz et al. 2013).
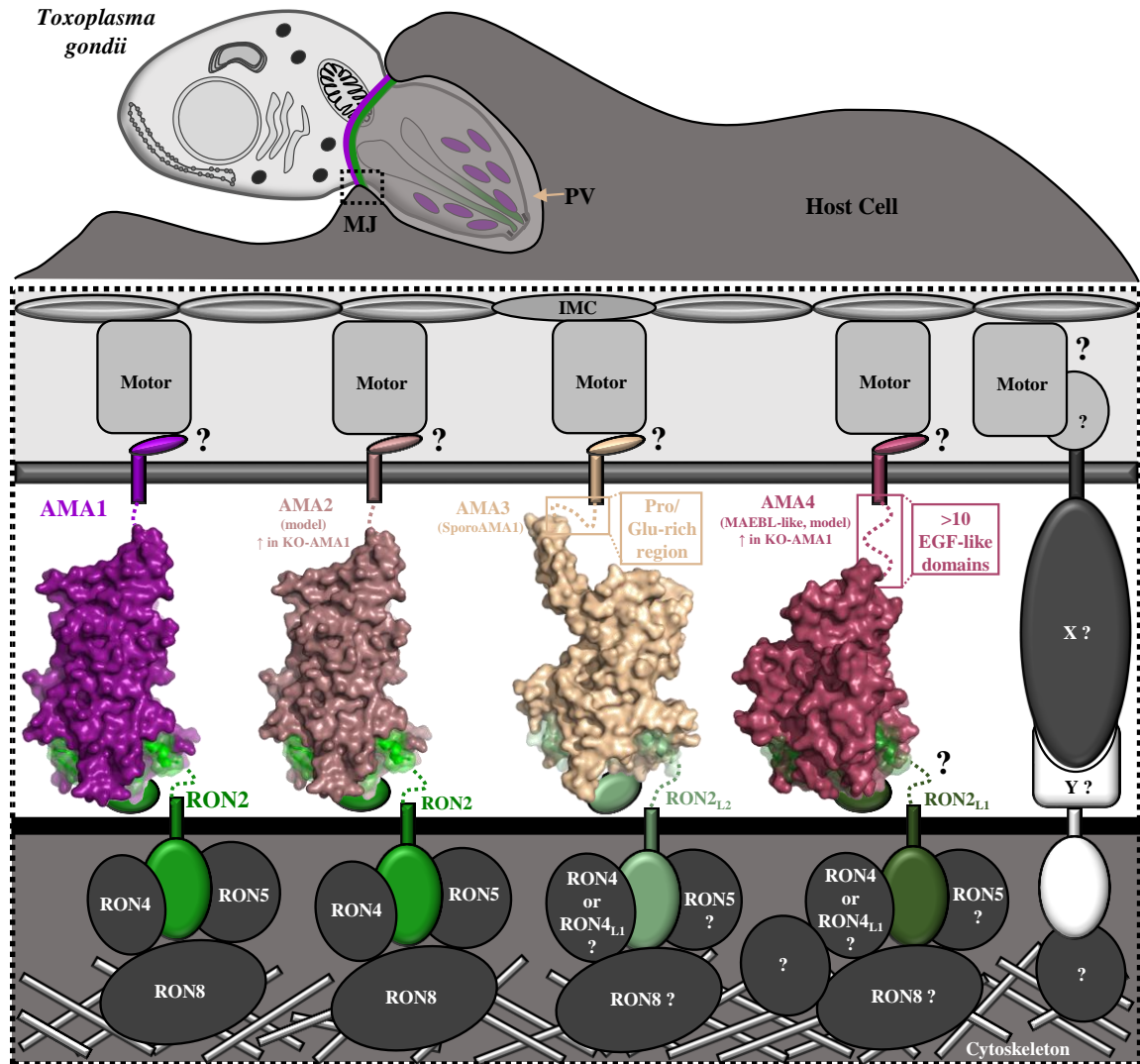
Figure 41: A refined model of the *T. gondii* MJ components showing plasticity and redundancy in the AMA-RON2 extracellular link.

Top – Schematic of *T. gondii* parasite invading a host cell. Bottom – Zoom-in of the composition of proteins with the ability to structure the MJ in different life cycle stages of *T. gondii*. IMC, Inner Membrane Complex. X and Y indicate possible currently unidentified proteins with the capability to structure the MJ. AMA2 and AMA4 are homology models based on *Tg*AMA1 (Lamarque, Roques et al. 2014).

The identification of AMA and RON2 families, and more importantly of diverse compensatory AMA-RON2 assemblies, highlights the biological importance of the AMA-RON2 pairing; this ensemble of AMA-RON2 pairings likely enables invasion across an

array of life cycle stages and cellular niches, while also providing parasites under attack by the immune system with alternate pathways for an invasion event that is essential for their survival as obligate intracellular parasites. Overall, revealing the redundant and compensatory AMA-RON2 based mechanisms for MJ dependent invasion in the resourceful apicomplexan parasites reinforces targeting AMA-RON2 complexes for the design and development of invasion inhibitory therapeutics.

5.3 Assessing the relevance of AMA-RON2 structural data to the development of broadly reactive therapeutics

Many of the current therapeutics targeting apicomplexan parasites are problematic due to the lack of broad-range efficacy and emergence of drug-resistant parasite forms, clearly demonstrating the need for development of novel therapeutics (Section 1.3). Given the obligate intracellular nature of the apicomplexan parasites, inhibiting the parasites' ability to recognize, attach to, or invade host cells through the MJ before they are sequestered in the intracellular environment presents attractive novel targets for therapeutics. The only widely conserved step in this invasion sequence at the molecular level is MJ formation and progression, where AMA and RON2 family members form binary complexes that are essential for host cell invasion. From a reductionist experimental approach, we now have high-resolution insight into the detailed mechanism of AMA-RON2 assembly, but can this structural insight be translated into a form that is useful for guiding or supporting the development of small molecule drugs or for enhancing vaccine efficiency?

The intricacy of the apicomplexans strongly argues against a single protein approach to disease prevention or treatment, but the inclusion of an AMA-directed therapeutic that broadly prevents RON2 binding would be highly beneficial (Macraild, Anders et al. 2011, Kortagere 2012). Targeting the AMAs is generally considered to be more feasible than RON2s, as AMAs are presented on the parasite surface for longer periods of time, whereas RON2s are only minimally exposed to the environment after translocation to the host cell and before binding an AMA. Accordingly, regions on the surface of AMA proteins that are both conserved across the phylum and required for RON2 coordination could provide excellent targets for the development of broadly reactive small molecule competitive inhibitors and vaccines (Lim, Debono et al. 2013).

Prior to the start of this work, despite the progress made in identifying the components of the MJ, a structural map of the critical extracellular components was lacking, limiting potential insight into the mechanism of assembly that might be able to guide the development of therapeutics. With the characterization of *Tg*AMA1-*Tg*RON2sp, *Tg*AMA3-*Tg*RON2$_{L2}$D3 (*Tg*SporoAMA1-*Tg*SporoRON2D3), and *Pf*AMA1-*Pf*RON2sp reported in this dissertation, combined with the previously determined AMA1 structures [*Pf*AMA1 (Bai, Becker et al. 2005), *Pv*AMA1 (Pizarro, Vulliez-Le Normand et al. 2005), *Tg*AMA1 (Crawford, Tonkin et al. 2010)], and the structures of AMA1 from *Babesia divergens* (*Bd*AMA1) and *Neospora caninum* (*Nc*AMA1) (Tonkin, Crawford et al. 2013), a wealth of data is now available to identify regions of the AMA-RON2 interface that are conserved and required for complex formation. Only limited cross-reactivity has been shown between AMA1 from one genus and RON2 from another (Tonkin, Crawford et al. 2013), but the recent determination of the cross-reactive complex of *Tg*AMA1 with *Eimeria tenella* RON2D3 (*Et*RON2D3) provides the first structural insight into the basis for plasticity enabling cross-genera reactivity (Tonkin, Brossier et al. In prep). In addition, the structure of *Pf*AMA1 with the invasion inhibitory R1 peptide provides insight into how non-native synthetic peptides might bind and occupy the AMA apical groove (Vulliez-Le Normand, Tonkin et al. 2012).

Viewing all of the co-structures of AMA in complex with different molecules, a few regions of interest start to emerge as the most commonly occupied by residues with similar properties (Figure 42). The RON2 N-terminal helix shows variable modes of interaction with AMA surface loops, and the tip of the cystine loop clearly shows the most diverse array of binding interactions. Burial into the most groove-central pocket exposed by removal of the DII loop is a common anchor point (*Tg*RON2sp Pro1309, *Et*RON2D3 Pro1274, *Tg*RON2$_{L2}$ Lys1010, *Pf*RON2sp1 Pro2033, R1-major Leu-P6) that is likely required to maintain the DII loop in its displaced conformation (Figure 42, box 1); although proline seems to be preferred, there is a considerable level of plasticity observed, as the same region can accommodate proline, leucine, lysine, and even tryptophan (apo *Tg*AMA1 and *Nc*AMA1 structures show a DII loop Trp buried in this same pocket (Crawford, Tonkin et al. 2010, Tonkin, Crawford et al. 2013)). Most of the co-structures show the ligand bracing against the back wall of the groove; *Tg*RON2sp Ile1328, *Et*RON2D3 Val1293,

*Pf*RON2sp1 Val2054, and R1-major Leu-P8 all occupy the same shallow surface pocket (Figure 42, box 2), but *Tg*RON2$_{L2}$ does not coordinate in this region as the C-terminal portion of the peptide exits the groove. A third clear region of structural convergence is the disulfide bond (*Tg*RON2sp Cys1313 and Cys1323, *Et*RON2D3 Cys1278 and Cys1288 [partial occupancy disulfide], *Tg*RON2$_{L2}$ Cys1015 and Cys1026, *Pf*RON2sp1 Cys2037 and Cys2049), which is seated in a crevice between the same two AMA loops in nearly the same position in each AMA-RON2 co-structure (Figure 42, box 3). Often, a hydrophobic residue immediately following the first cysteine lifts the disulfide away from the base of the groove (*Tg*RON2sp Val 1314, *Et*RON2D3 Leu1279, *Pf*RON2sp1 Phe2038); while the linear R1 peptide obviously does not maintain a disulfide bond in this position, it does still occupy the base of the groove with a phenylalanine residue (Phe-P12) (Figure 42, box 3). Overall, significant diversity can be found in how different molecules access the apical groove of AMA, but there are some common themes in three central groove regions that are consistently occupied.
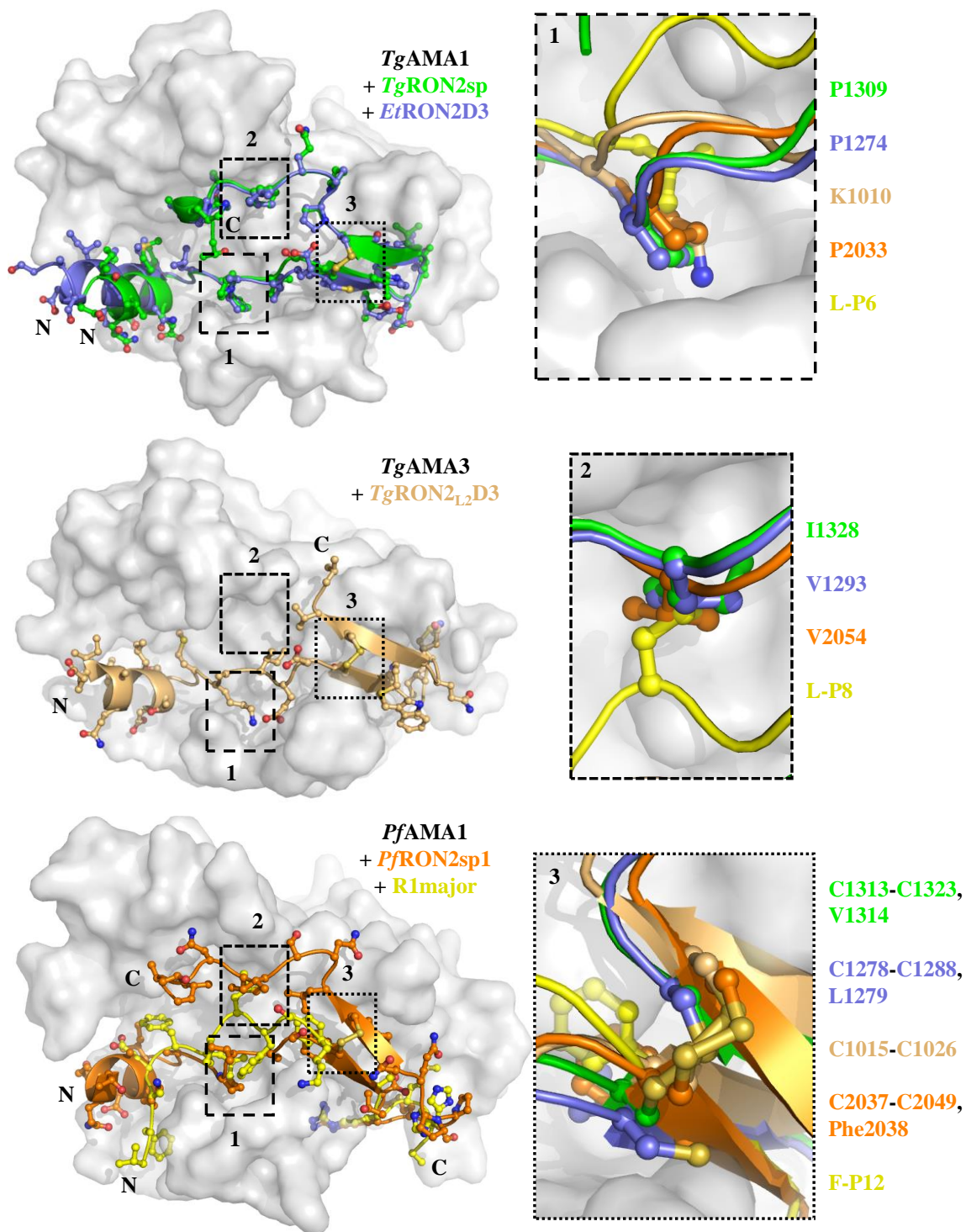
Figure 42: Occupation of the AMA apical groove by RON2 and R1.

