A central protein complex essential for Invasion in *Toxoplasma gondii*

Mirko Singer^{1*}, Kathrin Simon¹, Ignasi Forné², Markus Meissner^{1*}

¹ Experimental Parasitology, Faculty of Veterinary Medicine, Ludwig-Maximilians-University (LMU) Munich, Germany.

² Protein Analysis Unit, Biomedical Center (BMC), Faculty of Medicine, Ludwig-Maximilians-University (LMU) Munich, Großhaderner Straße 9, 82152, Martinsried, Germany.

* To whom correspondence should be addressed: Mirko.Singer@lmu.de, Markus.Meissner@lmu.de

Apicomplexa are obligate intracellular parasites. While most species are restricted to specific hosts and cell types, *Toxoplasma gondii* can invade every nucleated cell derived from warm-blooded animals. This broad host range suggests that this parasite can recognize multiple host cell ligands or structures, leading to the activation of a central protein complex, which should be conserved in all apicomplexans. During invasion, the unique secretory organelles (micronemes and rhoptries) are sequentially released and several micronemal proteins have been suggested to be required for host cell recognition and invasion. However, to date only few micronemal proteins have been demonstrated to be essential for invasion.

Cysteine Repeat Modular Proteins (CRMPs) are a family of apicomplexan specific proteins. In *Toxoplasma gondii*, two CRMPs are present in the genome. The Kringle domain containing protein (CRMPA) and the GCC2 GCC3 domain containing protein (CRMPB). Here we demonstrate that both proteins form a complex that contains additional micronemal proteins. Disruption of this complex results in a block of rhoptry secretion and parasites being unable to invade the host cell. In conclusion, this complex is a central invasion complex conserved in all apicomplexans.

Introduction

The apicomplexan parasite *Toxoplasma gondii* is capable to actively invade every nucleated vertebrate cell. How the parasite is capable to achieve such an extraordinary broad host range is unknown. While it has been suggested that the parasite expresses a large repertoire of receptors derived from unique secretory organelles (micronemes) that can bind to a great variety of surface markers¹, most micronemal proteins can be deleted without significantly affecting invasion *in vitro*. It is possible that the parasite can detect highly conserved surface proteins, glycosylated proteins or even more

conserved properties of cells like surface charge or membrane components of the plasma membrane. Each invasion event requires the discharge of the rhoptries, specialized secretory organelles found in all invasive stages of apicomplexans, and the establishment of the junctional complex, through which the parasite invades². Since this step is conserved in all apicomplexan parasites, we hypothesized the presence of a central protein complex, acting upstream of rhoptry secretion and being required for invasion into different host cells.

Host cell invasion is a stepwise process³. Upon initial contact with the host cell, micronemes are

secreted. The micronemes contain a huge variety of micronemal proteins that also appear to function in a stepwise process from gliding motility, host cell recognition, activation of rhoptry secretion and force transmission during parasite entry. Surprisingly, only few micronemal proteins have been demonstrated to be critical for host cell invasion, which might be due to redundancies within the arsenal of micronemal proteins, as suggested for AMA1^{4,5}. Signaling cascades resulting in secretion of micronemes are conserved between *Toxoplasma* and *Plasmodium*⁶. The same is true for the main components of the gliding and invasion machinery⁷.

In the case of *Toxoplasma gondii*, the micronemal protein Mic8 appears to act upstream of rhoptry secretion⁸ and it was suggested to be a receptor triggering secretion of the rhoptries. However, to date neither interaction partners, nor host cell ligands could be identified that specifically interact with this protein, making its exact function enigmatic. Intriguingly, although crucial for host cell invasion by *Toxoplasma*, no direct homolog can be found in *Plasmodium*. The related protein Mic7 was recently characterized, and is important, but not essential for invasion of *T. gondii*⁹. Again, no direct homolog of MIC7 can be identified in *Plasmodium*.

Recently, a genome wide screen allowed the identification of the claudin-like apicomplexan microneme protein (CLAMP) that appears to have a role in formation of the tight junction through which the parasite invades¹⁰. Unlike MIC7 and MIC8, CLAMP is highly conserved in apicomplexans and therefore might play a central role in the initiation of rhoptry secretion. Furthermore, comparative analysis of specialized secretory organelles across the superphylum of alveolates led to the identification of unique structures that appear to be conserved, such as a central fusion rosette that was identified in *Paramecium*¹¹, *Tetrahymena*¹² and later also Coccidia, leading to the speculation that these organelles are sharing the same evolutionary history¹³. Analysis of secretory mutants of *P. tetraurelia* led to the identification of Nd proteins, such as Nd6 or Nd9 that are required for rosette formation and trichocyst secretion^{14,15}. Interestingly, the armadillo repeat protein Nd9p¹⁶ and RCC1-like protein Nd6p¹⁷, identified in *Paramecium* are also highly conserved in apicomplexan parasites, where they appear to play the same, conserved role in rosette formation and rhoptry secretion¹⁸.

More recently the rhoptry secretion system has been characterized with electron tomography in multiple species of apicomplexa^{19,20,21}, but integration of this structural information and single protein function is still in its infancy.

In *Plasmodium*, stage specific invasion factors and their receptors have been identified. In merozoites of *P. falciparum* the essential complex of RH5, RIPR and CyRPA is binding basigin on the erythrocytes surface via RH5 to trigger rhoptry secretion²². In *P. berghei* sporozoites both CD36 and CD52 are essential for invasion into hepatocytes, binding CD81²³. Of all these, RH5 is restricted to the Laveranian subgenus in *Plasmodium*²⁴, all other members have no homologs outside of *Plasmodium*. The switch from semi-selective to cell-type-selective invasion in the evolution of haemosporidia²⁵ suggests that restriction to a limited host cell repertoire is a secondary trait. Again, this would predict that species and stage specific receptors are modulating a central invasion complex.

In *P. berghei*, four Cysteine Repeat Modular Proteins (CRMPs) have been shown to have diverse phenotypes in the mosquito stage, including oocyst egress, salivary gland invasion and hepatocyte invasion^{26,27}. Individual gene deletion results in mild growth phenotypes of blood stage parasites in both *P. berghei* and *P. falciparum*^{28,29}.

Here we focused on the function of CRMPs in *Toxoplasma gondii*, since both copies were predicted to be essential in a genome wide screen¹⁰, are predicted to be multiple transmembrane domain (TM) proteins and form a separate cluster in a recently performed hyperplexed localization of organelle proteins by isotope tagging (hyperLOPIT) analysis³⁰.

We performed a bioinformatic analysis and identified Cysteine Repeat Modular Proteins (CRMPs) as a family of apicomplexan specific proteins that are widely conserved within the phylum. Both CRMPs are specifically required for rhoptry secretion and subsequent host cell invasion. Finally, we show that CRMPs are present in a multi-subunit complex with at least two additional, essential micronemal proteins, Mic15 and TSP1, which are required for rhoptry secretion and invasion.

Figure 1: Phylogeny and domain architecture of Cysteine Repeat Modular Proteins. **A** Phylogenetic relationship of CRMPs is shown after alignment of the entire coding sequence, bootstrap analysis (100x) supported most branches. **B** Domain architecture of selected CRMPs. Domains were predicted with SMART³¹. The conserved core containing the transmembrane domains was defined after the amino acid alignment and is indicated with a grey box. Some large N-terminal parts containing only internal repeats are not shown. Several domains are indicated: CLECT (c-type Lectin or carbohydrate recognition domain), ERC (Ephrin-receptor like), EGF-like (Epidermal growth factor-like), TM (transmembrane domain), Kringle (Kringle domain), EGF (Epidermal growth factor), coil coil (alpha helical coil coil domains), d1eq1a_ (apolipoprotein III), He_PIG (putative Immunoglobulin-like fold), d1ds9a_ (outer arm of dynein light chain), FU (Furin-like repeats), PdH1 (parallel beta helix repeat).

Results

The family of CRMPs in apicomplexan parasites

We performed a phylogenetic analysis of apicomplexan CRMPs and looked at the domain architecture (Figure 1 A and B, Figure S1). As local conservation is very low and many CRMPs have long insertions (Figure S1 C), we defined a “conserved core” which includes two EGF-like domains and the transmembrane domains.

Alignment of this core region resulted in comparable results (see Figure S1 A). The phylogenetic tree suggests that CRMPA of Coccidia groups with CRMP1 and CRMP2 (the CRMPA clade) of Haemosporidia and CRMPB groups with CRMP3 and CRMP4 (the CRMPB clade) of Haemosporidia (Figure 1 A). We could find two putative CRMPs in Cryptosporidia but only one CRMP in the recently sequenced marine Gregarine species *Porospora gigantea*³², just as in *Gregarina niphandrodes*.

