Conservation of the *Toxoplasma* conoid proteome in *Plasmodium* reveals a cryptic conoid feature that differentiates between blood- and vector-stage zoites

Ludek Koreny[1], Mohammad Zeeshan[2], Konstantin Barylyuk[1], Declan Brady[2], Huiling Ke[1], Rita Tewari[2], Ross F. Waller[1]

[1] Department of Biochemistry, University of Cambridge, Cambridge, CB2 1QW, UK
[2] School of Life Sciences, Queens Medical Centre, University of Nottingham, Nottingham, NG7 2UH, UK

Abstract

The apical complex is the instrument of invasion used by apicomplexan parasites, and the conoid is a conspicuous feature of this apparatus found throughout this phylum. The conoid, however, is believed to be heavily reduced or missing from *Plasmodium* species and other members of the Class Aconoidasida. Relatively few conoid proteins have previously been identified, making it difficult to address how conserved this feature is throughout the phylum, and whether it is genuinely missing from some major groups. Moreover, parasites such as *Plasmodium* species cycle through multiple invasive forms and there is the possibility of differential presence of the conoid between these stages. We have applied spatial proteomics and high-resolution microscopy to develop a more complete molecular inventory and understanding of the organisation of conoid-associated proteins in the model apicomplexan *Toxoplasma gondii*. These data revealed molecular conservation of all conoid substructures throughout Apicomplexa, including *Plasmodium* and even in allied Myzozoa, such as *Chromera*. We reporter-tagged and observed the expression and location of several conoid proteins in the malaria model *P. berghei* and revealed equivalent structures in all zoite forms, as well as evidence of molecular differentiation between blood-stage merozoites and the ookinetes and sporozoites of the mosquito vector. Collectively we show that the conoid is a conserved apicomplexan element at the heart of the invasion mechanisms of these highly successful and often devastating parasites.

Introduction

It is difficult to imagine a more insidious intrusion upon an organism's integrity than the penetration and occupation of intracellular spaces by another foreign organism. Apicomplexan parasites are masters of this transgression through actively seeking, binding to, and invading the cellular milieu of suitable animal hosts. From here, they manipulate and exploit these cells to promote their growth and onward transmission to other cells and other hosts. The impacts of these infection cycles include major human and animal diseases, such as malaria, toxoplasmosis and cryptosporidiosis in humans, and a spectrum of other diseases in livestock and wild animals (Montoya and Liesenfeld 2004; Striepen 2013; World Health Organization 2018; Rashid et al. 2019). The course of both human history and evolution has been shaped by these ubiquitous marauding parasites.

Key to the successful parasitic strategies of apicomplexans is the apical complex—a specialisation of the cell apical cortex that is the coordination centre of the interaction and penetration of host cells (Dos Santos Pacheco et al. 2020). Most of the apicomplexan cell is encased in a pellicle structure of flattened membrane vesicles beneath the plasma membrane, as are all members of supergroup Alveolata including dinoflagellates and ciliates (Gould et al. 2008). These 'alveoli' sacs are supported

by robust proteinaceous networks, and collectively this inner membrane complex (or IMC, as it is called in apicomplexans) provides shape and protection to the cell, as well as other functions such as gliding motility in apicomplexans by IMC-anchored motors (Frénal, Dubremetz, et al. 2017). The IMC, however, is a general obstruction to other cellular functions that occur at the plasma membrane, including exocytosis and endocytosis (Harding and Meissner 2014). Thus, the apical complex has evolved alongside the IMC to provide a location for these functions. When apicomplexans attack their host's cells the apical complex is the site of exocytosis of both initial host- recognition and -binding molecules, and then subsequent secretion of molecules is injected into the host to create a platform in its plasma membrane for parasite penetration (Kats et al. 2008; Lebrun et al. 2014). In infections such as those that humans suffer from, upon host cell invasion further exocytosed molecules create a modified environment in the host cell that facilitate the parasites protection from the immune system, growth, and reproduction. In many gregarine apicomplexans, on the other hand, only partial penetration of the host occurs and the parasite endocytoses host cytosol via their embedded apical complex (Schrével et al. 2016). Near relatives of Apicomplexa, such as colpodellids and some dinoflagellates, similarly feed on prey and host cells through their apical complex—such is the antiquity of this cell feature (Mylnikov 2009; Okamoto and Keeling 2014). The apical complex is thus a coordination of the cell cytoskeleton that defines an available disc of plasma membrane that is unobscured by the IMC, and vesicular trafficking machinery for delivery to and exchange with the extracellular environment. A protuberance of the cell at the apical complex also provides mechanical properties to this important site.

Functional studies of the apical complex have occurred in select experimentally amenable taxa, mostly *Toxoplasma* and *Plasmodium*, but a mechanistic understanding of this cell feature or its conservation is still in its infancy. Rather, the apical complex is most broadly understood from ultrastructural studies, and apical rings are the basis of this structure. An apical polar ring (APR) acts as a microtubule organising centre (MTOC) for the subpellicular microtubules, and a second APR coordinates the apical margin of the IMC (Russell and Burns 1984; Dubremetz and Ferguson 2009; Morrissette and Gubbels 2014). Within this opening created by the APRs are further rings, a notable one being the 'conoid' that is conspicuous throughout much of Apicomplexa (Scholtyseck et al. 1970). The conoid is an open conical structure of variable length and cone pitch. It interacts intimately with secretory organelles including micronemes, rhoptries and other vesicles that penetrate and occupy its lumen (Paredes-Santos et al. 2012; Suarez et al. 2019). An open conoid (or 'pseudoconoid') in the related *Perkinsus* even has microneme-like vesicles physically tethered to it, and in coccidia a pair of intraconoidal microtubules is also lined by a chain of microvesicles (Perkins 1976; Paredes-Santos et al. 2012). In gregarines endocytosis occurs through the conoid aperture (Schrével et al. 2016). Thus, while the APRs appears to play chiefly structural organising roles, the conoid is closely associated with the events and routes of vesicular trafficking, delivery, and in some cases uptake.

In most examples the conoid has a fibrous presentation by electron microscopy, a trait that is chiefly attributed to the presence of tubulin (Hu et al. 2002; Morrissette and Gubbels 2014). In the *Toxoplasma* conoid, this tubulin forms unusual open fibres with a comma-shaped profile (Hu et al. 2002). The ancestral state of conoid tubulin, however, is likely canonical microtubules as seen in gregarines, *Chromera*, and other apicomplexan relatives (Mylnikov 2009; Obornik et al. 2011; Schrével et al. 2016). It is unclear if the modified tubulin fibres of the *Toxoplasma* conoid arose specifically within coccidia, or are more widespread in apicomplexans due to the limits of resolution or preservation of this dense structural feature. Nevertheless, this tubulin component demonstrates a degree of plasticity of the conoid structure. Electron microscopy also reveals further distinct rings at the apical end of the conoid, the preconoidal rings (also called apical rings), where the conoid is in closest contact and perhaps interaction with the cell plasma membrane (Nichols and Chiappino 1987; Dubremetz and Ferguson 2009). Electron microscopy does not reveal any direct attachment

fibres or structures from the conoid to the IMC. However, in *Toxoplasma* it is known that at least one protein links the conoid directly to the APR(s) (Katris et al. 2014) and thus that there is molecular architecture beyond that observed by ultrastructure.

A conspicuous structural deviation to the apical complex in Apicomplexa is the predicted loss of the conoid in some groups, a state enshrined within the class name Aconoidasida. This class contains two important groups: Haemosporida, such as *Plasmodium* spp., and Piroplasmida. Aconoidasida are considered to have either completely lost the conoid (e.g. *Babesia*, *Theileria*), or at least lost it from multiple zoite stages, e.g. *Plasmodium* spp. stages other than the ookinete. However, even while conoids have been attributed to micrographs of ookinete stages in some *Plasmodium* spp., in other studies these are alternatively labelled as apical polar rings (Aikawa 1967; Scholtyseck et al. 1970; Hanssen et al. 2013), and the prevailing understanding of many is that a conoid was lost outright.

The uncertainty over whether the conoid is present in Aconoidasida is a product of two problems. One is that we have little insight into the function of the conoid, so the consequences of its 'loss' are difficult to predict. The other is that we still know relatively little of the molecular composition of the conoid that would allow the objective testing for the presence of a homologous structure (Dos Santos Pacheco et al. 2020). The conspicuous ultrastructure of conoids such as those of coccidia draw attention to tubulin being a major element, however it is known that there are other conoid proteins responsible for its assembly, stability, and function during invasion (Katris et al. 2014; Graindorge et al. 2016; Leung et al. 2017; Long, Anthony, et al. 2017; Long, Brown, et al. 2017; Nagayasu et al. 2017; Back et al. 2020; Leung et al. 2020; Tosetti et al. 2020). To test if a homologous conoid cell feature is present in Aconoidasida, but cryptic by traditional microscopy techniques, fuller knowledge of the molecules that characterise this feature in a 'classic' conoid model are needed. In our study we have sought such knowledge for the *Toxoplasma gondii* conoid using multiple proteomic approaches. We then asked if these conoid-specific proteins are present in similar locations within Aconoidasida, using the model *Plasmodium berghei* to investigate each of its zoite forms, ookinetes, sporozoites and merozoites. In doing so we address the question of what common machinery underpins the mechanisms of invasion and host exploitation that are central to these parasites' lifestyles and impact.

Results

Spatial proteomic methods identify new candidate conoid proteins.

To expand our knowledge of the proteins that contribute to conoid structure and function we applied multiple spatial proteomics discovery approaches. The primary method used was hyperplexed Localisation of Organelle Proteins by Isotope Tagging (hyperLOPIT) that we have recently applied to extracellular *T. gondii* tachyzoites (Christoforou et al. 2016; Barylyuk et al. 2020). This approach entailed generating hyperLOPIT datasets from three independent samples. In each sample, mechanically disrupted cells were dispersed on density gradients, and the distinct abundance distribution profiles formed by different subcellular structures and compartments were used to identify proteins that belong to common cellular niches. Combining the data from the three experiments provided enhanced discriminating power of protein location assignments, and from the 3832 proteins that were measured in each experiment we reported 63 proteins assigned to one of the two apical protein clusters, *apical 1* and *apical 2*, above a 99% probability threshold, and another 13 assigned to these clusters but below this high-confidence cut-off. These two clusters were verified as comprising proteins specific to the structures associated with the conoid, apical polar ring, and apical cap of the IMC (Barylyuk et al. 2020). In addition to the 3832 proteins used in this high-resolution spatial proteome, a further 1013 proteins were quantified in either only two or one of the hyperLOPIT datasets. While assignment of these proteins had less data support, a further 16

proteins were assigned to the *apical* clusters from analysis of either the pairs of LOPIT experiments, or the single experiments. From these analyses, 92 proteins in total were assigned as putative apical proteins across these hyperLOPIT samples (Table S1).

Of the 92 putative apical protein assignments by hyperLOPIT, 13 had been validated as being located to the very apex of the cell during our hyperLOPIT study (Barylyuk et al. 2020), 23 by us or others previously, and 21 proteins were known to be specific to the apical cap or other IMC elements (Table S1 and refs therein). This left a further 35 new protein candidates for which there was no independent validation of their apical location. To bolster support for conoid-specific location we applied a second spatial proteomic strategy, proximity dependent biotinylating and pulldown, or BioID. We made three apical BioID 'baits' by endogenous gene 3' fusions with coding sequence for the promiscuous biotin-ligase BirA*. Two baits were known conoid markers: SAS6-like (SAS6L), a protein previously attributed to the preconoidal ring in *T. gondii* (de Leon et al. 2013) but by super-resolution imaging we observe across the entire conoid length (see below); and RNG2 where the C-terminus of this large protein is anchored to the APRs that is in close proximity to the midpoint of the conoid in intracellular parasites (Katris et al. 2014). A third bait protein is an otherwise uncharacterised Membrane Occupation and Recognition Nexus (MORN) domain-containing protein (TGME49_245710) that we call MORN3. In an independent study of MORN proteins, we identified MORN3's location as being throughout the IMC but with greatest abundance in a band at the apical cap, although excluded from the very apex where the conoid is located (Fig 1a). Using these three BioID baits we rationalised that SAS6L and RNG2 proximal proteins would be enriched for those near the conoid, and MORN3 proximal proteins would be enriched for apical cap and IMC proteins but not for conoid-proximal proteins.

T. gondii cell lines expressing each of the BioID bait proteins were grown for 24 hours in host cells in elevated exogenous biotin. Streptavidin-detection of biotinylated proteins on Western blots showed unique banding patterns of each cell line and when compared to wildtype controls (cells lacking Bir* fusions) (Fig 1b). Biotinylated proteins from each cell line were then purified on a streptavidin matrix and analysed by mass spectrometry. Proteins enriched ≥3-fold compared to the control or detected in the bait cell lines but not in the control, are indicated in Table S1. Of the hyperLOPIT-assigned *apical* proteins, 25 were also detected by BioID with both SAS6L and RNG2 but not MORN3, and these included known conoid-associated proteins (e.g. MyoH, CPH1, CIP2, CIP3, SAS6L, RNG2). Seven proteins were BioID-detected by MORN3 but not SAS6-L or RNG2, and these are all known apical cap or IMC proteins (AC4, AC8, AC9, AC10, AAP5, IMC9, IMC11). These data indicate that the BioID spatial proteomics indeed enrich for apical proteins, with the differences between SAS6L/RNG2 and MORN3 labelling offering a level of discrimination for conoid-associated proteins when compared to apical cap proteins.