Left – semi-transparent grey surface of AMA (top – *Tg*AMA1; middle – *Tg*AMA3/SporoAMA1; bottom – *Pf*AMA1) in complex with RON2 or R1 shown as ball-

and-stick side chains on cartoon backbone. Right – zoom-in on three closely-apposed regions of interest, with overlayed complexes shown on a *Tg*AMA1 background for clarity.

Initially, two regions of the AMA1 surface were identified as attractive targets for drug design, both based primarily off the co-structures of *Pf*AMA1 with invasion inhibitory antibodies and peptides: (1) the Arg knob-in-hole interaction observed with mAb 1F9, IgNAR14I-1, and R-P15 (Figure 31B), and (2) a deep cavity between an AMA1 DI surface loop and the tip of the DII loop (Macraild, Anders et al. 2011). With the knowledge we now have, it is clear that the first pocket would suffer from extreme species specificity, while the second pocket would require pinning the relatively mobile DII loop in its bound state and would not be conserved between different apicomplexan genera. More recently, a set of three relatively well conserved AMA surface pockets was described (Tonkin, Crawford et al. 2013), and now these pockets can be further defined based on the knowledge of how different molecules dock into the AMA apical groove (Figure 42) combined with the newly identified AMA2 in *T. gondii* (Lamarque, Roques et al. 2014 Under review at Nature Comms). The three pockets can be described as Pocket 1, which is exposed by the ejection of the DII loop; Pocket 2, which is more of a backstop along the least mobile loops of the apical groove edge; and Pocket 3, which is on the opposite side of the apical groove as the DII loop, but importantly does not extend down to the specificity-determining RON2 cystine loop binding region (Figure 43, Table 12).
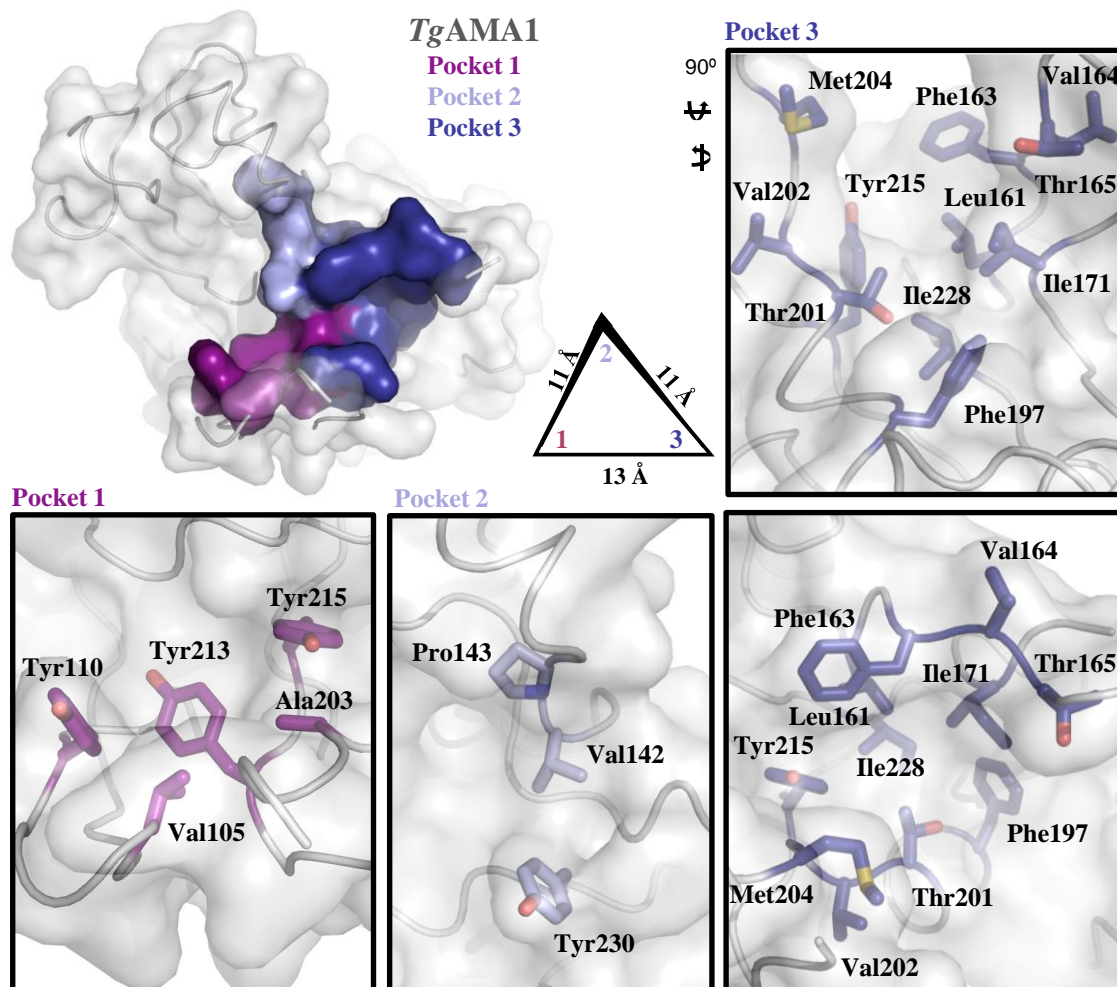
Figure 43: Identification of conserved pockets on the AMA apical surface.

Top Left – Surface of *Tg*AMA1 shown as a transparent grey surface over tube-style backbone. Residues defining Pocket 1 are shown as a deep purple solid surface; Pocket 2 (backstop), light blue; Pocket 3, dark blue. Inset, triangle representing the distance between the pocket centers, with thickened lines indicating projection out of the page as pocket 2 sits in a higher plane than Pockets 1 and 3. Pockets 1, 2, 3 are enlarged with a transparent surface over coloured stick representation of the residues making up each pocket.

Table 12: Residues forming the conserved apical surface pockets of AMA proteins. Residues with structurally equivalent positions are horizontally aligned.

*Tg*AMA1	*Tg*AMA2[a]	*Tg*AMA3	*Nc*AMA1	*Bd*AMA1	*Pf*AMA1	*Pv*AMA1
Pocket 1						
Val105	Leu84	NA[b]	Val99	Val126	Val137	Val82
Tyr110	Tyr89	Tyr134	Tyr104	His131	Tyr142	Tyr87
Ala203	Ala194	Ala224	Ala197	Pro225	Pro226	ND[c]
Tyr213	Tyr209	Tyr235	Tyr207	Tyr233	Tyr234	Tyr179
Tyr215	Tyr211	Tyr237	Tyr209	Tyr235	Tyr236	His181
Pocket 2 (backstop)						
Val142	Val123	Ala164	Ile136	Ile157	Val169	Val114
Pro143	Pro124	Pro165	Pro137	Ser158	Ala170	Ala115
Tyr230	Tyr226	Ser252	Ser224	Phe250	Tyr251	Tyr196
Pocket 3						
Leu161	Met141	Asn181	Met155	Phe169	Phe183	Phe128
Phe163	Met143	Gln183	Phe157	Pro170	Pro185	Ala131
Val164[d]	His144[d]	Val184[d]	Val158[d]	ND[c]	Thr186	Asn132
Thr165	Gln145	Tyr185	Thr159	ND[c]	Pro188	His134
Ile171	Leu148	Phe191	Ile165	Leu195	Met190	Ile135
Phe197	Trp188	Tyr218	Phe191	Tyr219	His220	His165
Thr201	Thr192	Thr222	Thr195	Ile223	Met224	Phe169
Val202[d]	Ser193[d]	Ile223[d]	Thr196[d]	Ile224[d]	Ile225[d]	Val170[d]
Met204	His195	Val225	Met198	Ala226	Asp227	ND[c]
Tyr215	Tyr211	Tyr237	Tyr209	Tyr235	Tyr236	His181
Ile228	Ile224	Val250	Ile222	Ile248	Ile249	Met194

[a] Homology model based on *Tg*AMA1 structure

[b] Not applicable (NA) due to shortened loop

[c] Not determined (ND) due to flexible, un-modeled loop

[d] Participation in pocket formation by backbone only

Pocket 1 is the most conserved of the three pockets, with relatively small hydrophobic residues lining the edge of the pocket (*Tg*AMA1 Val105 and Ala203), except for *Tg*AMA3 where a shortened apical loop results in a more open Pocket 1 (Figure 43, Table 12). A trio of tyrosines form the highly conserved core of Pocket 1, with only two histidine substitutions observed; while this substitution may not seem conservative, the two tyrosines which allow substitutions primarily function in aromatic stacking interactions, which could be maintained by an aromatic histidine residue. Pocket 2 appears to be more

of a backstop than a true pocket, as it extends up out of the base of the apical groove. The loops on this side of the groove are not nearly as flexible as the loops on the other side, and the added height creates an effective structural brace without stringent requirements for residue conservation; the residue at the tip of the backstop is small, the center residue is hydrophobic, the base residue generally has hydrogen bonding capability, and there are no charged residues in any of the three positions. Pocket 3 covers the largest surface area, and is less well conserved, but is of prime importance, as access to the highly conserved Pocket 1 requires ejection of the AMA DII loop, which would require a strong foothold elsewhere on the surface, particularly when targeting *Tg*AMA1 or *Nc*AMA1 that have a deeply integrated DII loop. Pocket 3 is generally hydrophobic and the sides are defined by flexible loops that move upon ligand binding (as evidenced by a number of non-determined residues in Table 12), and may represent a capacity for an induced fit binding mechanism. Interestingly, the tyrosine that was originally thought to represent the center of the apical groove (*Pf*AMA1 Tyr251 (Collins, Withers-Martinez et al. 2009) or *Tg*AMA1 Tyr230) is only a component of the backstop, while a neighboring tyrosine (*Pf*AMA1 Tyr236 or *Tg*AMA1 Tyr215) is an essential component of both Pockets 1 and 3, providing a central border between the two. The only divergence away from this Tyr is His181 in *Pv*AMA1, and the imidazole ring overlays well with the aromatic Tyr ring, thereby still providing a similar pocket border. The centers of the three pockets are closely positioned on the surface, leading to a triangular orientation with both Pockets 1 and 3 relatively equidistant (11 Å) from backstop 2 that sits in a higher plane, and approximately 13 Å between Pockets 1 and 3 (Figure 43); these dimensions are compatible with attainable dimensions of small molecule drugs. An exciting proof-of-concept study recently identified a small molecule inhibitor of the *Pf*AMA1-*Pf*RON2 complex that is able to block merozoite invasion in the low micromolar range specifically by blocking MJ formation; while it has not yet been detailed how these molecules block the AMA1-RON2 interaction, the cross-strain reactivity observed and preliminary docking experiments suggest the possibility that these compounds are occupying the AMA1 hydrophobic groove in the conserved pockets identified above and thereby preventing RON2 binding (Srinivasan, Yasgar et al. 2013).