Interestingly, in both *Theileria annulata* and *Babesia bovis*, CRMP1 and CRMP2 are next to each other on the same chromosome, with a central DNA sequence that is identical in both, suggesting a conserved gene duplication events predating their split into separate species. Recent data from *Babesia bovis* suggests that one gene of each group (CRMPA clade and CRMPB clade) is expressed in bloodstages and the other one in kinetes³³. In *P. berghei* single gene deletions of all four CRMPs have been generated^{26,27} and in *P. falciparum* phenotypic scores (Mutant fitness score) of (-0,19) to (-1,59) have been assigned²⁹. In *Chromera velia*, a free-living apicomplexan ancestor, many very diverse CRMPs can be found^{34,35}. Together, this analysis demonstrates a huge diversity of apicomplexan CRMPs, with the occurrence of independent gene losses and duplications. However, at least one copy of CRMPA clade and CRMPB clade appears to be maintained throughout the evolution of apicomplexan parasites. The two CRMPs of *T. gondii*, TGGT1_261080 and TGGT1_292020 were identified as critical for the lytic cycle of the parasite in a genome wide CRISPR/Cas9 screen with a phenotypic score of -3,95 and -3,64 respectively¹⁰. Both proteins contain 9 transmembrane domains. While TGGT1_261080 contains a Kringle domain (CRMPA), TGGT1_292020 is a GCC2 and GCC3 domain containing protein (CRMPB).

CRMPA and CRMPB show a unique subcellular localization

For localization of CRMPs, we used Cas9-YFP assisted selection marker free tagging in the recipient strain Δ Ku80-DiCre³⁶ (Figure S2). We generated two parasites lines with both proteins C-terminally tagged (RH-CRMPA-3HA^{floxed} CRMPB-sYFP2 and RH-CRMPB-3HA^{floxed} CRMPA-sYFP2) (Figure 2 A, Figure S4 B). The proteins appear to localize throughout the secretory pathway of the parasite with considerable retention time in the ER, with a clear accumulation close to the apical conoidal micronemes but distinct from the micronemes associated with the subpellicular microtubules where no significant co-localization with MIC2 or MIC8³⁷ was evident (Figure 2 B). In good agreement, LOPIT (Localization of Organelle Proteins by Isotope Tagging)³⁰ predicts them as

micronemal but places them in clusters next to the ER, distinct from other micronemal proteins (TAGM-MAP is micronemal with probability 0, MAGM-MCMC is micronemal with probability 1) (Figure S1 B). The localization throughout the secretory pathway of the parasite was confirmed by colocalization with additional micronemal, Golgi and ER markers (see below). Time lapse analysis of parasites expressing Halo-labeled CRMPs (RH-CRMPB-Halo CRMPA-3HA^{floxed} and RH-Halo-CRMPA-3HA^{floxed} CRMPB-sYFP2) support a dynamic localization throughout the secretory pathway (Supplementary Movie 1).

Next, we wished to analyze the behavior of CRMPB and CRMPA within the secretory pathway and during host cell invasion and inserted an additional tag in the extracellular part of the protein (Figure S2). To determine if these proteins are delivered to the plasma membrane, we analyzed parasites expressing CRMPB with an extracellular sYFP2 tag (RH-sYFP2-CRMPB-3HA^{floxed} CRMPA-sYFP2). For this, parasites were inoculated on host cells in the presence of a GFP antibody which was visualized post fixation after permeabilization of the parasite, resulting in staining of CRMPB that was at the plasma membrane during the experiment. (Figure 2 C). This analysis revealed that in parasites that had recently invaded, CRMPB accumulates near the basal pole of the parasite or in the proximity, potentially the site of invasion. These data suggest that CRMPB is secreted to the plasma membrane.

Furthermore, analysis of freshly invaded parasites expressing double-tagged CRMPA and tagged CRMPB (RH-Halo-CRMPA-3HA^{floxed} CRMPB-sYFP2) revealed an accumulation of CRMPB and CRMPA near the basal end of the parasite irrespective if the protein is tagged C-terminally (cytosolic tail) or if the tag has been placed in the extracellular part of the protein (Figure 2 D, Figure S4 A, C and D), suggesting that these proteins are accumulating at the tight junction and are consequently excluded from the parasitophorous vacuole membrane during invasion. Therefore, we analyzed the localization of CRMPs of freshly invaded parasites relative to the tight junction using Ron2, tagged with SNAP². In some cases, we found CRMPs at or outside of the site marked by Ron2 where the tight junction closed (Figure 2 E). Live imaging (Supplementary movies 2) suggests that

Figure 2: Subcellular localization of CRMPB and CRMPA using STED microscopy. **A** Colocalization of CRMPA and CRMPB. RH-CRMPA-3HA^{flox} CRMPB-sYFP2 was stained with for HA (Star635P) and sYFP2 (Atto594). A colocalization map using the multiplication of both signals shows a high degree of colocalization within the whole cell. A parasite line with swapped tags shows comparable localization (Figure S4). **B** Colocalization of RH-CRMPB-3HA^{flox} stained for HA (Star635P) with micronemal proteins Mic2 and Mic8. **C** A pulse invasion assay of parasites with extracellular tagged CRMPB (RH-sYFP2-CRMPB-3HA^{flox} CRMPA-sYFP2) was performed in the presence of α -GFP antibodies. Post invasion parasites were fixed, permeabilized and the α -GFP antibody was visualized. See schematics of experimental design on the right. YFP signal is all CRMPs of the parasite. **D** Parasites with extracellular CRMPA Halo (RH-Halo-CRMPA-3HA^{flox} CRMPB-sYFP2) were labeled with Halo-Janelia 646 and fixed after pulse invasion in the presence of propionol. The putative invasion site is marked with an arrow. Tagging schematics of experimental design on the right. **E** Pulsed invasion of RH-CRMPB-3HA^{flox} CRMPA-sYFP2 Ron2-SNAP and RH-CRMPA-3HA^{flox} CRMPB-sYFP2 Ron2-SNAP after labeling with SNAP-Cell. Relative position of the basal Ron2 signal at the site where the parasitophorous vacuole closes and the CRMP signal were scored as colocalization, CRMP outside of Ron2 or only internal CRMP signal. Scale bars are 5 μ m.

CRMPB and CRMPA are present mainly at the basal end of moving and invading tachyzoites. A weaker signal was also observed at the apical tip in some instances. We have however no indication which of

the observed localizations is functional for invasion. In one instance, the ring formation at the tight junction is visible, suggesting that CRMPs are excluded at the tight junction. We also observed trail

deposition in RH-Halo-CRMPA but not in C-terminally labeled RH-CRMPB-Halo. This could be an indication that in a subset of the plasma membrane fraction the extracellular part is proteolytically cleaved.

CRMPs are essential for invasion

For the functional characterization of CRMPB, we used the DiCre-System³⁸ that allows conditional excision of a gene of interest. In a first attempt, we inserted a LoxP site up- and downstream of CRMPB. However, gene excision rate was around 10%, complicating a reliable characterization of the resulting phenotype (data not shown). Given the enormous size of CRMPB (27 kbp), we concluded that excision of the full-length gene via DiCre is inefficient. We previously inserted a LoxP site directly into the coding region of a gene (the Mic8 extracellular floxed parasite strain used as a positive control in this study (Figure S2 D)). Inspired by work in *P. falciparum*³⁹, we decided to use an existing intron of CRMPB to insert a LoxP-site. As branchpoints are generally rather close to the acceptor site at the 3' end of the intron^{40,41,42}, we inserted the LoxP site 58 bp downstream of the 5'-end of the intron (Figure S2 A). Induction with rapamycin resulted in efficient excision of three cysteine repeats and all transmembrane domains of CRMPB, as evidenced by plaque assay, IFA and confirmation of correct excision by genotyping (Figure 3 A and D, Figure S2 E).

In the case of CRMPA, we also introduced the LoxP site into an intron (Figure S2 B, E). Here only the last two transmembrane domains and the C-terminal tail are disrupted upon DiCre induction. Similarly to CRMPB, we achieved efficient excision upon treatment with rapamycin. Disruption of either gene resulted in parasites incapable of forming any plaques in an HFF monolayer, similar to positive control parasites (Mic8 extracellular domain) (Figure 3 A). We could still detect residual protein after 48 h and invasion was only slightly inhibited. In contrast, 72 hours after induction both proteins are undetectable in immunofluorescence assays (Figure 3 D). Consequently, all assays presented here were performed 72 hours after induction. CRMPB and

CRMPA did not have any measurable impact on gliding motility. Nevertheless, host cell invasion was severely reduced compared to wild type (Figure 3 B, C). Since this phenotype appeared similar to the phenotypes observed for CLAMP and MIC8, where parasites are unable to secrete their rhoptry content and consequently fail to establish a tight junction, we analyzed vacuole formation. Indeed, vacuole formation was significantly reduced, suggesting a block in invasion prior to rhoptry secretion. In contrast, intracellular development was not affected (Figure 3 C). In conclusion, CRMPA and CRMPB appear to act at a similar time during the invasion pathway as MIC8 and CLAMP. Analysis of the localization of Ron2-SNAP post invasion also supported a role of CRMPs in rhoptry secretion (Figure S6). We only observed a single invasion event each of parasites depleted of CRMPA and CRMPB that showed normal Ron2-SNAP localization (rhoptry staining and signal at the basal pole). All other parasites remained extracellular and Ron2-SNAP localization was within the tachyzoites, suggesting that they did not secrete any rhoptry content. We therefore speculated that CRMPA and CRMPB might act together as part of a multi-subunit complex that is required to trigger rhoptry secretion.