Validation of conoid proteins and determination of substructural location

We confirmed the identification of new apical complex proteins in the region of the conoid in the hyperLOPIT study by endogenous 3'-tagging of candidate genes with reporters (Barylyuk et al. 2020). Imaging by wide-field fluorescence microscopy showed 13 of these proteins confined to a single small punctum at the extreme apex of the cell (Table 1). To test if our expanded hyperLOPIT analysis, including proteins with weaker hyperLOPIT *apical* assignment support, contained further proteins at the apical tip, five further candidates were tagged by the same method. All of these proteins were observed to show the same extreme apical location (Fig S1, Table 1).

The conoid of *T. gondii* is a small structure (~400 x 500 nm) in close proximity to the APRs and preconoidal rings, so widefield fluorescence microscopy is less able to distinguish between proteins at any of these specific structures. To determine the specific locations of our conoid-proximal proteins we employed 3D-SIM super-resolution imaging in cells co-expressing either SAS6L or RNG2

with C-terminal epitope reporters. By 3D-SIM we observe the V5 epitope-tagged version of SAS6L to locate to the entire barrel of the conoid (Fig 2a), rather than above the conoid in the region of the preconoidal rings as was previously reported for YFP-tagged SAS6L (de Leon et al. 2013). RNG2 C-terminal reporters locate to one of the APRs that in *T. gondii* are too close to each other to distinguish by light microscopy, but are nevertheless markers of the extreme apex of the IMC (Dubremetz and Ferguson 2009; Katris et al. 2014). Thus, these two markers provide definition of the relative positions of the new proteins.

Three of the new apical proteins, which were detected by BioID with baits SAS6L and RNG2 but not MORN3, all co-localised to the SAS6L pattern of the full body of the conoid (Fig 2a, Table 1). These labelled structures are situated within the APRs defined by the RNG2 marker. A further three proteins, with this same pattern of BioID detection, all formed a ring at the posterior end, or base, of the conoid (Fig 2b, Table 1). These rings were posterior to the APRs (RNG2) consistent with the retracted position of the conoid in intracellular parasites. We also examined the high-resolution location of one protein detected by all three BioID baits, and this protein colocalised with the RNG2 marker (Fig 3a, Table 1) consistent with the convergence of the three bait proteins at the APRs. We also examined further apical proteins that showed mixed BioID detection, including absence of any BioID positive signals. Four of these proteins formed rings at positions anterior to the conoid, and a further two proteins formed puncta in the same anterior location but were too small to resolve further (Fig 3b, Table 1).

Collectively, these super-resolution data confirm the identities of several new conoid-associated proteins with distinct protein positions at various sites on the conoid, and the enrichment for further conoid proteins by our spatial proteomic methods. During the hyperLOPIT validation of proteins assigned as *Plasma Membrane – peripheral 2* (on the cytosolic leaflet), one protein (MORN2) was found to be enriched at an apical dot (Barylyuk et al. 2020). Given the close proximity of the conoid apex to the plasma membrane, and unknown molecular interactions between these cell structures that might occur here, we examined the location of MORN2 in high resolution (3D-SIM). MORN2 was seen as a small ring anterior to the conoid with a discernible gap between it and the SAS6L-labelled conoid body (Fig 3c). This location is consistent with MORN2 being associated with the plasma membrane and potentially forming a continuum of the annular structures through from the conoid, preconoidal rings, and to apical plasma membrane.

Evolutionary conservation of Toxoplasma conoid proteins throughout Myzozoa

Using the expanded knowledge of conoid-associated proteins determined in this study, and previously identified conoid-associated proteins, we then asked the following questions. How conserved is the *T. gondii* conoid proteome in other apicomplexans and related groups within the Myzozoa (chromerids, dinoflagellates)? Is there evidence of conoid loss from some Aconoidasida taxa as previously suggested? Presence or absence of conoid protein orthologues was tested for in genomic data by two methods. In the first, eukaryote-wide orthogroups defined by OrthoFinder from 79 evolutionary divergent taxa (Barylyuk et al. 2020) were used to test for presence/absence of myzozoan orthologues in the orthogroups of *T. gondii* conoid proteins. The second method used reciprocal sequence searches including BLASTP and iterative HMMER searches. Orthologues identified by either method for select myzozoan taxa are shown in Table 1.

The orthology analysis shows that *T. gondii* conoid proteins are most highly conserved in other coccidia (*Sarcocystis*, *Eimeria*) where most orthologues (~80%) can be detected, although some are still seen to be *Toxoplasma*-specific. In other major apicomplexan groups, represented by *Plasmodium* spp., Piroplasmida, *Cryptosporidium* spp., *Gregarina*, approximately half of the conoid proteins are found, and this is also the case for the nearest apicomplexan relatives the chromerids. These reduced sets of identifiable conoid proteins could represent protein loss in these taxa, gain of novel proteins in

the coccidia, or rapid evolution of relatively unconstrained primary protein sequence of some of these proteins resulting in failure of orthologue detection. Nevertheless, there is evidence of substantial conservation of the proteins associated with the conoid throughout all groups, including members of the Aconoidasida that show similar presence of orthologues to other known conoid-containing taxa (Table 1). The conserved proteins include those associated with all structural components of the *T. gondii* conoid—conoid body, conoid base, and preconoidal rings. Indirect evidence that these proteins continue to function in conoid-type structures, rather than being repurposed for other cell functions, is provided by the absence of most of these proteins in the dinoflagellate *Symbiodinium microadriaticum* (Table 1). This dinoflagellate is a coral symbiont and has apparently lost its feeding apparatus (peduncle) as a homologue of the apical complex structures. There is further evidence of broadly conserved apical proteins detected by our spatial proteomic approaches, but these remain to have their specific apical location determined (Table S1).

Conoid proteins locate to apical rings in Plasmodium zoites

To test if orthologues of *T. gondii* conoid-associate proteins occur in equivalent apical structures in *Plasmodium*, nine proteins were initially selected for reporter tagging in *P. berghei* (Table 1). This model provided ready access to all three invasive zoite forms of the parasite: the ookinete and sporozoite forms produced in the mosquito vector, and the merozoite form of the mammalian blood-staged infection. The nine proteins represented the three sites associated with the conoid (body, base, and preconoidal rings) and one APR protein. GFP fusions of these proteins were initially observed in the large ookinete form by live cell widefield fluorescence imaging, and an apical location was seen for all (Fig 4). Eight of these proteins were resolved only as a dot or short bar at the extreme apical end of the cell, whereas the APR orthologue presented as an apical cap. In addition, a tenth putative conoid protein (ICAP16) identified by Sidik et al (2016), that was not detected in our hyperLOPIT data but was by RNG2-BioID, also has a *Plasmodium* orthologue (Table 1). Live reporter-tagging of this protein also reveals an apical punctum (Fig 4). The locations of all ten of these proteins are consistent with equivalent functions in the apical complex to those of *T. gondii*.

To further resolve the location of the *P. berghei* apical proteins, 3D-SIM was performed on fixed cells for six proteins representing the different presentations found in *T. gondii*. The *P. berghei* orthologue of the conoid body protein (PBANKA_0310700) was resolved as a ring at the cell apex, and this structural presentation was also seen for orthologues of the conoid base (PBANKA_1216300) and preconoidal rings (PBANKA_1347000) (Fig 5). Further, two orthologues that are unresolved preconoidal puncta in *T. gondii* are seen in in *P. berghei* to mimic this presentation either as an apical punctum (PBANKA_1025300) or a barely resolved small ring (PBANKA_1313300) (Fig 5). The apical polar ring orthologue (PBANKA_1334800) that showed a broader cap of signal by widefield imaging was revealed as a ring of much larger diameter than the rings of the conoid orthologues (Fig 5). Furthermore, short spines that radiate from this ring in a posterior direction are resolved that account for the cap-like signal at lower resolution. The location of this protein is consistent with an APR function, although more elaborate in structure than what is seen in *T. gondii* (see Fig 3a). Finally, the ICAP16 orthologue (PBANKA_1419000) also resolved as a ring (Fig 5) consistent with its predicted conoid location in *T. gondii* (Sidik et al. 2016). In all cases examined, the location and structures formed by the *Plasmodium* orthologues phenocopied those of *T. gondii*, strongly suggestive of conservation of function.

Conoid-type structures are present but compositionally distinct between vector and mammalian Plasmodium zoite forms

The presence of a possible homologous conoid structure in *Plasmodium* has been previously attributed to the ookinete stage, but this structure is widely considered to be absent from blood-stage merozoites. With our new markers for components of an apparent conoid in *P. berghei*, and also the new APR protein (PBANKA_1334800), we tested for presence of these proteins in the other

zoite stages: sporozoites and merozoites. To date we have tested for detectable presence of seven of these proteins in sporozoites, and eight in merozoites. In sporozoites all proteins tested for are detected at the cell apex (Fig 6a) and super-resolution imaging of three of these again showed either a ring or unresolved punctum (Fig 6b).

In merozoites, of the eight proteins tested for, only six have been detected in this alternative zoite form of the parasite. Notably, the conoid body orthologue (PBANKA_0310700) and the APR ring protein (PBANKA_1334800) are absent in this cell form (Table 1), suggesting differences in apical complex construction beyond even the conoid. However, all six of the other conoid-associated orthologues are present in merozoites, each forming an apical punctum juxtaposed to the nucleus consistent with apical location (Fig 7). These data support conservation of conoid constituents in the apical complex of both sporozoites and merozoites, but either a reduction in the complexity of this structure in merozoites or the possible substitution for other proteins that are yet to be identified.

Discussion:

The definition often used for the conoid in Apicomplexa is based on the observation of the tubulin-rich fibres observable by electron microscopy. But this cell feature is actually part of a continuum of structures better described as the 'conoid complex' (Dos Santos Pacheco et al. 2020). This complex starts at the apical limits of the IMC coordinated by the two APRs: one at the anterior end of subpellicular microtubules, the other and the apical end of the IMC alveoli (Nichols and Chiappino 1987; Brockley Paterson and Desser 1989; Guttery et al. 2012). The conoid body is tethered within these APRs and its anterior end continues as a series or apical rings (preconoidal rings) towards the apical plasma membrane. Proteins have been previously located to all of these 'substructures' including some linking one substructure to the next (Dos Santos Pacheco et al. 2020). Indeed, the apparent spatial separation of compartments by ultrastructure is smaller than the size of many of the individual molecules that build it (Katris et al. 2014). Thus, at a molecular level it is unclear what the limits of any one substructure are, if this is even a biologically relevant notion.

In this study we have provided a substantial expansion of knowledge of the molecular composition of the conoid complex in *T. gondii* and the architecture of these molecules within this structure. Previous identification of conoid complex proteins used methods including subcellular enrichment, correlation of transcriptional expression, and proximity tagging (BioID) (Hu et al. 2006; Long, Anthony, et al. 2017; Long, Brown, et al. 2017). Amongst these datasets many components have been identified, although often with a high false positive rate. We have found the hyperLOPIT strategy to be a powerful approach for enriching in proteins specific to the apex of the cell, and BioID has further refined identification of proteins specific to the conoid complex region. Collectively, we now know of some 52 proteins that locate at the conoid complex (Table 1), with a dataset of many further proteins that await closer inspection (Table S1). Moreover, we have used high-resolution microscopy to define the specific locations of many, and these show dedicated locations from the APRs through to proteins tethered to the plasma membrane. These data reveal a molecularly complex feature well beyond its tubulin element.

The conservation of a large part of the conoid complex proteome throughout Apicomplexa argues that this is a cell feature maintained throughout this phylum. Conservation includes proteins from all substructural elements—conoid body, conoid base, APRs, preconoidial rings and puncta—suggesting maintenance of this structure in its entirety rather than only select elements. Gene knockout studies in *T. gondii* indicate that these conserved elements play key roles in the viability of these parasites (Table 1). Moreover, the conservation of the conoid complex proteome extends to apicomplexan sister taxa such as *Chromera* which supports previous hypotheses that similarity of the ultrastructure

of ‘apical complexes’ in these myzozoan allies represent genuine homologous structures. While the current sample of dinoflagellates indicates substantial absence of this feature, this could be secondary loss in this coral symbiont.

