Published in final edited form as:

Biochim Biophys Acta. 2017 February ; 1865(2): 208–219. doi:10.1016/j.bbapap.2016.11.010.

***Toxoplasma gondii* Immune Mapped Protein 1 is anchored to the inner leaflet of the plasma membrane and adopts a novel protein fold**

Yonggen Jia^{#1}, Stefi Benjamin^{#2}, Qun Liu³, Yingqi Xu², Sunil Kumar Dogga¹, Jing Liu³, Stephen Matthews^{2,*}, and Dominique Soldati-Favre^{1,*}

¹Department of Microbiology and Molecular Medicine, CMU, University of Geneva, Rue Michel-Servet 1, CH-1211 Geneva

²Department of Life Sciences, Imperial College London, South Kensington, London, SW7 2AZ UK

³Key Laboratory of Animal Epidemiology and Zoonosis, Ministry of Agriculture, College of Veterinary Medicine, China Agricultural University, Beijing 100193, China

[#] These authors contributed equally to this work.

Abstract

Immune mapped protein 1 (IMP1) was first identified as a protective antigen in *Eimeria maxima* and described as vaccine candidate and invasion factor in *Toxoplasma gondii*. We show here that TgIMP1 localizes to the inner leaflet of plasma membrane via dual acylation. Mutations either in the N-terminal myristoylation or palmitoylation (G2 and C5) sites relocalize TgIMP1 to the cytosol. The first 11 amino acids are sufficient for PM targeting and the presence of lysine (K7) is critical. Disruption of *TgIMP1* gene by double homologous recombination revealed no invasion defect or any measurable alteration in the lytic cycle of tachyzoites. Following immunization with TgIMP1 DNA vaccine, mice challenged with either wild type or IMP1-ko parasites showed no significant difference in protection. The sequence analysis identified a structured C-terminal domain that is present in a broader family of IMP1-like proteins conserved across the Apicomplexa phylum. We present the solution structure of this domain determined from NMR data and describe a new protein fold not seen before.

Keywords

Toxoplasma gondii; acylation; invasion; vaccine; NMR; protein structure

Introduction

The phylum of Apicomplexa comprises medical and veterinary important protozoan parasites including *Plasmodium*, *Toxoplasma*, *Eimeria*, *Neospora*, *Scarcocystis*, *Babesia*, *Theileria* and *Cryptosporidium* species. The limited repertoire of efficacious drugs and the emergence of drug resistance have considerably hampered the control of these parasites and

*Corresponding authors: dominique.soldati-Favre@unige.ch, s.j.matthews@imperial.ac.uk.

boosted the research towards the development of vaccines [1]. In this context, an Immune mapped protein 1 (IMP1) was identified as an antigen eliciting protective immunity against chicken coccidian *Eimeria maxima*, by combining parasite genetics and selective barriers with population-based genetic fingerprinting [2]. Recent studies in *Toxoplasma gondii* and *Neospora caninum* reported that TgIMP1 and NcIMP1 provided protection following immunization with DNA vaccine in the mouse model [3, 4]. IMP1 lacks conserved domains with known function except for the predicted myristoylation and palmitoylation sites at the N-terminus, which are presumed to confer membrane association to the protein. Antibodies raised against NcIMP1 localized it to the plasma membrane (PM) and were shown to have inhibitory effects on host cell invasion [4].

In this study, we show that the acylation sites in the N-terminus of TgIMP1 are involved in anchoring the protein to the cytoplasmic face of the plasma membrane. Given this internal localization, we have revisited the reported roles of IMP1 in invasion and as vaccine candidate by generating parasites lacking TgIMP1. Furthermore we have identified a conserved globular domain at the C-terminus TgIMP1 that represents a broader protein family present in all Apicomplexans. We also elucidate the high-resolution structure by solution state NMR and reveal a new protein fold.

Results

IMP1 possesses a globular C-terminal domain conserved across all Apicomplexans

Bioinformatic analysis indicated that IMP1 is restricted to the coccidian-subgroup of the Apicomplexa phylum, but belongs to a broader family of proteins that include closely related IMP1-like proteins that group into six distinct phylogenetic clusters (IMP1 to IMP6). (Fig. 1). The amino acids sequences alignment of all proteins used for the phylogeny is presented in Fig S1. TgIMP1 and other members of the IMP1 cluster possess conserved N-terminal myristoylation and palmitoylation sites followed by an extended linker region and a conserved C-terminal domain (termed the IMP1-like domain). In contrast, members in other groups lack the residues predicted for N-terminal acylation. While the proteins clustering in IMP1, 5 and 6 are found restricted to the coccidian sub-group of Apicomplexa, the members of IMP4 are present across all members of the phylum except in *Cryptosporidium species*. The conservation of the globular domain across the Apicomplexans defines this broader family of IMP1-like proteins and may share conserved or similar functions.

IMP1 is localized to the inner leaflet of the plasma membrane in *Toxoplasma gondii*

In eukaryotic cells, protein post-translational modification by dual acylation such as myristoylation and palmitoylation, ensures membrane association for those proteins lacking a transmembrane domain. The conserved glycine followed by cysteine residues (MGXXCS/T) at the N-terminus of IMP1 implies that this protein is likely membrane-bound through dual acylation (myristoylation and palmitoylation). To determine the subcellular distribution of TgIMP1 in tachyzoites, a second copy coding for the C-terminally epitope-tagged TgIMP1 was expressed stably under the control of a tubulin promoter. Additionally, to detect the expression and localization of endogenous TgIMP1, a parasite line was obtained in which a triple Ty-tag was inserted at the C-terminus of the gene by single

homologous recombination at the endogenous *TgIMP1* locus in RH ku80-ko strain. Expression of endogenous, TgIMP1, TgIMP1-3Ty-KI (knock-in) and second copy TubTgIMP1-Ty assessed by western blot analysis using either rabbit anti-IMP1 [3] or mouse anti-Ty antibodies (Fig. 2A).