After assembling all the currently characterized AMA structures and AMA-RON2 co-structures, it becomes clear that the most divergent pairing is *Tg*AMA3-*Tg*RON2$_{L2}$,

although characterization of the *Tg*AMA4-*Tg*RON2$_{L1}$ co-structure will likely show even more extreme receptor-ligand co-evolution. It is intriguing that maximal diversity is observed between AMA-RON2 pairs within a single parasite species and not between pairings from distantly related apicomplexans. This observation could be explained by one or more of three scenarios: (1) high stringency must be maintained to prevent intra-species cross-reactivity of AMA-RON2 pairs, although the reason for this is unclear, or (2) a single parasite containing highly divergent AMA-RON2 pairs would be more viable in the situation that one of the pairs is blocked from functionality, or (3) invasion of different cell types by different life cycle stages places selective pressure on a particular AMA-RON2 pair (section 4.5). Given that the apicomplexan parasites have established such sophisticated molecular strategies to ensure active host cell invasion, comprehensive targeting of the available invasion mechanisms will be incredibly beneficial to the design and development of a highly efficient therapeutic that is capable of broadly preventing or treating apicomplexan infections.

5.4 Future explorations

Investigation of the AMA1-RON2 extracellular link holds further interest in terms of drug and vaccine development; co-structures of AMA1 with invasion inhibitory small molecules and ternary complexes of AMA1-RON2 with protective antibodies would be very informative for structure-based drug design initiatives and antibody development. As AMA4 and MAEBL are likely to play a role in redundant pathways for MJ-mediated invasion, structural characterization of these proteins from *T. gondii* and *P. falciparum* is also of high priority. Although significant progress has been made in understanding the components of the extracellular link of the MJ complex, there are still likely to be additional parasite-host cell linking proteins that remain to be identified and that would serve as excellent additions to new strategic drugs and vaccines. Furthermore, very little is conclusively known extending away from the extracellular link to the parasite cytosol and host cell cytosol components of the MJ. On the parasite side, the role of AMA in generating signals that are transmitted to the parasite interior and the linkage of AMA or another transmembrane adhesive protein to the motor complex to provide the motive force for invasion are two unresolved, and highly debated, issues. On the host cell side, the

stoichiometry of the RON complex, the presence of RON2 and RON4 isoforms in the RON complex, the identity of host cell derived RON binding partners, and the mechanism by which host proteins are filtered by the MJ are all unknown. Furthermore, discerning the role of AMA and the RONs in parasites that do not appear to invade through an MJ mechanism – such as *Theileria* species (Shaw 2003) – would provide valuable insight into the functional plasticity of these important molecules. While we now have a structural foothold to build up the entire MJ complex, there is still much work to be done before we will have a robust model of the complete MJ complex that includes the variety of possible proteins that extend from the parasite motor through the extracellular environment to the host cell anchor.

Bibliography

Achur, R. N., I. Kakizaki, S. Goel, K. Kojima, S. V. Madhunapantula, A. Goyal, M. Ohta, S. Kumar, K. Takagaki and D. C. Gowda (2008). "Structural interactions in chondroitin 4-sulfate mediated adherence of *Plasmodium falciparum* infected erythrocytes in human placenta during pregnancy-associated malaria." Biochemistry **47**(47): 12635-12643.

Aikawa, M. (1977). "Variations in structure and function during the life cycle of malarial parasites." Bulletin of the World Health Organization **55**(2-3): 139-156.

Aikawa, M., L. H. Miller, J. Johnson and J. Rabbege (1978). "Erythrocyte entry by malarial parasites. A moving junction between erythrocyte and parasite." The Journal of Cell Biology **77**(1): 72-82.

Ajzenberg, D., N. Cogne, L. Paris, M. H. Bessieres, P. Thulliez, D. Filisetti, H. Pelloux, P. Marty and M. L. Darde (2002). "Genotype of 86 *Toxoplasma gondii* isolates associated with human congenital toxoplasmosis, and correlation with clinical findings." The Journal of Infectious Diseases **186**(5): 684-689.

Ajzenberg, D., H. Yera, P. Marty, L. Paris, F. Dalle, J. Menotti, D. Aubert, J. Franck, M. H. Bessieres, D. Quinio, H. Pelloux, L. Delhaes, N. Desbois, P. Thulliez, F. Robert-Gangneux, C. Kauffmann-Lacroix, S. Pujol, M. Rabodonirina, M. E. Bougnoux, B. Cuisenier, C. Duhamel, T. H. Duong, D. Filisetti, P. Flori, F. Gay-Andrieu, F. Pratlong, G. Nevez, A. Totet, B. Carme, H. Bonnabau, M. L. Darde and I. Villena (2009). "Genotype of 88 *Toxoplasma gondii* isolates associated with toxoplasmosis in immunocompromised patients and correlation with clinical findings." The Journal of Infectious Diseases **199**(8): 1155-1167.

Alexander, D. L., S. Arastu-Kapur, J. F. Dubremetz and J. C. Boothroyd (2006). "*Plasmodium falciparum* AMA1 binds a rhoptry neck protein homologous to *Tg*RON4, a component of the moving junction in *Toxoplasma gondii*." Eukaryotic Cell **5**(7): 1169-1173.

Alexander, D. L., J. Mital, G. E. Ward, P. Bradley and J. C. Boothroyd (2005). "Identification of the moving junction complex of *Toxoplasma gondii*: a collaboration between distinct secretory organelles." PLoS Pathogens **1**(2): e17.

Allison, A. C. (1954). "Protection afforded by sickle-cell trait against subtertian malareal infection." British Medical Journal **1**(4857): 290-294.

Anders, R. F., P. E. Crewther, S. Edwards, M. Margetts, M. L. Matthew, B. Pollock and D. Pye (1998). "Immunisation with recombinant AMA-1 protects mice against infection with *Plasmodium chabaudi*." Vaccine **16**(2-3): 240-247.

Araujo, F. G. and J. S. Remington (1987). "Toxoplasmosis in immunocompromised patients." European Journal of Clinical Microbiology **6**(1): 1-2.

Bai, T., M. Becker, A. Gupta, P. Strike, V. J. Murphy, R. F. Anders and A. H. Batchelor (2005). "Structure of AMA1 from *Plasmodium falciparum* reveals a clustering of polymorphisms that surround a conserved hydrophobic pocket." <u>Proceedings of the National Academy of Sciences of the United States of America</u> **102**(36): 12736-12741.

Bannister, L. H., J. M. Hopkins, A. R. Dluzewski, G. Margos, I. T. Williams, M. J. Blackman, C. H. Kocken, A. W. Thomas and G. H. Mitchell (2003). "*Plasmodium falciparum* apical membrane antigen 1 (*Pf*AMA-1) is translocated within micronemes along subpellicular microtubules during merozoite development." <u>Journal of Cell Science</u> **116**(Pt 18): 3825-3834.

Bannister, L. H., J. M. Hopkins, R. E. Fowler, S. Krishna and G. H. Mitchell (2000). "A brief illustrated guide to the ultrastructure of *Plasmodium falciparum* asexual blood stages." <u>Parasitology Today</u> **16**(10): 427-433.

Bannister, L. H., J. M. Hopkins, R. E. Fowler, S. Krishna and G. H. Mitchell (2000). "Ultrastructure of rhoptry development in *Plasmodium falciparum* erythrocytic schizonts." <u>Parasitology</u> **121**(Pt 3): 273-287.

Bargieri, D. Y., N. Andenmatten, V. Lagal, S. Thiberge, J. A. Whitelaw, I. Tardieux, M. Meissner and R. Menard (2013). "Apical membrane antigen 1 mediates apicomplexan parasite attachment but is dispensable for host cell invasion." <u>Nature Communications</u> **4**: 2552.

Barragan, A., F. Brossier and L. D. Sibley (2005). "Transepithelial migration of *Toxoplasma gondii* involves an interaction of intercellular adhesion molecule 1 (ICAM-1) with the parasite adhesin MIC2." <u>Cellular Microbiology</u> **7**(4): 561-568.

Barragan, A. and L. D. Sibley (2002). "Transepithelial migration of *Toxoplasma gondii* is linked to parasite motility and virulence." <u>The Journal of Experimental Medicine</u> **195**(12): 1625-1633.

Battye, T. G., L. Kontogiannis, O. Johnson, H. R. Powell and A. G. Leslie (2011). "iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM." <u>Acta Crystallographica Section D: Biological Crystallography</u> **67**(Pt 4): 271-281.

Baum, J. and A. F. Cowman (2011). "Biochemistry. Revealing a parasite's invasive trick." <u>Science</u> **333**(6041): 410-411.

Baum, J., A. T. Papenfuss, B. Baum, T. P. Speed and A. F. Cowman (2006). "Regulation of apicomplexan actin-based motility." <u>Nature Reviews Microbiology</u> **4**(8): 621-628.

Besteiro, S., J. F. Dubremetz and M. Lebrun (2011). "The moving junction of apicomplexan parasites: a key structure for invasion." <u>Cellular Microbiology</u> **13**(6): 797-805.

Besteiro, S., A. Michelin, J. Poncet, J. F. Dubremetz and M. Lebrun (2009). "Export of a *Toxoplasma gondii* rhoptry neck protein complex at the host cell membrane to form the moving junction during invasion." PLoS Pathogens **5**(2): e1000309.

Binka, F. N., A. Kubaje, M. Adjuik, L. A. Williams, C. Lengeler, G. H. Maude, G. E. Armah, B. Kajihara, J. H. Adiamah and P. G. Smith (1996). "Impact of permethrin impregnated bednets on child mortality in Kassena-Nankana district, Ghana: a randomized controlled trial." Tropical Medicine and International Health **1**(2): 147-154.

Boothroyd, J. C. and J. F. Dubremetz (2008). "Kiss and spit: the dual roles of *Toxoplasma* rhoptries." Nature Reviews Microbiology **6**(1): 79-88.

Bosch, J., C. A. Buscaglia, B. Krumm, B. P. Ingason, R. Lucas, C. Roach, T. Cardozo, V. Nussenzweig and W. G. Hol (2007). "Aldolase provides an unusual binding site for thrombospondin-related anonymous protein in the invasion machinery of the malaria parasite." Proceedings of the National Academy of Sciences of the United States of America **104**(17): 7015-7020.

Boulanger, M. J., M. L. Tonkin and J. Crawford (2010). "Apicomplexan parasite adhesins: novel strategies for targeting host cell carbohydrates." Current Opinion in Structural Biology **20**(5): 551-559.

Bowden, T. A., A. R. Aricescu, R. J. Gilbert, J. M. Grimes, E. Y. Jones and D. I. Stuart (2008). "Structural basis of Nipah and Hendra virus attachment to their cell-surface receptor ephrin-B2." Nature Structural and Molecular Biology **15**(6): 567-572.

Bowie, W. R., A. S. King, D. H. Werker, J. L. Isaac-Renton, A. Bell, S. B. Eng and S. A. Marion (1997). "Outbreak of toxoplasmosis associated with municipal drinking water. The BC *Toxoplasma* Investigation Team." Lancet **350**(9072): 173-177.

Bradley, P. J., C. Ward, S. J. Cheng, D. L. Alexander, S. Coller, G. H. Coombs, J. D. Dunn, D. J. Ferguson, S. J. Sanderson, J. M. Wastling and J. C. Boothroyd (2005). "Proteomic analysis of rhoptry organelles reveals many novel constituents for host-parasite interactions in *Toxoplasma gondii*." The Journal of Biological Chemistry **280**(40): 34245-34258.

Brissonnoel, A., P. Trieucuot and P. Courvalin (1988). "Mechanism of Action of Spiramycin and Other Macrolides." Journal of Antimicrobial Chemotherapy **22**: 13-23.

Brydges, S. D., G. D. Sherman, S. Nockemann, A. Loyens, W. Daubener, J. F. Dubremetz and V. B. Carruthers (2000). "Molecular characterization of *Tg*MIC5, a proteolytically processed antigen secreted from the micronemes of *Toxoplasma gondii*." Molecular and Biochemical Parasitology **111**(1): 51-66.

Brymora, A., V. A. Valova and P. J. Robinson (2004). "Protein-protein interactions identified by pull-down experiments and mass spectrometry." Current Protocols in Cell Biology **Chapter 17**: Unit 17.15.

Burg, J. L., D. Perelman, L. H. Kasper, P. L. Ware and J. C. Boothroyd (1988). "Molecular analysis of the gene encoding the major surface antigen of *Toxoplasma gondii*." Journal of Immunology **141**(10): 3584-3591.

Camilleri, M., K. Madsen, R. Spiller, B. Greenwood-Van Meerveld and G. N. Verne (2012). "Intestinal barrier function in health and gastrointestinal disease." Neurogastroenterology and Motility **24**(6): 503-512.