The interpretation of a conoid being absent in *Plasmodium* stages, and in other Aconoidasida, mostly stems from the lack of a conspicuous extended fibrous structure as seen in coccidia, *Cryptosporidium* spp. and gregarines. A tubulin component of the conoid, however, is reported from other Families of Order Haemosporida such as in *Leucocytozoon* and *Haemoproteus* ookinetes although in dramatically reduced form (Atkinson 1989; Brockley Paterson and Desser 1989). In *Leucocytozoon* only a few microtubules are seen in longitudinal section by TEM and with some difficulty due to the surrounding density of other molecules. With any further reduction in the tubulin component in *Plasmodium* spp., or other Aconoidasida, detection of conoid tubulin by ultrastructural approaches would be even more challenged, and the dominance of the subpellicular microtubules confounds visualisation by specific staining of a potential highly reduced tubulin component.

Regardless of the outstanding question of whether there is a cryptic presence of tubulin in *Plasmodium* conoid-complexes, our data here provide strong support for other compositional elements of this structure being present in all *Plasmodium* zoite forms, and we have previously also shown SAS6L and Myosin B to locate to apical rings in *Plasmodium* (Wall et al. 2016; Wall et al. 2019). Reduction or even loss of the tubulin component seen in coccidia might reflect different mechanical properties and needs for these cells within their target host environments. It is presumed that the conoid in *Toxoplasma* provides a mechanical role in invasion, including the motility of the conoid that can be actively protruded from the cell apex. It is poorly known if this motility feature is seen more widely in coccidia or other apicomplexans, but it might be an adaptation in *Toxoplasma* to the tremendously wide range of hosts and cell types that its zoites must invade, including penetrating thick mucilaginous layers at the gut epithelium. A reduction in this mechanical element of the *Plasmodium* conoid complex likely reflects either the different invasion challenges presented by its its hosts, or different solutions that these parasites have evolved. Such evolutionary change in the composition and architecture of the conoid complex is further supported by differentiation of its proteome between the different *Plasmodium* zoite forms. We observe the blood-stage merozoites to be further reduced, or modified, when compared to ookinetes and sporozoites, and we previously observed SAS6L to be absent from merozoites also, but present in ookinetes and sporozoites (Wall et al. 2016). This differentiation includes the absence of an APR protein in merozoites that in ookinetes forms an elaborate ring with radial posterior projections. We speculate that these proteins delineate so-called ‘tines’, translucent columns seen by TEM that occupy a dilation between the apical portion of the IMC and the subpellicular microtubules seen in *Plasmodium* ookinetes as well as other apicomplexans (Paterson et al. 1988; Brockley Paterson and Desser 1989; Guttery et al. 2012). This APR protein presents an even more elaborate structure than that seen in *Toxoplasma* where it forms a simple ring. Perhaps this is a *Plasmodium* adaptation in lieu of the pronounced extendibility of the conoid complex displayed by *Toxoplasma*. In *Plasmodium* these compositional differences have likely been produced by the different zoite stages’ invasion requirements between merozoites entering erythrocytes and the multiple penetration events for the ookinetes to reach the mosquito gut basal lamina, or the sporozoites to reach this vector’s salivary glands and then mammalian hepatocytes (Vlachou et al. 2006). Ookinete and sporozoite apical complexes might be under further selection by their need for high invasion competence. Each stage represents a transmission bottleneck with success among only one or few parasites required for onward transmission. Increased investment in a robust and reliable apparatus might be justified at these important transmission points.

Evidence of conserved elements of the conoid and conoid-complex throughout Apicomplexa, despite differences in construction and ultrastructure, raise the question of what are the functions of this structure and are they common throughout the phylum? Indeed, even in *Toxoplasma* the function of the conoid is relatively poorly understood. Most studies of individual protein elements have identified molecules required for its the assembly and stability (Leung et al. 2017; Long, Anthony, et al. 2017; Nagayasu et al. 2017; Back et al. 2020; Leung et al. 2020; Tosetti et al. 2020). But other studies have implicated roles in control processes, including activating motility and exocytosis, both of which are requirements for invasion as well as host egress events (Katris et al. 2014; Graindorge et al. 2016; Long, Brown, et al. 2017). Indeed, the conoid is intimately associated with both exocytic vesicles and the apex of the plasma membrane, and this is a common trait throughout not just Apicomplexa but other myzozoans including perkinsids and dinoflagellates (Perkins 1976; Larsen 1988; Paredes-Santos et al. 2012; Suarez et al. 2019). The presence of conserved domains in the conoid-complex proteome is enriched for those that mediate protein-protein interactions as well as responses to regulatory molecules (e.g. Ca^{2+}, cyclic nucleotides) or regulatory protein modifications, and these features are seen in many of the proteins conserved widely amongst taxa (Table 1). This speaks to the conoid complex comprising an ordered regulatory platform for control of vesicle trafficking, fusion and fission, as well as initiation of cell motility. Such a feature as this seems unlikely to be superfluous and lost in these parasites so heavily dependent on mediating complex interactions with hosts through this portal in an otherwise obstructed elaborate cell pellicle. Recognising the common components and importance of the conoid complex throughout Apicomplexa is highly relevant to understanding the mechanisms of invasion and host interaction and the pursuit of better drugs and intervention strategies to combat the many diseases that they cause.

Acknowledgments:

We are grateful to Mike Deery who performed the LC-MS/MS analysis of peptide samples and Julie Howard Murkin for data processing of the BioID LC-MS/MS, both at the Cambridge Centre for Proteomics, and Emilie Daniel for technical assistance with *P. berghei* cell line generation. This work was supported by the Medical Research Council (MR/M011690/1 to R.F.W.; G0900109, G0900278, and MR/K011782/1 to R.T.) the Wellcome Trust through Investigator Award 214298/Z/18/Z to R.F.W, the Isaac Newton Trust – Leverhulme Trust through Early Career Fellowship ECF-2015-562 to K.B., and the Biotechnology and Biological Sciences Research Council (BB/N017609/1) to R.T.

Table 1: *Toxoplasma* conoid-associated proteins and presence of their orthologues in related taxa.

[a]*T. gondii* ME49	Protein Name	[b]Known Localization	* Ref for Localization	Proteomics				[d]Phylogenetic distribution of orthologues									[a]*P. berghei* ANKA	[e]zoite stage			[f]CRISPR KO Score	[g]Conserved Domains
				[c]LOPIT	BioID			*Sarcocystis neurona*	*Eimeria* spp.	*Plasmodium* spp.	*Theileria* spp.	*Babesia* spp.	*Cryptosporidium* spp.	*Gregarina niphandrodes*	Chromerids	*Symbiodinium microadriaticum*		Ookinete	Sporozoite	Merozoite		
					SAS6-like	RNG2	MORN3															
222350		conoid body	LOPIT, 1, TS	+	●	●		●	●	●	●	●	●	●			1229900					
274120		conoid body	1, TS	+	●	●		●	●	●			●	●	●		0310700	●	●	○	0.64	
291880		conoid body	LOPIT, TS	+	●	●		●	●	●							0616200				1.77	
244470	RNG2	conoid+APR	7, 8, TS	+	●	●		●		●			●		●						-4.21	
256030	DCX	conoid body	2, 3	+	●	●	●	●	●	●	●	●	●	●	●		1232600				-5.03	UBQ, p25-α
295450	DIP13	conoid body	6					●	●	●		●	●		●	●	1141900				0.67	
301420	SAS6L	conoid body	4, 5, TS	+	●	●		●	●	●	●	●	●	●	●	●	1414900	●	●		-1.62	SAS6_N
226040	CAM3	conoid body	9	+					●												-3.25	EF
243250	MyoH	conoid body	10	+	●	●		●	●		●	●	●		●						-3.94	Myo, Cal, RCC1
246930	CAM1	conoid body	11, 10	+					●				●								1.09	EF
262010	CAM2	conoid body	11, 9	+				●	●						●						-0.81	EF
246720		conoid base	LOPIT, 1, TS	+	●	●		●	●	●			●		●	●	0109800	●	●	●	0.24	EF
258090		conoid base	LOPIT, 1, TS	+	●	●		●	●	●	●	●	●		●		1216300	●	?	?	-1.34	
266630	CPH1	conoid base	LOPIT, 1, TS	+	●	●		●	●	●	●	●	●	●	●		0620600				-4.16	ANK
208340		PCR	LOPIT, TS	+	●	●		●	●	●	●	●	●		●		0907700	●	?	●	-0.81	PH-like
253600		PCR	LOPIT, TS	+				●	●	●	●	●	●	●			0713200	●	●	?	-2.4	
306350		PCR	LOPIT, TS	+				●	●	●	●	●	●	●	●		1347000	●	●	●	-0.84	
250840	MLC3	PCR	10, TS	+		●		●	●		●	●	●	●							-1.91	EF
292120	MORN2	PCR	LOPIT, TS					●	●				●	●	●	●					-0.04	MORN
250340	Centrin 2	PCR+AA	12, 11, 13	+				●	●	●	●	●	●	●	●		1310400				-4.41	EF
219070		PCP	TS	+				●	●	●	●		●				1025300	●	?	●	-2.2	EF, Crp, CAP_ED
274160		PCP	TS	+				●	●	●	●	●	●		●		1313300	●	●	●	-2.8	
239300	ICMAP1	ICMT	14	+				●	●												-0.74	
267370	Kinesin A	APR	15	+	●	●		●	●	●				●	●	●					-2.7	Kinesin
320030		APR	LOPIT, 1, TS	+	●	●	●		●	●							1334800	●	●	○	-0.19	
243545	RNG1	APR	7, 16																		2.54	
315510	APR1	APR	15	+	●	●			●												-0.05	
202120	ICAP16	apex	18			●		●	●	●	●	●	●				1419000	●	●	●	-2.1	PH-like
206430	FRM1	apex	17	+				●	●	●	●	●	●	●			1245300				-3.24	TPR
209890	ICAP4	apex	18					●	●	●	●	●	●	●	●	●	1439000				-4.84	
216080	AKMT	apex	19		●	●		●	●	●	●	●	●	●	●		0932500				-4.3	SET
225020	CIP3	apex	1	+	●	●		●	●	●			●		●		1309800				-2.78	
227000		apex	LOPIT, 1	+	●	●		●	●	●				●			0510100				-3.17	
291020	MyoL	apex	20					●		●	●	●		●	●		1435500				-1.83	Myo head, RCC1
310070	AAMT	apex	21					●	●	●			●		●	●	1318900				-1.22	Methyltrans
312630	GAC	apex	22					●	●	●	●	●	●	●			1137800	●	●	●	-3.53	
315780	MLC7	apex	10	+					●	●		●	●		●	●	0514800				-0.12	EF
210810	CAP1	apex	23						●						●						-0.73	
226990		apex	1	+	●	●															1.41	
234250	CIP1	apex	1		●	●			●								1423000				-2.02	
234270		apex	1	+	●			●													-0.44	
239560	MyoE	apex	20					●	●												0.11	Myo head
252880	CRMP	apex	23	+	●	●		●													-2.35	Ax_dyn_light
254870		apex	1					●	●												0.64	TerD
255895		apex	1	+	●	●		●	●												0.23	
257300	CIP2	apex	1	+	●	●		●	●												-2.49	
278780		apex	LOPIT	+	●	●	●		●												-2.77	UBQ
284620		apex	TS	+				●	●												-1.02	
295420		apex	LOPIT, 1	+	●	●		●	●												-1.57	TerD
297180		apex	LOPIT	+	●	●		●	●												-1.52	
311260	MLC5	apex	10	+				●	●		●	●	●								-0.33	EF
313780		apex	1	+	●	●															0.71	

[a] Proteins with location data by microscopy in this and the hyperLOPIT study shown in red.

[b] Known localization defined as 'apex' when low resolution imaging only has identified an apical punctum at the apex of the cell. PCR, preconoidal ring; AA, apical annuli; PCP, preconoidal punctum; ICMT, intraconoidal microtubules; APR, apical polar ring.

[c] Proteins with data in the hyperLOPIT datasets are indicated with '+'.

[d] Presence of orthologue by: ●, orthogroup analysis previously described (Barylyuk et al. 2020); ●, reverse BLASTP analysis (Dos Santos Pacheco et al. 2020) but absent in the orthogroup analysis; ●, reverse interative hmmer search but not in the other analyses.
[e] CRISPR KO phenotype scores where more strongly negative scores indicate increasing detrimental competitive growth in in vitro culture conditions (Sidik et al. 2016).
[f] Conserved domain abbreviations: EF, EF-hand; Myo, myosin; Cal, Calmodulin-binding motifs; TPR, Tetratricopeptide repeat; Methyltrans, Methyltransferase; Ax_dyn_light, Axonemal dynein light chain.
[*] References for localization data: TS, this study; LOPIT, (Barylyuk et al. 2020); 1, (Long, Anthony, et al. 2017); 2, (Nagayasu et al. 2017); 3, (Leung et al. 2020); 4, (de Leon et al. 2013); 5, (Wall et al. 2016); 6, (Lévêque et al. 2016); 7, (Tosetti et al. 2020); 8, (Katris et al. 2014); 9, (Long, Brown, et al. 2017); 10, (Graindorge et al. 2016); 11, (Hu et al. 2006); 12, (Lentini et al. 2019); 13, (Leung et al. 2019); 14, (Heaslip et al. 2009); 15, (Leung et al. 2017); 16, (Tran et al. 2010); 17, (Tosetti et al. 2019); 18, (Sidik et al. 2016); 19, (Heaslip et al. 2011); 20, (Frénal, Jacot, et al. 2017); 21, (Engelberg et al. 2020); 22, (Jacot et al. 2016); 23, (Gould et al. 2011); 24, (Skariah et al. 2012).