Many protein complexes are destabilized once one component is removed. Therefore, we analyzed the interdependence of CRMPA and CRMPB by investigating their respective localization and relative expression levels upon disruption of the putative partner (Figure 3 D, E). We measured total fluorescent intensity of individual parasites analyzing protein abundance with and without excision of the Ha tagged protein (Figure 3 E). Removal of CRMPB results in significant lower abundance of CRMPA. In contrast, removal of CRMPA had only slight effects on the protein levels of CRMPB. This suggests that the presence of CRMPB is required for CRMPA, but not vice versa. Potentially this discrepancy can be explained by expression of residual protein after rapamycin induction. The positioning of the intron results in loss of all TMs for CRMPB, but only the last two TMs of CRMPA. Together, these data indicate that both proteins interact. We therefore decided to perform a detailed analysis of this putative complex.

Figure 3: Phenotypical analysis of CRMPs. **A** Plaque formation after 7 days of CRMPB and CRMPA, both with LoxP in the Intron (Figure S2). Mic8 floxed extracellular domain (RH-Mic8extrac.^{flox}) is shown as positive control. Representative plaque assays are shown. **B** Gliding motility was assessed on FBS coated dishes 3 days post induction using rapamycin. The speed of tachyzoites moving for 10 or more consecutive seconds is shown. Statistical analysis: two-tailed student's t-test. **C** Invasion, evacuole formation and intracellular replication was measured 72 h post induction (48 h for Mic8 control). Standard deviation between three independent replicates is shown. Statistical analysis: two-tailed student's t-test. ** p<0.01. *** p<0.001. **D** Localization of triple tagged CRMPs (RH-SNAP-CRMPB-3HA^{flox} CRMPA-sYFP2 and RH-Halo-CRMPA-3HA^{flox} CRMPB-sYFP2) +/- rapamycin is shown after 72 hours. Extracellular parasites are imaged to avoid host cell background in the analysis. Channels are normalized to -rapamycin. Note that the 3HA tag is excised after rapamycin induction, whereas the extracellular tag of CRMPA is not (see Figure S2) and tagging schematics

below, also indicating the part of the protein excised on the genomic level. **E** Tachyzoites shown in **D** are analyzed for total fluorescence. Each circle represents one tachyzoite, the color indicates + / - rapamycin induction. Position on the x-axis indicates relative 3HA fluorescence, position on the y-axis the relative presence of the extracellular tag and circle size the relative presence of the other CRMP (sYFP2). Each measurement is normalized to the average intensity within all tachyzoites analyzed of the no rapamycin group. Statistical analysis: two-tailed student's t-test. All comparison of + and - rapamycin are $p < 0.0001$, except CRMPB YFP + / - R: $p = 0.23$. Scale bars are 5 μm .

CRMPA and CRMPB are in a complex with at least two additional invasion factors

To test if both CRMPs are interacting and form a complex and to identify additional interaction partners, we performed mass spectrometry on three separate parasite lines (Figure 4). First, we performed pulldown of RH-CRMPB-3HA^{flox} CRMPA-sYFP2, RH-CRMPA-3HA^{flox} CRMPB-sYFP2 and DiCre parasites using magnetic beads coated with α -Ha antibodies. Only few proteins were enriched in both pulldowns compared to the wild type control, Mic15 (TGGT1_247195) and both CRMPs themselves (Figure 4 A, B, D). To achieve a higher sensitivity, we tagged CRMPB with TurboID at the C-terminus (RH-CRMPB-TurboID)⁴³. Biotin labeling was performed for 30 minutes on extracellular tachyzoites (Figure 4 C). Several proteins were enriched in comparison to wild type tachyzoites. In good agreement with the pulldown, CRMPB, CRMPA and Mic15 were identified as top hits, with several additional proteins being enriched (Figure 4 D). Based on their prediction by Lopit as plasma membrane-associated (plasma membrane peripheral 2 for TGGT1_247000 and TGGT1_285500) or micronemal (Mic15 and TSP1) and their negative phenotypic score, we shortlisted the tetratricopeptide repeat-containing protein (TGGT1_247000), thrombospondin type 1 domain-containing protein (TSP1, TGGT1_277910), a protein with three TSP1 repeats (Mic15, TGGT1_247195) and the uncharacterized protein with four transmembrane domains (TGGT1_285500) for further characterization. Mic15 and TSP1 are predicted to contain a single transmembrane domain, TSP1 lacks an N-terminal signal peptide. The protein TGGT1_247000 contains a single N-terminal transmembrane domain with the rest of the protein predicted to be non-cytosolic, 285000 contains four predicted transmembrane domains. The localization analysis of the CRMP complex in tachyzoites suggest that only a small fraction is present at the plasma

membrane for a short time window during invasion. Thus, experimental conditions to identify potential interaction partners of the CRMPs during invasion were modified. We generated the transgenic parasites RH-CRMPA-TurboID and RH-CRMPB-TurboID and designed an invasion specific biotinylation assay. For this, each strain was incubated on an HFF monolayer either in invasion buffer with biotin or in ENDO buffer blocking invasion with biotin for 30 minutes. Subsequently mass spectrometry of biotinylated proteins was performed. No protein was significantly enriched in the invasion sample compared to the control, for both CRMPs analyzed individually and combined (Supplementary Table 1). We did not detect Mic7, Mic8, CLAMP, ND6, ND9, TgNdP1, TgNdP2, Ferlin 2, TGGT1_277840, TGGT1_253570, all of which have been previously linked to invasion, in any dataset for CRMPs-TurboID or pulldown.

TSP1 and MIC15 are required for host cell invasion

We performed a colocalization analysis of MIC15, TSP1, TGGT1_285500 and TGGT1_247000 by tagging them C-terminally with HALO in the RH-CRMPB-HA^{flox} parasite line (Figure 5 A, C, Figure S5 B, C). Additionally, we looked at HALO localization and relative signal strength three days post rapamycin induced excision of CRMPB. Mic15 and TSP1 perfectly colocalized with CRMPB (Figure 5B) and in the absence of CRMPB were partially retained within the ER (Figure 5 A, C, Figure S5 A). In contrast, TGGT1_285500 localized to very few vesicles, which did not change in the absence of CRMPB (Figure S5 B). TGGT1_247000 localized mainly to the ER, with a prominent enrichment at a basal sub-compartment of the ER. Deletion of CRMPB did not result in changes of localization (Figure S5 C). We also performed time laps recordings with all HALO tagged proteins and tagged CRMPB with HALO for reference (Supplementary video 1). Together these data

Figure 4: Identification of interaction partners. **A** Experimental outline of mass spectrometry-based protein identification. Parasites with C-terminally tagged 3HA were used for direct pulldown experiments (depicted on the left, results shown in **B**). In a separate experiment, C-terminally tagged CRMPB with TurboID was used in an attempt to identify more transient interaction partners of CRMPB post secretion. Here pulldown with streptavidin coupled beads was performed after incubation of extracellular tachyzoites with biotin for 30 min. **B** Interactors of CRMPA 3HA and CRMPB 3HA. The normalized relative log 2 protein intensity of CRMPB 3HA - WT are blotted on the x-axis, while blotted on the y-axis for CRMPA-3HA - WT. Baits are highlighted in blue. Correspondence between numbers in the plot and protein names is listed in **D**. **C** Transient interactors of CRMPB. Proteins enriched in the CRMPB TurboID sample are present in the upper right area of the graph. Proteins indicated with name are those further characterized. **D** List of proteins highlighted in **B** and **C**.

Figure 5: Localization of potential interaction partners. **A** Localization of Mic15 and CRMPB (RH-CRMPB-HA^{flox} Mic15-Halo) labeled with α -HA Atto594 and Halo stained with Janelia fluor 646. On the right localization 72h post induction with rapamycin is shown. **B** Box blot of colocalization analysis of proteins with CRMPB 3HA, CRMPB data point is colocalization of CRMPA-3HA with CRMPB-sYFP2. Intracellular parasites were stained with α -HA Alexa 488. Mic15, TSP1, TGGT1_285500 and TGGT1_247000 were labeled pre-fixation with Janelia 646, CRMPA was labeled

with α -GFP atto594, Mic2, Mic8 and Amal were labeled with respective antibodies and stained with Star635p. Single vacuoles of deconvolved stacks were analyzed in all planes with the color2 plugin and the Pearson's R value of 10 vacuoles is shown (box shows 25%-75%, median indicated by line, average by little white box, outliers black dots, error bars show standard deviation). **C** Localization of TSP1 and CRMPB (RH-CRMPB-HA^{flox} TSP1-Halo) as in **A**. **D** Localization of Mic15 and Gap45 in RH Mic15-3HA^{flox} induced + and - 96 h rapamycin. **E** Localization of TSP1 and GAP45 in RH-TSP1-3HA^{flox} induced + and - 96 h rapamycin. **F** Plaque assay of floxed Mic15 and TSP1. Scale bar of the entire plaque assay is 5 mm, scale bar of the higher resolution area is 500 μ m. **G** Invasion assay of RH-Mic15-3HA^{flox} and RH-TSP1-3HA^{flox} 72h post induction. All scale bars except **G** are 5 μ m. Statistical analysis: two-tailed student's t-test. *** p<0,001.

suggest that TGGT1_247000 and TGGT1_285500 are not part of the CRMP complex.