On SDS-PAGE, TubTgIMP1-Ty and endogenous TgIMP1-3Ty migrate considerably slower than the predicted molecular mass (44.9 kDa and 47.8 kDa, respectively), possibly due to the post-modification and a relative abundance of negatively charged amino acids [5]. Indirect immunofluorescence assay (IFA) established that TubTgIMP1-Ty is predominantly localized to the parasite pellicle (Fig. 2B). Despite a weak endogenous level of expression, IFA unambiguously confirmed the same pellicular targeting (Fig. 2B lower panel).

In Alveolates including the Apicomplexa, the pellicle is composed of a set of flattened cisternae forming the inner membrane complex (IMC) underneath the plasma membrane (PM) [6, 7]. TgIMP1 was not detected in the forming IMC of the nascent daughter cells during the process parasite division, suggesting that the protein is anchored into the PM[8]. To confirm that TgIMP1 indeed localized to the PM rather than the IMC, fresh egressed extracellular parasites were treated with aerolysin. This vacuolating toxin dramatically separates the PM from the underlying IMC, enabling differential localization of these closely apposed membrane systems [9, 10]. Following aerolysin-treatment, TgIMP1-Ty was separated from GAP45 but colocalized with the GPI-anchored SAG1, two markers that are used for IMC and PM, respectively (Fig. 2C).

To determine if TgIMP1 was localized to the extracellular or cytoplasmic face of the PM, freshly egressed parasites were incubated with trypsin. On intact parasites, trypsin could only access and digest polypeptides exposed on parasite surface. In contrast to the surface protein SAG1, TgIMP1, as well as the intracellular protein profilin, were fully resistant to the proteinase treatment demonstrating that the TgIMP1 is anchored to inner surface of the PM (Fig. 2D).

N-terminal acylation is required for plasma membrane localization of TgIMP1

All IMP1 homologues in Coccidia harbor putative myristoylation and palmitoylation motifs at their N-termini (Fig. 3A). CSS-Palm 4.0 (<http://csspalm.biocuckoo.org/online.php>) server was used to predict palmitoylation. A score of 38.63 cysteine in position 5 with a cutoff of 4.22 (high threshold), suggesting that is very likely for this cysteine to be palmitoylated.

To determine the importance of acylation for TgIMP1 localization to the pellicle, both the glycine and cysteine were mutated to alanine residues. The resulting TubTgIMP1 (GC-AA)-Ty, was cytoplasmic indicating that the dual acylation is critical for TgIMP1 localization to the PM (Fig. 3C). By western blot analysis the apparent molecular mass of the mutant was slightly smaller than the wild type protein indicative of a loss of post-translational lipid modification (Fig. 3B).

The first eleven amino acids of TgIMP1 are sufficient for plasma membrane targeting

Alignment of the N-terminus amino acid sequence of available IMP1 homologues in coccidian parasites show that each possess a conserved glycine at position two, a cysteine at

position five (predicted to be myristoylated and palmitoylated, respectively), as well as a basic amino acid lysine residue at position seven (Fig. 3A). To define the sequence determinant for PM localization of TgIMP1, a minimal C-terminal mCherry-Ty fusion to the first 11 amino acids of TgIMP1 was shown to be sufficient to target the chimeric protein to the PM (Fig. 3C). Since mCherry-Ty lacks cysteine residues, no additional palmitoylation could participate in its PM localization. Subsequently, the individual mutation of Cys5 and Lys7 residues were also shown to perturb PM targeting. However, while these mutations are anticipated to affect palmitoylation, myristoylation at the Gly2 still occurs and confers enhanced affinity to membranes. Concordantly, these two mutants showed nuclear exclusion and enrichment in intracellular membranes, likely including the ER (Fig. 3C). The membrane association of the two mutants (C5A and K7A) was further assessed by fractionation. The two mutants were detected both in supernatants and pellets. The K7A mutant was entirely soluble in the presence of 1% Triton X-100, indicating the conserved K7 residue plays an important role in the N-terminus myristoylation (Fig. 3D). Collectively, these results suggest that the first 11 amino acids of TgIMP1 contains the determinant for PM targeting domain via acylation.

The solution structure of the globular IMP1-like domain reveals a novel fold

The N-terminal myristoylation and palmitoylation sites identified in these studies as crucial for IMP1 PM localization of IMP1, but absent in IMP2. Furthermore, IMP1 possesses an extended N-terminus that is predicted to a flexible linker connecting a conserved C-terminus. This C-terminal region is conserved across both IMP1 and IMP2 families suggesting that it contains a globular domain that likely contributes to a possibly shared function, we term this region the 'IMP1-like domain'. To explore this further we expressed recombinant TgIMP1, subjected it to proteolytic digestion and analyzed the products by SDS-PAGE (Fig. 4A). A stable C-terminal fragment of approximately 25 kDa was observed, which we predicted to represent the conserved IMP1-like domain. To characterize the foldedness of TgIMP1 further we recorded a 1D ^1H NMR spectrum. The presence of high-field shifted methyl groups between 0.5 and -0.5 ppm (Fig. 4B and inset) indicates the presence of an ordered, globular domain, however the spectrum is dominated by a large highly disordered region, which is indicated by the sharp aliphatic peaks between 0.5 and 3.5 ppm (Fig. 4B). This region was attributed to the N-terminus carrying the myristoylation and palmitoylation sites. Truncation of the N-terminus and subsequent evaluation by NMR localized the folded domain globular within TgIMP1 residues encompassing S176 to A400, which indeed corresponds to the predicted IMP1-like domain identified from the sequence alignment (Fig. 4C). We then embarked on structural studies using multidimensional NMR. Attempts to complete a full backbone resonance assignment of a $^{13}\text{C}/^{15}\text{N}/^2\text{H}$ -labelled sample of the TgIMP1 IMP1-like domain were unsuccessful as several amide resonance were absent in the spectrum, which indicates significant conformation exchange that would preclude a high resolution structure determination (Fig. 4D). To circumvent this, we explored our NMR studies on several homologues of TgIMP1, in which the IMP1-like domain is present and conserved. One such example that yielded dramatically improved NMR spectra was IMP2 from *Plasmodium falciparum* (PfIMP2). Studies have shown that PfIMP2 is expressed in the late schizont stages of *P. falciparum* suggesting that the IMPs

may also play a role in invasion [11]. In *T. gondii*, TgIMP2.1 is a cytosolic protein (Fig. S2A) with an apparent molecular weight of ~26 kDa by western blot analysis (Fig. S2B).