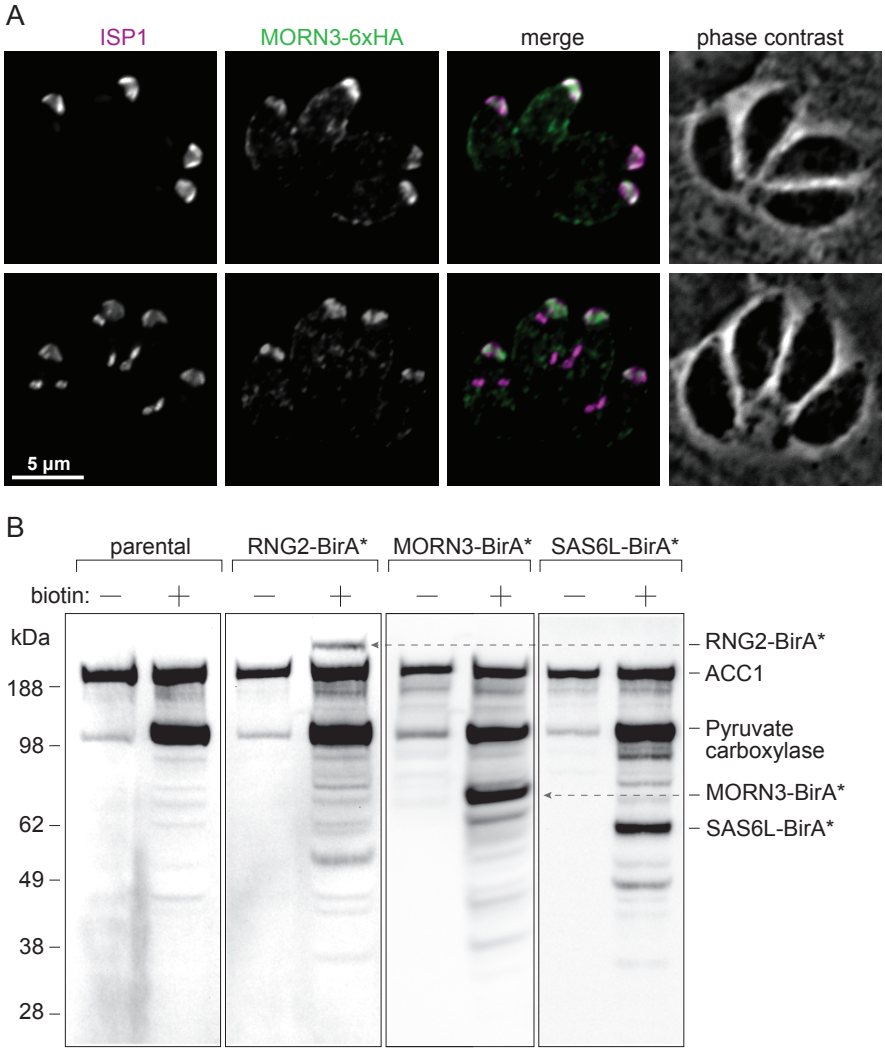

Figure 1: Apically targeted BioID using baits RNG2, SAS6L and new apical cap protein MORN3.
A. Immuno-detection of HA-tagged MORN3 in *T. gondii* intracellular parasites co-stained for apical cap-marker ISP1. Upper panels show mature cells, lower panels show internal daughter pellicle buds forming at the early stages of endodyogeny. Scale bar = 5 µm.
B. Steptavidin-detection of biotinylated proteins after 24 hours of growth either with or without elevated biotin. Native biotinylated proteins ACC1 and pyruvate carboxylase are seen in the parental control (lacking BirA*) and in the BioID bait cell lines. Additional biotinylated proteins are seen in each of the bait cell lines, including self-biotinylation of the bait fusion.

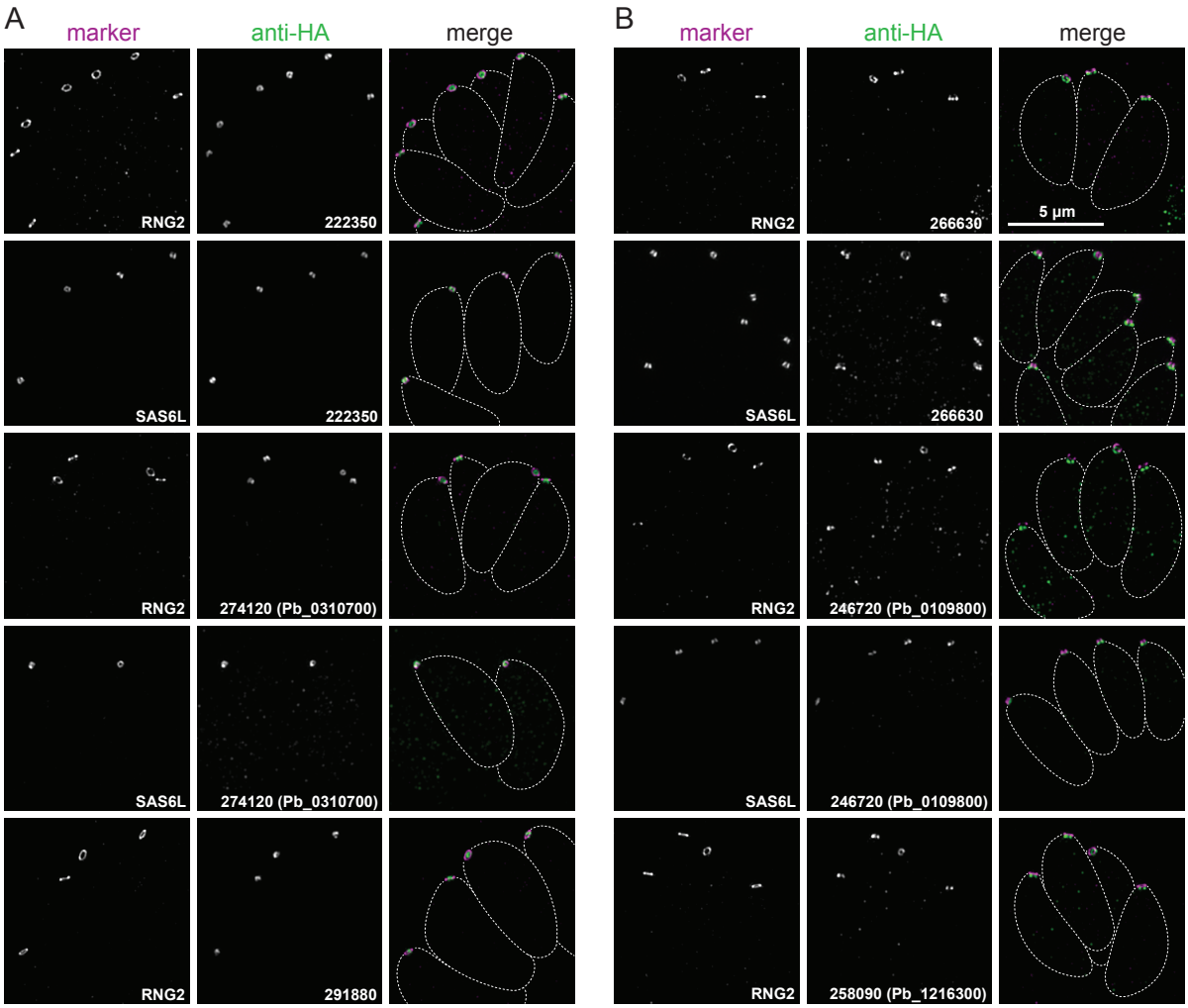

Figure 2: Super-resolution imaging of *T. gondii* proteins at the conoid body and base.
Immuno-detection of HA-tagged conoid proteins (green) in cells co-expressing either APR marker RNG2 or conoid marker SAS6L (magenta). All images are at the same scale, scale bar = 5 µm.
A. Proteins specific to the conoid body.
B. Proteins specific to the conoid base.

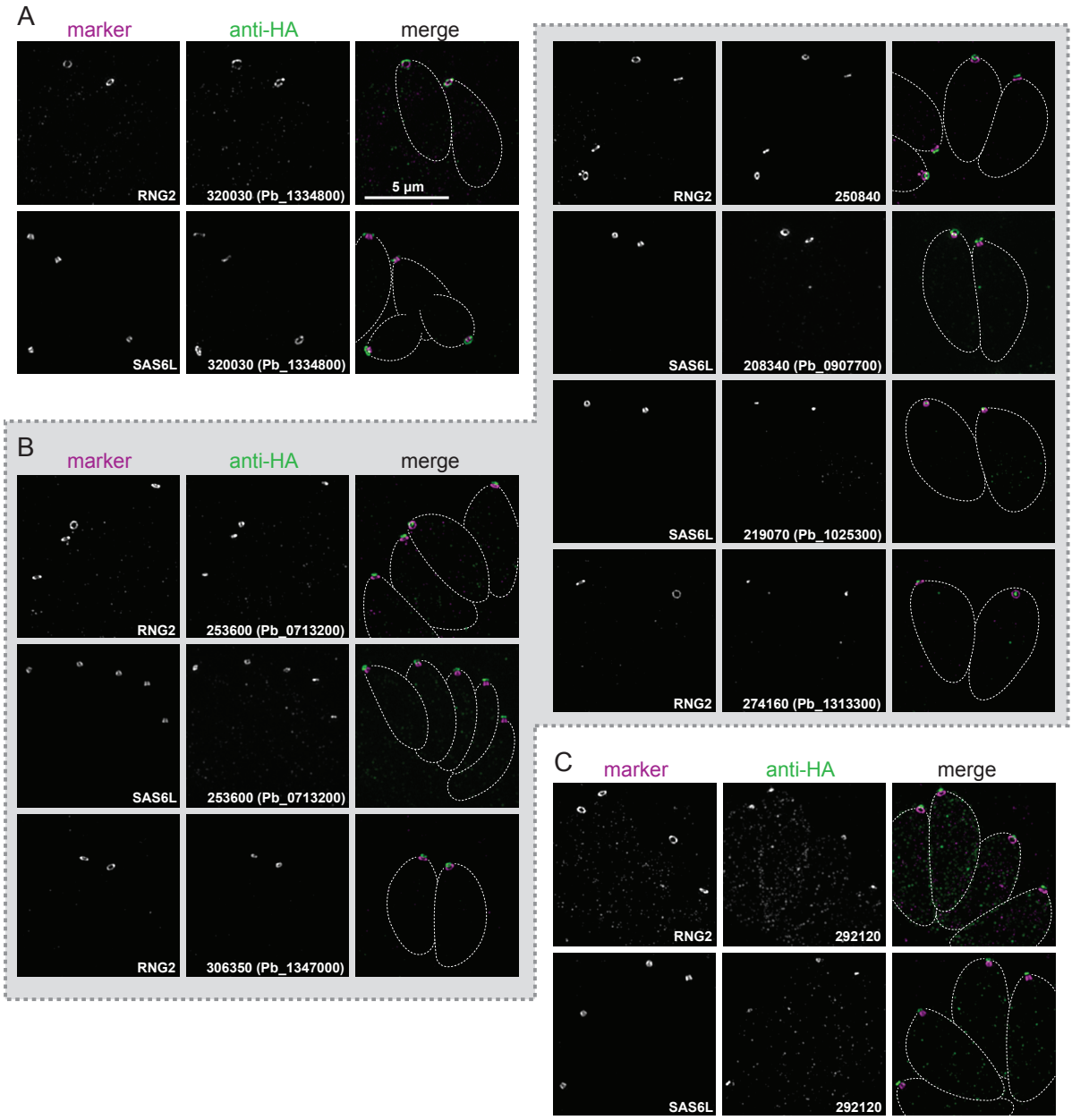

Figure 3: Super-resolution imaging of *T. gondii* proteins at the APR, preconoidal rings, and apical plasma membrane.

Immuno-detection of HA-tagged proteins as for Figure 2.