Next, we generated conditional mutants for these candidates, using the DiCre-system, as described above. In case of Mic15, the LoxP site was inserted into an intron, resulting in the excision of the last TSP1 domain and the transmembrane domain (Figure S2 C). While deletion of TGG1_285000 resulted in the formation of slightly smaller plaques when compared to non-induced parasites, the number of plaques was the same. Further analysis indicated a delay in replication and defects in intracellular development (Figure S5 D). Thus, a functional link to CRMPs seems unlikely. Similarly, deletion of tetratricopeptide protein (TGG1_247000), although critical for parasite propagation, demonstrated a strong defect in intracellular development (Figure S5 E).

In contrast, both TSP1 and Mic15 did form very small plaques, demonstrating that they are critical for parasite growth. Further analysis confirmed their critical role in host cell invasion (Figure 5 D - G). This is in good agreement of them being part of the same complex, which depends on the presence of all subunits in order to facilitate efficient host cell invasion.

Structure prediction of the invasion complex

No crystal structure of any homolog of the CRMPs has been experimentally solved so far. In the recently published AlphaFold Protein structure database^{44,45}, only the structure of *P. falciparum* CRMP2, DICDI (*Dictyostelium discoideum*) Q86I19 and two additional DICDI homologues are partially predicted. While the extracellular part of the proteins is structurally different, the 9 transmembrane domains and some directly associated amino acids resolve quite similar in the predicted structure. We used colabfold, a google colab notebook using the AlphaFold algorithm⁴⁶ to predict the core region of CRMPB and CRMPA, as well as the multimer structure with both TMs⁴⁷. Both transmembrane

domains were predicted to form a heterodimer with each other with a very low inter-complex PAE (predicted alignment error) and a high confidence measure (pLDDT). Prediction of the directly adjacent C-terminal region resulted in a heterodimer with mixed confidence values (Figure 6 A).

Prediction of upstream region did not result in reliable structures of CRMPA and CRMPB as a complex (low score and low pLDDT score for CRMPB) (Figure S7 A). We modeled the entire proteins in fragments with Robetta⁴⁸ and assembled CRMPA. Many individual domains can be identified that do not predict via sequence analysis. One example is a single six-bladed-beta-propeller that was found in both CRMPA and CRMPB as well as in PfCRMP2 (Figure S7 A).

Prediction of the complete complex at once is not possible due to the huge size of the proteins. Both TMs folded very similar to PfCRMP2 as well as each other and assembled as an almost symmetric heterodimer. Control runs of homodimers of CRMPA with CRMPA or CRMPB with CRMPB did not result in an interaction site (Figure S7 B). We also modelled the TMs of PbCRMP1-4. Only PbCRMP1 with PbCRMP3 and PbCRMP2 with PbCRMP4 resulted in a heterodimer structure. PbCRMP1/4 and PbCRMP2/3 did not form dimers when modelled (Figure S7 C). Structural mapping of PbCRMP 1 to 4 onto the TgCRMPA/B heterodimer mapped PbCRMP1 and 2 onto TgCRMPA and PbCRMP3 and 4 onto TgCRMPB as the phylogeny suggests (data not shown).

Mic15 is a typical type 1 single pass transmembrane domain protein with a relatively short cytosolic tail (179 aa) as most classical micronemal proteins. It contains a long internal repeat and three extracellular thrombospondin type 1 domains. TSP1 contains a single TSP1 repeat and a potential transmembrane domain near the C-terminus. Mic15 is distantly related to TRP1 of *Plasmodium berghei*, which has been linked to oocyst egress and salivary gland invasion of sporozoites⁴⁹. The structure of PfTRP1 was predicted by the AlphaFold database. We also predicted the c-terminal end of Mic15 and TSP1 using colabfold. Both TRP1 and Mic15 structures

Figure 6: Structural modeling of CRMPs. Multimer structures are predicted with Colabfold for TgCRMPA (blue) and TgCRMPB (orange). **A** TgCRMPs form a heterodimer with their TMs (1), shown in side and front view, PAE score defines how sure the algorithm is to predicts the relative special distance of any amino acid, the pLDDT score defines how precise local structures are predicted. Sequence alignment blots (coverage) indicate the number of sequences that could be aligned to each position. (2) prediction of the C-terminal part (last amino acids missing). The PAE and pLDDT scores vary over both fragments. **B** Rhoptry secretion sequential input model. Invasion components linked to rhoptry secretion that are at or in direct proximity to the plasma membrane are highlighted. Based on the hierarchy of evolutionary conservation we expect that the signal from the external CRMPA/B complex to the internal ND6/ND9 complex is central to rhoptry secretion, leading to downstream interaction with the apical vesicle. External species-specific receptors like Mic8 and Mic7 as well as accessory components of CRMPA/B are expected to modulate the CRMPs. As no direct interaction has been shown yet, potential mediators are unknown. In contrast to ND6/ND9 and other internal components, all secreted components are only transiently present at the apical part of the plasma membrane.

contain a small domain directly upstream of the TM domain that is not defined, partially also conserved in TSP1. Using the structural search of the Dali server⁵⁰ we found high homology with the domain 1

of aerolysin, involved in binding of N-linked oligosaccharides of GPI-anchored proteins^{51,52}.

Discussion

During invasion of the host cell, apicomplexan parasites sequentially secrete their unique secretory organelles. First the micronemes are released and subsequently, upon contact with the host cell, rhoptry secretion is triggered, leading to the formation of a tight interaction, the moving junction, through which the parasite invades the host cell. While some of the signaling mechanisms to microneme secretion have been described⁶, the mechanisms and signaling cascades leading to rhoptry discharge are still enigmatic. While some components, such as MIC8⁸, CLAMP¹⁰ or Nd-proteins¹⁸ have been identified, their mode of action is unknown. Interestingly, MIC8 is only conserved in coccidia, while CLAMP homologs exist in all apicomplexans and Nd-proteins appear to be conserved throughout the superphylum of alveolates. These findings suggest that the basic mechanism involved in rhoptry secretion is conserved, but adaptation to different lifestyles and host species required the evolution of specific invasion factors to ensure correct rhoptry discharge at the right time during invasion.

Here we describe an invasion complex essential for rhoptry discharge that consists of both, a highly conserved core, CRMPA and CRMPB, and the more unique accessory components Mic15 and TSP1. In our experiments, functional disruption of either CRMPs, MIC15 or TSP1 led to a significant block in invasion, but not gliding motility. Similar to previously described rhoptry secretion factors^{8,10,18}, deletion of the CRMPA or CRMPB resulted in a block of rhoptry secretion.

While the CRMPs are present in all apicomplexa, the copy numbers are not identical and general sequence conservation is low. It is striking that in *Chromera velia* and even *Paramecium tetraurelia* a highly expanded repertoire of CRMPs can be found in the genome. It is not clear if this difference in numbers of CRMPs compared to apicomplexa results from copy number expansion in free living species or gene loss in early apicomplexa. Many apicomplexa, including coccidia have only two members of CRMPs. In *Plasmodium* species as well as *Babesia* and *Theileria* we could identify four members of CRMPs. In *P. berghei* and *P.*

falciparum, these are found on different chromosomes, in the genome of *Babesia* and *Theileria* species, gene duplication is still visible within the genome. It seems likely that the more diverse protein family plays a role in various biological processes related to environmental sensing in *C. velia* and *P. tetraurelia* and only one or two members specialized into their essential role for rhoptry secretion within apicomplexa. In *Cryptosporidium*, two CRMPs are also present in the genome, however these are far more variable compared to other apicomplexa.

Our data demonstrate that the CRMPs complex is secreted to the plasma membrane during invasion. Low expression and a low secretion rate did not permit to localize its exact position at the plasma membrane during invasion, but an accumulation at the posterior pole in moving tachyzoites and post-invasion could be detected. While one would expect that a surface receptor required for rhoptry secretion is localized at the apical tip where the sensing of a suitable host cell would result in productive invasion, we only detect a relatively weak accumulation of the CRMP complex at the tip. As no signaling pathway has been identified and no drug found that triggers rhoptry secretion, one could speculate that rhoptry secretion might be triggered by direct protein-protein interactions. In contrast to classical micronemal proteins, the C-terminal tails of CRMPB and CRMPA are much longer (565 and 478 amino acids respectively) and could act as platforms for protein-protein interactions leading to rhoptry secretion. However, our proteomic approaches did not reveal any clear candidates. The high abundance of differentially phosphorylated amino acids at the C-terminus (24 for CRMPB, 10 for CRMPA)⁵³ suggests that protein-protein interaction could already be primed in extracellular tachyzoites. In the RH5, RIPR and CyRPA complex, two micronemal and a rhoptry protein (Rh5) form a complex prior to rhoptry secretion^{22,54}. Similar of what we expect for the CRMPs complex within Apicomplexa, only parts of this complex are conserved in all Plasmodium species²⁴.

Another open question is how CRMP integrates with the function of other essential rhoptry-secretion factors, such as Mic8 and CLAMP. While CLAMP could trigger the last step in tight junction

formation preceding Ron2-AMA1 interaction, Mic8 seems to be part of the same step in rhoptry secretion as the CRMP complex, without any evidence for direct interaction between both. Similarly, all attempts to identify downstream signaling molecules that interact with the complex were unsuccessful. If these interactions are as transient as the speed of the entire invasion process suggests, it will be very challenging to identify them via standard proteomic approaches and therefore transient interactions with CLAMP or Nd proteins cannot be ruled out. Potentially a mutant that allows interaction but blocks invasion could be helpful, if signaling persists in this case.

