Dissecting the interface between apicomplexan parasite and host cell: Insights from a divergent AMA-RON2 pair

Michelle L. Parker, Diana M. Penarete-Vargas, Phineas T. Hamilton, Amandine Guérin, Jitender P. Dubey, Steve J. Perlman, Furio Spano, Maryse Lebrun, Martin J. Boulanger

SI Methods and Materials

Ethics statement. This study was conducted according to European Union guidelines for handling laboratory animals. Immunizations for antibody production in rabbits was conducted at the CRBM animal house (Montpellier) and approved by the Committee on the Ethics of Animal Experiments (Languedoc-Roussillon, Montpellier) (Permit Number: D34-172-4, delivered on 20/09/2009). Immunizations for antibody production in mice was carried out at the Istituto Superiore di Sanità and authorized by the Italian Ministry of Health, according to Legislative Decree 116/92 that implemented the European Directive 86/609/EEC.

Evolutionary Analysis. To retrieve homologs of *T. gondii* AMA1 and RON2 from apicomplexans, we constructed a custom database of predicted protein sequences from EuPathDB (http://eupathdb.org/eupathdb/). Annotated protein sets for 17 apicomplexan parasites (including *Plasmodium*, *Toxoplasma*, *Neospora*, *Eimeria*, *Theileria*, *Babesia*, and *Cryptosporidium*) were downloaded and queried with AMA1 and RON2 amino acid sequences from *T. gondii* using the iterative jackHMMER algorithm of HMMER (1) (inclusion E-value 10^{-12}). AMA and RON2 sequences were aligned using MUSCLE (2) in Geneious (Biomatters Ltd) and after visual inspection poorly aligning sequences were discarded. *Plasmodium* sequences homologous to RON2 and annotated as CLAGs were discarded for clarity. Alignments were

cleaned of gaps using a partial deletion threshold of 75%, and phylogenetic trees constructed in MEGA5 (3) after model selection with ProTest3 (4) (WAG+G+I for AMAs and LG+G+I for RON2s) with 1000 bootstrap replicates.

Predicted protein annotations and species and strain identifiers correspond to accessions from EuPathDB: PF3D7, *Plasmodium falciparum* 3D7; PVX, *P. vivax* Sal1; PY17X, *P. yoelii yoelii* 17X; PBANKA, *P. berghei* ANKA; BBM, *Babesia microti* RI; BBOV, *B. bovis* T2Bo; BEWA, *Theileria equi* WA; TP01, *T. parva* Muguga; TA, *T. annulata* Ankara; TOT, *T. orientalis* Shintoku; TGME49, *Toxoplasma gondii* ME49; NCLIV, *Neospora caninum* LIV; ETH, *Eimeria tenella* Houghton.

Cloning, protein production, and purification. *Tg*AMA4 DIDIIEGF1 (TGME49_294330; Ser58 to Asp553) was produced recombinantly in insect cells and purified as previously described (5). The tandem EGF domains and Cys-rich regions were predicted using ProSite (6). Selenomethionine labeled protein was generated by infecting *Tni* cells in ESF-921 media (Expression Systems; Davis, CA), followed by exchange into methionine deficient media after 7 h and addition of 100 – 200 mg/L of selenomethionine after a further 7.5 h. Cells were harvested 72 h post infection and purified as described for the native protein.

A construct encoding domain 3 (D3) of *Tg*RON2$_{L1}$ (TGME49_294400; Gln1292 to Ser1324) was synthesized by GenScript and cloned into a modified pET32a vector and produced in *E. coli* BL21 cells as a TRX fusion. *Tg*RON2$_{L1}$D3-TRX was purified on its own or in complex with *Tg*AMA4 using established protocols (7). *Tg*RON2$_{L1}$D3-TRX was used for ITC experiments, while samples co-purified with *Tg*AMA4 and cleaved from TRX were used for crystallographic experiments. Selenomethionine-labeled protein was produced in *E. coli* 834 cells using established protocols (8) and purified as described for the native protein. Constructs

encoding *Tg*RON2$_{L1}$D3 double mutants for ITC (Asn1296Ala/Pro1309Ala and Cys1307Ser/Cys1313Ser) were synthesized by GenScript, and produced as TRX fusions using the same protocol as for the native protein.

The pGEX-*Tg*RON2$_{L1}$D3 plasmid (5) was used as a template to generate *Tg*RON2$_{L1}$ mutants using QuikChangeII (Agilent 200523). A fragment of *Tg*RON2$_{L1}$ (Leu539 to Tyr983) and a fragment of *Tg*AMA3 (Asn393 to Ser566) were cloned into pGEX-5X-1 and produced in *E. coli* BL21. GST-tagged proteins were produced as described previously (5).

Primers are listed in Table S3 and all plasmids were sequenced.

Production of anti-*Tg*AMA4, anti-*Tg*RON2$_{L1}$ and anti-*Tg*AMA3 sera. Rabbits were immunized with 100 µg of recombinant *Tg*AMA4 DIDIIEGF1 via subcutaneous injection on days 1, 7 and 28 in 400 µL of PBS. Anti-*Tg*RON2$_{L1}$ antibodies (against *Tg*RON2$_{L1}$ fragment Leu539-Tyr983) and anti-*Tg*AMA3 antibodies (against *Tg*AMA3 fragment Asn393 to Ser566) were raised by intraperitoneal injection of BALB/c mice with 50 µg of recombinant protein on days 1 (complete Freund's adjuvant), 28 (incomplete Freund's adjuvant) and 49 (PBS).

Isothermal titration calorimetry. Purified *Tg*AMA4, *Tg*RON2$_{L1}$D3-TRX, *Tg*RON2$_{L1}$D3(N1296A/P1309A)-TRX, *Tg*RON2$_{L1}$D3(C1307S/C1313S)-TRX and TRX were dialyzed against 20 mM Hepes pH 7.5, 150 mM NaCl overnight at 4 °C. All ITC experiments were carried out at 25 °C on a MicroCal ITC$_{200}$ Instrument (Malvern). The sample cell contained 0.2 mL of 12 µM *Tg*AMA4, and the TRX-fused peptide (120 µM) was added in 19 injections of 2 µL each. TRX was injected as a negative control. The data were processed using Origin software (MicroCal) and the dissociation constant (K_D) determined using a one-site model.

Crystallization and X-ray data collection. Crystals of *Tg*AMA4DIDIIEGF1 (20 mg/mL) were grown at 18 °C in 0.1 M Bis-Tris pH 6.0, 25% PEG3350 and cryoprotected in paratone.

Diffraction data were collected on beamline 08ID-1 at the Canadian Light Source (CLS) at a wavelength of 1.0332 Å. Crystals of *Tg*AMA4+*Tg*RON2$_{L1}$D3 (18 mg/mL) were grown in 0.2 M Lithium sulfate, 0.1 M Tris-HCl pH 8.5, 25% PEG3350 and cryoprotected in paratone. Selenomethionine protein crystallized in 0.2 M Ammonium sulfate, 0.1 M Hepes pH 7.8, 25% PEG3350 and crystals were cryoprotected in reservoir solution supplemented with 12.5% glycerol. Diffraction data were collected on beamline 08ID-1 at CLS at a wavelength of 0.9794 Å.

Data processing, structure determination and refinement. Diffraction data were processed to 2.05 Å (*Tg*AMA4 - two molecules in the asymmetric unit that superimpose with an rmsd of 0.46 Å over 496 Cαs), 2.1 Å (*Tg*AMA4SeMet-*Tg*RON2$_{L1}$D3SeMet), or 1.53 Å (*Tg*AMA4-*Tg*RON2$_{L1}$D3 - two complexes in the asymmetric unit superimpose with an rmsd of 0.19 Å over 466 Cαs and display a conformationally flexible C-terminal EGF domain) resolution using Imosflm (9) and Aimless (10). The structure of *Tg*AMA4-*Tg*RON2$_{L1}$D3 was solved by Selenium single wavelength anomalous dispersion. A total of 36 high confidence Se sites were identified and refined using the ShelxC/D/E pipeline (11), which enabled building and registering of nearly 80% of the backbone using buccaneer (12) in the CCP4 suite of programs (13). *Tg*AMA4 and native *Tg*AMA4-*Tg*RON2$_{L1}$D3 structures were solved by molecular replacement using a single *Tg*AMA4 chain from the Se-phased model in Phaser (14). Model building and selection of solvent atoms were performed in COOT (15) and the models refined using Phenix.refine (16) in Phenix (17). Structural validation was performed with MolProbity (18). Ramachandran plots showed greater than 96% of residues in the most favored conformations. For each dataset, 5% of reflections were set aside for calculation of R_{free}. Data collection and refinement statistics are listed in Table S1.

Parasite cultures. Oocysts of strain EGS (19) were provided by Jitender Dubey. Sporozoite excystation was performed by mechanical and enzymatic disruption of the cyst wall. First oocysts were vortexed in Hank's Balanced Salt Solution (HBSS) with 1 mm glass beads for 2 min and incubated at 37 °C for 20 min in HBSS containing 0.25% Trypsin (Lonza) and 0.75% Sodium tauroglycocholate (Merck). Free sporozoites were checked by microscopy and washed twice in 10% heat-inactivated fetal bovine serum HBSS.

Western blot. Tachyzoites and sporozoites were separated by 10% SDS PAGE, transferred to nitrocellulose membranes and blocked with 5% non-fat dry milk in TNT buffer (140 mM NaCl, 15 mM Tris, 0.05% Tween20) for 1 h. Membranes were incubated with primary antibodies at the dilutions listed in Table S4 in 5% non-fat dry milk-TNT. After five washes with TNT buffer the membranes were incubated with alkaline phosphatase conjugated secondary antibodies and revealed with 5-bromo-4-chloro-3-indolyl phosphate/nitro blue tetrazolium (BCIP/NBT; Sigma).

Immunofluorescence assay on sporozoites. Confluent human foreskin fibroblast monolayers grown on 24-well plates containing 12 mm coverslips were infected with 5 x 10^5 excysted sporozoites for 1 h or 48 h and then fixed in 4% paraformaldehyde (PAF) in PBS for 30 min. After three washes in PBS, cells were permeabilized for 10 min with 0.1% Triton X-100 in PBS and then blocked for 30 min with 10% fetal bovine serum in PBS. The cells were then stained with primary antibodies diluted in 2% PBS for 1 h, washed three times in PBS and incubated with secondary antibodies diluted in 2% PBS for 1 h. Antibody dilutions are listed in Table S4. The coverslips were mounted with Immunomount (Calbiochem) and the images were acquired using a Zeiss Axioimager epifluorescence microscope with a Zeiss Axiocam MRm CCD camera driven by the Axiovision software (Zeiss), at the Montpellier RIO imaging facility.

Sporozoite invasion inhibition assay. Freshly excysted ME49 sporozoites were resuspended in DMEM containing 2% fetal bovine serum/10 mM HEPES. Following a 20 min incubation at 4 °C with 200 µg/ml of GST or *Tg*RON2$_{L1}$D3-GST, parasites were added to HFF monolayers in 8-well chamber slides in the presence of the recombinant proteins and allowed to invade for 30 min at 37 °C. After 4% formaldehyde fixation and extensive washing, extracellular and intracellular sporozoites were differentially stained by immunofluorescence using anti-SRS28 (formerly known as sporoSAG) rabbit and mouse antisera. Counts of intracellular sporozoites were carried out in triplicate on 30 microscopic fields for each condition using a 63x objective.

Plasmid constructs, parasite transfection and selection of transformants. The pLIC-KO-AMA4 and the pLIC-RON2$_{L1}$-HA$_3$ plasmids were obtained by ligation independent cloning procedures using the pLIC-HA$_3$-CAT plasmid (20). The first 1300 bp of the *Tg*AMA4 coding sequence and a genomic fragment corresponding to the 2000 bp of the C-term of *Tg*RON2$_{L1}$ were amplified with primers ML1821/ML1822 and ML813/ML814, respectively (Table S3). The KO-AMA1/KO-AMA2 cell line was transfected with pLIC-KO-AMA4 or pLIC-RON2$_{L1}$-HA$_3$ plasmids after linearization with *NruI* and *StuI*, respectively. For each transfection, clones were selected under chloramphenicol selection and isolated by limiting dilution cloning and screened by PCR for correct vector integration.

ELISA. ELISAs were performed as described previously (5). Briefly, Maxisorp 96 well plates were coated overnight with 1 µg/ml of recombinant *Tg*AMA4 and after washes and saturation with PBS-1% bovine serum albumin recombinant *Tg*RON2$_{L1}$D3-GST proteins were incubated at increasing concentrations for 1 h. Detection was performed with rat anti-GST and goat anti-rat IgG horseradish peroxidase conjugate (Invitrogen). The absorbance was read at 450 nm after addition of SIGMAFAST-OPD substrate.

Tachyzoite invasion inhibition assay. Invasion assays were performed as described previously (5). Briefly, $5x10^6$ freshly released tachyzoites were incubated at 4 °C for 20 min and allowed to invade during 5 min at 38 °C. For the inhibition assays, 200 µg/ml of GST or *Tg*RON2$_{L1}$D3-GST peptide were added during the 20 min of incubation and the period of invasion. After 4% paraformaldehyde fixation, immunofluorescence assay was performed using anti-SRS29B (also known as SAG1) and anti-ROP1 antibodies. Counts of intracellular parasites were performed on 30 microscopic fields for each triplicated condition. The assays were performed independently at least three times.

Homology modeling. *Tg*AMA3 Pro/Val/Glu-rich region and select *Tg*AMA4 EGF/Cys-rich domains were modeled using iTASSER (21). For *Tg*AMA3, a semi-extended, kinked model was chosen, as the generated models varied from completely extended to compact. For *Tg*AMA4 EGF/Cys rich domains were predicted based on ProSite analysis (6), and models were estimated from a template of the first *Tg*AMA4 EGF domain (PDB ID: 4Z81) that represents the general size and shape of each module.

The structural model for *Pf*MAEBL M2 (Asn589 – Val1008; PF3D7_1147800) was generated using Modeller 9v8 through the Chimera interface (22, 23), based on a *Tg*AMA4 DI-DII template (PDB ID: 4Z80), with which it shares 29% identity in this region. The final model of *Pf*MAEBL M2 was chosen based on its low value of the normalized Discrete Optimized Protein Energy value (zDOPE), energy minimized using Chimera (22), and validated by visual inspection and MolProbity (18) and ProQ (24) (rating: very good model).

Fig. S1. Sequence alignment of structurally characterized AMA proteins. Sequences of DI and DII from *Tg*AMA1 (TGME49_255260), *Nc*AMA1 (NCLIV_028680), *Bd*AMA1 (ACC96234), *Pf*AMA1 (PF3D7_1133400), *Pv*AMA1 (PVX_092275), *Tg*AMA3 (TGME49_315730), and *Tg*AMA4 (TGME49_294330) were aligned in MEGA6.0 (25) using MUSCLE (2) and illustrated

in ESPript (26). Minor adjustments were made manually to anchor the alignment on the Cys residues and known structural features. DI and DII loops are annotated and colored as in Fig. 3D. Disulfides that are part of the AMA1-type core are numbered and indicated with a grey starburst above the alignment; disulfides that are novel in *Tg*AMA4 are lettered and indicated with a yellow starburst below the alignment.

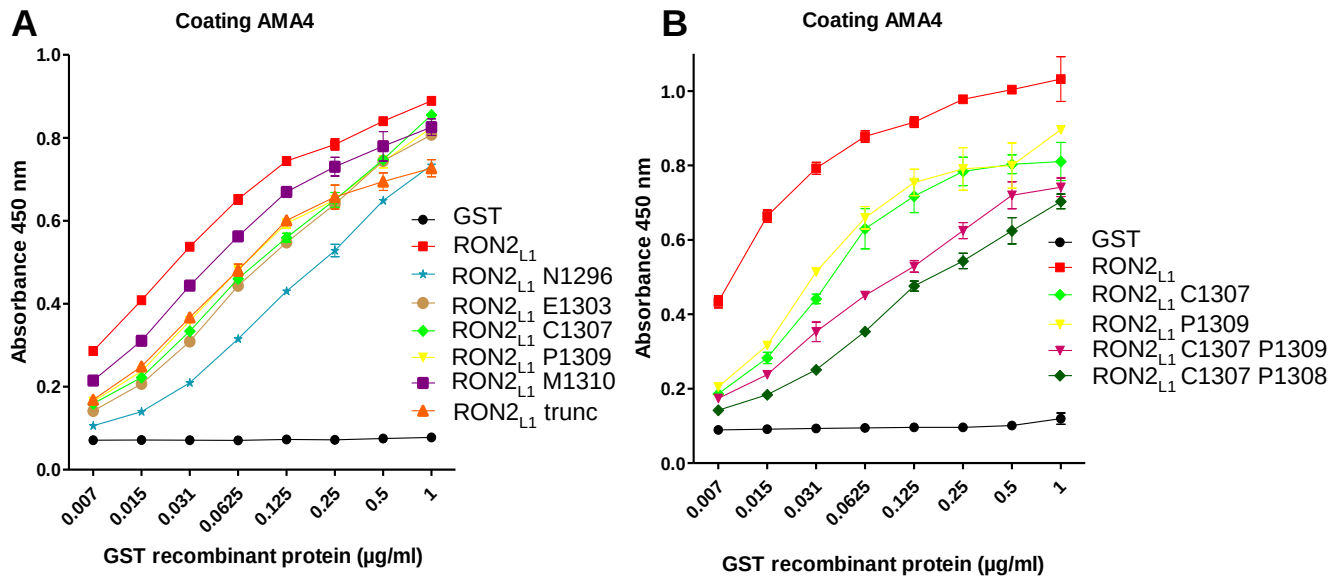
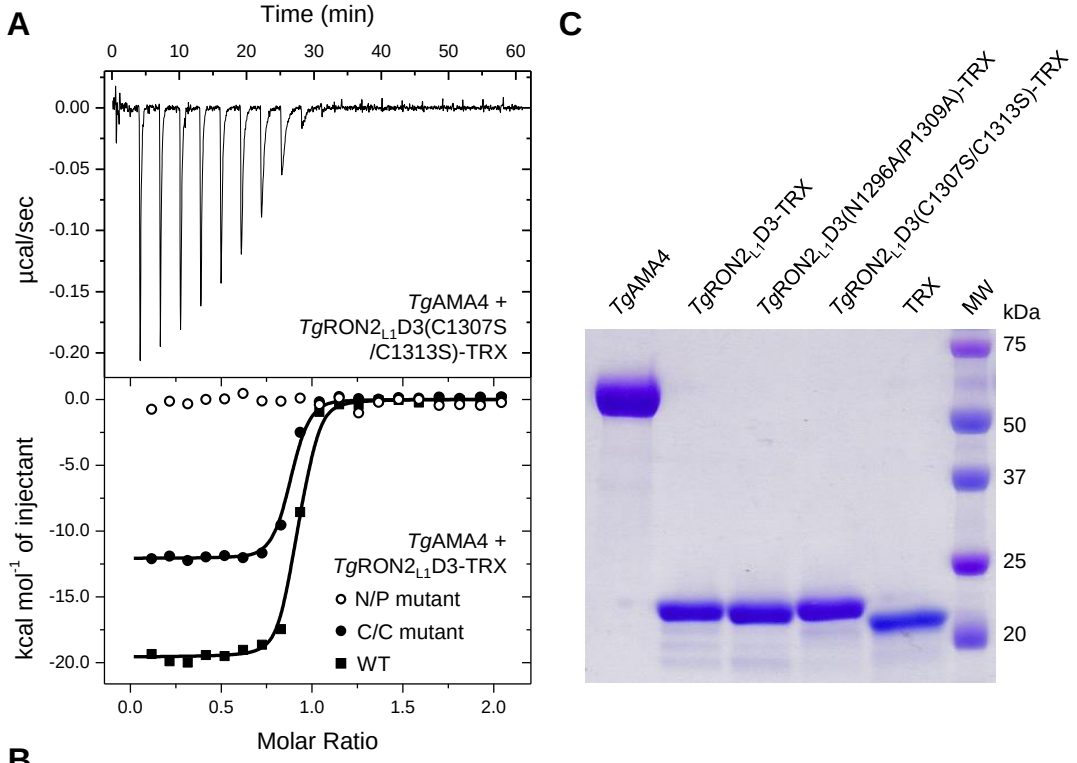


Fig. S2. *Tg*AMA4-*Tg*RON2$_{L1}$D3 form a composite interface with several residues important for their global interaction. ELISA performed in plates coated with recombinant *Tg*AMA4 DIDIIEGF1 protein testing either the wild type or the mutated *Tg*RON2$_{L1}$D3-GST proteins. GST was used as a control. Values represent means ± SD, n=2, from a representative experiment out of 3 independent assays. *(A)* Single mutants. *(B)* Cystine loop interrogation.



B

Protein	Ligand	K_D (nM)	Molar Ratio ($RON2_{L1}$/AMA4)	ΔH (kcal/mol)	-TΔS (kcal/mol)	ΔG (kcal/mol)
***Tg*AMA4**	*Tg*$RON2_{L1}$D3	12.0 ± 2.1	0.87 ± 0.01	-19.6 ± 0.1	8.8	-10.8
	Tg$RON2_{L1}$D3 N/P	ND	---	---	---	---
	Tg$RON2_{L2}$D3 C/C	11.0 ± 2.2	0.83 ± 0.01	-12.2 ± 0.1	1.3	-10.9
	TRX	ND	---	---	---	---

ND, interaction not detected

Fig. S3. *Tg*$RON2_{L1}$D3 double mutations disrupt the interaction or alter the binding mode with *Tg*AMA4. *(A)* ITC thermogram of *Tg*AMA4 interacting with native and mutant *Tg*$RON2_{L1}$D3-TRX. Upper panel shows heats recorded for the interaction of *Tg*AMA4 with the double cystine mutant of *Tg*$RON2_{L1}$D3-TRX. Lower panel of integrated data shows the lack of detectable interaction of *Tg*AMA4 with the Asn/Pro double mutant of *Tg*$RON2_{L1}$D3-TRX and the altered binding profile with the double Cys mutant compared to the wild-type (WT) sequence. *(B)* Table of ITC results for experiments performed at 25 °C. *(C)* Commassie Blue stained SDS-PAGE gel of purified recombinant protein samples used for ITC analysis.

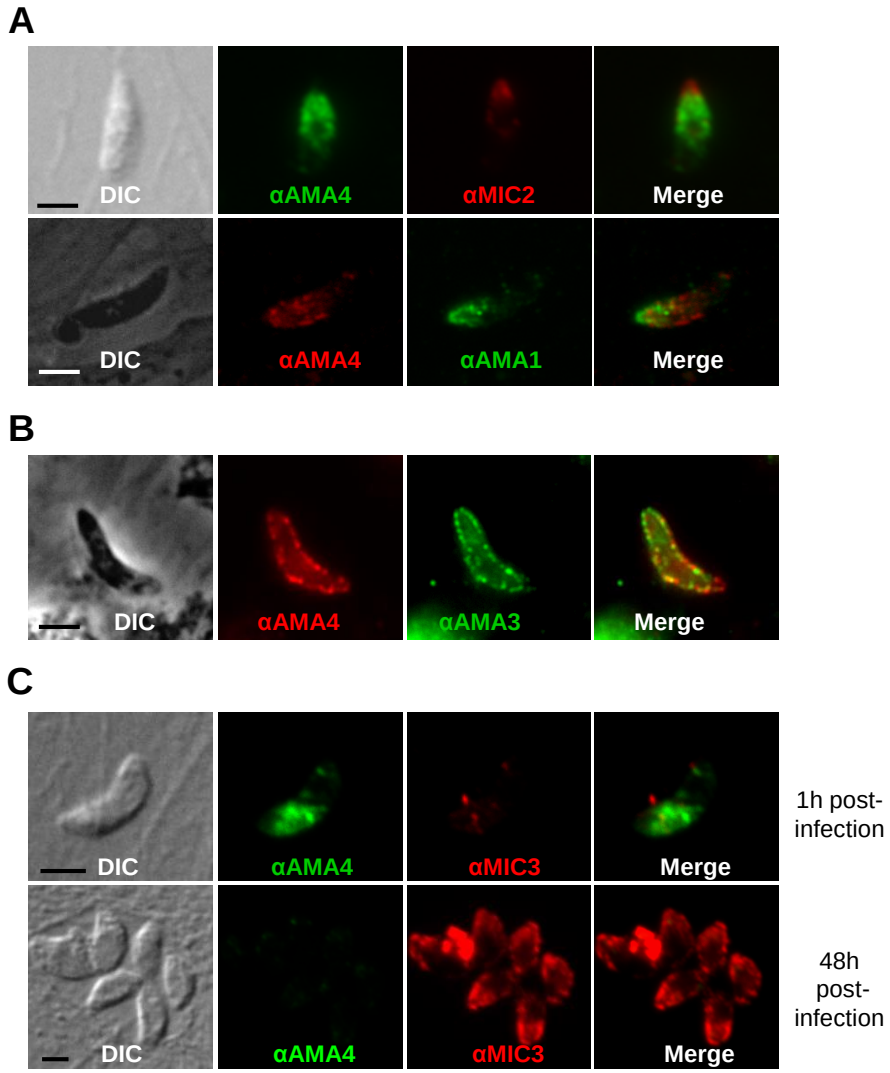


Fig. S4. *(A)* Immunofluorescence on Triton X-100 permeabilized sporozoites one hour post-infection reveals partial colocalization with microneme markers *Tg*MIC2 (top) and *Tg*AMA1 (bottom). This observation is consistent with the previously observed localization pattern for *Tg*AMA3, which might be explained by the fact that sporozoites tend to have many more micronemes than tachyzoites, and the staining pattern could reflect distinct subpopulations of micronemes in the sporozoite stage (7). DIC, differential interference contrast. Scale bars, 5 µm. *(B)* Immunofluorescence on extracellular non-permeabilized sporozoites reveals the

redistribution of *Tg*AMA4 on the entire surface of the parasite, a characteristic shared by most micronemal proteins, including the sporozoite *Tg*AMA3 protein. *(C)* Immunofluorescence after one hour (top; sporozoite stage) or 48 hours (bottom; tachyzoite stage) infection of HFFs cells with sporozoites using the anti-*Tg*AMA4 and anti-*Tg*MIC3 antibodies. Note that the tachyzoite microneme protein *Tg*MIC3 is not expressed in sporozoite, but is present in the micronemes after conversion to tachyzoite, and that *Tg*AMA4 behaves in exactly the inverse pattern, being expressed in sporozoite and absent in tachyzoite.

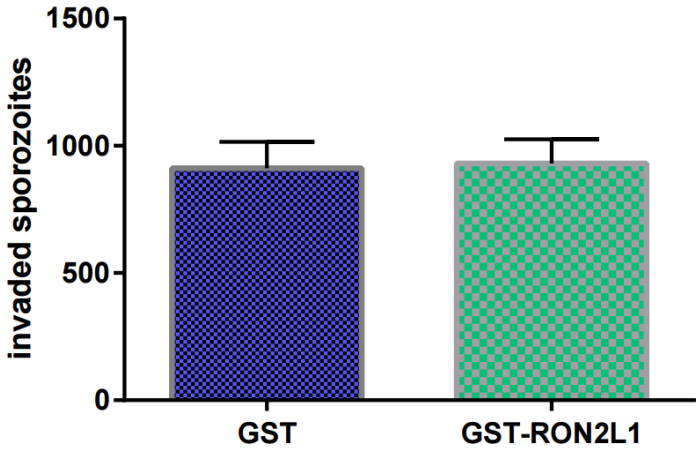


Fig. S5. Invasion assay of ME49 sporozoites into HFF cells in the presence of 200 µg/ml of GST or *Tg*RON2$_{L1}$D3-GST. No inhibitory effect was observed in three independent experiments.

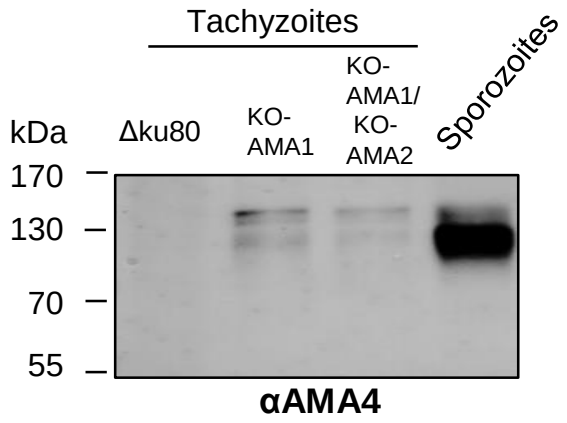


Fig. S6. Western blot on one million tachyzoites from Δ*ku80* (parental strain), KO-AMA1 and KO-AMA1/KO-AMA2, and of sporozoites using anti-*Tg*AMA4 antibodies. Note that in tachyzoites *Tg*AMA4 displayed three bands, and that the lower one becomes dominant in sporozoite. Whether this reflects different alternative splicing, proteolytic cleavages or post-translational changes between the tachyzoite and sporozoite stages will need further investigations.

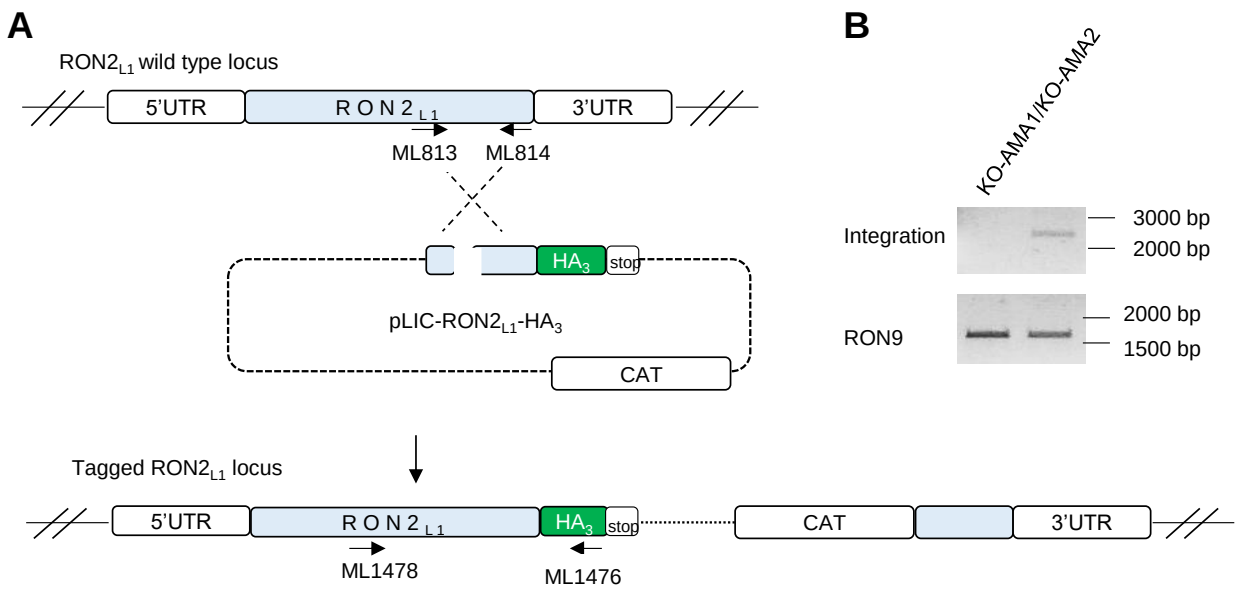


Fig. S7. C-terminal HA_3 tagging of *Tg*$RON2_{L1}$ locus in KO-AMA1/KO-AMA2 parasites. *(A)* Scheme of the pLIC-$RON2_{L1}$-HA plasmid used to introduce three HA epitopes in the $RON2_{L1}$ locus by single homologous recombination. *(B)* PCR to verify pLIC-$RON2_{L1}$-HA integration in KO-AMA1/KO-AMA2 transfected and non-transfected parasites using primers ML1478 and ML1476. The amplification of the *Tg*RON9 locus is shown as a control.

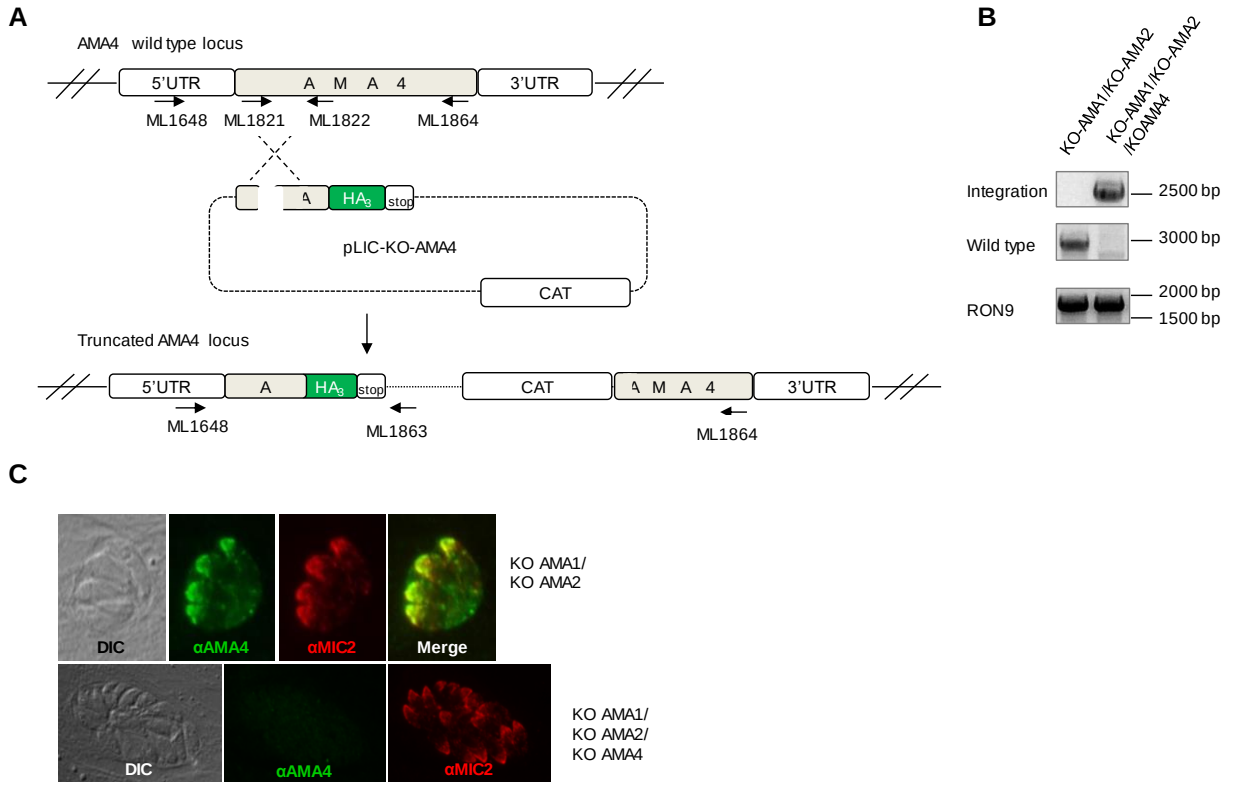


Fig. S8. Disruption of the *Tg*AMA4 locus in KO-AMA1/KO-AMA2 parasites. *(A)* Scheme of the *Tg*AMA4 disruption strategy. The pLIC-KO-AMA4 plasmid containing the first 1300 bp of the AMA4 coding sequence, three HA epitopes and the chloramphenicol acetyltransferase resistance was used to truncate the *Tg*AMA4 wild type locus. After single homologous recombination the pLIC-KO-AMA4 plasmid is integrated giving to the truncated *Tg*AMA4 locus. *(B)* PCR to verify pLIC-KO-AMA4 integration in KO-AMA1/KO-AMA2 transfected and non-transfected parasites using primers ML1648 and ML1863. The amplification of the wild type locus (with primers ML1648 and ML1864) and RON9 locus are also shown as controls. *(C)* IFAs on intracellular KO-AMA1/AMA2 and KO-AMA1/AMA2/AMA4 strains using anti-*Tg*AMA4 antibodies show the depletion of *Tg*AMA4 in the triple functional mutant and reveals the specificity of the anti-*Tg*AMA4 antibodies.

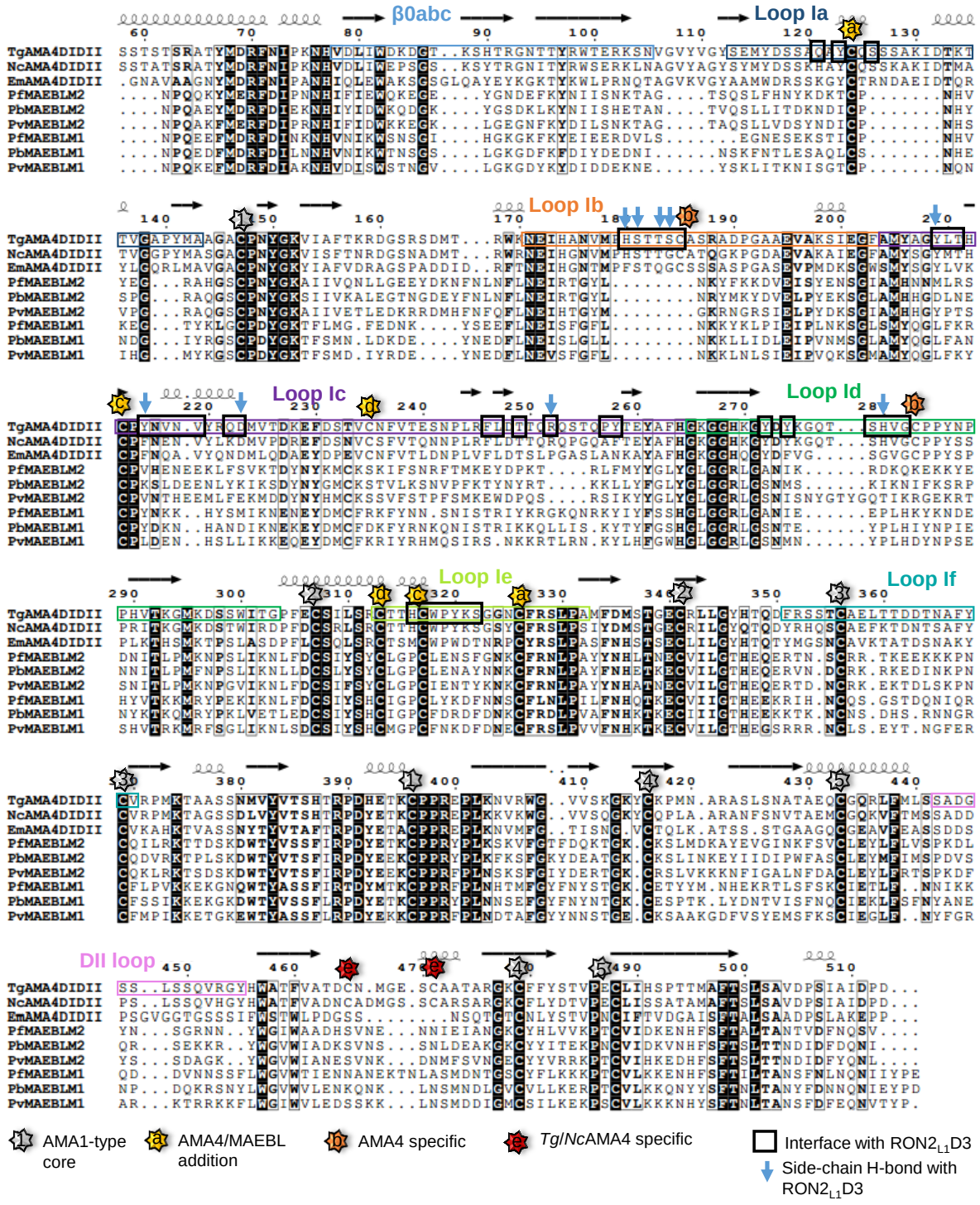


Fig. S9. Sequence alignment of *Tg*AMA4 with representative AMA/MAEBL family members. Sequences of DI and DII from *Tg*AMA4 (TGME49_294330), *Nc*AMA4 (NCLIV_001350), *Eimeria maxima* AMA4 (*Em*AMA4, EMWEY_00022320), *Pf*MAEBL M1 and M2

(PF3D7_1147800.1), *Pb*MAEBL M1 and M2 (PBANKA_090130.1), and *Pv*MAEBL M1 and M2 (PVX_092975) were aligned in MEGA6.0 (25) using MUSCLE (2) and illustrated in ESPript (26). Minor adjustments were made manually to anchor the alignment on the Cys residues. DI and DII loops are annotated and colored as in Fig. 3D. Disulfides that are part of the AMA1-type core are numbered and indicated with a grey starburst; disulfides that are conserved in AMA4/MAEBL family members are lettered and indicated with a yellow starburst; disulfide unique to *Eimeriorina* AMA4s is lettered and indicated with an orange starburst; disulfide unique to *Tg*/*Nc*AMA4 is lettered and indicated with a red starburst. *Tg*AMA4 residues at the interface with *Tg*RON2$_{L1}$ (>1 Å^2 BSA) are highlighted with a thick black box; residues that form side-chain dependent hydrogen bonds are indicated with blue arrows.

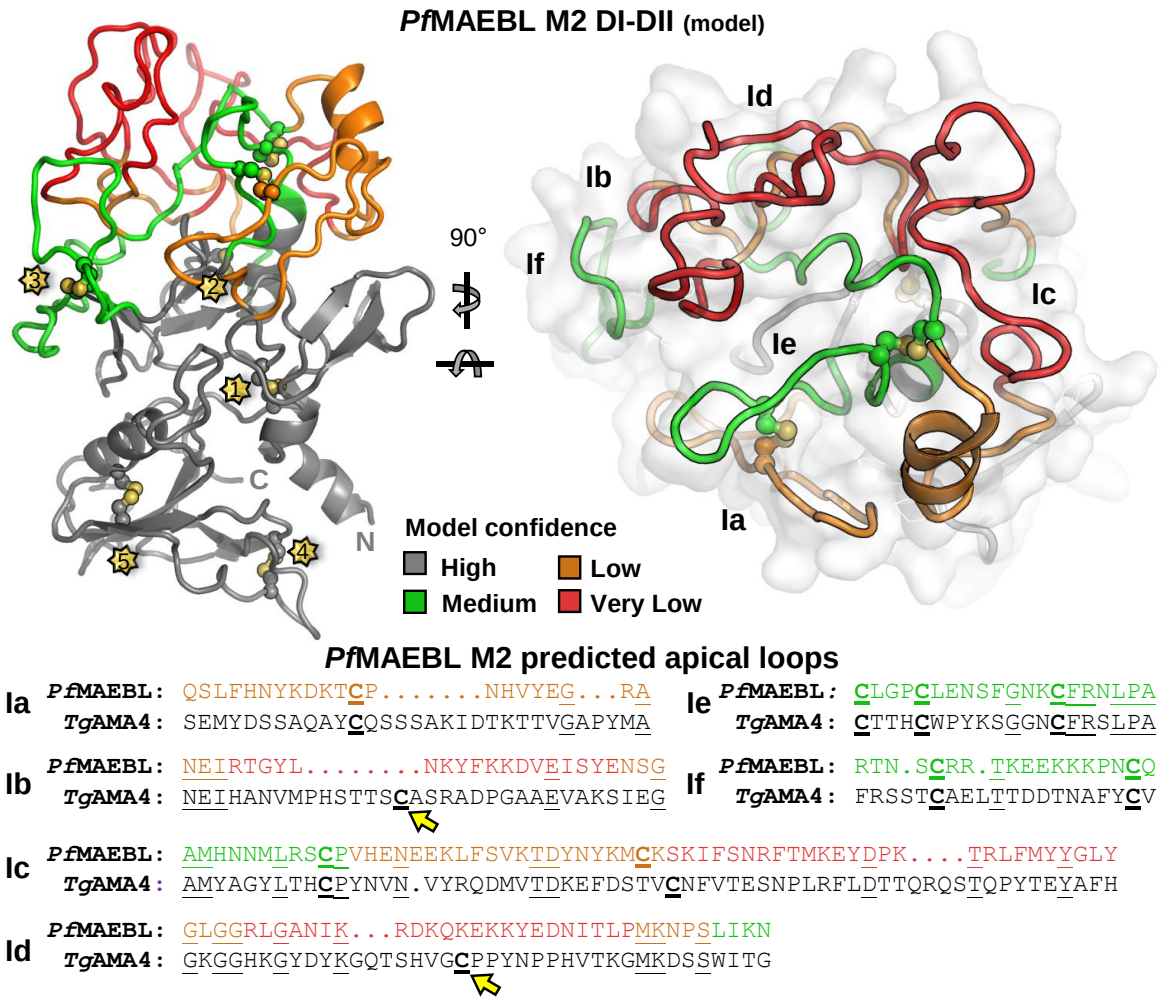


Fig. S10. Homology modeling of *Pf*MAEBL M2 suggests a well conserved core, but a restructured apical surface. *Left* - Cartoon representation of *Pf*MAEBL M2 homology model, based on the *Tg*AMA4 DI/DII structure, colored on a scale of confidence regarding the three-dimensional positions from grey (high confidence) to red (very low confidence). The grey core is consistent with MAEBLs likely adopting two vertically stacked Plasminogen/Apple/Nematode (PAN) domains. The conserved length and cysteine-rich nature of loops Ie and If suggest similar arrangements between *Tg*AMA4 and *Pf*MAEBL M2. However, the remaining four DI loops have deletions of up to 30% and low sequence identity with *Tg*AMA4 resulting in reduced model

confidence. Loops Ia and Ic are predicted to be disulfide-pinned to Loop Ie, but the organization of the two loops apart from the anchored residues is unclear. Perhaps most importantly, the disulfide bond in *Tg*AMA4 DI that pins Loop Ib and Id to form the *Tg*RON2$_{L1}$D3 cystine loop binding pocket is not conserved in MAEBL M1 or M2. Disulfide bonds are shown in ball and stick and colored by element; conserved disulfides of the structural core are indicated by a yellow starburst with numbers correlating to Fig. 2. *Right* – Apical view of the *Pf*MAEBL M2 model, shown in the same orientation as Fig. 3D, with a semi-transparent white surface. *Bottom* – Sequence alignments of the *Tg*AMA4 apical loops with the corresponding predicted loop regions of *Pf*MAEBL M2 extracted from a multiple sequence alignment of AMA4/MAEBL proteins (Fig. S7); the yellow arrows indicate the Loop Ib-Id cysteines of *Tg*AMA4 that are absent in *Pf*MAEBL M2. Note by comparison with Fig. S7 that several of the most important *Tg*AMA4 residues responsible for coordinating *Tg*RON2$_{L1}$D3 are located in DI loop insertions that are not conserved in MAEBL M1 or M2 (e.g. Loop Ib).

Supplementary Tables

Table S1. Data collection and refinement statistics.

	*Tg*AMA4	*Tg*AMA4SeMet - *Tg*RON2$_{L1}$D3SeMet	*Tg*AMA4 - *Tg*RON2$_{L1}$D3
A. Data collection statistics			
Spacegroup	$P2_1$	$P6_{1/5}$	$P6_{1/5}$
Cell dimensions			
a, b, c (Å)	39.2, 202.6, 72.8	121.6, 121.6, 143.6	120.6, 120.6, 141.7
α, β, γ (°)	90, 100.85, 90	90, 90, 120	90, 90, 120
Resolution (Å)	50.65-2.05 (2.16-2.05)*	84.92-2.10 (2.15-2.10)	37.18-1.53 (1.56-1.53)
Measured reflections	164,443 (20,939)	2,240,997 (120,078)	1,327,154 (54,976)
Unique reflections	66,314 (9,231)	70,037 (4,530)	175,481 (8,622)
Redundancy	2.5 (2.3)	32.0 (26.5)	7.6 (6.4)
Completeness (%)	95.4 (91.3)	100.0 (100.0)	99.9 (99.4)
$I/\sigma(I)$	6.6 (2.1)	17.6 (7.1)	10.5 (2.6)
R_{merge}	0.095 (0.502)	0.156 (0.465)	0.102 (0.458)
$CC_{1/2}$	0.989 (0.666)	0.998 (0.971)	0.996 (0.853)
B. Refinement statistics			
Spacegroup	$P2_1$		$P6_5$
Resolution (Å)	49.09-2.05		37.18-1.53
R_{work} / R_{free}	0.172/0.209		0.153/0.172
No. of atoms			
Protein (A/B/C/D)	3835/3800		3608/3613/225/217
Glycerol/Sulfate	12		24/15
Water	677		1361
B-factors (Å^2)			
Protein (A/B/C/D)	32.7/38.2		18.3/19.8/27.7/32.1
Glycerol/Sulfate	33.1		29.2/26.8
Water	37.8		32.7
r.m.s. deviation from ideality			
Bond lengths (Å)	0.003		0.006
Bond angles (°)	0.782		1.083

*Highest resolution shell is shown in parenthesis

Table S2. Interactions at the *Tg*AMA4-*Tg*RON2$_{L1}$ interface (chains A-C). Residues chosen for mutation in this study are shown in bold; deleted residues shown in italics; BSA, buried surface area.

***Tg*RON2$_{L1}$D3 feature**	***Tg*RON2$_{L1}$**	**BSA (>15 Å^2)**	***Tg*RON2$_{L1}$**	***Tg*AMA4**	**Distance (Å)**
N-terminal coil	Ile1293	36			
	Val1294	68	Val1294 [O]	Ser322 [N]	2.88
	Gln1295	37			
Helix	**Asn1296**	97	**Asn1296** [N]	Tyr320 [O]	2.89
			Asn1296 [Nδ2]	Trp318 [O]	2.86
			Asn1296 [Oδ1]	Tyr320 [N]	2.92
	Gln1297	105	Gln1297 [N]	Asp223 [Oδ2]	2.89
	Ser1298	86	Ser1298 [O]	Tyr209 [OH]	3.11
			Ser1298 [Oγ]	His316 [O]	2.65
	Ser1299	51	Ser1299 [Oγ]	His180 [Nε2]	2.84
	Ala1301	46			
	Pro1302	38			
	Glu1303	44	**Glu1303** [Oε1]	Thr183 [Oγ1]	2.61
			Glu1303 [Oε2]	Ser181 [Oγ]	2.79
			Ser1305 [Oγ]	Tyr215 [OH]	2.66
Cystine loop (**Cys1307** to **Cys1313**)	**Pro1308**	55	**Pro1308** [O]	Ser184 [Oγ]	3.22
			Pro1308 [O]	His280 [Nε2]	2.94
	Pro1309	103			
	Met1310	164	**Met1310** [O]	Thr211 [N]	3.17
	Gly1311	20			
	Ile1312	113			
C-terminal coil	Met1314	52			
	Gly1316	43	*Gly1316* [O]	Arg252 [NH2]	2.79
	Ile1318	36			
	Pro1321	46			
	Ile1322	65			

Table S3. Primers used in this study.

Primer name	Primer sequence	Construct /PCR
*Tg*AMA4 forward	ACATGACCATGGGAAGCAGCACAAGC	*Tg*AMA4 DIDIIEGF1
*Tg*AMA4 EGF1 reverse	CTGTCTGGCGGCCGCATCGCACCTTTCTCCGGTG	
*Tg*RON2$_{L1}$F.539	TATGGGATCCAGCTCTGCCACAGAATGGGCGA	*Tg*RON2$_{L1}$ (L539-Y983)
*Tg*RON2$_{L1}$R.983	CATATCAGTAGCGAGAGTCGTTGGTAACGT	
*Tg*AMA3.F.393	TATGGGATCCCGAACTGGGCGAACTTTTACCT	*Tg*AMA3 (N393-S566)
*Tg*AMA3.R.566	GATCTCAGCTGCCTTCTTTCTCCACAGTCT	
ML1821 forward	TACTTCCAATCCAATTTAATGCACGGTCAGGAAAA TCCTACAAGC	pLIC-KO-AMA4-CAT
ML1822 reverse	TCCTCCACTTCCAATTTTAGCGTTACCGCCACTTTT GTATGGC	
ML813 forward	TACTTCCAATCCAATTTAATGCGCGGAAAAAGATG GACGCATCCTCC	pLIC-RON2$_{L1}$-HA3-CAT
ML814 reverse	TCCTCCACTTCCAATTTTAGCCAGTTTATCGAAAC GAGCAAGCAG	
ML1648 forward	CCAGTCGACACCTAG	PCR for pLIC-KO-AMA4-CAT integration
ML1863 reverse	CATTCGCAGACGGATCGTCT	
ML1864 reverse	CAGGGAAGTGAAGGCCATGG	PCR for AMA4 wild type amplification
ML 1478 forward	GAAAAGGACAATTCCCAACCAG	PCR for pLIC-RON2$_{L1}$-HA3-CAT integration
ML 1476 reverse	CAGCGTAGTCCGGGACGTCGTAC	
ML1957 forward	CGCCAAGCTGCTCTGAGCCTGGACGATCTGCTCA	*Tg*RON2$_{L1}$ N1297A
ML1958 reverse	TGAGCAGATCGTCCAGGCTCAGAGCAGCTTGGCG	
ML1959 forward	CAGCCACTCAACGCGGGCGCCAAGC	*Tg*RON2$_{L1}$ E1303A
ML1960 reverse	GCTTGGCGCCCGCGTTGAGTGGCTG	
ML196 1 forward	GCCCATGGGCGGGGCGCCACTCAACTCG	*Tg*RON2$_{L1}$ C1307A
ML1962 reverse	CGAGTTGAGTGGCGCCCCGCCCATGGGC	
ML1963 forward	CAAATGCCCATGGCCGGGCAGCCACTC	*Tg*RON2$_{L1}$ P1309A
ML1964 reverse	GAGTGGCTGCCCGGCCATGGGCATTTG	
ML1965 forward	CCATGCAAATGCCCGCGGGCGGGCAGCCAC	*Tg*RON2$_{L1}$ M1310A
ML1966 reverse	GTGGCTGCCCGCCCGCGGGCATTTGCATGG	
ML1967 forward	GCATTTGCATGGACCCTCCGGCGACCGC	*Tg*RON2$_{L1}$D3 truncation
ML1968 reverse	GCGGTCGCCGGAGGGTCCATGCAAATGC	
ML2060 forward	AAATGCCCATGGGCGCGGCGCCACTCAACTCGGG	*Tg*RON2$_{L1}$ C1307A/P1308A
ML2061 reverse	CCCGAGTTGAGTGGCGCCGCGCCCATGGGCATTT	
ML2062 forward	CAAATGCCCATGGCCGGGGCGCCACTCAACTCG	*Tg*RON2$_{L1}$ C1307A/P1309A
ML2063 reverse	CGAGTTGAGTGGCGCCCCGGCCATGGGCATTTG	

Table S4. Antibodies used in this study.

Antibody	Western blot dilution	IFA dilution
Rabbit anti-*Tg*AMA4 (this study)	1:1000	1:1000
Mouse anti-*Tg*RON2$_{L1}$ (this study)	1:1000	1:100
Mouse Mab anti-*Tg*AMA1 B3.90 (27)	-	1:1000
Mouse anti-*Tg*MIC2 T34A11 (28)	-	1:100
Mouse Mab anti-*Tg*MIC3 T4 2F3 (29)	-	1:200
Rabbit anti-*Tg*RON9 (30) PEST	-	1:200
Mouse MAb T5 3E2 anti-*Tg*ROP5 (31)	1:1000	-
Rat anti-HA (Roche, clone 3F10)	1:100	-
Rabbit anti-*Tg*ROP1 S2b (32)	-	1:1000
Mouse MAb 2E5 anti-*Tg*SAG1 (33)	-	1:1000

Supplementary References

1. Johnson LS, Eddy SR, & Portugaly E (2010) Hidden Markov model speed heuristic and iterative HMM search procedure. *BMC Bioinformatics* 11:431.
2. Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 32(5):1792-1797.
3. Tamura K, *et al.* (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* 28(10):2731-2739.
4. Darriba D, Taboada GL, Doallo R, & Posada D (2011) ProtTest 3: fast selection of best-fit models of protein evolution. *Bioinformatics* 27(8):1164-1165.
5. Lamarque MH, *et al.* (2014) Plasticity and redundancy among AMA-RON pairs ensure host cell entry of Toxoplasma parasites. *Nat Commun* 5:4098.
6. Sigrist CJ, *et al.* (2013) New and continuing developments at PROSITE. *Nucleic Acids Res* 41(Database issue):D344-347.
7. Poukchanski A, *et al.* (2013) Toxoplasma gondii sporozoites invade host cells using two novel paralogues of RON2 and AMA1. *PLoS One* 8(8):e70637.
8. Tonkin ML, Beck JR, Bradley PJ, & Boulanger MJ (2014) The inner membrane complex sub-compartment proteins critical for replication of the apicomplexan parasite Toxoplasma gondii adopt a pleckstrin homology fold. *J Biol Chem* 289(20):13962-13973.
9. Battye TG, Kontogiannis L, Johnson O, Powell HR, & Leslie AG (2011) iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. *Acta Crystallogr D Biol Crystallogr* 67(Pt 4):271-281.
10. Evans PR & Murshudov GN (2013) How good are my data and what is the resolution? *Acta Crystallogr D Biol Crystallogr* 69(Pt 7):1204-1214.
11. Sheldrick GM (2010) Experimental phasing with SHELXC/D/E: combining chain tracing with density modification. *Acta Crystallogr D Biol Crystallogr* 66(Pt 4):479-485.
12. Cowtan K (2008) Fitting molecular fragments into electron density. *Acta Crystallogr D Biol Crystallogr* 64(Pt 1):83-89.
13. Winn MD, *et al.* (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr* 67(Pt 4):235-242.
14. McCoy AJ (2007) Solving structures of protein complexes by molecular replacement with Phaser. *Acta Crystallogr D Biol Crystallogr* 63(Pt 1):32-41.
15. Emsley P, Lohkamp B, Scott WG, & Cowtan K (2010) Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* 66(Pt 4):486-501.
16. Afonine PV, *et al.* (2012) Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr D Biol Crystallogr* 68(Pt 4):352-367.
17. Adams PD, *et al.* (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* 66(Pt 2):213-221.
18. Chen VB, *et al.* (2010) MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr* 66(Pt 1):12-21.
19. Paredes-Santos TC, *et al.* (2013) Spontaneous cystogenesis in vitro of a Brazilian strain of Toxoplasma gondii. *Parasitol Int* 62(2):181-188.
20. Huynh MH & Carruthers VB (2009) Tagging of endogenous genes in a Toxoplasma gondii strain lacking Ku80. *Eukaryot Cell* 8(4):530-539.

21. Zhang Y (2008) I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics* 9:40.
22. Pettersen EF, *et al.* (2004) UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem* 25(13):1605-1612.
23. Eswar N, *et al.* (2006) Comparative protein structure modeling using Modeller. *Curr Protoc Bioinformatics* Chapter 5:Unit 5 6.
24. Wallner B & Elofsson A (2003) Can correct protein models be identified? *Protein Sci* 12(5):1073-1086.
25. Kumar S, Nei M, Dudley J, & Tamura K (2008) MEGA: a biologist-centric software for evolutionary analysis of DNA and protein sequences. *Briefings in Bioinformatics* 9(4):299-306.
26. Gouet P, Courcelle E, Stuart DI, & Metoz F (1999) ESPript: analysis of multiple sequence alignments in PostScript. *Bioinformatics* 15(4):305-308.
27. Donahue CG, Carruthers VB, Gilk SD, & Ward GE (2000) The Toxoplasma homolog of Plasmodium apical membrane antigen-1 (AMA-1) is a microneme protein secreted in response to elevated intracellular calcium levels. *Mol Biochem Parasitol* 111(1):15-30.
28. Achbarou A, *et al.* (1991) Characterization of microneme proteins of Toxoplasma gondii. *Mol Biochem Parasitol* 47(2):223-233.
29. Garcia-Reguet N, *et al.* (2000) The microneme protein MIC3 of Toxoplasma gondii is a secretory adhesin that binds to both the surface of the host cells and the surface of the parasite. *Cell Microbiol* 2(4):353-364.
30. Lamarque MH, *et al.* (2012) Identification of a new rhoptry neck complex RON9/RON10 in the Apicomplexa parasite Toxoplasma gondii. *PLoS One* 7(3):e32457.
31. El Hajj H, Lebrun M, Fourmaux MN, Vial H, & Dubremetz JF (2007) Inverted topology of the Toxoplasma gondii ROP5 rhoptry protein provides new insights into the association of the ROP2 protein family with the parasitophorous vacuole membrane. *Cell Microbiol* 9(1):54-64.
32. Besteiro S, Michelin A, Poncet J, Dubremetz JF, & Lebrun M (2009) Export of a Toxoplasma gondii rhoptry neck protein complex at the host cell membrane to form the moving junction during invasion. *PLoS Pathog* 5(2):e1000309.
33. Couvreur G, Sadak A, Fortier B, & Dubremetz JF (1988) Surface antigens of Toxoplasma gondii. *Parasitology* 97 (Pt 1):1-10.