A. Protein specific to the apical polar ring (APR).

B. Proteins specific to structures at the anterior end of the conoid body.

C. A peripheral protein (MORN2) of the inner leaflet of the plasma membrane.

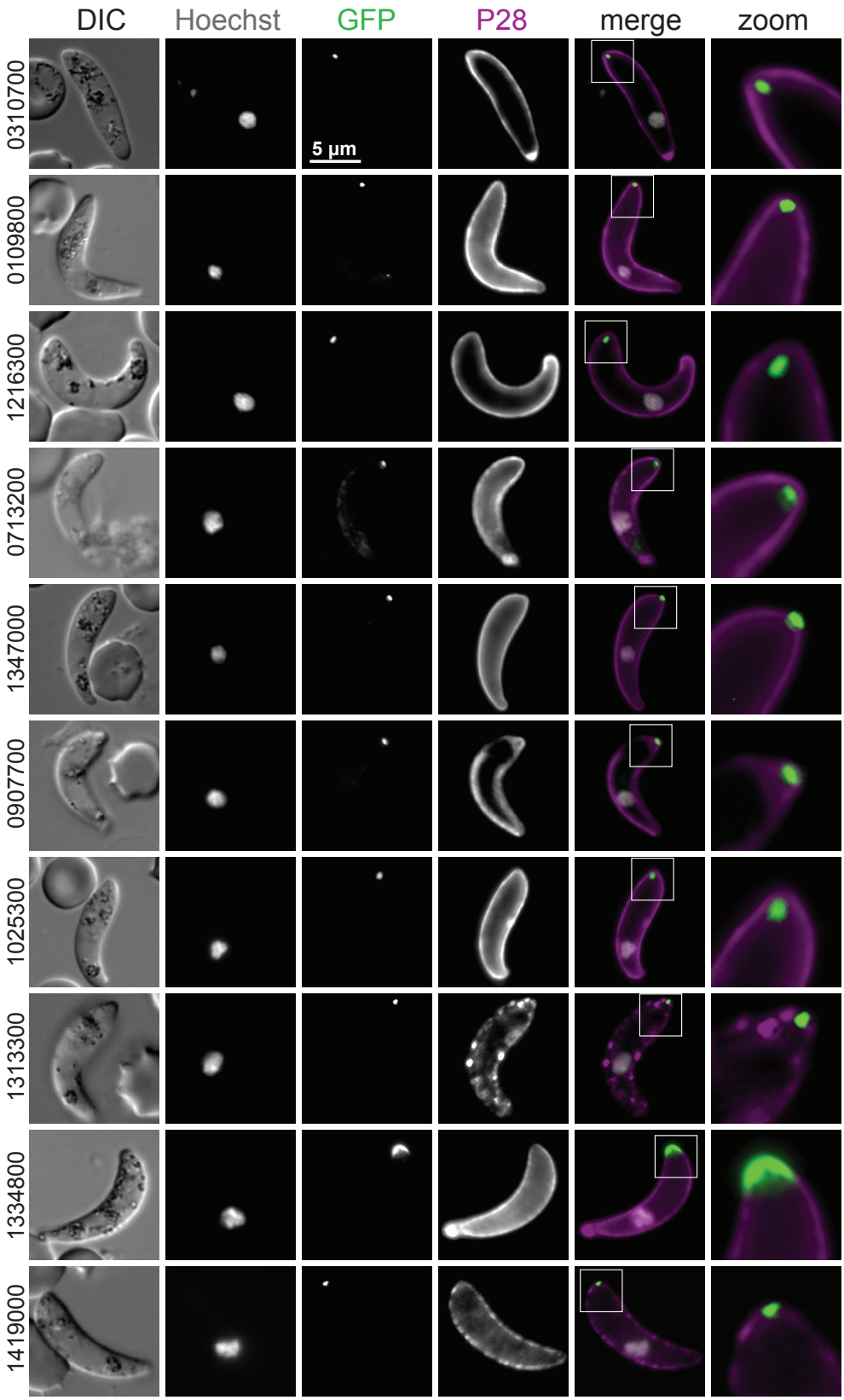

Figure 4: Live cell imaging of *P. berghei* ookinetes expressing GFP fusions of conoid complex orthologues.

Widefield fluorescence imaging showing GFP (green), Hoechst 33342-stained DNA (grey), and live cy3-conjugated antibody staining of ookinete surface protein P28 (magenta). All images are at the same scale, scale bar = 5 µm shown.

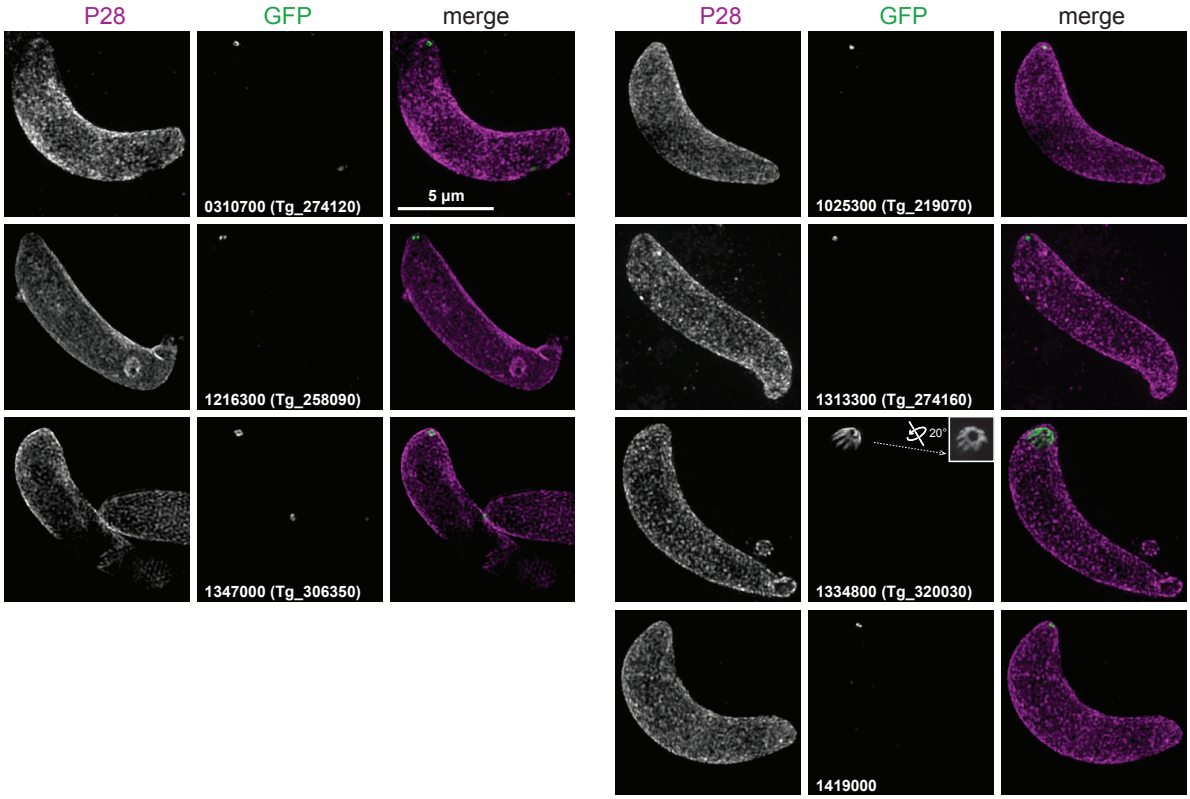

Figure 5: Super-resolution imaging of *P. berghei* conoid complex orthologues in ookinetes.
GFP-fused conoid complex proteins (green) shown with the cell surface stained for ookinete surface protein P28 (magenta). Inset for APR protein (1334800) shows rotation of the 3D-reconstruction to view the parasite apex face on. All images are at the same scale, scale bar = 5 µm.

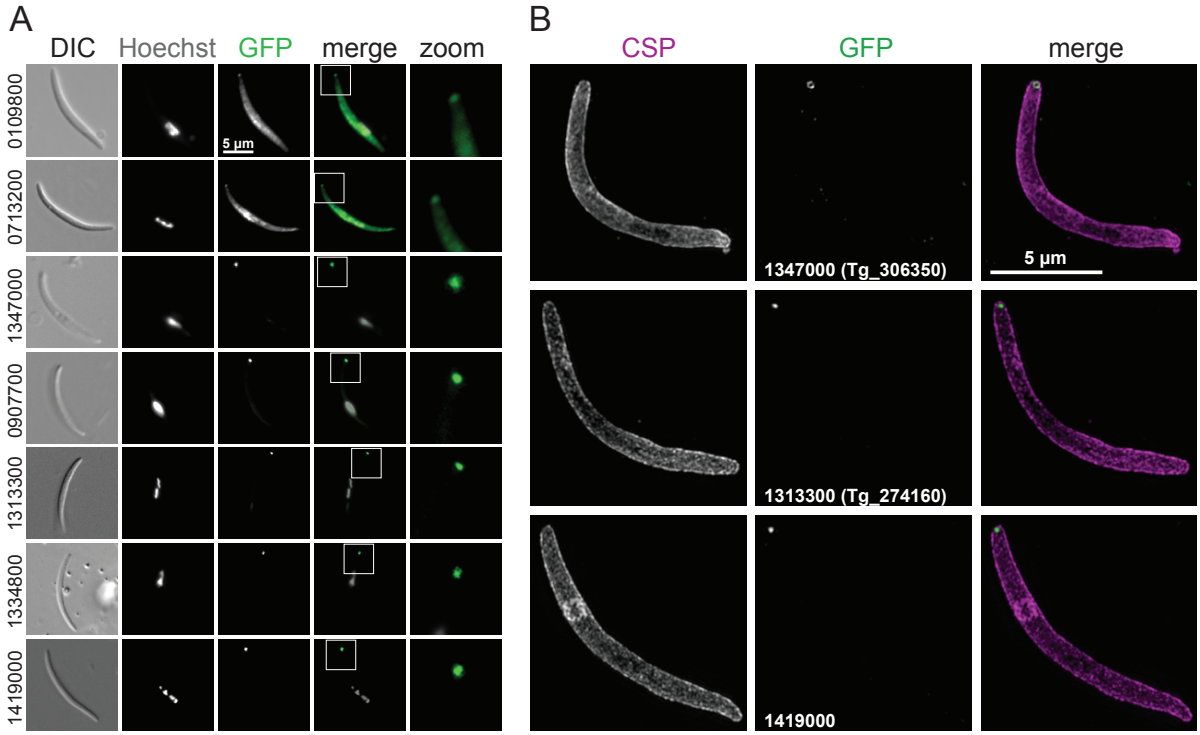

Figure 6: Live cell widefield, and super-resolution imaging of *P. berghei* sporozoites expressing GFP fusions of conoid complex orthologues.

A. Widefield fluorescence imaging showing GFP (green) and Hoechst 33342-stained DNA (grey). All images are at the same scale, scale bar = 5 µm shown.

B. Super-resolution imaging of GFP-fused conoid complex proteins (green) in fixed cells shown with the cell surface stained for sporozoite surface protein 13.1 (magenta). All images are at the same scale, scale bar = 5 µm shown.

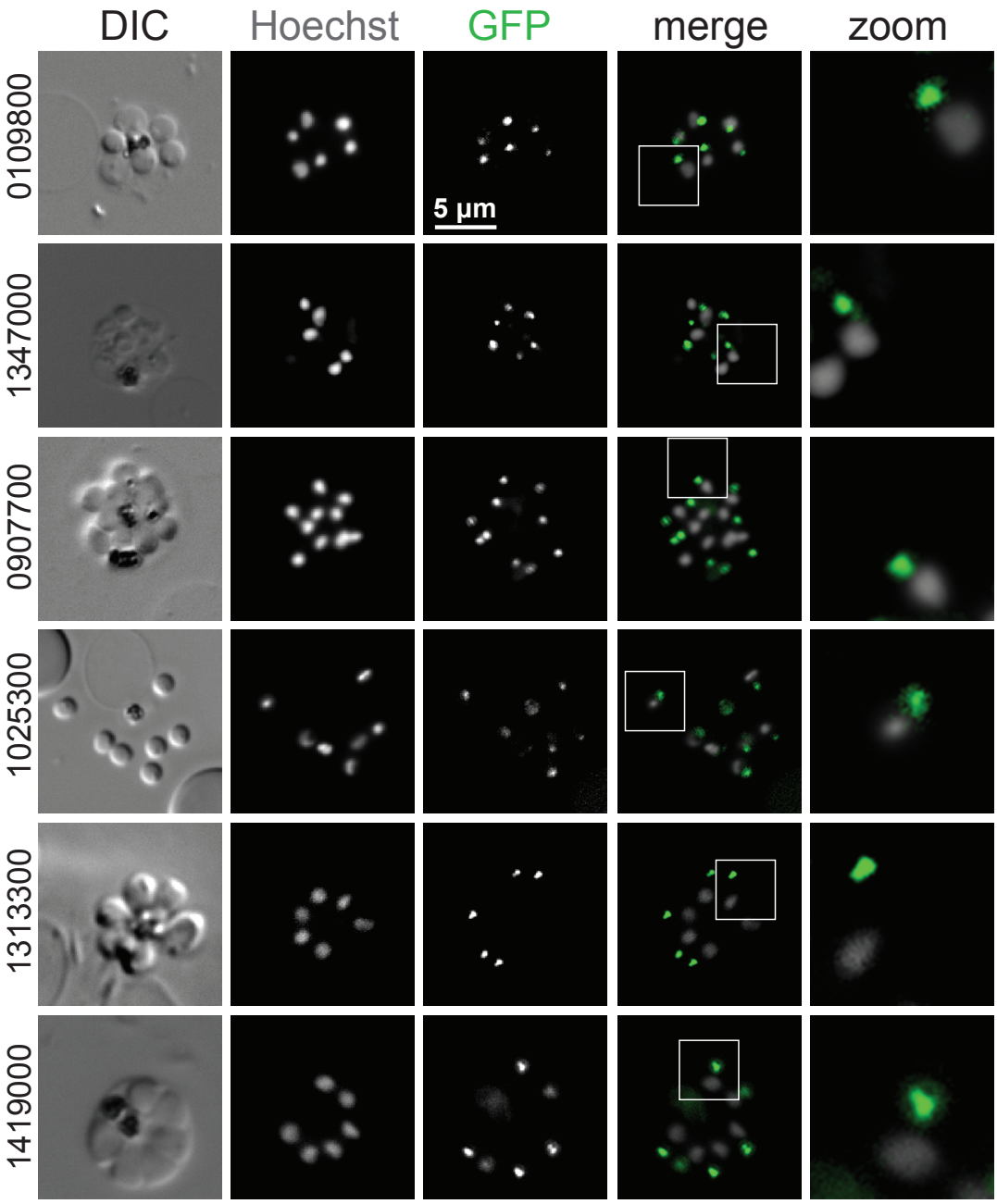

Figure 7: Live cell imaging of *P. berghei* merozoites expressing GFP fusions of conoid complex orthologues.