In addition to the core complex consisting of CRMPA and CRMPB, we identified MIC15 and TSP1 as additional components that are both critical for invasion. The closest ortholog of Mic15 in *Plasmodium* species is the thrombospondin related protein (PBANKA_0707900), which seems to have a similar function in invasion⁴⁹. They share a TSP1 domain with TgTSP1 but sequence conservation is very low.

Since deletions of CRMPA, CRMPB and MIC15 results in the same phenotype, expression and localization of CRMPA depends on CRMPB and both pulldown and TurboID experiments demonstrated a direct interaction, we used structure predictions to obtain additional information regarding this complex. Colabfold suggests that the most likely interaction site of both CRMPs are the TMs, as is also suggested by the interdependence expression experiment in Figure 3. We also observed a specific interaction partner of CRMPA clade and CRMPB clade heterodimer formation in *Plasmodium berghei*, whereas this was not the case for *Babesia bovis*, where most of the TMs of both members of the CRMPA clade are identical and thus interchangeable. We fitted all the structures of the CRMP TMs individually on our TgCRMPs multimer structure. All structure predictions from the CRMPA clade reliably mapped to TgCRMPA, and all B clade to CRMPB. We also modeled the co-structure of the C-terminus and the two proteins seem to interact, although the IDDT score was low. It seems feasible that structure prediction could be used to generate an intracellular *Toxoplasma* interaction network as has already been done for yeast⁵⁵. A subset analysis around known invasion factors might already help to identify putative

interaction surfaces that can then be tested experimentally. Multimer structure prediction can also be used to find putative interaction partners on the host cell surface, currently complicated by the fact that protein modifications like glycation is not predicted and might be essential for interaction (the same is true for phosphorylation of the C-terminus). Final confirmation of the CRMPs complex, and especially how Mic15 and TSP1 interact might be achieved by a cryo-EM structure in the future.

The most interesting questions are however still open. What do the extracellular parts of the CRMPs bind to? With what does the C-terminus interact and most importantly, how is rhoptry secretion dependent and triggered by at least three seemingly independent factors?

How can we functionally integrate previously known components of rhoptry secretion with the CRMP complex? While at this point highly speculative, based on their conservation throughout evolution, from being conserved in alveolate (Nd-proteins and CRMPs), in apicomplexa (CLAMP) down to being specific for a family (MIC8), we favor a model, where the highly conserved Nd-proteins represent a central platform that is triggered by the action of the CRMP-core complex in all alveolates. From there on adaptations lead to the specialization for invasion, including CLAMP. Species specific unique proteins like Mic8 or RH5 and accessory CRMP components like MIC15 and TSP1 are required for species-specific activation. Consequently, we would favor a model for the sequential action of the rhoptry secretion factors, where (in the case of *T. gondii*) Mic8 acts upstream of CRMPs, which in turn activate the Nd complex (Figure 6 B).

In summary, we present here a new essential invasion complex, where the CRMPs are the core components that associate with accessory proteins. Removal of a single CRMP results in functional disruption of this invasion complex. The main internal interaction interface seems to be the 9 transmembrane domains of each CRMP, resulting in heterodimer formation which is required for the stability of the complex.

Methods

Sequence analysis and phylogenetic tree generation

Sequences were collected using BLAST of the CRMP1-4 of *P. berghei* (<https://blast.ncbi.nlm.nih.gov/>). Sequence alignment was generated with wasabi using the PRANK algorithm (<https://www.ebi.ac.uk/goldman-srv/webprank/>)⁵⁶. Sequence analysis and phylogenetic tree generation was performed with MegaX⁵⁷. Local alignment model (Figure S1 C) was generated with geneious prime 2021.1.1 and domain models were generated in Adobe illustrator with assistance of EMBL smart (<http://smart.embl-heidelberg.de/>)³¹ and transmembrane domain prediction (<http://www.cbs.dtu.dk/services/TMHMM/>)⁵⁸.

Cloning of DNA constructs

For C-terminal tagging, we generated a small library of tags which contain the same linker sequence GCTAAAATTGGAAGTGGAGGA (coding for AKIGSGGR), followed by a STOP codon and a LoxP site (TAAATAACTTCGTA TAGCATACATTATACGAAGTTAT). Several tags were cloned into pGem or Puc19 using standard methods to allow multiple tagging with a single primer pair. Tagging sequences for SYP2, TurboID and SNAP were a kind gift from Dorus Dadella (Addgene_22878), Alice Ting (Addgene_107169) and from Won Do Heo (Addgene_58370)^{59,43,60}. All primers used in this study were ordered from Thermo Fischer Scientific or Integrated DNA Technologies (IDT) (see supplementary Table 2).

Culture of T. gondii

Parasite culture (*T. gondii* tachyzoites RH-strain) was performed onto human foreskin fibroblasts (ATCC; SCRC-1041). Cells were cultured at 37 °C and 5% CO₂ with DMEM (Sigma, D6546), supplemented with 10% FBS (BioSell FBS.US.0500), 4 mM L-Glutamate (Sigma, G7513) and 20 µg/ml gentamicin (Sigma G1397).

Generation of transgenic parasites

Genetic modification of all parasites used in this study was performed using CRISP/CAS9 guided double strand break in the RH Ku80 DiCre strain³⁶. For the gRNA design EuPaGDT (<http://grna.ctegd.uga.edu/>) was used to select a gRNA sequence⁶¹. Cutting sites downstream of the STOP were preferentially used rather than recodonizing the coding region to protect the repair template from gRNA binding. Downstream cutting sites resulted in the removal of the sequence between the STOP to the cut-site after repair. The gRNA sequence was cloned into and transiently expressed from a Cas9-YFP vector⁶². All tags used in the lab are flanked by the same linker sequence upstream and by a STOP-LoxP sequence downstream. On both sides 50 bp of homology was added via PCR to generate the repair templates for homologous recombination using Q5 polymerase (NEB). For internal tagging, the reverse primer was tag specific omitting the STOP-LoxP. Upstream or Intron-LoxP sequences were integrated using a 100 bp single oligo, consisting of 33 bp homology on both sides of the LoxP sequence. For transfection, 10 – 12 µg per Cas9 plasmid, purified PCR product of 200 µl reaction volume and 2-5 µl of LoxP repair oligo (100 µM) were pooled and purified using a PCR purification Kit (Blirt). Freshly lysed parasites were transfected using the Amaxa 4D-Nucleofactor system (Lonza) P3 primary cells kit, programme FI-158. 24 to 48 h post transfection, parasites were mechanically released, filtered with a 3 µm filter and sorted for YFP expression into a 96 well plate (5 and 10 events; 48 wells each) (FACSARIA III, BD Biosciences). Individual plaques were screened by PCR and all modified loci were sequenced to confirm correct modification. For multiple modification, subsequent rounds of transfection, sorting and genotyping took place. For primer sequences and tags as well as parasites generated see supplementary methods.

Labeling and pulldown

For dataset 1: direct pulldown, parasites (RH-CRMPA-3HA, RH-CRMPB-3HA and RH-DiCre) were mechanically released from host cells, filtered and washed with PBS, then stored at -80 °C. For

dataset 2, biotin labelling: parasites (RH-CRMPB-TurboID or RH-DiCre) were mechanically released and filtered, then parasites were incubated at 37 °C for 30 min in 1 ml of DMEM media. Samples were centrifuged at 1500 g for 5 min and washed with PBS (0°C) three times, then stored at -80 °C. For dataset 3, parasites (RH-CRMPA-TurboID or RH-CRMPB-TurboID) parasites were resuspended in 2 ml of ENDO buffer with 150 µM of Biotin or in HBSS (with 25 mM HEPES, 1% FCS and 150 µM Biotin) on a 6 cm dish of HFF. After incubation the dish was scratched and centrifuged in 50 ml of ice-cold PBS at 1500 g for 10 min at 0 °C. Supernatant was discarded, and 2 identical washes performed followed by resuspension in 1 ml of PBS and a final spin after transfer into an 1,5 ml tube. Sample was stored at -80 °C.

For HA pulldown, parasite pellet was lysed for 30 min on ice in 1 ml of RIPA buffer (0,5% sodium deoxycholate, 150 mM NaCl, 1mM EDTA, 0,1% SDS, 50 mM Tris-HCL pH 8,0, 1% Triton-X-100 with 1:100 protease and phosphatase inhibitor cocktail, Sigma). Now 50 µL (0.25mg) of Pierce Anti-HA magnetic beads and 350 µL of 0.05% TBS-T were vortexed gently. Beads were placed into a magnetic stand to collect the beads against the side of the tube and supernatant was removed. Was with 1 ml of TBS-T for 1 min, remove supernatant. Now the lysed sample was added (spin and only use supernatant) and incubated for 30 min with gentle mixing. Beads were washed with 300 µl of TBS-T, then 300 µl of ultrapure water.