The quality of the NMR spectra for PfIMP2 facilitated the complete sequential assignment using heteronuclear multidimensional NMR methods [12]. Subsequent side-chain assignment facilitated an automated analysis of the nuclear Overhauser data and the solution structure determination using the ARIA protocol [13]. The final structural statistics are summarized in Table 1. The 10 lowest energy-minimized structures out of 100 calculated exhibited root-mean-square deviations (RMSDs) from the average structure of 0.25 Å for all backbone atoms and about 0.58 Å for all non-hydrogen atoms (Fig. 5A).

The structure of PfIMP2 contains a central long helix (α 3) that is sandwiched between two β -sheets; one comprising five β -strands (β 1- β 5) and the other four (β 6- β 9; Fig. 5B). Two other major helices, α 1 and α 5, are located within inter-strand loops β 3- β 4 and β 8- β 9, respectively. Two short helices are present at the end of the large helix α 3 and at the C-terminus α 4. Interestingly, the structure displays a striking internal symmetry. The protein sequence contains a single cysteine (C71) residue located in the second α -helix which is conserved across the IMP2 family (Fig. 5C). Searches for protein structure similarities using a of available servers [14], principally DALI and PDBeFold, failed to reveal any statistically significant similarity to protein domains of known structure confirming the identification of a new fold. Predicted phosphorylation sites residues include S92 and S93, which are located near the sixth beta strand (Fig. 5C). Based on the structure of PfIMP2, S92 does not seem to be available for phosphorylation due as it points to the interior of the structure, whereas S93 is solvent accessible. The structure of the IMP1-like domain of TgIMP1 could be readily homology modeled based on the structure of PfIMP2. A comparison of the structures of TgIMP1 and PfIMP2 shows that TgIMP1 has three insertions that are absent from PfIMP2 (Fig. 5D and Fig. S3). These insertions occur within loop regions of the IMP1-like domain and are predicted to possess some helical secondary structure. The homology model predictions are consistent with available NMR chemical shifts and are conserved across equivalent coccidian proteins suggesting that they may have some functional significance. The analysis of the solvent-exposed surface properties of the PfIMP2 reveals a prominent surface comprising exposed hydrophobic residues delineated by significant surface charge (Fig. 6A-C; middle and right). Although this observation would suggest a role for this region in a functional interaction, presumably with another parasite component, the surface conservation is limited in places, which may either suggest different target ligands or that this surface is not essential for function (Fig. 6D).

The disruption of TgIMP1 has no phenotypical effect on the tachyzoites

As C-terminus domains from the IMP1 and IMP2 have no homology to other proteins with known function a double homologous recombination strategy was designed to delete the 5' coding region of the gene (Fig. 7A). The disruption of *TgIMP1* was confirmed by genomic PCR analysis (Fig. 7B) and the absence of TgIMP1 expression was assessed by western blot using anti-TgIMP1 antibodies (Fig. 7C).

The phenotypic consequences of TgIMP1 deletion were first examined by a plaque assay, which recapitulate all the steps of the parasite's lytic cycle including host cell attachment,

invasion, intracellular proliferation and egress from infected cell. No obvious difference in the size of the plaques was observed between parental parasites (Ku80-ko) and IMP1-ko (Fig. 7D). Furthermore by performing more specific and quantitative assays, no significant defect in intracellular growth and replication was observed (Fig. 7E). In addition and contrasting with the reports in the literature[4], the parasites lacking TgIMP1 showed no impairment in invasion (Fig. 7F), suggesting that TgIMP1 does not have a role in host cell invasion.

TgIMP1 confers no significant immune protection

Parasite lacking TgIMP1 were first examined for possible alteration of virulence in the mouse model of infection. The RH strain used as background to disrupt TgIMP1 is a type I strain that exhibits a severe virulence phenotype. All mice were sacrificed on day 8 following peritoneal injection of 500 parental (Ku80-ko) or IMP1-ko parasites (Fig. 8A).

Previous study indicated TgIMP1 DNA vaccine could elicit protective immunity against *T. gondii* tachyzoites challenge. The parasites lacking TgIMP1 were used here to determine the specificity of this DNA vaccine toward TgIMP1. When comparing the protection efficacy of vaccinated mice challenged with parental strain (Ku80-ko) or with IMP1-ko parasites no difference in immune protection could be observed (Fig. 8B, C).

Discussion

This study aimed to explore the role and importance in *T. gondii* of a conserved protein previously described as an immunity protective protein in *E. maxima*[2]. TgIMP1 is shown here to localize to the parasite pellicle and more precisely on the inner leaflet of the PM and likely via the dual acylation. Presumably TgIMP1 is co-translationally modified by N-myristoylation, a process of replacement of the initiator methionine residue with an amide bond, and post-translationally by S-acylation where palmitic acid is added to cysteine residue at position 5. This interpretation is further supported by the identification of TgIMP1 in the recently reported palmitome of *T. gondii* [15, 16]. The dual acylation in the N-terminal of TgIMP1 is essential for its plasma membrane localization. Besides the two conserved residues for myristoylation (G2) and palmitoylation (C5), the lysine at position 7 (K7) is also highly conserved. Lysine at the position 7 has previously been reported as a critical determinant for myristoylation in viral and mammalian proteins [17]. Myristoylation or another features that enhanced affinity to membrane is a prerequisite for palmitoylation and required for initial membrane interaction. Expectedly, mutation of K7/A in the first N11 amino acid sequence of TgIMP1 abolished its plasma membrane localization, but mutation did not depletion its membrane-binding properties, indicating the rules of substrate specificity varies in different organism [18]. The protein S-acyl transferase (PAT) implicated in the anchoring of TgIMP1 to the PM is not known, however, the repertoire of PAT in *T. gondii* has been fully characterized and it is plausible that PAT located at the PM such as TgDHHC4 or TgDDHC13 would be involved [19].