Cao, J., O. Kaneko, A. Thongkukiatkul, M. Tachibana, H. Otsuki, Q. Gao, T. Tsuboi and M. Torii (2009). "Rhoptry neck protein RON2 forms a complex with microneme protein AMA1 in *Plasmodium falciparum* merozoites." Parasitology International **58**(1): 29-35.

Carme, B., F. Bissuel, D. Ajzenberg, R. Bouyne, C. Aznar, M. Demar, S. Bichat, D. Louvel, A. M. Bourbigot, C. Peneau, P. Neron and M. L. Darde (2002). "Severe acquired toxoplasmosis in immunocompetent adult patients in French Guiana." Journal of Clinical Microbiology **40**(11): 4037-4044.

Carruthers, V. and J. C. Boothroyd (2007). "Pulling together: an integrated model of *Toxoplasma* cell invasion." Current Opinion in Microbiology **10**(1): 83-89.

Carruthers, V. B. (1999). "Armed and dangerous: *Toxoplasma gondii* uses an arsenal of secretory proteins to infect host cells." Parasitology International **48**(1): 1-10.

Carruthers, V. B. and L. D. Sibley (1997). "Sequential protein secretion from three distinct organelles of *Toxoplasma gondii* accompanies invasion of human fibroblasts." European Journal of Cell Biology **73**(2): 114-123.

Carruthers, V. B. and Y. Suzuki (2007). "Effects of *Toxoplasma gondii* infection on the brain." Schizophrenia Bulletin **33**(3): 745-751.

Carter, R., K. N. Mendis, L. H. Miller, L. Molineaux and A. Saul (2000). "Malaria transmission-blocking vaccines--how can their development be supported?" Nature Medicine **6**(3): 241-244.

Cavalier-Smith, T. (1993). "Kingdom protozoa and its 18 phyla." Microbiologcal Reviews **57**(4): 953-994.

Charron, A. J. and L. D. Sibley (2004). "Molecular partitioning during host cell penetration by *Toxoplasma gondii*." Traffic **5**(11): 855-867.

Chen, V. B., W. B. Arendall, 3rd, J. J. Headd, D. A. Keedy, R. M. Immormino, G. J. Kapral, L. W. Murray, J. S. Richardson and D. C. Richardson (2010). "MolProbity: all-atom structure validation for macromolecular crystallography." Acta Crystallographica Section D: Biological Crystallography **66**(Pt 1): 12-21.

Chesne-Seck, M. L., J. C. Pizarro, B. Vulliez-Le Normand, C. R. Collins, M. J. Blackman, B. W. Faber, E. J. Remarque, C. H. Kocken, A. W. Thomas and G. A. Bentley (2005).

"Structural comparison of apical membrane antigen 1 orthologues and paralogues in apicomplexan parasites." Molecular and Biochemical Parasitology **144**(1): 55-67.

Coley, A. M., A. Gupta, V. J. Murphy, T. Bai, H. Kim, M. Foley, R. F. Anders and A. H. Batchelor (2007). "Structure of the malaria antigen AMA1 in complex with a growth-inhibitory antibody." PLoS Pathogens **3**(9): 1308-1319.

Coley, A. M., K. Parisi, R. Masciantonio, J. Hoeck, J. L. Casey, V. J. Murphy, K. S. Harris, A. H. Batchelor, R. F. Anders and M. Foley (2006). "The most polymorphic residue on *Plasmodium falciparum* apical membrane antigen 1 determines binding of an invasion-inhibitory antibody." Infection and Immunity **74**(5): 2628-2636.

Collins, C. R., C. Withers-Martinez, G. A. Bentley, A. H. Batchelor, A. W. Thomas and M. J. Blackman (2007). "Fine mapping of an epitope recognized by an invasion-inhibitory monoclonal antibody on the malaria vaccine candidate apical membrane antigen 1." The Journal of Biological Chemistry **282**(10): 7431-7441.

Collins, C. R., C. Withers-Martinez, F. Hackett and M. J. Blackman (2009). "An inhibitory antibody blocks interactions between components of the malarial invasion machinery." PLoS Pathogens **5**(1): e1000273.

Colovos, C. and T. O. Yeates (1993). "Verification of protein structures: patterns of nonbonded atomic interactions." Protein Science **2**(9): 1511-1519.

Contini, C. (2008). "Clinical and diagnostic management of toxoplasmosis in the immunocompromised patient." Parassitologia **50**(1-2): 45-50.

Cortes, A., M. Mellombo, I. Mueller, A. Benet, J. C. Reeder and R. F. Anders (2003). "Geographical structure of diversity and differences between symptomatic and asymptomatic infections for *Plasmodium falciparum* vaccine candidate AMA1." Infection and Immunity **71**(3): 1416-1426.

Cortina-Borja, M., H. K. Tan, M. Wallon, M. Paul, A. Prusa, W. Buffolano, G. Malm, A. Salt, K. Freeman, E. Petersen, R. E. Gilbert and T. European Multicentre Study on Congenital (2010). "Prenatal treatment for serious neurological sequelae of congenital toxoplasmosis: an observational prospective cohort study." PLoS Medicine **7**(10): e1000351.

Cowman, A. F. and B. S. Crabb (2006). "Invasion of red blood cells by malaria parasites." Cell **124**(4): 755-766.

Crawford, J., M. L. Tonkin, O. Grujic and M. J. Boulanger (2010). "Structural characterization of apical membrane antigen 1 (AMA1) from *Toxoplasma gondii*." Journal of Biological Chemistry **285**(20): 15644-15652.

Crompton, P. D., S. K. Pierce and L. H. Miller (2010). "Advances and challenges in malaria vaccine development." The Journal of Clinical Investigation **120**(12): 4168-4178.

De Salvador-Guillouet, F., D. Ajzenberg, S. Chaillou-Opitz, M. C. Saint-Paul, B. Dunais, P. Dellamonica and P. Marty (2006). "Severe pneumonia during primary infection with an atypical strain of *Toxoplasma gondii* in an immunocompetent young man." The Journal of Infection **53**(2): e47-50.

Deans, J. A., T. Alderson, A. W. Thomas, G. H. Mitchell, E. S. Lennox and S. Cohen (1982). "Rat monoclonal antibodies which inhibit the *in vitro* multiplication of *Plasmodium knowlesi*." Clinical and Experimental Immunology **49**(2): 297-309.

Deans, J. A., A. M. Knight, W. C. Jean, A. P. Waters, S. Cohen and G. H. Mitchell (1988). "Vaccination trials in rhesus monkeys with a minor, invariant, *Plasmodium knowlesi* 66 kD merozoite antigen." Parasite Immunology **10**(5): 535-552.

Desmonts, G. and J. Couvreur (1974). "Congenital toxoplasmosis. A prospective study of 378 pregnancies." The New England Journal of Medicine **290**(20): 1110-1116.

Desmonts, G., F. Daffos, F. Forestier, M. Capella-Pavlovsky, P. Thulliez and M. Chartier (1985). "Prenatal diagnosis of congenital toxoplasmosis." Lancet **1**(8427): 500-504.

Diskin, R., P. M. Marcovecchio and P. J. Bjorkman (2010). "Structure of a clade C HIV-1 gp120 bound to CD4 and CD4-induced antibody reveals anti-CD4 polyreactivity." Nature Structural and Molecular Biology **17**(5): 608-613.

Dobrowolski, J. M. and L. D. Sibley (1996). "*Toxoplasma* invasion of mammalian cells is powered by the actin cytoskeleton of the parasite." Cell **84**(6): 933-939.

Donahue, C. G., V. B. Carruthers, S. D. Gilk and G. E. Ward (2000). "The *Toxoplasma* homolog of *Plasmodium* apical membrane antigen-1 (AMA-1) is a microneme protein secreted in response to elevated intracellular calcium levels." Molecular and Biochemical Parasitology **111**(1): 15-30.

Donald, R. G., D. Carter, B. Ullman and D. S. Roos (1996). "Insertional tagging, cloning, and expression of the *Toxoplasma gondii* hypoxanthine-xanthine-guanine phosphoribosyltransferase gene. Use as a selectable marker for stable transformation." Journal of Biological Chemistry **271**(24): 14010-14019.

Dondorp, A. M., S. Yeung, L. White, C. Nguon, N. P. J. Day, D. Socheat and L. von Seidlein (2010). "Artemisinin resistance: current status and scenarios for containment." Nature Reviews Microbiology **8**: 272-280.

Dubey, J. P. (1998). "Advances in the life cycle of *Toxoplasma gondii*." International Journal for Parasitology **28**(7): 1019-1024.

Dubey, J. P. (2004). "Toxoplasmosis - a waterborne zoonosis." Veterinary Parasitology **126**(1-2): 57-72.

Dubey, J. P. and J. K. Frenkel (1972). "Cyst-induced toxoplasmosis in cats." The Journal of Protozoology **19**(1): 155-177.

Dubey, J. P., D. S. Lindsay and C. A. Speer (1998). "Structures of *Toxoplasma gondii* tachyzoites, bradyzoites, and sporozoites and biology and development of tissue cysts." Clinical Microbiology Reviews **11**(2): 267-299.

Dubey, J. P., C. A. Speer, S. K. Shen, O. C. Kwok and J. A. Blixt (1997). "Oocyst-induced murine toxoplasmosis: life cycle, pathogenicity, and stage conversion in mice fed *Toxoplasma gondii* oocysts." The Journal of Parasitology **83**(5): 870-882.

Dubremetz, J. F., N. Garcia-Réguet, V. Conseil and M. N. Fourmaux (1998). "Invited review: Apical organelles and host-cell invasion by Apicomplexa." International Journal for Parasitology **28**(7): 1007-1013.

Dutta, S., L. S. Dlugosz, D. R. Drew, X. Ge, D. Ababacar, Y. I. Rovira, J. K. Moch, M. Shi, C. A. Long, M. Foley, J. G. Beeson, R. F. Anders, K. Miura, J. D. Haynes and A. H. Batchelor (2013). "Overcoming antigenic diversity by enhancing the immunogenicity of conserved epitopes on the malaria vaccine candidate apical membrane antigen-1." PLoS Pathogens **9**(12): e1003840.

Dutta, S., S. Y. Lee, A. H. Batchelor and D. E. Lanar (2007). "Structural basis of antigenic escape of a malaria vaccine candidate." Proceedings of the National Academy of Sciences of the United States of America **104**(30): 12488-12493.

El-Nawawi, F. A., M. A. Tawfik and R. M. Shaapan (2008). "Methods for inactivation of *Toxoplasma gondii* cysts in meat and tissues of experimentally infected sheep." Foodborne Pathogens and Disease **5**(5): 687-690.

Elbez-Rubinstein, A., D. Ajzenberg, M. L. Darde, R. Cohen, A. Dumetre, H. Yera, E. Gondon, J. C. Janaud and P. Thulliez (2009). "Congenital Toxoplasmosis and Reinfection during Pregnancy: Case Report, Strain Characterization, Experimental Model of Reinfection, and Review." Journal of Infectious Diseases **199**(2): 280-285.

Emsley, P. and K. Cowtan (2004). "Coot: model-building tools for molecular graphics." Acta Crystallographica Section D: Biological Crystallography **60**(Pt 12): 2126-2132.

Escalante, A. A., H. M. Grebert, S. C. Chaiyaroj, M. Magris, S. Biswas, B. L. Nahlen and A. A. Lal (2001). "Polymorphism in the gene encoding the apical membrane antigen-1 (AMA-1) of *Plasmodium falciparum.* X. Asembo Bay Cohort Project." Molecular and Biochemical Parasitology **113**(2): 279-287.

Eswar, N., B. Webb, M. A. Marti-Renom, M. S. Madhusudhan, D. Eramian, M. Y. Shen, U. Pieper and A. Sali (2006). "Comparative protein structure modeling using Modeller." Current Protocols in Bioinformatics **Chapter 5**: Unit 5 6.

Evans, P. (2006). "Scaling and assessment of data quality." Acta Crystallographica Section D: Biological Crystallography **62**(Pt 1): 72-82.

Faber, B. W., E. J. Remarque, C. H. Kocken, P. Cheront, D. Cingolani, F. Xhonneux, M. Jurado, M. Haumont, S. Jepsen, O. Leroy and A. W. Thomas (2008). "Production, quality

control, stability and pharmacotoxicity of cGMP-produced *Plasmodium falciparum* AMA1 FVO strain ectodomain expressed in *Pichia pastoris*." Vaccine **26**(48): 6143-6150.

Fauquenoy, S., W. Morelle, A. Hovasse, A. Bednarczyk, C. Slomianny, C. Schaeffer, A. Van Dorsselaer and S. Tomavo (2008). "Proteomics and glycomics analyses of N-glycosylated structures involved in *Toxoplasma gondii*--host cell interactions." Molecular and Cellular Proteomics **7**(5): 891-910.