For TurboID pulldown, the pellet was lysed for 30 min on ice in 1 ml of RIPA buffer with 1% Triton-X-100 and 1:100 protease inhibitor. Per sample 50 µl (100 µl for the invasion specific second experiment) of beads (Dynabeads MyOne Streptavidin T1, Invitrogen) were washed 3 times with 1 ml of PBS using the magnet system. Beads were responded with cleared lysed parasites in RIPA buffer and incubated 1 h at RT while gently mixing. Beads were washed 5 times with 1 ml RIPA buffer without Triton. Beads were washed three times with 50mM NH₄HCO₃ and 10% of the volume was separated for western blot controls.

Mass spectrometry

Beads were incubated with 10 ng/µL trypsin in 1 M urea 50mM NH₄HCO₃ for 30 minutes, washed with 50mM NH₄HCO₃ and the supernatant digested overnight (ON) in presence of 1mM DTT. Digested peptides were alkylated and desalted prior to LC-MS analysis.

For LC-MS/MS purposes, desalted peptides were injected in an Ultimate 3000 RSLCnano system (Thermo), separated in a 15-cm analytical column (75 µm ID with ReproSil-Pur C18-AQ 2.4 µm from Dr. Maisch) with a 50-min gradient from 4 to 40% acetonitrile in 0.1% formic acid. The effluent from the HPLC was directly electrosprayed into a Qexactive HF (Thermo) operated in data dependent mode to automatically switch between full scan MS and MS/MS acquisition. Survey full scan MS spectra (from m/z 375–1600) were acquired with resolution R=60,000 at m/z 400 (AGC target of 3x10⁶). The 10 most intense peptide ions with charge states between 2 and 5 were sequentially isolated to a target value of 1x10⁵, and fragmented at 27% normalized collision energy. Typical mass spectrometric conditions were: spray voltage, 1.5 kV; no sheath and auxiliary gas flow; heated capillary temperature, 250 °C; ion selection threshold, 33.000 counts.

MaxQuant 1.6.14.0 was used to identify proteins and quantify by iBAQ with the following parameters:

Database UP000005641_Toxoplasma gondii_20201123.fasta ; MS tol, 10 ppm; MS/MS tol, 20 ppm Da; Peptide FDR, 0.1; Protein FDR, 0.01 min. peptide Length, 7; Variable modifications, Oxidation (M); Fixed modifications, Carbamidomethyl (C); Peptides for protein quantitation, razor and unique; Min. peptides, 1; Min. ratio count, 2. Identified proteins were considered as interaction partners of the bait if their MaxQuant iBAQ Z-score normalized values were as follows: $\log_2(\text{CRMPA-3HA}) - \log_2(\text{WT}) + \log_2(\text{CRMPB-3HA}) - \log_2(\text{WT}) > 1,8$. The TurboID dataset was blotted in a volcano blot. One CRMPB-TurboID dataset was excluded from the final analysis as PCA analysis indicated a high divergence from the other two replica. Additionally, proteins that have a

known function, are not expressed, have no localization data by lopit or a phenotypic score > -1 we excluded from further analysis. Proteins with nuclear, proteasomal, mitochondrial, rhoptry and apicoplast localization were also excluded. For display and analysis, the Perseus software⁶³ was used.

Microscopy assays

All microscopy was performed at a Leica DMI8 widefield microscope with a DF C9000 GTC camera or at the Abberior 3D STED microscope (STED microscopy). For live cell microscopy a heated chamber and 5% CO₂ was used.

Immunofluorescence labelling

Parasites stained with Halo or SNAP were labelled prior to fixation. SNAP was labelled with SNAP-Cell 647-SiR (from NEB Biolabs), Halo was labeled with Janelia Fluor 646 Halo ligand. Dyes were stained life for 30 min to 2 hours, washed 3 times with PBS and incubated for 15 min in media. Cells were fixed 15 min with 4% PFA in PBS, washed 3 times with PBS and blocked for 30 min with 2-3% BSA. If the sample was permeabilized, 0.2% Triton was added here. Primary antibody incubation was performed at RT for 1 h in blocking buffer, followed by 3 washes with PBS. The secondary antibody was incubated together with hoechst if desired in blocking buffer at RT for 1 hour, followed by 5 washes with PBS. Samples were imaged directly in PBS or mounted with Prolong Gold, Prolong Diamond or Prolong Glass (Thermofischer Scientific). Secondary antibodies used were Alexa488, Atto594 (Thermofischer Scientific), Abberior Star 580, Abberior 635P and Abberior StarRed (Abberior). Staining of Biotin labeled TurboID samples was performed with directly labeled Streptavidin (Alexa488 or Alexa594) (Thermofischer Scientific).

Plaque assay

Plaque assays were performed on confluent HFFs in 6 well-plates in triplicates. Of each parasite line, 1000 parasites were used per well, incubated for 7 days with 4 ml of media per well $\pm$ 50 nM rapamycin as previously described⁶⁴. A RH-DiCre

and RH-Mic8extrac.^{floxed} control were run in parallel. Overview images were acquired with a cell phone camera against an illuminated white back ground. Inserts were generated with the Leica DMI8 widefield microscope using a 10 x 10 field stitched image with 10% overlay.

Gliding assay

Parasites were induced with $\pm$ 50 nM rapamycin 3 days prior to gliding assay. Live-cell microscopy chambers (μ -Slide 8 well (Ibidi) were incubated with 50% FBS at RT for 1h. Freshly filtered parasites were incubated in HBSS with 1% FBS and 25 mM HEPES and directly transferred into the μ -slide 8 well and imaged in the preheated Leica DMI8 microscope with 1fps with a 20x air objective.

Invasion assay

Fixed invasion assay (red/green assay) was performed on HFFs cells on glass slides in 24 well plates. Parasites after $\pm$ rapamycin were mechanically released and filtered and adjusted to 3×10^6 parasites in 200 μ l of DMEM. Parasites were settled for 10 min at RT and incubated at 37 °C for 20 minutes, followed by fixation of 4% PFA for 15 min. Samples was washed 3x with PBS, blocked without Triton (3% BSA in PBS) for 1 h, stained with anti-Sag1 antibody⁶⁵, washed 3x with PBS and stained with anti-mouse Alexa-488. Sample was briefly fixed with PFA, washed 3x with PBS, blocked with Triton (3% BSA in PBS, 0.2% Triton-X-100) for 1 h, stained with anti-Gap45 antibody, washed 3x with PBS and stained with anti-rabbit Abberior-580.

For 24 h invasion/replication assay, 5×10^6 of freshly release and filtered tachyzoites were allowed to invade for 20 min. Treatment $\pm$ rapamycin was started 72 h before invasion. After 24 h, cells were fixed and stained as in the red/green assay.

Live cell imaging of invasion was performed in HBSS, 25 mM HEPES, pH 7.3, 1% FCS at the Leica DMI8. Parasites were mechanically released, filtered, resuspended in pre-warmed medium and

directly imaged for 20 min in a heated microscope chamber with 5% CO₂.

Evacuole assay

Freshly egressed and filtered tachyzoites 72 h after incubated with $\pm$ rapamycin were adjusted to 5×10^6 parasites in 250 μ l and incubated with 1 μ M of Cytochalasin D (CytoD) for 10 min at RT. Parasites were added to confluent HFF on glass slides in 24 well plates and incubated for 15 min at 37 °C. Cells were fixed and blocked in 3% BSA in PBS. Sample were stained with anti-ROP1 mAb T52A3 (Soldati-Favre lab) and anti-Sag1 (Abcam 138698).

STED microscopy

Super-resolution microscopy was performed at the Abberior 3D STED microscope, equipped with 3 color STED. Imaging settings were individually adjusted to the sample to maximize signal to noise and resolution and vary depending on signal strength. Generally, 60x60x250 nm sampling was performed for confocal mode, 30x30x250 nm for 2D STED and 60x60x60 nm for 3D STED.

Image analysis

Deconvolution of widefield images was performed with Huygens essentials (SVI). Microscopy images were analyzed using Fiji⁶⁶. Colocalization analysis was performed using the Coloc 2 plugin (https://github.com/fiji/Colocalisation_Analysis). The localized colocalization display was performed by multiplication of every single pixel between both channels. This helps to visualize if the colocalization is position independent or not, no quantification is performed.

For the relative signal analysis in Figure S5 A, the relative localization of Mic15 was quantified in the presence and absence of CRMPB. Apical relative signal is percentage of signal apical of the nucleus (area selected manually) of total signal within one tachyzoite. Circles represent individual tachyzoites, the center plane of a deconvolved stack was analyzed. In the post ER analysis, the nucleus was automatically thresholded using the Huang algorithm of the Hoechst staining (this generates a threshold surrounding the nucleus) and this threshold was applied onto the Halo signal, giving

rise to the ER signal. The remaining fluorescence intensity within the parasite is defined as the post ER signal. Each dot represents all planes of a vacuole of a deconvolved stack. In all analysis sample was normalized with a background control.