Widefield fluorescence imaging showing GFP (green) and Hoechst 33342-stained DNA (grey) with some parasites seen pre-egress from the erythrocyte and others post egress. All images are at the same scale, scale bar = 5 µm shown.

Supplemental Material

Table S1: *Toxoplasma* candidate apical proteins identified by hyperLOPIT and BioID methods, and presence of their orthologues in related taxa

Footnotes:

[a] Known localization defined as 'apex' when low resolution imaging only has identified an apical punctum at the apex of the cell. PCR, preconoidal ring; AA, apical annuli; PCP, preconoidal punctum; ICMT, intraconoidal microtubules; APR, apical polar ring; SPMT, subpellicular microtubules.

[b] LOPIT assignment strength: ●, Protein assigned to one of the two *apical* clusters at >99% probability in (before '/') the analysis of concatenated three or two datasets where the protein was full quantified or (after '/') the analysis of individual hyperLOPIT datasets; ●, as for above except apical assignment probability is below the 99% threshold; ●, assignment to a cluster other than *apical* in one non-catenated hyperLOPIT datasets.

[c] Presence of orthologue by: ●, orthogroup analysis previously described (Barylyuk et al. 2020); ●, reverse BLASTP analysis (Dos Santos Pacheco et al. 2020) but absent in the orthogroup analysis; ●, reverse interative hmmer search but not in the other analyses.

[d] CRISPR KO phenotype scores where more strongly negative scores indicate increasing detrimental competitive growth in in vitro culture conditions (Sidik et al. 2016).

[*] References for localization data: TS, this study; LOPIT, (Barylyuk et al. 2020); 1, (Long, Anthony, et al. 2017); 2, (Nagayasu et al. 2017); 3, (Leung et al. 2020); 4, (de Leon et al. 2013); 5, (Wall et al. 2016); 6, (Lévêque et al. 2016); 7, (Tosetti et al. 2020); 8, (Katris et al. 2014); 9, (Long, Brown, et al. 2017); 10, (Graindorge et al. 2016); 11, (Hu et al. 2006); 12, (Lentini et al. 2019); 13, (Leung et al. 2019); 14, (Heaslip et al. 2009); 15, (Leung et al. 2017); 16, (Tran et al. 2010); 17, (Tosetti et al. 2019); 18, (Sidik et al. 2016); 19, (Heaslip et al. 2011); 20, (Frénal, Jacot, et al. 2017); 21, (Engelberg et al. 2020); 22, (Jacot et al. 2016); 23, (Gould et al. 2011); 24, (Skariah et al. 2012); 25, (Chen et al. 2015); 26, (Chen et al. 2017); 27, (Beck et al. 2010); 28, (Anderson-White et al. 2011); 29, (Suvorova et al. 2015); 30, (Liu et al. 2013); 31, (Liu et al. 2016); 32, (Frénal et al. 2010).

Table S2: Oligonucleotide primers used for *P. berghei* gene tagging

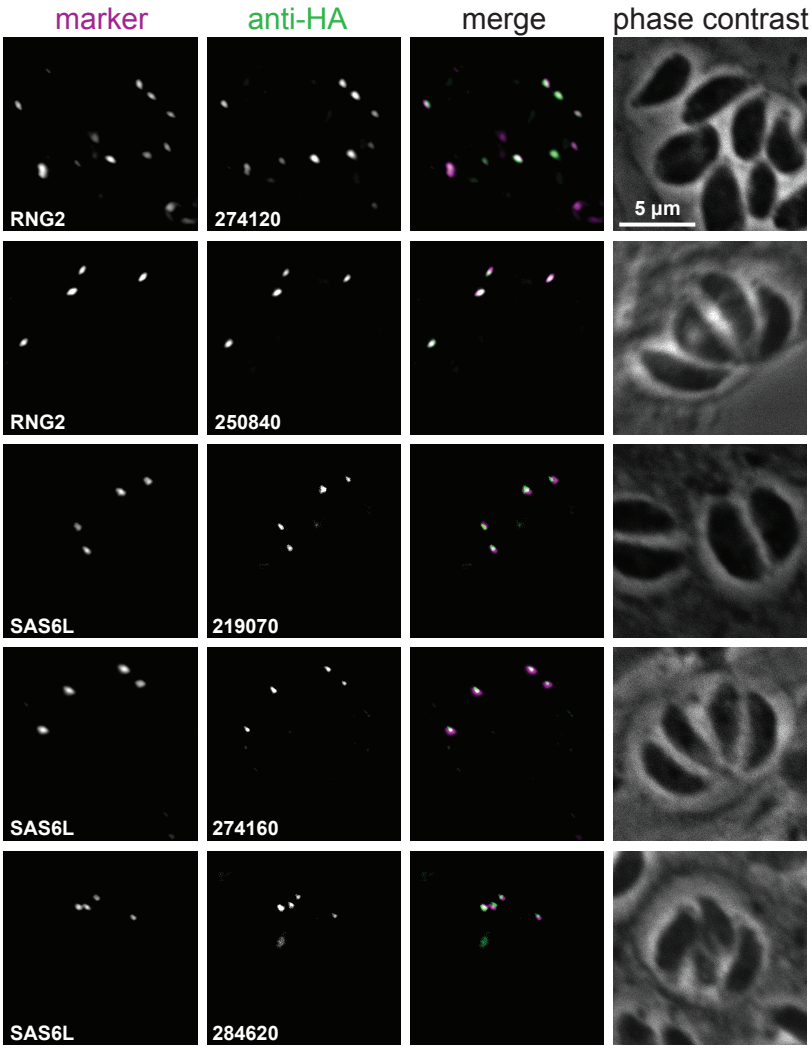

Figure S1: Immuno-detection of candidate conoid complex proteins, identified by hyperLOPIT and BioID approaches, in *T. gondii* intracellular parasites

Widefield fluorescence imaging of HA-tagged candidates (green) co-expressing either APR marker RNG2 or conoid marker SAS6L (magenta). All images are at the same scale, scale bar = 5 µm.

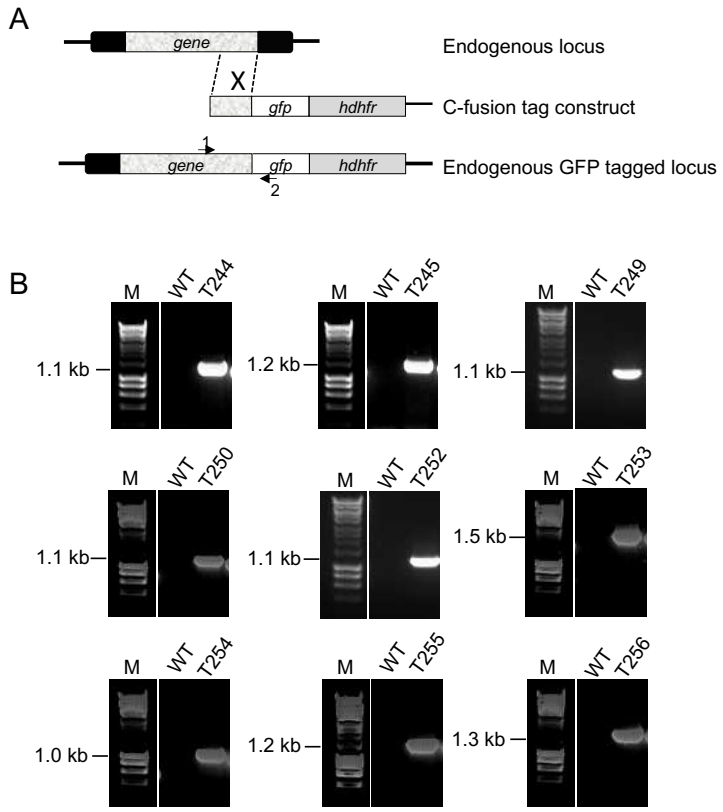

Figure S2:

A. Schematic representation of the endogenous gene locus, the constructs and the recombined gene locus.
B. Validation of correct gene tagging by diagnostic PCR

Materials and Methods:

Growth and generation of transgenic T. gondii

T. gondii tachyzoites from the strain RH and derived strains, including RH Δku80/TATi (Sheiner et al. 2011), were maintained at 37°C with 10% CO_2 growing in human foreskin fibroblasts (HFFs) cultured in Dulbecco's Modified Eagle Medium supplemented with 1% heat-inactivated fetal bovine serum, 10 unit ml^{-1} penicillin and 10 µg ml^{-1} streptomycin, as described elsewhere (Roos et al. 1994). When appropriate for selection, chloramphenicol was used at 20 µM and pyrimethamine at 1 µM. Scaling up of the parasite culture for hyperLOPIT experiments was done according to the method described by Roos et al. (1994).

BioID

Sample preparation: For the proximity biotinylation assay, we generated three different cell lines (*T. gondii* tachyzoites RH Δku80 TATi) by *in situ* genomic C-terminal tagging of one of the three bait proteins (SAS6L, RNG2 or MORN3) with the promiscuous bacterial biotin ligase, BirA*. Our protein tagging protocol is described in our previous work (Barylyuk et al. 2020). We then followed the BioID protocol of Chen et al. (Chen et al. 2015). The biotinylation of the proximal proteins by the BirA* enzyme was triggered by addition of 150 µM biotin into the ED1 growth media 24 h prior to parasite egress. The non-tagged parental cell line was used as a negative control for background biotinylation.

The BirA* activity was validated by a western blot detection of the biotinylated proteins. $1x10^7$ tachyzoites were lysed in the NuPAGE™ LDS Sample Buffer (ThermoFisher) and separated in the NuPAGE™ 4-12% Bis-Tris gel (ThermoFisher) and blotted on a nitrocellulose membrane, which was then blocked by 3% BSA for 1h. The membrane was then incubated in the presence of streptavidin-HRP conjugate (ThermoFisher; 1:1000 dilution in 1% BSA) for 1h, followed by five 5-minute washes in TBST buffer. The HRP chemiluminescent signal was visualised by the Pierce™ West Pico kit (ThermoFisher) and a digital camera.

For the proteomic analysis ~ 2×10^9 tachyzoites were harvested after 24-h biotinylation and egress. The parasites were separated from the host-cell debris by passing through 3 µm filter and washed 5x in phosphate-buffered saline. The cell pellets were lysed in RIPA buffer and the volume of each sample was adjusted to 1.8 mL and 5 mg of total protein content. 200 µL of unsettled Pierce™ Streptavidin Magnetic Beads (ThermoFisher) were first washed in RIPA buffer and then incubated with the cell lysates for 1h at room temperature with gentle rotation of the tubes. The beads were then washed 3x in RIPA, 1x in 2M UREA 10mM Tris-HCl pH 8.0 and again 3x in RIPA, reduced in 10 mM DTT 100 mM ammonium bicarbonate solution for 30 min at 56°C, alkylated in 55 mM iodoacetamide 100 mM ammonium bicarbonate for 45 min at room temperature in dark, and finally washed in 50 mM ammonium bicarbonate for 15 min at room temperature with gentle rotation. The peptides were digested off the beads by an overnight 37°C incubation in 1 µg of trypsin dissolved in 50 mM ammonium bicarbonate and the remaining peptides were extracted by 10% formic acid. Both peptide extractions were combined, desalted using C18 solid-phase extraction cartridges (SepPak C18, 100 mg sorbent, Waters), dried in a vacuum centrifuge (Labconco), and resuspended in 0.1% (vol.) aqueous formic acid.