Feng, Z. P., D. W. Keizer, R. A. Stevenson, S. Yao, J. J. Babon, V. J. Murphy, R. F. Anders and R. S. Norton (2005). "Structure and inter-domain interactions of domain II from the blood-stage malarial protein, apical membrane antigen 1." Journal of Molecular Biology **350**(4): 641-656.

Flegr, J. (2007). "Effects of *Toxoplasma* on human behavior." Schizophrenia Bulletin **33**(3): 757-760.

Flegr, J. (2013). "How and why *Toxoplasma* makes us crazy." Trends in Parasitology **29**(4): 156-163.

Fraser, T. S., S. H. Kappe, D. L. Narum, K. M. VanBuskirk and J. H. Adams (2001). "Erythrocyte-binding activity of *Plasmodium yoelii* apical membrane antigen-1 expressed on the surface of transfected COS-7 cells." Molecular and Biochemical Parasitology **117**(1): 49-59.

Frevert, U., I. Usynin, K. Baer and C. Klotz (2008). "*Plasmodium* sporozoite passage across the sinusoidal cell layer." Subcellular Biochemistry **47**: 182-197.

Fritz, H. M., P. W. Bowyer, M. Bogyo, P. A. Conrad and J. C. Boothroyd (2012). "Proteomic analysis of fractionated *Toxoplasma* oocysts reveals clues to their environmental resistance." PLoS One **7**(1): e29955.

Fritz, H. M., K. R. Buchholz, X. Chen, B. Durbin-Johnson, D. M. Rocke, P. A. Conrad and J. C. Boothroyd (2012). "Transcriptomic analysis of *Toxoplasma* development reveals many novel functions and structures specific to sporozoites and oocysts." PLoS One **7**(2): e29998.

Gaffar, F. R., A. P. Yatsuda, F. F. Franssen and E. de Vries (2004). "Erythrocyte invasion by *Babesia bovis* merozoites is inhibited by polyclonal antisera directed against peptides derived from a homologue of *Plasmodium falciparum* apical membrane antigen 1." Infection and Immunity **72**(5): 2947-2955.

Gardner, M. J., N. Hall, E. Fung, O. White, M. Berriman, R. W. Hyman, J. M. Carlton, A. Pain, K. E. Nelson, S. Bowman, I. T. Paulsen, K. James, J. A. Eisen, K. Rutherford, S. L. Salzberg, A. Craig, S. Kyes, M. S. Chan, V. Nene, S. J. Shallom, B. Suh, J. Peterson, S. Angiuoli, M. Pertea, J. Allen, J. Selengut, D. Haft, M. W. Mather, A. B. Vaidya, D. M. Martin, A. H. Fairlamb, M. J. Fraunholz, D. S. Roos, S. A. Ralph, G. I. McFadden, L. M. Cummings, G. M. Subramanian, C. Mungall, J. C. Venter, D. J. Carucci, S. L. Hoffman,

C. Newbold, R. W. Davis, C. M. Fraser and B. Barrell (2002). "Genome sequence of the human malaria parasite *Plasmodium falciparum*." Nature **419**(6906): 498-511.

Garrison, E., M. Treeck, E. Ehret, H. Butz, T. Garbuz, B. P. Oswald, M. Settles, J. Boothroyd and G. Arrizabalaga (2012). "A forward genetic screen reveals that calcium-dependent protein kinase 3 regulates egress in *Toxoplasma*." PLoS Pathogens **8**(11): e1003049.

Gaur, D. and C. E. Chitnis (2011). "Molecular interactions and signaling mechanisms during erythrocyte invasion by malaria parasites." Current Opinion in Microbiology **14**(4): 422-428.

Gilson, P. R. and B. S. Crabb (2009). "Morphology and kinetics of the three distinct phases of red blood cell invasion by *Plasmodium falciparum* merozoites." International Journal for Parasitology **39**(1): 91-96.

Giovannini, D., S. Spath, C. Lacroix, A. Perazzi, D. Bargieri, V. Lagal, C. Lebugle, A. Combe, S. Thiberge, P. Baldacci, I. Tardieux and R. Menard (2011). "Independent roles of apical membrane antigen 1 and rhoptry neck proteins during host cell invasion by Apicomplexa." Cell Host and Microbe **10**(6): 591-602.

Good, M. F. and D. L. Doolan (2010). "Malaria vaccine design: immunological considerations." Immunity **33**(4): 555-566.

Goodyer, L. I., A. M. Croft, S. P. Frances, N. Hill, S. J. Moore, S. P. Onyango and M. Debboun (2010). "Expert review of the evidence base for arthropod bite avoidance." Journal of Travel Medicine **17**(3): 182-192.

Gouet, P., E. Courcelle, D. I. Stuart and F. Metoz (1999). "ESPript: multiple sequence alignments in PostScript." Bioinformatics **15**: 305-308.

Gupta, A., T. Bai, V. Murphy, P. Strike, R. F. Anders and A. H. Batchelor (2005). "Refolding, purification, and crystallization of apical membrane antigen 1 from *Plasmodium falciparum*." Protein Expression and Purification **41**(1): 186-198.

Hansson, G. C. (2012). "Role of mucus layers in gut infection and inflammation." Current Opinion in Microbiology **15**(1): 57-62.

Harris, K. S., J. L. Casey, A. M. Coley, J. A. Karas, J. K. Sabo, Y. Y. Tan, O. Dolezal, R. S. Norton, A. B. Hughes, D. Scanlon and M. Foley (2009). "Rapid optimization of a peptide inhibitor of malaria parasite invasion by comprehensive N-methyl scanning." Journal of Biological Chemistry **284**(14): 9361-9371.

Harris, K. S., J. L. Casey, A. M. Coley, R. Masciantonio, J. K. Sabo, D. W. Keizer, E. F. Lee, A. McMahon, R. S. Norton, R. F. Anders and M. Foley (2005). "Binding hot spot for invasion inhibitory molecules on *Plasmodium falciparum* apical membrane antigen 1." Infection and Immunity **73**(10): 6981-6989.

Healer, J., S. Crawford, S. Ralph, G. McFadden and A. F. Cowman (2002). "Independent translocation of two micronemal proteins in developing *Plasmodium falciparum* merozoites." Infection and Immunity **70**(10): 5751-5758.

Hehl, A. B., C. Lekutis, M. E. Grigg, P. J. Bradley, J. F. Dubremetz, E. Ortega-Barria and J. C. Boothroyd (2000). "*Toxoplasma gondii* homologue of plasmodium apical membrane antigen 1 is involved in invasion of host cells." Infection and Immunity **68**(12): 7078-7086.

Henderson, K. A., V. A. Streltsov, A. M. Coley, O. Dolezal, P. J. Hudson, A. H. Batchelor, A. Gupta, T. Bai, V. J. Murphy, R. F. Anders, M. Foley and S. D. Nuttall (2007). "Structure of an IgNAR-AMA1 complex: targeting a conserved hydrophobic cleft broadens malarial strain recognition." Structure **15**(11): 1452-1466.

Henriquez, S. A., R. Brett, J. Alexander, J. Pratt and C. W. Roberts (2009). "Neuropsychiatric disease and *Toxoplasma gondii* infection." Neuroimmunomodulation **16**(2): 122-133.

Hodder, A. N., P. E. Crewther and R. F. Anders (2001). "Specificity of the protective antibody response to apical membrane antigen 1." Infection and Immunity **69**(5): 3286-3294.

Hodder, A. N., P. E. Crewther, M. L. Matthew, G. E. Reid, R. L. Moritz, R. J. Simpson and R. F. Anders (1996). "The disulfide bond structure of *Plasmodium* apical membrane antigen-1." Journal of Biological Chemistry **271**(46): 29446-29452.

Hoff, E. F., S. H. Cook, G. D. Sherman, J. M. Harper, D. J. Ferguson, J. F. Dubremetz and V. B. Carruthers (2001). "*Toxoplasma gondii*: molecular cloning and characterization of a novel 18-kDa secretory antigen, *Tg*MIC10." Experimental Parasitology **97**(2): 77-88.

Homer, M. J., I. Aguilar-Delfin, S. R. Telford, 3rd, P. J. Krause and D. H. Persing (2000). "Babesiosis." Clinical Microbiology Reviews **13**(3): 451-469.

Hotez, P. J. (2008). "Neglected infections of poverty in the United States of America." PLoS Neglected Tropical Diseases **2**(6): e256.

Howe, D. K. and L. D. Sibley (1995). "*Toxoplasma gondii* comprises three clonal lineages: correlation of parasite genotype with human disease." The Journal of Infectious Diseases **172**(6): 1561-1566.

Howell, S. A., C. Withers-Martinez, C. H. Kocken, A. W. Thomas and M. J. Blackman (2001). "Proteolytic processing and primary structure of *Plasmodium falciparum* apical membrane antigen-1." Journal of Biological Chemistry **276**(33): 31311-31320.

Hu, K., J. Johnson, L. Florens, M. Fraunholz, S. Suravajjala, C. DiLullo, J. Yates, D. S. Roos and J. M. Murray (2006). "Cytoskeletal components of an invasion machine--the apical complex of *Toxoplasma gondii*." PLoS Pathogens **2**(2): e13.

Igonet, S., B. Vulliez-Le Normand, G. Faure, M. M. Riottot, C. H. Kocken, A. W. Thomas and G. A. Bentley (2007). "Cross-reactivity studies of an anti-*Plasmodium vivax* apical membrane antigen 1 monoclonal antibody: binding and structural characterisation." Journal of Molecular Biology **366**(5): 1523-1537.

Jackson, M. H. and W. M. Hutchison (1989). "The prevalence and source of *Toxoplasma* infection in the environment." Advances in Parasitology **28**: 55-105.

Jones, J. L. and J. P. Dubey (2010). "Waterborne toxoplasmosis--recent developments." Experimental Parasitology **124**(1): 10-25.

Kabsch, W. (2010). "Xds." Acta Crystallographica Section D: Biological Crystallography **66**(Pt 2): 125-132.

Kappe, S. H., A. M. Vaughan, J. A. Boddey and A. F. Cowman (2010). "That was then but this is now: malaria research in the time of an eradication agenda." Science **328**(5980): 862-866.

Kariu, T., M. Yuda, K. Yano and Y. Chinzei (2002). "MAEBL is essential for malarial sporozoite infection of the mosquito salivary gland." The Journal of Experimental Medicine **195**(10): 1317-1323.

Kato, K., D. C. Mayer, S. Singh, M. Reid and L. H. Miller (2005). "Domain III of *Plasmodium falciparum* apical membrane antigen 1 binds to the erythrocyte membrane protein Kx." Proceedings of the National Academy of Sciences of the United States of America **102**(15): 5552-5557.

Keeley, A. and D. Soldati (2004). "The glideosome: a molecular machine powering motility and host-cell invasion by Apicomplexa." Trends in Cell Biology **14**(10): 528-532.

Khan, A., C. Jordan, C. Muccioli, A. L. Vallochi, L. V. Rizzo, R. Belfort, Jr., R. W. Vitor, C. Silveira and L. D. Sibley (2006). "Genetic divergence of *Toxoplasma gondii* strains associated with ocular toxoplasmosis, Brazil." Emerging Infectious Diseases **12**(6): 942-949.

Kim, K. and L. M. Weiss (2004). "*Toxoplasma gondii*: the model apicomplexan." International Journal for Parasitology **34**(3): 423-432.

Klonis, N., D. J. Creek, L. Tilley (2013). "Iron and heme metabolism in *Plasmodium falciparum* and the mechanism of action of artemisinins." Current Opinion in Microbiology **16**(6): 722-727.

Kocken, C. H., C. Withers-Martinez, M. A. Dubbeld, A. van der Wel, F. Hackett, A. Valderrama, M. J. Blackman and A. W. Thomas (2002). "High-level expression of the malaria blood-stage vaccine candidate *Plasmodium falciparum* apical membrane antigen 1 and induction of antibodies that inhibit erythrocyte invasion." Infection and Immunity **70**(8): 4471-4476.

Kortagere, S. (2012). "Screening for small molecule inhibitors of *Toxoplasma gondii*." Expert Opinion on Drug Discovery **7**(12): 1193-1206.

Krissinel, E. and K. Henrick (2007). "Inference of macromolecular assemblies from crystalline state." Journal of Molecular Biology **372**(3): 774-797.

Ladda, R., M. Aikawa and H. Sprinz (1969). "Penetration of erythrocytes by merozoites of mammalian and avian malarial parasites." Journal of Parasitology **55**(3): 633-644.