Structure prediction

Multimer structure prediction was performed using Colabfold^{46,45,67,47}

(<https://github.com/sokrypton/ColabFold>). We used the mmseqs2 notebooks, both v.12 and v.13 (v. 1.3 always runs 5 models resulting in a higher failure rate to finish the prediction before timeout). The transmembrane region of CRMPA, CRMPB and Mic15 were predicted alone and in combination, as a control also CRMPB with CRMPB and CRMPA with CRMPA as dimer prediction. Same was performed for the sequences upstream and downstream of the TMs. Additionally, single sequences were submitted to Robetta (using Rosettafold (<https://rosetta.bakerlab.org/>)⁴⁸). To define the TMs of the CRMPs, we used the pdb file of PfCRMP2 from the AlphaFold database and mapped the amino acids to the alignment to define the TMs for all other CRMPs. Only for CRMPB this had to be manually adjusted due to bad alignment quality upstream of the TMs. The joined structure of TgCRMPA modeled with RobettaFold, Modeling was performed of 9 fragments, with 60 amino acid overlap. Structures were assembled using Swiss-pdb Viewer.

Data analysis

Graphs are generated with MS Excel or Origin pro (OriginLab). All other image analysis was performed using Fiji. Figures were generated with Adobe Illustrator.

Acknowledgment

We would like to thank colleagues for sharing of primary antibodies and plasmids used in this study. We thank Sonja Härtle and Marina Kohn for access to the FACS facility and expert support. Hirdesh Kumar assisted us in evaluation of the modeled protein structures. M.S. was funded by the Volkswagen Foundation.

Author contributions

M.S. and M.M. conceived the study and designed the experiments. M.S. and K.S. generated the parasite lines used in this study. M.S. performed most experiments. M.S. and I.F. performed and analyzed the mass spectrometry experiments. M.S. and M.M. analyzed the data and wrote the manuscript.

Competing interests

The authors declare no competing interest.

Additional data

Supplementary figures S1-S7

Supplementary data file

Supplementary movie 1: Intracellular imaging

Endogenous tagging of CRMPA in the extracellular part (RH-Halo-CRMPA-3HA^{flox}) or C-terminal tagging with Halo (all other proteins) shows that

they are very dynamic during intracellular development.

Supplementary movie 2: Localisation of CRMPA and CRMPB during invasion

Invasion of Halo labelled (Janelia 646) CRMPA and CRMPB into HFF cells (RH-Halo-CRMPA-3HA^{flox} RH-CRMPB-Halo). Note how part of the protein is visible at the basal end of the parasite, both prior to invasion and after. In one event signal is observed at the tight junction.

Supplementary table 1: DNA oligonucleotides used in this study

Supplementary table 2: Pulldown and TurboID mass spectrometry data

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE⁶⁸ partner repository with the identifier **PXD031649**.

References

- 1 Carruthers, V. B. Armed and dangerous: *Toxoplasma gondii* uses an arsenal of secretory proteins to infect host cells. *Parasitol Int* **48**, 1-10, doi:10.1016/s1383-5769(98)00042-7 (1999).
- 2 Besteiro, S., Dubremetz, J. F. & Lebrun, M. The moving junction of apicomplexan parasites: a key structure for invasion. *Cell Microbiol* **13**, 797-805, doi:10.1111/j.1462-5822.2011.01597.x (2011).
- 3 Carruthers, V. & Boothroyd, J. C. Pulling together: an integrated model of *Toxoplasma* cell invasion. *Curr Opin Microbiol* **10**, 83-89, doi:10.1016/j.mib.2006.06.017 (2007).
- 4 Bargieri, D. Y. *et al.* Apical membrane antigen 1 mediates apicomplexan parasite attachment but is dispensable for host cell invasion. *Nat Commun* **4**, 2552, doi:10.1038/ncomms3552 (2013).
- 5 Lamarque, M. H. *et al.* Plasticity and redundancy among AMA-RON pairs ensure host cell entry of *Toxoplasma* parasites. *Nat Commun* **5**, 4098, doi:10.1038/ncomms5098 (2014).
- 6 Bisio, H. & Soldati-Favre, D. Signaling Cascades Governing Entry into and Exit from Host Cells by *Toxoplasma gondii*. *Annu Rev Microbiol* **73**, 579-599, doi:10.1146/annurev-micro-020518-120235 (2019).
- 7 Frenal, K., Dubremetz, J. F., Lebrun, M. & Soldati-Favre, D. Gliding motility powers invasion and egress in Apicomplexa. *Nat Rev Microbiol* **15**, 645-660, doi:10.1038/nrmicro.2017.86 (2017).
- 8 Kessler, H. *et al.* Microneme protein 8-a new essential invasion factor in *Toxoplasma gondii*. *J Cell Sci* **121**, 947-956, doi:10.1242/jcs.022350 (2008).

- 9 Broncel, M. *et al.* Profiling of myristoylation in *Toxoplasma gondii* reveals an N-myristoylated protein important for host cell penetration. *Elife* **9**, doi:10.7554/eLife.57861 (2020).
- 10 Sidik, S. M. *et al.* A Genome-wide CRISPR Screen in *Toxoplasma* Identifies Essential Apicomplexan Genes. *Cell* **166**, 1423-1435 e1412, doi:10.1016/j.cell.2016.08.019 (2016).
- 11 Satir, B. H. & Oberg, S. G. Paramecium fusion rosettes: possible function as Ca²⁺ gates. *Science* **199**, 536-538, doi:10.1126/science.341312 (1978).
- 12 Satir, B., Schooley, C. & Satir, P. Membrane fusion in a model system. Mucocyst secretion in *Tetrahymena*. *J Cell Biol* **56**, 153-176, doi:10.1083/jcb.56.1.153 (1973).
- 13 Porchet-Hennere, E. & Nicolas, G. Are rhoptries of *Coccidia* really extrusomes? *J Ultrastruct Res* **84**, 194-203, doi:10.1016/s0022-5320(83)90130-2 (1983).
- 14 Plattner, H. *et al.* Genetic dissection of the final exocytosis steps in *Paramecium tetraurelia* cells: cytochemical determination of Ca²⁺-ATPase activity over performed exocytosis sites. *J Cell Sci* **46**, 17-40 (1980).
- 15 Beisson, J., Cohen, J., Lefort-Tran, M., Pouphile, M. & Rossignol, M. Control of membrane fusion in exocytosis. Physiological studies on a *Paramecium* mutant blocked in the final step of the trichocyst extrusion process. *J Cell Biol* **85**, 213-227, doi:10.1083/jcb.85.2.213 (1980).
- 16 Froissard, M., Keller, A. M. & Cohen, J. ND9P, a novel protein with armadillo-like repeats involved in exocytosis: physiological studies using allelic mutants in *paramecium*. *Genetics* **157**, 611-620, doi:10.1093/genetics/157.2.611 (2001).
- 17 Gogendeau, D., Keller, A. M., Yanagi, A., Cohen, J. & Koll, F. Nd6p, a novel protein with RCC1-like domains involved in exocytosis in *Paramecium tetraurelia*. *Eukaryot Cell* **4**, 2129-2139, doi:10.1128/EC.4.12.2129-2139.2005 (2005).
- 18 Aquilini, E. *et al.* An Alveolata secretory machinery adapted to parasite host cell invasion. *Nat Microbiol* **6**, 425-434, doi:10.1038/s41564-020-00854-z (2021).
- 19 Mageswaran, S. K. *et al.* In situ ultrastructures of two evolutionarily distant apicomplexan rhoptry secretion systems. *Nat Commun* **12**, 4983, doi:10.1038/s41467-021-25309-9 (2021).
- 20 Martinez, M. *et al.* In situ structure and priming mechanism of the rhoptry secretion system in *Plasmodium* revealed by cryo-electron tomography. *bioRxiv*, 2022.2001.2011.475861, doi:10.1101/2022.01.11.475861 (2022).
- 21 Segev-Zarko, L.-a. *et al.* Ionophore-stimulation promotes re-organization of the invasion machinery of *Toxoplasma gondii*. *bioRxiv*, 2022.2001.2012.476068, doi:10.1101/2022.01.12.476068 (2022).
- 22 Volz, J. C. *et al.* Essential Role of the PfRh5/PfRipr/CyRPA Complex during *Plasmodium falciparum* Invasion of Erythrocytes. *Cell Host Microbe* **20**, 60-71, doi:10.1016/j.chom.2016.06.004 (2016).
- 23 Manzoni, G. *et al.* *Plasmodium* P36 determines host cell receptor usage during sporozoite invasion. *Elife* **6**, doi:10.7554/eLife.25903 (2017).
- 24 Knuepfer, E. *et al.* Divergent roles for the RH5 complex components, CyRPA and RIPR in human-infective malaria parasites. *PLoS Pathog* **15**, e1007809, doi:10.1371/journal.ppat.1007809 (2019).
- 25 Valkiūnas, G. *Avian malaria parasites and other haemosporidia*. (CRC Press, 2005).
- 26 Thompson, J. *et al.* *Plasmodium* cysteine repeat modular proteins 1-4: complex proteins with roles throughout the malaria parasite life cycle. *Cell Microbiol* **9**, 1466-1480,