Peptide analysis by liquid chromatography and mass spectrometry: Peptide fractions were analysed in a Q Exactive hybrid quadrupole-Orbitrap mass spectrometer (Thermo Fisher Scientific Inc., Waltham, MA, USA) coupled on-line with a nanoflow ultrahigh-performance liquid chromatography (nano-UPLC) system Dionex UltiMate 3000 RSLCnano (Thermo Fisher Scientific Inc., Waltham, MA, USA) through an EASY-Spray nano-electrospray ionization (nano-ESI) source (Thermo Fisher Scientific Inc., Waltham, MA, USA). Samples were loaded onto an Acclaim PepMap 100 C18 trapping column (300 µm i.d. × 5 mm length, 5 µm particle size, 100 Å pore size; Thermo Fisher

Scientific) from the UltiMate 3000 autosampler with 0.1% aqueous formic acid for 3 minutes at a flow rate of 10 µL/min. After this period, the column valve was switched to allow elution of peptides from the pre-column onto the analytical column. The trapped peptides were then separated on an analytical reverse-phase nano EASY-Spray RSLC C18 column (75 µm i.d. × 50 cm length, 2 µm particle size, 100 Å pore size; Thermo Fisher Scientific) at a flow rate of 300 nL/min using a gradient elution program. Solvent A was 0.1% (vol.) formic acid in water (HPLC-grade, Fisher). Solvent B was 0.1% (vol.) formic acid in water-acetonitrile (HPLC-grade, Rathburn) blend (20% : 80%, vol.). The linear gradient employed was 2-40% B in 30 minutes. Further wash and equilibration steps gave a total run time of 60 minutes.

The peptides eluting from the nano-LC column were ionized by applying a positive voltage of 2.1 kV to the nano-ESI emitter electrode. The mass spectrometer was operated in the data-dependent acquisition mode (DDA). The precursor survey mass spectra (MS1) were acquired between 380 and 1500 *m/z* at a 70,000 resolution with an AGC target of 1e6 and a maximum C-trap fill time of 250 ms. The top twenty most intense precursor ions were selected in the quadrupole using an isolation window of 1.5 *m/z* and subjected to fragmentation in the HCD mode at a normalised collision energy of 25. A dynamic exclusion limit of 20 s was applied. The fragment ion spectra (MS2, or MS/MS) were acquired between 100 and 1600 m/z at a 17,500 resolution with an AGC target of 5e4 and a maximum C-trap fill time of 120 ms. The raw data were managed in XCalibur v2.1 (Thermo Fisher Scientific).

LC-MS/MS data processing and protein identification: The raw LC-MS/MS data were interpreted, and peak lists were generated using Proteome Discoverer 1.4 (Thermo Fisher Scientific). The peak lists were exported as Mascot generic files and searched against the annotated protein sequences of *Toxoplasma gondii* strain ME49 (ToxoDB.org release 29, retrieved on 20 December 2016; 8,322 sequences) and common contaminant proteins (cRAP, 115 sequences) with Mascot v2.6.0 (Matrix Science) assuming digestion by trypsin with up to two missed cleavage sites. The precursor and fragment ion tolerances were set to 5 ppm and 0.1 Da, respectively. Carbamidomethylation of cysteine was set as a static modification. Oxidation of methionine and deamidation of asparagine and glutamine were allowed as variable modifications. Scaffold (version Scaffold_4.7.5, Proteome Software Inc., Portland, OR) was used to validate MS/MS based peptide and protein identifications. Peptide identifications were accepted if they could be established at greater than 95.0% probability by the Peptide Prophet algorithm (Keller et al. 2002) with Scaffold delta-mass correction. Protein identifications were accepted if they could be established at greater than 95.0% probability and contained at least 2 identified peptides. Protein probabilities were assigned by the Protein Prophet algorithm (Nesvizhskii et al. 2003). Proteins that contained similar peptides and could not be differentiated based on MS/MS analysis alone were grouped to satisfy the principles of parsimony. Total spectrum counts were used to calculate the enrichment of proteins in the BioID pulldown samples relative to the wild-type control sample. The contaminant proteins were not considered for the analysis.

hyperLOPIT Spatial proteomics

The hyperLOPIT spatial proteomics data were acquired and analysed as described in (Barylyuk et al. 2020). Three independent hyperLOPIT datasets contained 4,189, 4,547, and 4,292 proteins identified at FDR < 5% based on at least one unique high-confidence peptide (FDR < 1%) and quantified across all ten TMT quantification channels. These three 10plex datasets were considered individually. In addition, the data on proteins shared across pairs or all three individual datasets were concatenated affording three 20plex (4,021, 3,916, and 4,078 proteins for $1 \cap 2$, $1 \cap 3$, and $2 \cap 3$, respectively) and one 30plex (3,832 proteins), thus, yielding seven datasets in total. For each dataset, the normalised protein abundance distribution profiles were derived from the TMT reporter ion intensity values. A Bayesian machine-learning classification method based on t-augmented Gaussian mixture models

with maximum *a posteriori* prediction (TAGM-MAP) (Crook et al. 2018) was applied to probabilistically attribute proteins to 26 subcellular niches defined by 718 marker proteins. The protein localisation probability was calculated as the probability of a protein belonging to the most likely subcellular class and not being an outlier. Assignments at a localisation probability above 99% were accepted.

Immunofluorescence microscopy

T. gondii-infected HFF monolayers grown on glass coverslips were fixed with 2% formaldehyde at room temperature for 15 min, permeabilized with 0.1% TritonX-100 for 10 min and blocked with 2% BSA for 1 h. The coverslips were then incubated with a primary antibody for 1 h, followed by 1 h incubation with a secondary antibody. Coverslips were mounted using ProLong® Diamond Antifade Mountant with DAPI (Invitrogen). Images were acquired using a Nikon Eclipse Ti widefield microscope with a Nikon objective lens (Plan APO, 100x/1.45 oil), and a Hamamatsu C11440, ORCA Flash 4.0 camera.

3D-Structured Illumination Microscopy (3D-SIM) was implemented on a DeltaVision OMX V4 Blaze (Applied Precision) with samples prepared as for widefield immunofluorescence assay (IFA) microscopy expect High Precision coverslips (Marienfeld Superior, No1.5H with a thickness of 170 µm ± 5 µm) were used in cell culture and Vectashield (Vector Laboratories) was used as mounting reagent. Samples were excited using 405, 488 and 594 nm lasers and imaged with a 60x oil immersion lens (1.42 NA). The structured illumination images were reconstructed in softWoRx software version 6.1.3 (Applied Precision). All fluorescence images were processed using Image J software (http://rsbweb.nih.gov./ij/).

Ethics statement

The animal work performed in the UK passed an ethical review process and was approved by the United Kingdom Home Office. Work was carried out in accordance with the United Kingdom 'Animals (Scientific Procedures) Act 1986' and in compliance with 'European Directive 86/609/EEC' for the protection of animals used for experimental purposes. The permit number for the project licence is 40/3344.

Generation of transgenic Plasmodium berghei and genotype analyses

To observe the location of *P. berghei* proteins, the C-terminus of the gene was tagged with green fluorescent protein (GFP) sequence by single crossover homologous recombination. To generate the GFP-tag line, a region of these genes downstream of the ATG start codon was amplified, ligated to p277 vector, and transfected as described previously (Guttery et al. 2012). The p277 vector contains the human *dhfr* cassette, conveying resistance to pyrimethamine. A schematic representation of the endogenous gene locus, the constructs and the recombined gene locus can be found in Fig S2. A list of primers used to amplify these genes can be found in Table S2. For the parasites expressing a C-terminal GFP-tagged protein, diagnostic PCR was used with primer 1 (Int primer) and primer 2 (ol492) to confirm integration of the GFP targeting construct. The primer details can be found in Table S2**.**

P. berghei phenotype analyses

Blood containing approximately 50,000 parasites of the GFP-tagged lines was injected intraperitoneally (i.p.) into mice to initiate infections. Parasitaemia was monitored by microscopy on Giemsa stained thin smears. Four to five days post infection blood was collected with 10-15% parasitaemia and allowed to further develop in schizont and ookinete culture medium as described previously (PMID: 29138437). The schizonts/merozoites and ookinetes were examined after 24 h of culture with a Zeiss AxioImager M2 microscope (Carl Zeiss, Inc) fitted with an AxioCam ICc1 digital camera. To examine the location of these proteins in sporozoites, 30–50 *Anopheles stephensi* SD 500

mosquitoes were allowed to feed for 20 min on anaesthetized, infected mice who's asexual parasitaemia had reached 15% and were carrying comparable numbers of gametocytes as determined on Giemsa stained blood films. On day 21 post-feeding, 20 mosquitoes were dissected, and their guts and salivary glands crushed separately in a loosely fitting homogenizer to release sporozoites, which were then used for imaging.

References:

Aikawa M. 1967. Ultrastructure of the pellicular complex of *Plasmodium fallax*. J Cell Biol 35:103–113.

Anderson-White BR, Ivey FD, Cheng K, Szatanek T, Lorestani A, Beckers CJ, Ferguson DJP, Sahoo N, Gubbels M-J. 2011. A family of intermediate filament-like proteins is sequentially assembled into the cytoskeleton of *Toxoplasma gondii*. Cell Microbiol 13:18–31.

Atkinson CT. 1989. Ultrastructure of the ookinetes of *Haemoproteus meleagridis* (Haemosporina: Haemoproteidae). J. Parasitol. 75:135–141.

Back PS, O'Shaughnessy WJ, Moon AS, Dewangan PS, Hu X, Sha J, Wohlschlegel JA, Bradley PJ, Reese ML. 2020. Ancient MAPK ERK7 is regulated by an unusual inhibitory scaffold required for *Toxoplasma* apical complex biogenesis. PNAS 117:12164–12173.

Barylyuk, K., Koreny, L., Ke, H., Butterworth, S., Crook, O.M., Lassadi, I., Gupta, V., Tromer, E.C., Mourier, T., Stevens, T.J., Breckels, L.M., Pain, A., Lilley, K.S., and Waller, R.F. (2020) A subcellular atlas of *Toxoplasma* reveals the functional context of the proteome. bioRixv https://doi.org/10.1101/2020.04.23.057125

Beck JR, Rodriguez-Fernandez IA, de Leon JC, Huynh M-H, Carruthers VB, Morrissette NS, Bradley PJ. 2010. A novel family of *Toxoplasma* IMC proteins displays a hierarchical organization and functions in coordinating parasite division. PLOS Pathog 6:e1001094.

Brockley Paterson W, Desser SS. 1989. The polar ring complex in ookinetes of *Leucocytozoon simondi* (Apicomplexa: Haemosporina) and evidence for a conoid in haemosporidian ookinetes. Eur J Protistol 24:244–251.

Chen AL, Kim EW, Toh JY, Vashisht AA, Rashoff AQ, Van C, Huang AS, Moon AS, Bell HN, Bentolila LA, et al. 2015. Novel components of the *Toxoplasma* inner membrane complex revealed by BioID. MBio 6:e02357–14.

Chen AL, Moon AS, Bell HN, Huang AS, Vashisht AA, Toh JY, Lin AH, Nadipuram SM, Kim EW, Choi CP, et al. 2017. Novel insights into the composition and function of the *Toxoplasma* IMC sutures. Cellular Microbiology 19:e12678.

Christoforou A, Mulvey CM, Breckels LM, Geladaki A, Hurrell T, Hayward PC, Naake T, Gatto L, Viner R, Martinez Arias A, et al. 2016. A draft map of the mouse pluripotent stem cell spatial proteome. Nat Commun 7:8992.

Crook OM, Mulvey CM, Kirk PDW, Lilley KS, Gatto L. 2018. A Bayesian mixture modelling approach for spatial proteomics. PLoS Comput. Biol. 14:e1006516.

de Leon JC, Scheumann N, Beatty W, Beck JR, Tran JQ, Yau C, Bradley PJ, Gull K, Wickstead B, Morrissette NS. 2013. A SAS-6-like protein suggests that the *Toxoplasma* conoid complex evolved from flagellar components. Euk Cell 12:1009–1019.

Dos Santos Pacheco N, Tosetti N, Koreny L, Waller RF, Soldati-Favre D. 2020. Evolution, Composition, Assembly, and Function of the Conoid in Apicomplexa. Trends in Parasitology:1–17.

Dubremetz J-F, Ferguson DJP. 2009. The role played by electron microscopy in advancing our understanding of *Toxoplasma gondii* and other apicomplexans. Int J Parasitol 39:883–893.

Engelberg K, Chen C-T, Bechtel T, Sánchez Guzmán V, Drozda AA, Chavan S, Weerapana E, Gubbels M-J. 2020. The apical annuli of *Toxoplasma gondii* are composed of coiled-coil and signalling proteins embedded in the inner membrane complex sutures. Cellular Microbiology 22:e13112.

Frénal K, Dubremetz J-F, Lebrun M, Soldati-Favre D. 2017. Gliding motility powers invasion and egress in Apicomplexa. Nat Rev Microbiol 15:645–660.

Frénal K, Jacot D, Hammoudi P-M, Graindorge A, Maco B, Soldati-Favre D. 2017. Myosin-dependent cell-cell communication controls synchronicity of division in acute and chronic stages of *Toxoplasma gondii*. Nat Commun 8:15710.