Lal, A. A., M. A. Hughes, D. A. Oliveira, C. Nelson, P. B. Bloland, A. J. Oloo, W. E. Hawley, A. W. Hightower, B. L. Nahlen and V. Udhayakumar (1996). "Identification of T-cell determinants in natural immune responses to the *Plasmodium falciparum* apical membrane antigen (AMA-1) in an adult population exposed to malaria." Infection and Immunity **64**(3): 1054-1059.

Lamarque, M., S. Besteiro, J. Papoin, M. Roques, B. Vulliez-Le Normand, J. Morlon-Guyot, J. F. Dubremetz, S. Fauquenoy, S. Tomavo, B. W. Faber, C. H. Kocken, A. W. Thomas, M. J. Boulanger, G. A. Bentley and M. Lebrun (2011). "The RON2-AMA1 interaction is a critical step in moving junction-dependent invasion by apicomplexan parasites." PLoS Pathogens **7**(2): e1001276.

Lamarque, M. H., M. Roques, M. Kong-Hap, M. L. Tonkin, G. Rugarabamu, J. B. Marq, D. M. Penarete-Vargas, M. J. Boulanger, D. Soldati-Favre and M. Lebrun (2014). "Plasticity and redundancy among AMA-RON pairs ensure host cell entry of Toxoplasma parasites." Nature Communications **5**: 4098.

Laurens, M. B., M. A. Thera, D. Coulibaly, A. Ouattara, A. K. Kone, A. B. Guindo, K. Traore, I. Traore, B. Kouriba, D. A. Diallo, I. Diarra, M. Daou, A. Dolo, Y. Tolo, M. S. Sissoko, A. Niangaly, M. Sissoko, S. Takala-Harrison, K. E. Lyke, Y. Wu, W. C. Blackwelder, O. Godeaux, J. Vekemans, M. C. Dubois, W. R. Ballou, J. Cohen, T. Dube, L. Soisson, C. L. Diggs, B. House, J. W. Bennett, D. E. Lanar, S. Dutta, D. G. Heppner, C. V. Plowe and O. K. Doumbo (2013). "Extended safety, immunogenicity and efficacy of a blood-stage malaria vaccine in malian children: 24-month follow-up of a randomized, double-blinded phase 2 trial." PloS One **8**(11): e79323.

Lebrun, M., A. Michelin, H. El Hajj, J. Poncet, P. J. Bradley, H. Vial and J. F. Dubremetz (2005). "The rhoptry neck protein RON4 re-localizes at the moving junction during *Toxoplasma gondii* invasion." Cellular Microbiology **7**(12): 1823-1833.

Lee, E. F., S. Yao, J. K. Sabo, W. D. Fairlie, R. A. Stevenson, K. S. Harris, R. F. Anders, M. Foley and R. S. Norton (2011). "Peptide inhibitors of the malaria surface protein, apical membrane antigen 1: Identification of key binding residues." Biopolymers **95**(5): 354-364.

Levine, N. D. (1988). "Progress in taxonomy of the Apicomplexan protozoa." Journal of Protozoology **35**(4): 518-520.

Leykauf, K., M. Treeck, P. R. Gilson, T. Nebl, T. Braulke, A. F. Cowman, T. W. Gilberger and B. S. Crabb (2010). "Protein kinase A dependent phosphorylation of apical membrane

antigen 1 plays an important role in erythrocyte invasion by the malaria parasite." <u>PLoS Pathogens</u> **6**(6): e1000941.

Li, F., A. Dluzewski, A. M. Coley, A. Thomas, L. Tilley, R. F. Anders and M. Foley (2002). "Phage-displayed peptides bind to the malarial protein apical membrane antigen-1 and inhibit the merozoite invasion of host erythrocytes." <u>Journal of Biological Chemistry</u> **277**(52): 50303-50310.

Lim, S. S., C. O. Debono, C. A. MacRaild, I. R. Chandrashekaran, O. Dolezal, R. F. Anders, J. S. Simpson, M. J. Scanlon, S. M. Devine, P. J. Scammells and R. S. Norton (2013). "Development of inhibitors of *Plasmodium falciparum* apical membrane antigen 1 based on fragment screening." <u>Australian Journal of Chemistry</u> **66**(12): 1530-1536.

Liu, Q., L. D. Singla and H. Zhou (2012). "Vaccines against *Toxoplasma gondii*: status, challenges and future directions." <u>Human Vaccines and Immunotherapeutics</u> **8**(9): 1305-1308.

Lovett, J. L. and L. D. Sibley (2003). "Intracellular calcium stores in *Toxoplasma gondii* govern invasion of host cells." <u>Journal of Cell Science</u> **116**(Pt 14): 3009-3016.

Luo, Y., E. A. Frey, R. A. Pfuetzner, A. L. Creagh, D. G. Knoechel, C. A. Haynes, B. B. Finlay and N. C. Strynadka (2000). "Crystal structure of enteropathogenic *Escherichia coli* intimin-receptor complex." <u>Nature</u> **405**(6790): 1073-1077.

Lyons, R. E., R. McLeod and C. W. Roberts (2002). "*Toxoplasma gondii* tachyzoite-bradyzoite interconversion." <u>Trends in Parasitology</u> **18**(5): 198-201.

MacPherson, G. G., M. J. Warrell, N. J. White, S. Looareesuwan and D. A. Warrell (1985). "Human cerebral malaria. A quantitative ultrastructural analysis of parasitized erythrocyte sequestration." <u>The American Journal of Pathology</u> **119**(3): 385-401.

Macraild, C. A., R. F. Anders, M. Foley and R. S. Norton (2011). "Apical membrane antigen 1 as an anti-malarial drug target." <u>Current Topics in Medicinal Chemistry</u> **11**(16): 2039-2047.

Marsh, K., D. Forster, C. Waruiru, I. Mwangi, M. Winstanley, V. Marsh, C. Newton, P. Winstanley, P. Warn, N. Peshu and et al. (1995). "Indicators of life-threatening malaria in African children." <u>The New England Journal of Medicine</u> **332**(21): 1399-1404.

Maubon, D., D. Ajzenberg, M. P. Brenier-Pinchart, M. L. Darde and H. Pelloux (2008). "What are the respective host and parasite contributions to toxoplasmosis?" <u>Trends in Parasitology</u> **24**(7): 299-303.

McCoy, A. J., R. W. Grosse-Kunstleve, P. D. Adams, M. D. Winn, L. C. Storoni and R. J. Read (2007). "Phaser crystallographic software." <u>Journal of Applied Crystallography</u> **40**(Pt 4): 658-674.

McMorrow, M. L., M. Aidoo and S. P. Kachur (2011). "Malaria rapid diagnostic tests in elimination settings--can they find the last parasite?" Clinical Microbiology and Infection **17**(11): 1624-1631.

Meira, C. S., J. E. Vidal, T. A. Costa-Silva, N. Frazatti-Gallina and V. L. Pereira-Chioccola (2011). "Immunodiagnosis in cerebrospinal fluid of cerebral toxoplasmosis and HIV-infected patients using *Toxoplasma gondii* excreted/secreted antigens." Diagnostic Microbiology and Infectious Disease **71**(3): 279-285.

Michel, R., K. Schupp, W. Raether and F. W. Bierther (1980). "Formation of a close junction during invasion of erythrocytes by *Toxoplasma gondii* in vitro." International Journal for Parasitology **10**(4): 309-313.

Miller, L. H., D. I. Baruch, K. Marsh and O. K. Doumbo (2002). "The pathogenic basis of malaria." Nature **415**(6872): 673-679.

Miller, M. A., W. A. Miller, P. A. Conrad, E. R. James, A. C. Melli, C. M. Leutenegger, H. A. Dabritz, A. E. Packham, D. Paradies, M. Harris, J. Ames, D. A. Jessup, K. Worcester and M. E. Grigg (2008). "Type X *Toxoplasma gondii* in a wild mussel and terrestrial carnivores from coastal California: new linkages between terrestrial mammals, runoff and toxoplasmosis of sea otters." International Journal for Parasitology **38**(11): 1319-1328.

Mital, J., M. Meissner, D. Soldati and G. E. Ward (2005). "Conditional expression of *Toxoplasma gondii* apical membrane antigen-1 (*Tg*AMA1) demonstrates that *Tg*AMA1 plays a critical role in host cell invasion." Molecular Biology of the Cell **16**(9): 4341-4349.

Mitchell, G. H., A. W. Thomas, G. Margos, A. R. Dluzewski and L. H. Bannister (2004). "Apical membrane antigen 1, a major malaria vaccine candidate, mediates the close attachment of invasive merozoites to host red blood cells." Infection and Immunity **72**(1): 154-158.

Miura, K., R. Herrera, A. Diouf, H. Zhou, J. Mu, Z. Hu, N. J. MacDonald, K. Reiter, V. Nguyen, R. L. Shimp, Jr., K. Singh, D. L. Narum, C. A. Long and L. H. Miller (2013). "Overcoming allelic specificity by immunization with five allelic forms of *Plasmodium falciparum* apical membrane antigen 1." Infection and Immunity **81**(5): 1491-1501.

Montero, E., M. Rodriguez, Y. Oksov and C. A. Lobo (2009). "*Babesia divergens* apical membrane antigen 1 and its interaction with the human red blood cell." Infection and Immunity **77**(11): 4783-4793.

Mordue, D. G., N. Desai, M. Dustin and L. D. Sibley (1999). "Invasion by *Toxoplasma gondii* establishes a moving junction that selectively excludes host cell plasma membrane proteins on the basis of their membrane anchoring." The Journal of Experimental Medicine **190**(12): 1783-1792.

Mordue, D. G., S. Hakansson, I. Niesman and L. D. Sibley (1999). "*Toxoplasma gondii* resides in a vacuole that avoids fusion with host cell endocytic and exocytic vesicular trafficking pathways." Experimental Parasitology **92**(2): 87-99.

Morisaki, J. H., J. E. Heuser and L. D. Sibley (1995). "Invasion of *Toxoplasma gondii* occurs by active penetration of the host cell." Journal of Cell Science **108**(Pt 6): 2457-2464.

Mullis, K. B. and F. A. Faloona (1987). "Specific synthesis of DNA in vitro via a polymerase-catalyzed chain reaction." Methods in enzymology **155**: 335-350.

Murshudov, G. N., A. A. Vagin and E. J. Dodson (1997). "Refinement of macromolecular structures by the maximum-likelihood method." Acta Crystallographica Section D: Biological Crystallography **53**(Pt 3): 240-255.

Nair, M., M. G. Hinds, A. M. Coley, A. N. Hodder, M. Foley, R. F. Anders and R. S. Norton (2002). "Structure of domain III of the blood-stage malaria vaccine candidate, *Plasmodium falciparum* apical membrane antigen 1 (AMA1)." Journal of Molecular Biology **322**(4): 741-753.

Narum, D. L., V. Nguyen, Y. Zhang, J. Glen, R. L. Shimp, L. Lambert, I. T. Ling, K. Reiter, S. A. Ogun, C. Long, A. A. Holder and R. Herrera (2008). "Identification and characterization of the *Plasmodium yoelii Py*P140/RON4 protein, an orthologue of *Toxoplasma gondii* RON4, whose cysteine-rich domain does not protect against lethal parasite challenge infection." Infection and Immunity **76**(11): 4876-4882.

Narum, D. L., S. A. Ogun, A. W. Thomas and A. A. Holder (2000). "Immunization with parasite-derived apical membrane antigen 1 or passive immunization with a specific monoclonal antibody protects BALB/c mice against lethal *Plasmodium yoelii yoelii* YM blood-stage infection." Infection and Immunity **68**(5): 2899-2906.

Narum, D. L. and A. W. Thomas (1994). "Differential localization of full-length and processed forms of PF83/AMA-1 an apical membrane antigen of *Plasmodium falciparum* merozoites." Molecular and Biochemical Parasitology **67**(1): 59-68.

Navaza, J. (2001). "Implementation of molecular replacement in AMoRe." Acta Crystallographica Section D: Biological Crystallography **57**(Pt 10): 1367-1372.

Nichols, B. A., M. L. Chiappino and G. R. O'Connor (1983). "Secretion from the rhoptries of *Toxoplasma gondii* during host-cell invasion." Journal of Ultrastructure Research **83**(1): 85-98.

Oakes, R. D., D. Kurian, E. Bromley, C. Ward, K. Lal, D. P. Blake, A. J. Reid, A. Pain, R. E. Sinden, J. M. Wastling and F. M. Tomley (2013). "The rhoptry proteome of *Eimeria tenella* sporozoites." International Journal for Parasitology **43**(2): 181-188.