- doi:10.1111/j.1462-5822.2006.00885.x (2007).
- 27 Douradinha, B. *et al.* Plasmodium Cysteine Repeat Modular Proteins 3 and 4 are essential for malaria parasite transmission from the mosquito to the host. *Malar J* **10**, 71, doi:10.1186/1475-2875-10-71 (2011).
- 28 Bushell, E. *et al.* Functional Profiling of a Plasmodium Genome Reveals an Abundance of Essential Genes. *Cell* **170**, 260-272 e268, doi:10.1016/j.cell.2017.06.030 (2017).
- 29 Zhang, M. *et al.* Uncovering the essential genes of the human malaria parasite Plasmodium falciparum by saturation mutagenesis. *Science* **360**, doi:10.1126/science.aap7847 (2018).
- 30 Barylyuk, K. *et al.* A Comprehensive Subcellular Atlas of the Toxoplasma Proteome via hyperLOPIT Provides Spatial Context for Protein Functions. *Cell Host Microbe* **28**, 752-766 e759, doi:10.1016/j.chom.2020.09.011 (2020).
- 31 Letunic, I. & Bork, P. 20 years of the SMART protein domain annotation resource. *Nucleic Acids Res* **46**, D493-D496, doi:10.1093/nar/gkx922 (2018).
- 32 Boisard, J. *et al.* Marine gregarine genomes illuminate current understanding of apicomplexan glideosome. *bioRxiv*, 2021.2008.2009.454911, doi:10.1101/2021.08.09.454911 (2021).
- 33 Ueti, M. W. *et al.* Transcriptome dataset of Babesia bovis life stages within vertebrate and invertebrate hosts. *Data Brief* **33**, 106533, doi:10.1016/j.dib.2020.106533 (2020).
- 34 Woo, Y. H. *et al.* Chromerid genomes reveal the evolutionary path from photosynthetic algae to obligate intracellular parasites. *Elife* **4**, e06974, doi:10.7554/eLife.06974 (2015).
- 35 Templeton, T. J. & Pain, A. Diversity of extracellular proteins during the transition from the 'proto-apicomplexan' alveolates to the apicomplexan obligate parasites. *Parasitology* **143**, 1-17, doi:10.1017/S0031182015001213 (2016).
- 36 Hunt, A. *et al.* Differential requirements for cyclase-associated protein (CAP) in actin-dependent processes of Toxoplasma gondii. *Elife* **8**, doi:10.7554/eLife.50598 (2019).
- 37 Kremer, K. *et al.* An overexpression screen of Toxoplasma gondii Rab-GTPases reveals distinct transport routes to the micronemes. *PLoS Pathog* **9**, e1003213, doi:10.1371/journal.ppat.1003213 (2013).
- 38 Andenmatten, N. *et al.* Conditional genome engineering in Toxoplasma gondii uncovers alternative invasion mechanisms. *Nat Methods* **10**, 125-127, doi:10.1038/nmeth.2301 (2013).
- 39 Jones, M. L. *et al.* A versatile strategy for rapid conditional genome engineering using loxP sites in a small synthetic intron in Plasmodium falciparum. *Sci Rep* **6**, 21800, doi:10.1038/srep21800 (2016).
- 40 Zhang, X. *et al.* Branch point identification and sequence requirements for intron splicing in Plasmodium falciparum. *Eukaryot Cell* **10**, 1422-1428, doi:10.1128/EC.05193-11 (2011).
- 41 Neuveglise, C., Marck, C. & Gaillardin, C. The intronome of budding yeasts. *C R Biol* **334**, 662-670, doi:10.1016/j.crv.2011.05.015 (2011).
- 42 Taggart, A. J., DeSimone, A. M., Shih, J. S., Filloux, M. E. & Fairbrother, W. G. Large-scale mapping of branchpoints in human pre-mRNA transcripts in vivo. *Nat Struct Mol Biol* **19**, 719-721, doi:10.1038/nsmb.2327 (2012).
- 43 Branon, T. C. *et al.* Efficient proximity labeling in living cells and organisms with TurboID. *Nat Biotechnol* **36**, 880-887, doi:10.1038/nbt.4201 (2018).
- 44 Varadi, M. *et al.* AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Res*, doi:10.1093/nar/gkab1061 (2021).

- 45 Jumper, J. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583-589, doi:10.1038/s41586-021-03819-2 (2021).
- 46 Mirdita, M. *et al.* ColabFold - Making protein folding accessible to all. *bioRxiv*, 2021.2008.2015.456425, doi:10.1101/2021.08.15.456425 (2021).
- 47 Evans, R. *et al.* Protein complex prediction with AlphaFold-Multimer. *bioRxiv*, 2021.2010.2004.463034, doi:10.1101/2021.10.04.463034 (2021).
- 48 Baek, M. *et al.* Accurate prediction of protein structures and interactions using a three-track neural network. *Science* **373**, 871-876, doi:10.1126/science.abj8754 (2021).
- 49 Klug, D. & Frischknecht, F. Motility precedes egress of malaria parasites from oocysts. *Elife* **6**, doi:10.7554/eLife.19157 (2017).
- 50 Holm, L. Using Dali for Protein Structure Comparison. *Methods Mol Biol* **2112**, 29-42, doi:10.1007/978-1-0716-0270-6_3 (2020).
- 51 Degiacomi, M. T. *et al.* Molecular assembly of the aerolysin pore reveals a swirling membrane-insertion mechanism. *Nat Chem Biol* **9**, 623-629, doi:10.1038/nchembio.1312 (2013).
- 52 Hong, Y. *et al.* Requirement of N-glycan on GPI-anchored proteins for efficient binding of aerolysin but not Clostridium septicum alpha-toxin. *EMBO J* **21**, 5047-5056, doi:10.1093/emboj/cdf508 (2002).
- 53 Treeck, M., Sanders, J. L., Elias, J. E. & Boothroyd, J. C. The phosphoproteomes of Plasmodium falciparum and Toxoplasma gondii reveal unusual adaptations within and beyond the parasites' boundaries. *Cell Host Microbe* **10**, 410-419, doi:10.1016/j.chom.2011.09.004 (2011).
- 54 Baum, J. *et al.* Reticulocyte-binding protein homologue 5 - an essential adhesin involved in invasion of human erythrocytes by Plasmodium falciparum. *Int J Parasitol* **39**, 371-380, doi:10.1016/j.ijpara.2008.10.006 (2009).
- 55 Humphreys, I. R. *et al.* Computed structures of core eukaryotic protein complexes. *Science* **374**, eabm4805, doi:10.1126/science.abm4805 (2021).
- 56 Loytynoja, A. & Goldman, N. webPRANK: a phylogeny-aware multiple sequence aligner with interactive alignment browser. *BMC Bioinformatics* **11**, 579, doi:10.1186/1471-2105-11-579 (2010).
- 57 Kumar, S., Stecher, G., Li, M., Knyaz, C. & Tamura, K. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Mol Biol Evol* **35**, 1547-1549, doi:10.1093/molbev/msy096 (2018).
- 58 Sonnhammer, E. L., von Heijne, G. & Krogh, A. A hidden Markov model for predicting transmembrane helices in protein sequences. *Proc Int Conf Intell Syst Mol Biol* **6**, 175-182 (1998).
- 59 Kremers, G. J., Goedhart, J., van Munster, E. B. & Gadella, T. W., Jr. Cyan and yellow super fluorescent proteins with improved brightness, protein folding, and FRET Forster radius. *Biochemistry* **45**, 6570-6580, doi:10.1021/bi0516273 (2006).
- 60 Lee, S. *et al.* Reversible protein inactivation by optogenetic trapping in cells. *Nat Methods* **11**, 633-636, doi:10.1038/nmeth.2940 (2014).
- 61 Peng, D. & Tarleton, R. EuPaGDT: a web tool tailored to design CRISPR guide RNAs for eukaryotic pathogens. *Microb Genom* **1**, e000033, doi:10.1099/mgen.0.000033 (2015).
- 62 Curt-Varesano, A., Braun, L., Ranquet, C., Hakimi, M. A. & Bougdour, A. The aspartyl protease TgASP5 mediates the export of the Toxoplasma GRA16 and GRA24 effectors into host cells. *Cell Microbiol* **18**, 151-167, doi:10.1111/cmi.12498 (2016).
- 63 Tyanova, S. *et al.* The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nat Methods* **13**,

- 731-740, doi:10.1038/nmeth.3901 (2016).
- 64 Egarter, S. *et al.* The toxoplasma Acto-MyoA motor complex is important but not essential for gliding motility and host cell invasion. *PLoS One* **9**, e91819, doi:10.1371/journal.pone.0091819 (2014).
- 65 Burg, J. L., Perelman, D., Kasper, L. H., Ware, P. L. & Boothroyd, J. C. Molecular analysis of the gene encoding the major surface antigen of *Toxoplasma gondii*. *J Immunol* **141**, 3584-3591 (1988).
- 66 Schindelin, J. *et al.* Fiji: an open-source platform for biological-image analysis. *Nat Methods* **9**, 676-682, doi:10.1038/nmeth.2019 (2012).
- 67 Mirdita, M., Steinegger, M. & Soding, J. MMseqs2 desktop and local web server app for fast, interactive sequence searches. *Bioinformatics* **35**, 2856-2858, doi:10.1093/bioinformatics/bty1057 (2019).
- 68 Perez-Riverol, Y. *et al.* The PRIDE database resources in 2022: a hub for mass spectrometry-based proteomics evidences. *Nucleic Acids Res* **50**, D543-D552, doi:10.1093/nar/gkab1038 (2022).