NMR data showed that the N-terminal region of TgIMP1 is highly dynamic and attached to a structured C-terminal domain of ~25kDa in size. Although the apparent disordered nature of this region could be due to the lack of myristoylation and palmitoylation because of

recombinant bacterial expression, it is more likely genuine and enables extension of the C-terminal domain away from the inner leaflet of the plasma membrane. The inherent flexibility of the linker would allow the C-terminal domain in IMP1 to sample many conformations and orientations for optimal interaction with a binding partner. This domain is conserved in a wider family of apicomplexan proteins including IMP1 and IMP2. The structure of this domain reveals a fold not previously observed in protein structures determined to date. Although this new structure fails to yield clear functional clues through similarity, the domain architecture provides a valuable addition to template libraries used in protein structure prediction. Furthermore, our structure provides some immediate suggestions for mutagenesis, once the specific cellular roles of IMP1 and IMP2 have been deduced.

Proteomics studies indicated that this expression of IMP1 and IMP2 proteins are elevated during invasion and therefore might participate directly host cell invasion and gliding mobility. Furthermore, previous study showed that antisera against NcIMP1 inhibited the *N. caninum* invasion in tissue culture [4]. However, invasion assay conducted with IMP1-ko parasites showed no significant invasion defect when compared to the parental Ku80-ko strain.

Previous work immunization using TgIMP1 as a DNA vaccine induced significant immune protection against *T. gondii* [3]. However the use of parasites lacking TgIMP1 failed to confirm that the previously reported was specific to TgIMP1. Importantly, studies on *T. gondii* and *E. tenella* have demonstrated that the efficiency and potency of the host immune response elicited by a certain antigen is considerably influenced by its subcellular localization [20, 21]. Although very few examples of protein myristoylation and palmitoylation could be exported to the outer surface of the cell membrane such as phosphoinositide-specific phospholipase C (PLC) in *Trypanosoma cruzi* [22], the fact that TgIMP1 is protected from trypsin degradation in intact extracellular parasites indicated that the protein remains inside the parasites. Intracellular toxoplasma proteins have been reported to induce protection [23] and it is also possible that IMP1 serves as an adjuvant or a potent stimulator of innate immune responses.

T. gondii isolates have been classified into three genetic types (I, II, and III) and isolates differ markedly in their virulence [24]. Though amino acid polymorphism of IMP1 occurs in two *E. maxima* strains, the comparison of the amino acid sequences of TgIMP1 from the three different types revealed no polymorphism in *T. gondii* RH parasite lacking TgIMP1 shown no alteration of virulence however, it is not possible to extrapolate this result to the other types of strains.

TgIMP1 is expressed in all life stages of *T. gondii* except in merozoites [25] and the raw data from *Eimeria* transcript Database indicates two expression peaks appearing in the stage of sporulated oocysts and during the second generation of the merozoites in *E. tenella* [26]. Although disruption of the *TgIMP1* in tachyzoites showed no loss of parasite fitness, a role in other stages as well as redundancy between members of this family of proteins cannot be excluded.

The function of the two classes IMP1 and IMP2, which possess the paralogous IMP1-like domain remain unknown, however its redundancy in the parasite genomes suggests an important role somewhere within the parasite lifecycle.

Materials and Methods

Cell culture and transfection of *Toxoplasma gondii*

T. gondii tachyzoites (RH-hxgprt-ko or RH-Ku80-ko) were maintained by serial passage in human foreskin fibroblasts (HFF) monolayer cells. Transfections were undertaken by electroporation as described previously [27]. Selection of transgenic parasites was performed with mycophenolic acid (MPA) and xanthine for HXGPRT gene selection [28] and stable clones were isolated by limiting dilution in 96 well plates.

Phylogenetic tree construction

Sequences of Apicomplexan orthologs of IMP related proteins were procured from EuPathDB and aligned using MUSCLE sequence alignment software [29]. The resulting sequence alignment was manually curated with BioEdit (<http://www.mbio.ncsu.edu/bioedit/bioedit.html>) to select for the IMP region and to remove uninformative alignments. The curated MUSCLE alignment was fed into PhyML [30] for phylogenetic tree construction, using LG model of amino acids substitution with NNI topology search. The sequences used in the phylogenetic analysis were listed in Supplementary figure 1.

Toxoplasma constructs

The primers for the TgIMPs were designed based on the TgIMP1 (TGME49_293470) and TgIMP2.1 (TGME49_276930) sequences obtained from the ToxoDB database and all the primers used in this study were listed in the supplemental data Table 1. The vector pTub8TgIMP1-Ty-HX was obtained by cloning of the TgIMP1 cDNA fragment into the *EcoRI* and *NsiI* sites in the pTubmycGFPPfmyoATailTy1. For the full length C-terminal 3Ty-tgging constructs pKI-TgIMP1-3Ty-HX and pKI-TgIMP2-3Ty-HX around 1.5 kb genomic fragments were amplified, digested with *KpnI* and *SbfI* and cloned into *KpnI* and *NsiI* sites of the pTub8MIC13-3Ty-HX construct [31]. Prior to transfection, plasmids were linearized in the middle of the PCR-amplified genomic fragments.

The nonacylatable mutant (pTubTgIMP1 (GC-AA)), which neither carries the myristoylation nor the palmitoylation sites, was obtained by PCR mutagenesis (primers listed in Table S1). The mCherry-Ty fusion construct mutants were verified by DNA sequencing. For the IMP1-ko vector, a genomic fragment of 1,481 bp corresponding of the 5' UTR of the *TgIMP1* gene was amplified by PCR and inserted into the *PacI* site of pKI TgIMP1-3Ty-HX. The direction of insertion and sequence were checked by restriction digestion and sequencing. The final IMP1-ko construct was linearized with *PacI* prior to transfection.

Western blot

Freshly released tachyzoites were harvested, washed once in PBS, solubilized directly in SDS-loading buffer, separated by electrophoresis in 12% polyacrylamide gels and

transferred. Western blots were incubated with respective mouse/rabbit antibodies in PBS, 0.05% Tween 20 and 5% skimmed milk power. After washes, the membrane was incubated with a peroxidase-conjugated goat anti-mouse or anti-rabbit antibody. Bound antibodies were visualized using the ECL plus system (GE Healthcare Bio-Sciences). Rabbit monoclonal anti-profilin (1:3,000) was used as loading control.