Olotu, A., G. Fegan, J. Wambua, G. Nyangweso, K. O. Awuondo, A. Leach, M. Lievens, D. Leboulleux, P. Njuguna, N. Peshu, K. Marsh and P. Bejon (2013). "Four-year efficacy of RTS,S/AS01E and its interaction with malaria exposure." The New England Journal of Medicine **368**(12): 1111-1120.

Ouattara, A., J. Mu, S. Takala-Harrison, R. Saye, I. Sagara, A. Dicko, A. Niangaly, J. Duan, R. D. Ellis, L. H. Miller, X. Z. Su, C. V. Plowe and O. K. Doumbo (2010). "Lack of allele-specific efficacy of a bivalent AMA1 malaria vaccine." Malaria Journal **9**: 175.

Pappas, G., N. Roussos and M. E. Falagas (2009). "Toxoplasmosis snapshots: global status of *Toxoplasma gondii* seroprevalence and implications for pregnancy and congenital toxoplasmosis." International Journal for Parasitology **39**(12): 1385-1394.

Petersen, E. and O. Liesenfeld (2007). Clinical Disease and Diagnostics. *Toxoplasma gondii* - The model apicomplexan: perspectives and methods. L. M. Weiss and K. Kim. London, Academic Press - Elsevier**:** 81-100.

Petersen, I., R. Eastman and M. Lanzer (2011). "Drug-resistant malaria: molecular mechanisms and implications for public health." FEBS Letters **585**(11): 1551-1562.

Peterson, M. G., V. M. Marshall, J. A. Smythe, P. E. Crewther, A. Lew, A. Silva, R. F. Anders and D. J. Kemp (1989). "Integral membrane protein located in the apical complex of *Plasmodium falciparum*." Molecular and Cellular Biology **9**(7): 3151-3154.

Pizarro, J. C., B. Vulliez-Le Normand, M. L. Chesne-Seck, C. R. Collins, C. Withers-Martinez, F. Hackett, M. J. Blackman, B. W. Faber, E. J. Remarque, C. H. Kocken, A. W. Thomas and G. A. Bentley (2005). "Crystal structure of the malaria vaccine candidate apical membrane antigen 1." Science **308**(5720): 408-411.

Polley, S. D. and D. J. Conway (2001). "Strong diversifying selection on domains of the *Plasmodium falciparum* apical membrane antigen 1 gene." Genetics **158**(4): 1505-1512.

Polley, S. D., T. Mwangi, C. H. Kocken, A. W. Thomas, S. Dutta, D. E. Lanar, E. Remarque, A. Ross, T. N. Williams, G. Mwambingu, B. Lowe, D. J. Conway and K. Marsh (2004). "Human antibodies to recombinant protein constructs of *Plasmodium falciparum* Apical Membrane Antigen 1 (AMA1) and their associations with protection from malaria." Vaccine **23**(5): 718-728.

Possenti, A., F. Fratini, L. Fantozzi, E. Pozio, J. P. Dubey, M. Ponzi, E. Pizzi and F. Spano (2013). "Global proteomic analysis of the oocyst/sporozoite of *Toxoplasma gondii* reveals commitment to a host-independent lifestyle." BMC Genomics **14**: 183.

Poukchanski, A., H. M. Fritz, M. L. Tonkin, M. Treeck, M. J. Boulanger and J. C. Boothroyd (2013). "*Toxoplasma gondii* sporozoites invade host cells using two novel paralogues of RON2 and AMA1." PLoS One **8**(8): e70637.

Proellocks, N. I., R. L. Coppel and K. L. Waller (2010). "Dissecting the apicomplexan rhoptry neck proteins." Trends in Parasitology **26**(6): 297-304.

Raymond, S. and L. Weintraub (1959). "Acrylamide gel as a supporting medium for zone electrophoresis." Science **130**(3377): 711.

Reed, M. B., S. R. Caruana, A. H. Batchelor, J. K. Thompson, B. S. Crabb and A. F. Cowman (2000). "Targeted disruption of an erythrocyte binding antigen in *Plasmodium falciparum* is associated with a switch toward a sialic acid-independent pathway of invasion." Proceedings of the National Academy of Sciences of the United States of America **97**(13): 7509-7514.

Reese, M. L., G. M. Zeiner, J. P. Saeij, J. C. Boothroyd and J. P. Boyle (2011). "Polymorphic family of injected pseudokinases is paramount in *Toxoplasma* virulence." Proceedings of the National Academy of Sciences of the United States of America **108**(23): 9625-9630.

Remarque, E. J., B. W. Faber, C. H. Kocken and A. W. Thomas (2008). "Apical membrane antigen 1: a malaria vaccine candidate in review." Trends in Parasitology **24**(2): 74-84.

Remarque, E. J., B. W. Faber, C. H. Kocken and A. W. Thomas (2008). "A diversity-covering approach to immunization with *Plasmodium falciparum* apical membrane antigen 1 induces broader allelic recognition and growth inhibition responses in rabbits." Infection and Immunity **76**(6): 2660-2670.

Richard, D., C. A. MacRaild, D. T. Riglar, J. A. Chan, M. Foley, J. Baum, S. A. Ralph, R. S. Norton and A. F. Cowman (2010). "Interaction between *Plasmodium falciparum* apical membrane antigen 1 and the rhoptry neck protein complex defines a key step in the erythrocyte invasion process of malaria parasites." Journal of Biological Chemistry **285**(19): 14815-14822.

Riglar, D. T., D. Richard, D. W. Wilson, M. J. Boyle, C. Dekiwadia, L. Turnbull, F. Angrisano, D. S. Marapana, K. L. Rogers, C. B. Whitchurch, J. G. Beeson, A. F. Cowman, S. A. Ralph and J. Baum (2011). "Super-resolution dissection of coordinated events during malaria parasite invasion of the human erythrocyte." Cell Host and Microbe **9**(1): 9-20.

Robben, P. M., D. G. Mordue, S. M. Truscott, K. Takeda, S. Akira and L. D. Sibley (2004). "Production of IL-12 by macrophages infected with *Toxoplasma gondii* depends on the parasite genotype." The Journal of Immunology **172**(6): 3686-3694.

Robert-Gangneux, F. and M. L. Darde (2012). "Epidemiology of and diagnostic strategies for toxoplasmosis." Clinical Microbiology Reviews **25**(2): 264-296.

Roberts, F., A. Kuo, L. Jones, R. McLeod and C. W. Roberts (2007). Ocular Toxoplasmosis: clinical features, pathology, pathogenesis, animal models, and immune responses. Toxoplasma: molecular and cellular biology. J. W. Ajioka and S. Soldati. Norfolk, UK, Horizon Bioscience**:** 59 - 91.

Ruff, M. D. (1999). "Important parasites in poultry production systems." Veterinary Parasitology **84**(3-4): 337-347.

Ruwende, C., S. C. Khoo, R. W. Snow, S. N. Yates, D. Kwiatkowski, S. Gupta, P. Warn, C. E. Allsopp, S. C. Gilbert, N. Peschu and et al. (1995). "Natural selection of hemi- and

heterozygotes for G6PD deficiency in Africa by resistance to severe malaria." Nature **376**(6537): 246-249.

Saeij, J. P., J. P. Boyle and J. C. Boothroyd (2005). "Differences among the three major strains of *Toxoplasma gondii* and their specific interactions with the infected host." Trends in Parasitology **21**(10): 476-481.

Saeij, J. P., J. P. Boyle, S. Coller, S. Taylor, L. D. Sibley, E. T. Brooke-Powell, J. W. Ajioka and J. C. Boothroyd (2006). "Polymorphic secreted kinases are key virulence factors in toxoplasmosis." Science **314**(5806): 1780-1783.

Santos, J. M., D. J. Ferguson, M. J. Blackman and D. Soldati-Favre (2011). "Intramembrane cleavage of AMA1 triggers *Toxoplasma* to switch from an invasive to a replicative mode." Science **331**(6016): 473-477.

Scallan, E., R. M. Hoekstra, F. J. Angulo, R. V. Tauxe, M. A. Widdowson, S. L. Roy, J. L. Jones and P. M. Griffin (2011). "Foodborne illness acquired in the United States--major pathogens." Emerging Infectious Diseases **17**(1): 7-15.

Schneider, C.A., W. S. Rasband, K. W. Eliceiri (2012). "NIH Image to ImageJ: 25 years of image analysis." Nature Methods **9**(7): 671-675.

Schubert, W. D., C. Urbanke, T. Ziehm, V. Beier, M. P. Machner, E. Domann, J. Wehland, T. Chakraborty and D. W. Heinz (2002). "Structure of internalin, a major invasion protein of *Listeria monocytogenes*, in complex with its human receptor E-cadherin." Cell **111**(6): 825-836.

Schwarzenbacher, R., A. Godzik, S. K. Grzechnik and L. Jaroszewski (2004). "The importance of alignment accuracy for molecular replacement." Acta Crystallographica Section D: Biological Crystallography **60**(Pt 7): 1229-1236.

Shaw, M. K. (2003). "Cell invasion by *Theileria* sporozoites." Trends in Parasitology **19**(1): 2-6.

Sheiner, L., J. M. Santos, N. Klages, F. Parussini, N. Jemmely, N. Friedrich, G. E. Ward and D. Soldati-Favre (2010). "*Toxoplasma gondii* transmembrane microneme proteins and their modular design." Molecular Microbiology **77**(4): 912-929.

Shen, B. and L. D. Sibley (2012). "The moving junction, a key portal to host cell invasion by apicomplexan parasites." Current Opinion in Microbiology **15**(4): 449-455.

Shen, B. and L. D. Sibley (2014). "*Toxoplasma* aldolase is required for metabolism but dispensable for host-cell invasion." Proceedings of the National Academy of Sciences of the United States of America **111**(9): 3567-3572.

Sibley, L. D. (2010). "How apicomplexan parasites move in and out of cells." Current Opinion in Biotechnology **21**(5): 592-598.

Silvie, O., J. F. Franetich, S. Charrin, M. S. Mueller, A. Siau, M. Bodescot, E. Rubinstein, L. Hannoun, Y. Charoenvit, C. H. Kocken, A. W. Thomas, G. J. Van Gemert, R. W. Sauerwein, M. J. Blackman, R. F. Anders, G. Pluschke and D. Mazier (2004). "A role for apical membrane antigen 1 during invasion of hepatocytes by *Plasmodium falciparum* sporozoites." Journal of Biological Chemistry **279**(10): 9490-9496.

Singh, S., M. M. Alam, I. Pal-Bhowmick, J. A. Brzostowski and C. E. Chitnis (2010). "Distinct external signals trigger sequential release of apical organelles during erythrocyte invasion by malaria parasites." PLoS Pathogens **6**(2): e1000746.

Snounou, G., S. Viriyakosol, W. Jarra, S. Thaithong and K. N. Brown (1993). "Identification of the four human malaria parasite species in field samples by the polymerase chain reaction and detection of a high prevalence of mixed infections." Molecular and Biochemical Parasitology **58**(2): 283-292.

Snow, R. W., C. A. Guerra, A. M. Noor, H. Y. Myint and S. I. Hay (2005). "The global distribution of clinical episodes of *Plasmodium falciparum* malaria." Nature **434**(7030): 214-217.

Soldati, D. and J. C. Boothroyd (1993). "Transient transfection and expression in the obligate intracellular parasite *Toxoplasma gondii*." Science **260**(5106): 349-352.

Speer, C. A., M. Tilley, M. E. Temple, J. A. Blixt, J. P. Dubey and M. W. White (1995). "Sporozoites of *Toxoplasma gondii* lack dense-granule protein GRA3 and form a unique parasitophorous vacuole." Molecular and Biochemical Parasitology **75**(1): 75-86.

Spring, M. D., J. F. Cummings, C. F. Ockenhouse, S. Dutta, R. Reidler, E. Angov, E. Bergmann-Leitner, V. A. Stewart, S. Bittner, L. Juompan, M. G. Kortepeter, R. Nielsen, U. Krzych, E. Tierney, L. A. Ware, M. Dowler, C. C. Hermsen, R. W. Sauerwein, S. J. de Vlas, O. Ofori-Anyinam, D. E. Lanar, J. L. Williams, K. E. Kester, K. Tucker, M. Shi, E. Malkin, C. Long, C. L. Diggs, L. Soisson, M. C. Dubois, W. R. Ballou, J. Cohen and D. G. Heppner, Jr. (2009). "Phase 1/2a study of the malaria vaccine candidate apical membrane antigen-1 (AMA-1) administered in adjuvant system AS01B or AS02A." PloS One **4**(4): e5254.

Srinivasan, P., W. L. Beatty, A. Diouf, R. Herrera, X. Ambroggio, J. K. Moch, J. S. Tyler, D. L. Narum, S. K. Pierce, J. C. Boothroyd, J. D. Haynes and L. H. Miller (2011). "Binding of *Plasmodium* merozoite proteins RON2 and AMA1 triggers commitment to invasion." Proceedings of the National Academy of Sciences of the United States of America **108**(32): 13275-13280.