Frénal K, Polonais V, Marq J-B, Stratmann R, Limenitakis J, Soldati-Favre D. 2010. Functional dissection of the apicomplexan glideosome molecular architecture. Cell Host and Microbe 8:343–357.
Gould SB, Kraft LGK, van Dooren GG, Goodman CD, Ford KL, Cassin AM, Bacic A, McFadden GI, Waller RF. 2011. Ciliate pellicular proteome identifies novel protein families with characteristic repeat motifs that are common to alveolates. Mol Biol Evol 28:1319–1331.
Gould SB, Tham W-H, Cowman AF, McFadden GI, Waller RF. 2008. Alveolins, a new family of cortical proteins that define the protist infrakingdom Alveolata. Mol Biol Evol 25:1219–1230.
Graindorge A, Frénal K, Jacot D, Salamun J, Marq J-B, Soldati-Favre D. 2016. The conoid associated motor MyoH is indispensable for *Toxoplasma gondii* entry and exit from host cells. PLOS Pathogens 12:e1005388.
Guttery DS, Poulin B, Ferguson DJP, Szöőr B, Wickstead B, Carroll PL, Ramakrishnan C, Brady D, Patzewitz E-M, Straschil U, et al. 2012. A unique protein phosphatase with kelch-like domains (PPKL) in *Plasmodium* modulates ookinete differentiation, motility and invasion. PLOS Pathogens 8:e1002948.
Hanssen E, Dekiwadia C, Riglar DT, Rug M, Lemgruber L, Cowman AF, Cyrklaff M, Kudryashev M, Frischknecht F, Baum J, et al. 2013. Electron tomography of *Plasmodium falciparum* merozoites reveals core cellular events that underpin erythrocyte invasion. Cellular Microbiology 15:1457–1472.
Harding CR, Meissner M. 2014. The inner membrane complex through development of *Toxoplasma gondii* and Plasmodium. Cellular Microbiology 16:632–641.
Heaslip AT, Ems-Mcclung SC, Hu K. 2009. TgICMAP1 Is a novel microtubule binding protein in *Toxoplasma gondii*. PLOS ONE 4:e7406.
Heaslip AT, Nishi M, Stein B, Hu K. 2011. The motility of a human parasite, *Toxoplasma gondii*, is regulated by a novel lysine methyltransferase. PLOS Pathogens 7:e1002201.
Hu K, Johnson J, Florens L, Fraunholz M, Suravajjala S, DiLullo C, Yates J, Roos DS, Murray JM. 2006. Cytoskeletal components of an invasion machine--the apical complex of *Toxoplasma gondii*. PLOS Pathog 2:e13.
Hu K, Roos DS, Murray JM. 2002. A novel polymer of tubulin forms the conoid of *Toxoplasma gondii*. J Cell Biol 156:1039–1050.
Jacot D, Tosetti N, Pires I, Stock J, Graindorge A, Hung Y-F, Han H, Tewari R, Kursula I, Soldati-Favre D. 2016. An apicomplexan actin-binding protein serves as a connector and lipid sensor to coordinate motility and invasion. Cell Host and Microbe 20:731–743.
Katris NJ, van Dooren GG, McMillan PJ, Hanssen E, Tilley L, Waller RF. 2014. The apical complex provides a regulated gateway for secretion of invasion factors in *Toxoplasma*. PLOS Pathogens 10:e1004074.
Kats LM, Cooke BM, Coppel RL, Black CG. 2008. Protein trafficking to apical organelles of malaria parasites - building an invasion machine. Traffic 9:176–186.
Keller A, Nesvizhskii AI, Kolker E, Aebersold R. 2002. Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. Anal. Chem. 74:5383–5392.
Larsen J. 1988. An ultrastructural study of *Amphidinium poecilochroum* (Dinophyceae), a phagotrophic dinoflagellate feeding on small species of cryptophytes. Phycologia 27:366–377.
Lebrun M, Carruthers VB, Cesbron-Delauw M-F. 2014. *Toxoplasma* secretory proteins and their roles in cell invasion and intracellular survival. In: Weiss LM, Kim K, editors. *Toxoplasma gondii* The Model Apicomplexan. Second Edition. Elsevier. pp. 389–453.
Lentini G, Dubois DJ, Maco B, Soldati-Favre D, Frénal K. 2019. The roles of Centrin 2 and Dynein Light Chain 8a in apical secretory organelles discharge of *Toxoplasma gondii*. Traffic 20:583–600.
Leung JM, He Y, Zhang F, Hwang Y-C, Nagayasu E, Liu J, Murray JM, Hu K. 2017. Stability and function of a putative microtubule-organizing center in the human parasite *Toxoplasma gondii*. Mol Biol Cell 28:1361–1378.
Leung JM, Liu J, Wetzel LA, Hu K. 2019. Centrin2 from the human parasite *Toxoplasma gondii* is required for its invasion and intracellular replication. J Cell Sci 132:jcs228791.
Leung JM, Nagayasu E, Hwang Y-C, Liu J, Pierce PG, Phan IQ, Prentice RA, Murray JM, Hu K. 2020. A doublecortin-domain protein of Toxoplasma and its orthologues bind to and modify the structure and organization of tubulin polymers. BMC Mol Cell Biol 21:8.
Lévêque MF, Berry L, Besteiro S. 2016. An evolutionarily conserved SSNA1/DIP13 homologue is a component of both basal and apical complexes of *Toxoplasma gondii*. Sci Rep 6:27809.
Liu J, He Y, Benmerzouga I, Sullivan WJ, Morrissette NS, Murray JM, Hu K. 2016. An ensemble of specifically targeted proteins stabilizes cortical microtubules in the human parasite *Toxoplasma gondii*. Mol Biol Cell 27:549–571.

Liu J, Wetzel L, Zhang Y, Nagayasu E, Ems-McClung S, Florens L, Hu K. 2013. Novel thioredoxin-like proteins are components of a protein complex coating the cortical microtubules of *Toxoplasma gondii*. Euk Cell 12:1588–1599.
Long S, Anthony B, Drewry LL, Sibley LD. 2017. A conserved ankyrin repeat-containing protein regulates conoid stability, motility and cell invasion in *Toxoplasma gondii*. Nat Commun 8:2236.
Long S, Brown KM, Drewry LL, Anthony B, Phan IQH, Sibley LD. 2017. Calmodulin-like proteins localized to the conoid regulate motility and cell invasion by *Toxoplasma gondii*. PLOS Pathogens 13:e1006379.
Montoya JG, Liesenfeld O. 2004. Toxoplasmosis. Lancet 363:1965–1976.
Morrissette N, Gubbels M-J. 2014. The Toxoplasma Cytoskeleton: *Structures*, Proteins and Processes. In: *Toxoplasma gondii* The Model Apicomplexan. Second Edition. Elsevier. pp. 455–503.
Mylnikov AP. 2009. Ultrastructure and phylogeny of colpodellids (Colpodellida, Alveolata). Biology bulletin 36:582–590.
Nagayasu E, Hwang Y-C, Liu J, Murray JM, Hu K. 2017. Loss of a doublecortin (DCX)-domain protein causes structural defects in a tubulin-based organelle of *Toxoplasma gondii* and impairs host-cell invasion.Chang F, editor. Mol Biol Cell 28:411–428.
Nesvizhskii AI, Keller A, Kolker E, Aebersold R. 2003. A statistical model for identifying proteins by tandem mass spectrometry. Anal. Chem. 75:4646–4658.
Nichols BA, Chiappino ML. 1987. Cytoskeleton of *Toxoplasma gondii*. J. Protozool. 34:217–226.
Obornik M, Vancová M, Lai D-H, Janouškovec J, Keeling PJ, Lukes J. 2011. Morphology and ultrastructure of multiple life cycle stages of the photosynthetic relative of apicomplexa, *Chromera velia*. Protist 162:115–130.
Okamoto N, Keeling PJ. 2014. The 3D structure of the apical complex and association with the flagellar apparatus revealed by serial TEM tomography in *Psammosa pacifica*, a distant relative of the Apicomplexa. PLOS ONE 9:e84653.
Paredes-Santos TC, de Souza W, Attias M. 2012. Dynamics and 3D organization of secretory organelles of *Toxoplasma gondii*. J. Struct. Biol. 177:420–430.
Paterson WB, Desser SS, Barta JR. 1988. Ultrastructural Features of the Apical Complex, Pellicle, and Membranes Investing the Gamonts of Haemogregarina magna (Apicomplexa: Adeleina) 1. Journal of Eukaryotic Microbiology 35:73–80.
Perkins FO. 1976. Zoospores of the oyster pathogen, *Dermocystidium marinum*. I. Fine structure of the conoid and other sporozoan-like organelles. Journal of Parasitology 62:959–974.
Rashid M, Rashid MI, Akbar H, Ahmad L, Hassan MA, Ashraf K, Saeed K, Gharbi M. 2019. A systematic review on modelling approaches for economic losses studies caused by parasites and their associated diseases in cattle. Parasitology 146:129–141.
Roos DS, Donald R, Morrissette NS, Moulton ALC. 1994. Molecular tools for genetic dissection of the protozoan parasite *Toxoplasma gondii*. In: Methods in Cell Biology. Vol. 45. European Journal of Cell Biology. pp. 27–63.
Russell DG, Burns RG. 1984. The polar ring of coccidian sporozoites: a unique microtubule-organizing centre. J Cell Sci 65:193–207.
Scholtyseck E, Mehlhorn H, Friedhoff K. 1970. The fine structure of the conoid of Sporozoa and related organisms. Z. Parasitenk 34:68–94.
Schrével J, Valigurová A, Prensier G, Chambouvet A, Florent I, Guillou L. 2016. Ultrastructure of *Selenidium pendula*, the type species of archigregarines, and phylogenetic relations to other marine Apicomplexa. Protist 167:339–368.
Sheiner L, Demerly JL, Poulsen N, Beatty WL, Lucas O, Behnke MS, White MW, Striepen B. 2011. A systematic screen to discover and analyze apicoplast proteins identifies a conserved and essential protein import factor. PLOS Pathog 7:e1002392.
Sidik SM, Huet D, Ganesan SM, Huynh M-H, Wang T, Nasamu AS, Thiru P, Saeij JPJ, Carruthers VB, Niles JC, et al. 2016. A genome-wide CRISPR screen in *Toxoplasma* identifies essential apicomplexan genes. Cell 166:1423–1435.e12.
Skariah S, Bednarczyk RB, McIntyre MK, Taylor GA, Mordue DG. 2012. Discovery of a novel *Toxoplasma gondii* conoid-associated protein important for parasite resistance to reactive nitrogen intermediates. J. Immunol. 188:3404–3415.
Striepen B. 2013. Parasitic infections: Time to tackle cryptosporidiosis. Nature 503:189–191.
Suarez C, Lentini G, Ramaswamy R, Maynadier M, Aquilini E, Berry-Sterkers L, Cipriano M, Chen AL, Bradley P, Striepen B, et al. 2019. A lipid-binding protein mediates rhoptry discharge and invasion in *Plasmodium falciparum* and *Toxoplasma gondii* parasites. Nat Commun 10:4041.

Suvorova ES, Francia M, Striepen B, White MW. 2015. A novel bipartite centrosome coordinates the apicomplexan cell cycle. PLOS Biology 13:e1002093.
Tosetti N, Dos Santos Pacheco N, Bertiaux E, Maco B, Bournonville L, Hamel V, Guichard P, Soldati-Favre D. 2020. Essential function of the alveolin network in the subpellicular microtubules and conoid assembly in *Toxoplasma gondii*. eLife 9:e56635.
Tosetti N, Dos Santos Pacheco N, Soldati-Favre D, Jacot D. 2019. Three F-actin assembly centers regulate organelle inheritance, cell-cell communication and motility in *Toxoplasma gondii*. eLife 8:684.
Tran JQ, De Leon JC, Li C, Huynh M-H, Beatty W, Morrissette NS. 2010. RNG1 is a late marker of the apical polar ring in *Toxoplasma gondii*. Cytoskeleton 67:586–598.
Vlachou D, Schlegelmilch T, Runn E, Mendes A, Kafatos FC. 2006. The developmental migration of *Plasmodium* in mosquitoes. Curr. Opin. Genet. Dev. 16:384–391.
Wall RJ, Roques M, Katris NJ, Koreny L, Stanway RR, Brady D, Waller RF, Tewari R. 2016. SAS6-like protein in *Plasmodium* indicates that conoid-associated apical complex proteins persist in invasive stages within the mosquito vector. Sci Rep 6:28604.
Wall RJ, Zeeshan M, Katris NJ, Limenitakis R, Rea E, Stock J, Brady D, Waller RF, Holder AA, Tewari R. 2019. Systematic analysis of *Plasmodium* myosins reveals differential expression, localisation, and function in invasive and proliferative parasite stages. Cellular Microbiology:e13082.
World Health Organization. 2018. World malaria report 2018.