Recombinant protein production

A DNA fragment encoding PfIMP2 (GenBank XP_001349200, PF3D7_0730400) was codon optimized for bacterial expression and synthesized (GeneArt, Invitrogen). This PfIMP2 sequence was subcloned into pNIC-ZB vector to express PfIMP2 as a fusion protein with TEV cleavable N-terminal His6 tag. pNIC-ZB was a gift from Opher Gileadi (Addgene plasmid # 26107). For protein expression, this vector was transformed into *Escherichia coli* RosettaTM2(DE3) (NEB, UK).

Transformants were grown overnight at 37°C in 100 ml M9 minimal medium with 50ng/ml kanamycin and 35 ng/ml chloramphenicol. This culture was clarified by centrifugation at 4500 rpm and the cell pellet resuspended in 2 l M9 minimal medium with 0.07 % ¹⁵NH₄Cl and 0.2 % ¹³C₆-glucose (Sigma) and kanamycin and chloramphenicol for ¹⁵N/¹³C-labeling of PfIMP2. Cell cultures at OD600 0.5–0.7 were induced with 1 mM IPTG and left to grow at 18°C overnight. Cells were then harvested by centrifugation, resuspended in ice-cold 50 mM Hepes buffer pH 8, 300 mM NaCl and lysed by sonication. The lysate was centrifuged at 16000 rpm for 40 min to obtain supernatant. The supernatant was purified by affinity chromatography using pre-packed Nickel column, 5 ml HisTrap FF (GE Healthcare, UK) in the AKTAprime (GE Healthcare, UK). The purified protein was cleaved with TEV protease to remove the N-terminal His tag and further purified by Gel filtration using a Superdex 75 Hiload 16/600 column (GE Healthcare, UK) with AKTAprime (GE Healthcare, UK). The purified protein was then concentrated and dialyzed into 50 mM Hepes, 250 mM NaCl and 10 mM DTT, pH 7.5. The final construct that was used to record NMR experiments contained a serine residue left from the TEV cleavage, followed by the complete amino acid sequence of PfIMP2. The molecular weight of this protein is 17.89 kDa. There are 159 residues in this construct but the methionine following the first serine residue is labelled as residue 1 throughout this article.

NMR assignment and solution structure calculation

NMR spectra were recorded at 303 K on Bruker DRX600 and DRX800 spectrometers equipped with cryo-probes. The chemical shifts of ¹HN, ¹⁵N, ¹³C_α, ¹³C_β and ¹³CO cross peaks were assigned using CBCA(CO)NH, HNCACB, HNCO and HN(CA)CO experiments[12, 32]. The aliphatic and aromatic side chain ¹H and ¹³C assignments were obtained from HBHA (CBCACO)NH, HCCH-TOCSY, (H)CC(CO)NHTOCSY and ¹H–¹³C NOESY-HMQC spectra. Three-dimensional ¹H-¹⁵N/¹³C NOESY-HMQC (mixing time 100 ms at 600 and 800 MHz) experiments provided the distance restraints for final structure calculation. The ARIA protocol (Linge et al, 2003) was used for expansion of the NOE assignment and final structure calculation. Dihedral angle restraints were derived from TALOS (Shen and Bax, 2013). The frequency window tolerance for assigning NOEs was ±0.05 and ±0.07 ppm for direct and indirect proton dimensions and ±0.7 and ±1.2 ppm for

nitrogen and carbon dimensions, respectively. The ARIA parameters were set to default values. The structure is deposited under PDB accession number is 5LG9 and the BRMB ID for the assignment is 26552. These are now explicitly stated in the manuscript.

Antibodies

Anti-GAP45, anti-Ty tag, anti-Profilin, anti-SAG1, anti-Catalase and anti-IMP1 antibodies were previously described[3, 9]. Immunoblots were visualized using a chemiluminescent substrate (Amersham, GE Healthcare). Anti-profilin antibodies were used to detect profilin as loading control.

Indirect immunofluorescence assay (IFA) and microscopy

Parasites infected HFF cells were fixed with 4% paraformaldehyde and 0.05% Glu in phosphate-buffered saline (PBS) for 15 min, followed by 5 min of neutralization with 0.1 M glycine in PBS. Fixed cells were permeabilized with 0.2% Triton X-100 in PBS (PBS-TX) for 20 min and then blocked with 2% bovine serum albumin (BSA). Intracellular growth and plaque assay were performed as previously described. Confocal images were taken with a Leica laser scanning confocal microscope (TCS-NT DM/IRB and SP2) using a 100x Plan-APO objective.

Trypsin treatment and aerolysine assay

IMC and plasma membrane separation was achieved by aerolysine treatment as previously described [33]. To further refine the localization of TgIMP1 at the pellicle, freshly released parasites were harvested, washed in PBS once and then were incubated with a 0.1% solution of trypsin at 37 °C for 0, 30 and 60 min. Following incubation, parasites were washed twice with cold PBS. The harvested parasites were solubilized directly in SDS-loading buffer and analyzed by Western blot.

Cell fractionation

Freshly released tachyzoites were harvested, washed once with PBS and then resuspended into PBS or PBS with 1% Triton-X-100. Parasites were lysed by freezing and thawing three times followed by sonication on ice. Pellet and soluble fractions were separated by centrifugation at 14,000 rpm for 30 min at 4°C. The final pellet was washed once with PBS and resuspended in 100 µL PBS containing the insoluble fraction. Equal amounts of all supernatants were analyzed by immunoblotting. Proteins were detected using anti-Ty antibodies, the cytosolic profilin was used as a control.

TgIMP1 DNA vaccine protection experiments

Female Kunming mice (5-6 weeks old) were purchased from Academy of Military Medical Sciences Laboratory Animal Center (Beijing, China). All mice were maintained under standard conditions and provided with rodent feed and water *ad libitum*. Prior to experiment, mice were acclimatized for one week. Animal experiments were conducted according to institutional guidelines for animal ethics. Mice (10/group) were injected intramuscularly with 100 µg of pcDNA-IMP1 plasmid DNA suspended in 100 µl sterile PBS in the thigh skeletal muscle. Control mice received PBS alone or empty pcDNA 3.1(+) vector for the

first injection. The second and third injections were carried out in 2 and 4 weeks post-primary injection with the same dosage. One week after the final injection, the mice of all groups were challenged intraperitoneally with 500 tachyzoites of the Ku80-ko or IMP1-ko strain.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by the Swiss National Foundation (FN3100A0-116722 to DSF, National Natural Science Foundation of China (31372424) to Q.L and in part by the Wellcome Trust (Senior Investigator Award 100280 and multiuser equipment grant 104833 to S.J.M.) DSF is HHMI senior international research scholar.

References

- [1]. Innes EA, Vermeulen AN. Vaccination as a control strategy against the coccidial parasites *Eimeria*, *Toxoplasma* and *Neospora*. *Parasitology*. 2006; 133(Suppl):S145–168. [PubMed: 17274844]
- [2]. Blake DP, Billington KJ, Copestake SL, Oakes RD, Quail MA, Wan KL, Shirley MW, Smith AL. Genetic Mapping Identifies Novel Highly Protective Antigens for an Apicomplexan Parasite. *PLoS pathogens*. 2011;7.
- [3]. Cui X, Lei T, Yang D, Hao P, Li B, Liu Q. *Toxoplasma gondii* immune mapped protein-1 (TgIMP1) is a novel vaccine candidate against toxoplasmosis. *Vaccine*. 2012; 30:2282–2287. [PubMed: 22310204]
- [4]. Cui X, Lei T, Yang DY, Hao P, Liu Q. Identification and characterization of a novel *Neospora caninum* immune mapped protein 1. *Parasitology*. 2012; 139:998–1004. [PubMed: 22405293]
- [5]. Shi Y, Mowery RA, Ashley J, Hentz M, Ramirez AJ, Bilgic B, Slunt-Brown H, Borchelt DR, Shaw BF. Abnormal SDS-PAGE migration of cytosolic proteins can identify domains and mechanisms that control surfactant binding. *Protein science : a publication of the Protein Society*. 2012; 21:1197–1209. [PubMed: 22692797]
- [6]. Ogino N, Yoneda C. The fine structure and mode of division of *Toxoplasma gondii*. *Arch Ophthalmol*. 1966; 75:218–227. [PubMed: 4159376]
- [7]. Dubremetz JF, Torpier G. Freeze fracture study of the pellicle of an eimerian sporozoite (Protozoa, Coccidia). *J Ultrastruct Res*. 1978; 62:94–109. [PubMed: 418187]
- [8]. Ouologuem DT, Roos DS. Dynamics of the *Toxoplasma gondii* inner membrane complex. *Journal of cell science*. 2014; 127:3320–3330. [PubMed: 24928899]
- [9]. Frenal K, Polonais V, Marq JB, Stratmann R, Limenitakis J, Soldati-Favre D. Functional Dissection of the Apicomplexan Glideosome Molecular Architecture. *Cell Host & Microbe*. 2010; 8:343–357. [PubMed: 20951968]
- [10]. Abrami L, Fivaz M, Glauser PE, Parton RG, van der Goot FG. A pore-forming toxin interacts with a GPI-anchored protein and causes vacuolation of the endoplasmic reticulum. *Journal of Cell Biology*. 1998; 140:525–540. [PubMed: 9456314]
- [11]. Aurrecoechea C, Brestelli J, Brunk BP, Dommer J, Fischer S, Gajria B, Gao X, Gingle A, Grant G, Harb OS, Heiges M, et al. PlasmoDB: a functional genomic database for malaria parasites. *Nucleic acids research*. 2009; 37:D539–543. [PubMed: 18957442]
- [12]. Benjamin S, Williams F, Kerry L, Matthews S. NMR assignment of the immune mapped protein 1 (IMP1) homologue from *Plasmodium falciparum*. *Biomol NMR Assign*. 2015; 9:393–395. [PubMed: 25947350]
- [13]. Rieping W, Habeck M, Bardiaux B, Bernard A, Malliavin TE, Nilges M. ARIA2: automated NOE assignment and data integration in NMR structure calculation. *Bioinformatics*. 2007; 23:381–382. [PubMed: 17121777]

- [14]. Mayr G, Domingues FS, Lackner P. Comparative Analysis of Protein Structure Alignments. *BMC Structural Biology*. 2007; 7:1–15. [PubMed: 17201922]
- [15]. Foe IT, Child MA, Majmudar JD, Krishnamurthy S, van der Linden WA, Ward GE, Martin BR, Bogoy M. Global Analysis of Palmitoylated Proteins in *Toxoplasma gondii*. *Cell Host Microbe*. 2015; 18:501–511. [PubMed: 26468752]
- [16]. Caballero MC, Alonso AM, Deng B, Attias M, de Souza W, Corvi MM. Identification of new palmitoylated proteins in *Toxoplasma gondii*. *Biochimica et biophysica acta*. 2016; 1864:400–408. [PubMed: 26825284]
- [17]. Kaplan JM, Mardon G, Bishop JM, Varmus HE. The first seven amino acids encoded by the v-src oncogene act as a myristylation signal: lysine 7 is a critical determinant. *Mol Cell Biol*. 1988; 8:2435–2441. [PubMed: 2841581]
- [18]. Traverso JA, Giglione C, Meinel T. High-throughput profiling of N-myristoylation substrate specificity across species including pathogens. *Proteomics*. 2013; 13:25–36. [PubMed: 23165749]
- [19]. Frenal K, Tay CL, Mueller C, Bushell ES, Jia Y, Graindorge A, Billker O, Rayner JC, Soldati-Favre D. Global analysis of apicomplexan protein S-acyl transferases reveals an enzyme essential for invasion. *Traffic*. 2013; 14:895–911. [PubMed: 23638681]
- [20]. Gregg B, Dzierzinski F, Tait E, Jordan KA, Hunter CA, Roos DS. Subcellular antigen location influences T-cell activation during acute infection with *Toxoplasma gondii*. *PLoS One*. 2011; 6:e22936. [PubMed: 21829561]
- [21]. Huang X, Zou J, Xu H, Ding Y, Yin G, Liu X, Suo X. Transgenic *Eimeria tenella* expressing enhanced yellow fluorescent protein targeted to different cellular compartments stimulated dichotomic immune responses in chickens. *J Immunol*. 2011; 187:3595–3602. [PubMed: 21876035]
- [22]. Martins VD, Okura M, Maric D, Engman DM, Vieira M, Docampo R, Moreno SNJ. Acylation-dependent Export of *Trypanosoma cruzi* Phosphoinositide-specific Phospholipase C to the Outer Surface of Amastigotes. *Journal of Biological Chemistry*. 2010; 285:30906–30917. [PubMed: 20647312]
- [23]. Sibley LD, Sharma SD. Ultrastructural localization of an intracellular *Toxoplasma* protein that induces protection in mice. *Infect Immun*. 1987; 55:2137–2141. [PubMed: 3305365]
- [24]. Howe DK, Sibley LD. *Toxoplasma gondii* comprises three clonal lineages: correlation of parasite genotype with human disease. *J Infect Dis*. 1995; 172:1561–1566. [PubMed: 7594717]
- [25]. Gajria B, Bahl A, Brestelli J, Dommer J, Fischer S, Gao X, Heiges M, Iodice J, Kissinger JC, Mackey AJ, Pinney DF, et al. ToxoDB: an integrated *Toxoplasma gondii* database resource. *Nucleic acids research*. 2008; 36:D553–556. [PubMed: 18003657]
- [26]. Rangel LT, Novaes J, Durham AM, Madeira AM, Gruber A. The *Eimeria* transcript DB: an integrated resource for annotated transcripts of protozoan parasites of the genus *Eimeria*. *Database (Oxford)*. 2013; 2013:bat006. [PubMed: 23411718]
- [27]. Soldati D, Boothroyd JC. Transient transfection and expression in the obligate intracellular parasite *Toxoplasma gondii*. *Science*. 1993; 260:349–352. [PubMed: 8469986]
- [28]. Donald RG, Carter D, Ullman B, Roos DS. Insertional tagging, cloning, and expression of the *Toxoplasma gondii* hypoxanthine-xanthine-guanine phosphoribosyltransferase gene. Use as a selectable marker for stable transformation. *J Biol Chem*. 1996; 271:14010–14019. [PubMed: 8662859]
- [29]. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic acids research*. 2004; 32:1792–1797. [PubMed: 15034147]
- [30]. Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Systematic biology*. 2010; 59:307–321. [PubMed: 20525638]
- [31]. Friedrich N, Santos JM, Liu Y, Palma AS, Leon E, Saouros S, Kiso M, Blackman MJ, Matthews S, Feizi T, Soldati-Favre D. Members of a novel protein family containing microneme adhesive repeat domains act as sialic acid-binding lectins during host cell invasion by apicomplexan parasites. *J Biol Chem*. 2010; 285:2064–2076. [PubMed: 19901027]

- [32]. Marchant J, Sawmynaden K, Saouros S, Simpson P, Matthews S. Complete resonance assignment of the first and second apple domains of MIC4 from *Toxoplasma gondii*, using a new NMRView-based assignment aid. *BIOMOLECULAR NMR ASSIGNMENTS*. 2008; 2:119–121. [PubMed: 19636884]
- [33]. Wichroski MJ, Melton JA, Donahue CG, Tweten RK, Ward GE. Clostridium septicum alpha-toxin is active against the parasitic protozoan *Toxoplasma gondii* and targets members of the SAG family of glycosylphosphatidylinositol-anchored surface proteins. *Infection and Immunity*. 2002; 70:4353–4361. [PubMed: 12117945]

Figure 1. Phylogenetic analysis of IMP1 homologues across all apicomplexan parasites
Phylogenetic tree based on sequences of Apicomplexan IMP related proteins cluster into six different groups, named IMP1-6. The tree was constructed using PhyML using LG model of amino acids substitution with NNI topology search, based on an amino acid alignment by MUSCLE. Values at tree nodes represent confidence values supporting nodal placement. The sequences are labeled with their accession numbers in EuPathDB.

Figure 2. Localization of TgIMP1 in tachyzoites.

(A) Western blotting analysis of Ku80-ko, TgIMP1-3Ty KI and TubTgIMP1-Ty.

(B) Immunofluorescence microscopic analysis of parasites expressing a second copy of TgIMP1-Ty using anti-Ty and the endogenous TgIMP1 using anti-TgIMP1, respectively (upper and middle panel). The pellicle localization of TgIMP1 was confirmed in the stain of TgIMP1-3Ty KI, using anti-Ty (lower panel).

(C) Extracellular tachyzoites expressing IMP1-Ty were incubated 30 min with aerolysin. After aerolysin treatment, the plasma membrane dramatically separated from the underlying IMC and the staining of TgIMP1 is co-localized with the plasma membrane marker SAG1.

(D) Freshly released parasites were incubated with Trypsin. After 30 and 60 mins proteinase trypsin digestion, SAG1, localized to the outer surface of the plasma membrane, was digested, while TgIMP1 kept intact as well as the cytosolic protein profilin.

Figure 3. The N-terminal 11 amino acids contain the minimal signals for TgIMP1 plasma membrane localization.

(A) The first 19 amino acids of TgIMP1 was aligned with its homologues in *Neospora caninum* (NCLIV_000430), *Hammondia hammondi* (HHA_293470), *Sarcocystis neurona* (SN3_00701045), *Eimeria tenella* (ETH_00030475) and *Eimeria maxima* (EMWEY_00011000). All of the homologues have a conserved glycine at position two predicted to be myristoylated, a cysteine at position five predicted to be palmitoylated and a lysine at position seven.

(B) Mutation of the residues predicted for myristoylation (G2) and palmitoylation (C5) to Alanine were generated by point mutation of the construct pTubTgIMP1 –Ty. After the mutation of the dual acylation sites, a size shift was observed in the mutant by western blot, due to lost of dual acylation.

(C) The TgIMP1 (GC-AA) mutation altered relocalized to cytoplasm, indicating a loss of membrane interaction due to the absence of acylation. (C middle) The N-terminus 11 amino acids of TgIMP1 were fused with mCherry-Ty and the chimera is still localized to the plasma membrane. Plasma-membrane localization of TgIMP1 was lost upon the mutation of Cys5 to Ala or the mutation of Lys 7 to Ala.

(D) The solubility of TgIMP1n11mCherry-Ty mutants was analyzed by fractionation assay. The mutation in the palmitoylation site (C5A) and in the conserved lysine (K7) made the protein mostly soluble.

Figure 4.

(A) SDS-PAGE gel obtained from running the samples from limited proteolysis of TgIMP1 IMP1 domain with Chymotrypsin. The lanes labelled M and C correspond to marker and control samples (TgIMP1 IMP1 domain), respectively. The time points at which samples were taken are labelled for the remaining lanes. Sequence showing of the locations of the domains and motifs in TgIMP1. The N-terminal plasma membrane localisation signal is colored blue, the myristoylation site is colored purple, palmitoylation site colored navy blue,

the linker region colored red and the IMP1 domain colored green. The cleavage sites of chymotrypsin, determined by sequencing, are colored black.

(B) Figure showing the 1D ^1H NMR spectrum of full length TgIMP1. The methyl region of the spectrum is zoomed in and shown on the top left. The spectrum shows that the protein has a large unfolded domain, which is masking the signal from the folded domain.

(C) 1D ^1H NMR spectrum of the isolated globular domain from TgIMP1, termed the IMP1-like domain. The high chemical shifts dispersion in the spectrum with peaks at <0.5 ppm and >8.5 ppm shows that this construct contains the ordered domain and the flexible region seen in full length TgIMP1 protein is absent.

(D) 2D ^1H - ^{15}N HSQC spectrum of the IMP1-like domain of TgIMP1 ($^{13}\text{C}/^{15}\text{N}/^2\text{H}$ labelled). The 3D data was incomplete with between 5-10% of amide peaks absent, therefore and comprehensive assignment with 100% confidence was not possible. Approximately 50% of the sequence was assigned confidently as shown.

(E) 2D ^1H - ^{15}N HSQC spectrum of the IMP domain of TgIMP1 ($^{13}\text{C}/^{15}\text{N}/^2\text{H}$ labelled). The residues that were assigned with 100% confidence are labelled in the spectrum. As the 3D data was incomplete, less than 50% of the sequence was assigned confidently.

Figure 5. Solution structure of PfIMP2 determined from NMR data

(A) Stereo representation of PfIMP2 NMR family of structures

(B) Front and top views of cartoon representation of PfIMP2. Secondary structure elements showing in red for α -helices, yellow for β -strands and green for loops, and are labelled according to sequence order

(C) Left - Cartoon representation of PfIMP2 showing the location of the single cysteine residue (C71 – pink stick representation). Right Cartoon representation of PfIMP2 showing the two predicted phosphorylation sites (S92 and S93) in pink.

(D) Cartoon representation of TgIMP1 structure, which was homology modelled based on the structure of PfIMP2 using Phyre3. The insertions present in TgIMP1 but absent in PfIMP2 are shown in light pink

Figure 6. Conserved surface hydrophobic and electrostatic properties of PfIMP2.

(A) Front, back and side views of cartoon representation of PfIMP2 structure.

(B) Front, back and side views of surface representation of PfIMP2 showing solvent exposed hydrophobic residues (Trp, Phe, Ile, Leu, Val, Met or Ala) shaded orange.

(C) Front, back and side views of electrostatic surface representation of PfIMP2. Blue indicates positively charged regions and red shows negative.

(D) Front, back and side views of ConSurf surface representation of PfIMP2 indicating the conservation of residues. Residues labelled yellow have insufficient data as calculation for

that site was performed on less than 10% of the sequences. Magenta indicate high conservation while cyan represents low.

Figure 7. Generation of a TgIMP1 knockout.

(A) Schematic representation of the strategy used to delete the *TgIMP1* N-terminus region by double homologous recombination. The knockout TgIMP1 construct contains 1,481 bp of the 5' un-translated region (5'UTR) and the 1,629 bp of its C-terminus coding sequence. Arrows represent the primers used for PCR reaction and the lengths of the PCR products are marked.

- (B) PCR analysis were performed on the IMP1-ko strain, showing that double homologous recombination occurs (primer combination AB, AC for the 5' region and FG, HG for 3' region) and the N-terminus of *TgIMP1* was deleted (primer combination DE).
- (C) Western blotting using antisera against recombinant TgIMP1 indicated the lost of the TgIMP1 expression in the IMP1-ko strain.
- (D) Plaque assay performed on an HFF monolayer inoculated with Ku80-ko or IMP1-ko parasites. 8 days after inoculation, the HFF was fixed and strained with crystal violet.
- (E) Freshly released parasites from Ku80-ko or IMP1-ko strain invade new HFF cell monolayer. Numbers of parasites per vacuoles were counted at 24 h after HFF cell infection. Values are means standard deviations ($\pm$ SD) for three independent experiments.
- (F) Parasites lacking IMP1 were not impaired in invasion.

Figure 8. Protection of BALB/c mice against Ku80-ko and IMP1-ko challenge.

(A) Virulence comparison between Ku80-ko and IMP1-ko. All mice were sacrificed on day 8 following peritoneal injection of 500 parental (Ku80-ko) or IMP1-ko parasites.

(B) Survival curves of immunized BALB/c mice after lethal challenge with 500 tachyzoites of Ku80-ko and IMP1-ko, 2 weeks after the last immunization.

Table 1

NMR structure calculation statistics data for PfIMP2

NMR restraints	
Distance constraints	
Total NOE	1683
Intraresidue	698
Interresidue	985
Sequential ($ i-j =1$)	443
Medium range ($ i-j <4$)	94
Long range ($ i-j >5$)	348
Total Dihedral angle Restraints	
Φ	136
Ψ	136
Structural Statistics	
Violations (mean and SD)	
Distance constraints (Å)	0.033 ± 0.004
Dihedral angle constraints (°)	1.40
Maximum dihedral angle violation (°)	0.67
Maximum distance constraint violation (Å)	0.54
Deviations from idealized geometry	
Bond length (Å)	0.004 ± 0.001
Bond angle (°)	0.52 ± 0.014
Impropers (°)	1474 ± 0.080
Average Pairwise rmsd^a(Å)	
Heavy	1.966
Backbone	2.317