Srinivasan, P., A. Yasgar, D. K. Luci, W. L. Beatty, X. Hu, J. Andersen, D. L. Narum, J. K. Moch, H. Sun, J. D. Haynes, D. J. Maloney, A. Jadhav, A. Simeonov and L. H. Miller (2013). "Disrupting malaria parasite AMA1-RON2 interaction with a small molecule prevents erythrocyte invasion." Nature Communications **4**: 2261.

Stowers, A. W., M. C. Kennedy, B. P. Keegan, A. Saul, C. A. Long and L. H. Miller (2002). "Vaccination of monkeys with recombinant *Plasmodium falciparum* apical membrane

antigen 1 confers protection against blood-stage malaria." Infection and Immunity **70**(12): 6961-6967.

Straub, K. W., S. J. Cheng, C. S. Sohn and P. J. Bradley (2009). "Novel components of the Apicomplexan moving junction reveal conserved and coccidia-restricted elements." Cellular Microbiology **11**(4): 590-603.

Suh, K. N., K. C. Kain and J. S. Keystone (2004). "Malaria." Canadian Medical Association Journal **170**(11): 1693-1702.

Suss-Toby, E., J. Zimmerberg and G. E. Ward (1996). "*Toxoplasma* invasion: the parasitophorous vacuole is formed from host cell plasma membrane and pinches off via a fission pore." Proceedings of the National Academy of Sciences of the United States of America **93**(16): 8413-8418.

Tamminga, C., M. Sedegah, S. Maiolatesi, C. Fedders, S. Reyes, A. Reyes, C. Vasquez, Y. Alcorta, I. Chuang, M. Spring, M. Kavanaugh, H. Ganeshan, J. Huang, M. Belmonte, E. Abot, A. Belmonte, J. Banania, F. Farooq, J. Murphy, J. Komisar, N. O. Richie, J. Bennett, K. Limbach, N. B. Patterson, J. T. Bruder, M. Shi, E. Miller, S. Dutta, C. Diggs, L. A. Soisson, M. R. Hollingdale, J. E. Epstein and T. L. Richie (2013). "Human adenovirus 5-vectored *Plasmodium falciparum* NMRC-M3V-Ad-PfCA vaccine encoding CSP and AMA1 is safe, well-tolerated and immunogenic but does not protect against controlled human malaria infection." Human Vaccines and Immunotherapeutics **9**(10): 2165-2177.

Tan, K. R., A. J. Magill, M. E. Parise, P. M. Arguin, C. Centers for Disease and Prevention (2011). "Doxycycline for malaria chemoprophylaxis and treatment: report from the CDC expert meeting on malaria chemoprophylaxis." The American Journal of Tropical Medicine and Hygiene **84**(4): 517-531.

Taylor, S., A. Barragan, C. Su, B. Fux, S. J. Fentress, K. Tang, W. L. Beatty, H. E. Hajj, M. Jerome, M. S. Behnke, M. White, J. C. Wootton and L. D. Sibley (2006). "A secreted serine-threonine kinase determines virulence in the eukaryotic pathogen *Toxoplasma gondii*." Science **314**(5806): 1776-1780.

Taylor, T. E., A. Borgstein and M. E. Molyneux (1993). "Acid-base status in paediatric *Plasmodium falciparum* malaria." Quaterly Journal of Medicine **86**(2): 99-109.

Tenter, A. M., A. R. Heckeroth and L. M. Weiss (2000). "*Toxoplasma gondii*: from animals to humans." International Journal for Parasitology **30**(12-13): 1217-1258.

Thera, M. A., O. K. Doumbo, D. Coulibaly, M. B. Laurens, A. Ouattara, A. K. Kone, A. B. Guindo, K. Traore, I. Traore, B. Kouriba, D. A. Diallo, I. Diarra, M. Daou, A. Dolo, Y. Tolo, M. S. Sissoko, A. Niangaly, M. Sissoko, S. Takala-Harrison, K. E. Lyke, Y. Wu, W. C. Blackwelder, O. Godeaux, J. Vekemans, M. C. Dubois, W. R. Ballou, J. Cohen, D. Thompson, T. Dube, L. Soisson, C. L. Diggs, B. House, D. E. Lanar, S. Dutta, D. G. Heppner, Jr. and C. V. Plowe (2011). "A field trial to assess a blood-stage malaria vaccine." The New England Journal of Medicine **365**(11): 1004-1013.

Thomas, A. W., J. A. Deans, G. H. Mitchell, T. Alderson and S. Cohen (1984). "The Fab fragments of monoclonal IgG to a merozoite surface antigen inhibit *Plasmodium knowlesi* invasion of erythrocytes." Molecular and Biochemical Parasitology **13**(2): 187-199.

Thompson, J. D., D. G. Higgins and T. J. Gibson (1994). "CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice." Nucleic Acids Research **22**(22): 4673-4680.

Tilley, M., M. E. Fichera, M. E. Jerome, D. S. Roos and M. W. White (1997). "*Toxoplasma gondii* sporozoites form a transient parasitophorous vacuole that is impermeable and contains only a subset of dense-granule proteins." Infection and Immunity **65**(11): 4598-4605.

Tonkin, M. L., J. Crawford, M. L. Lebrun and M. J. Boulanger (2013). "*Babesia divergens* and *Neospora caninum* apical membrane antigen 1 structures reveal selectivity and plasticity in apicomplexan parasite host cell invasion." Protein Science **22**(1): 114-127.

Tonkin, M. L., M. Roques, M. H. Lamarque, M. Pugniere, D. Douguet, J. Crawford, M. Lebrun and M. J. Boulanger (2011). "Host cell invasion by apicomplexan parasites: insights from the co-structure of AMA1 with a RON2 peptide." Science **333**(6041): 463-467.

Tordai, H., L. Banyai and L. Patthy (1999). "The PAN module: the N-terminal domains of plasminogen and hepatocyte growth factor are homologous with the apple domains of the prekallikrein family and with a novel domain found in numerous nematode proteins." FEBS Letters **12**(461): 63-67.

Torrey, E. F. and R. H. Yolken (2003). "*Toxoplasma gondii* and schizophrenia." Emerging Infectious Diseases **9**(11): 1375-1380.

Treeck, M., J. L. Sanders, J. E. Elias and J. C. Boothroyd (2011). "The phosphoproteomes of *Plasmodium falciparum* and *Toxoplasma gondii* reveal unusual adaptations within and beyond the parasites' boundaries." Cell Host and Microbe **10**(4): 410-419.

Treeck, M., M. Tamborrini, C. A. Daubenberger, T. W. Gilberger and T. S. Voss (2009). "Caught in action: mechanistic insights into antibody-mediated inhibition of *Plasmodium* merozoite invasion." Trends in Parasitology **25**(11): 494-497.

Treeck, M., S. Zacherl, S. Herrmann, A. Cabrera, M. Kono, N. S. Struck, K. Engelberg, S. Haase, F. Frischknecht, K. Miura, T. Spielmann and T. W. Gilberger (2009). "Functional analysis of the leading malaria vaccine candidate AMA-1 reveals an essential role for the cytoplasmic domain in the invasion process." PLoS Pathogens **5**(3): e1000322.

Triglia, T., J. Healer, S. R. Caruana, A. N. Hodder, R. F. Anders, B. S. Crabb and A. F. Cowman (2000). "Apical membrane antigen 1 plays a central role in erythrocyte invasion by *Plasmodium* species." Molecular Microbiology **38**(4): 706-718.

Tyler, J. S. and J. C. Boothroyd (2011). "The C-terminus of *Toxoplasma* RON2 provides the crucial link between AMA1 and the host-associated invasion complex." PLoS Pathogens **7**(2): e1001282.

Tyler, J. S., M. Treeck and J. C. Boothroyd (2011). "Focus on the ringleader: the role of AMA1 in apicomplexan invasion and replication." Trends in Parasitology **27**(9): 410-420.

Urquiza, M., J. E. Suarez, C. Cardenas, R. Lopez, A. Puentes, F. Chavez, J. C. Calvo and M. E. Patarroyo (2000). "*Plasmodium falciparum* AMA-1 erythrocyte binding peptides implicate AMA-1 as erythrocyte binding protein." Vaccine **19**(4-5): 508-513.

Valbuena, J., L. Rodriguez, R. Vera, A. Puentes, H. Curtidor, J. Cortes, J. Rosas and M. E. Patarroyo (2006). "Synthetic peptides from *Plasmodium falciparum* apical membrane antigen 1 (AMA-1) specifically interacting with human hepatocytes." Biochimie **88**(10): 1447-1455.

Vinson, V. (2011). "Built to Invade." Science Signalling **4**(183): ec209.

Vulliez-Le Normand, B., M. L. Tonkin, M. H. Lamarque, S. Langer, S. Hoos, M. Roques, F. A. Saul, B. W. Faber, G. A. Bentley, M. J. Boulanger and M. Lebrun (2012). "Structural and functional insights into the malaria parasite moving junction complex." PLoS Pathogens **8**(6): e1002755.

Wallner, B. and A. Elofsson (2003). "Can correct protein models be identified?" Protein Science **12**(5): 1073-1086.

Ward, G. E., L. H. Miller and J. A. Dvorak (1993). "The origin of parasitophorous vacuole membrane lipids in malaria-infected erythrocytes." Journal of Cell Science **106**(Pt 1): 237-248.

Waters, A. P., A. W. Thomas, J. A. Deans, G. H. Mitchell, D. E. Hudson, L. H. Miller, T. F. McCutchan and S. Cohen (1990). "A merozoite receptor protein from *Plasmodium knowlesi* is highly conserved and distributed throughout *Plasmodium*." Journal of Biological Chemistry **265**(29): 17974-17979.

Webster, J. P. (2007). "The effect of *Toxoplasma gondii* on animal behavior: playing cat and mouse." Schizophrenia Bulletin **33**(3): 752-756.

Winn, M. D., C. C. Ballard, K. D. Cowtan, E. J. Dodson, P. Emsley, P. R. Evans, R. M. Keegan, E. B. Krissinel, A. G. Leslie, A. McCoy, S. J. McNicholas, G. N. Murshudov, N. S. Pannu, E. A. Potterton, H. R. Powell, R. J. Read, A. Vagin and K. S. Wilson (2011). "Overview of the CCP4 suite and current developments." Acta Crystallographica Section D: Biological Crystallography **67**(Pt 4): 235-242.

Winter, G., A. R. Fersht, A. J. Wilkinson, M. Zoller and M. Smith (1982). "Redesigning enzyme structure by site-directed mutagenesis: tyrosyl tRNA synthetase and ATP binding." Nature **299**(5885): 756-758.

World Health Organization. (2010). Guidelines for the treatment of malaria. Appia, Geneva, World Health Organization Press.

World Health Organization. (2010). World malaria report. Geneva, Switzerland, World Health Organization.

Yan, J., B. Huang, G. Liu, B. Wu, S. Huang, H. Zheng, J. Shen, Z. R. Lun, Y. Wang and F. Lu (2013). "Meta-analysis of prevention and treatment of toxoplasmic encephalitis in HIV-infected patients." Acta Tropica **127**(3): 236-244.

Yolken, R. H. and E. F. Torrey (2008). "Are some cases of psychosis caused by microbial agents? A review of the evidence." Molecular Psychiatry **13**(5): 470-479.

Zhang, H., M. K. Compaore, E. G. Lee, M. Liao, G. Zhang, C. Sugimoto, K. Fujisaki, Y. Nishikawa and X. Xuan (2007). "Apical membrane antigen 1 is a cross-reactive antigen between *Neospora caninum* and *Toxoplasma gondii*, and the anti-*Nc*AMA1 antibody inhibits host cell invasion by both parasites." Molecular and Biochemical Parasitology **151**(2): 205-212.

Zintl, A., G. Mulcahy, H. E. Skerrett, S. M. Taylor and J. S. Gray (2003). "*Babesia divergens*, a bovine blood parasite of veterinary and zoonotic importance." Clinical Microbiology Reviews **16**(4): 622-636.

Zuccala, E. S., A. M. Gout, C. Dekiwadia, D. S. Marapana, F. Angrisano, L. Turnbull, D. T. Riglar, K. L. Rogers, C. B. Whitchurch, S. A. Ralph, T. P. Speed and J. Baum (2012). "Subcompartmentalisation of proteins in the rhoptries correlates with ordered events of erythrocyte invasion by the blood stage malaria parasite." PLoS One **7**(9): e46160.

Appendix

Copyright permissions for manuscripts adapted in this dissertation: