



Published in final edited form as:

Mol Microbiol. 2017 October ; 106(1): 1–21. doi:10.1111/mmi.13715.

The Serine/Threonine Phosphatases of Apicomplexan Parasites

Chunlin Yang¹ and Gustavo Arrizabalaga^{1,2,*}

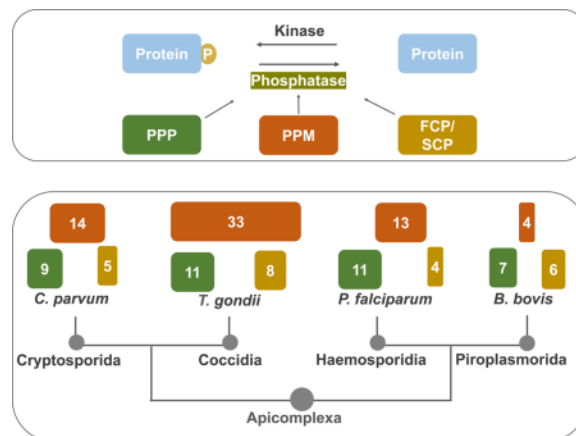
¹Department of Pharmacology and Toxicology, Indiana University School of Medicine, Indianapolis, IN, USA

²Department of Microbiology and Immunology, Indiana University School of Medicine, Indianapolis, IN, USA

SUMMARY

The balance between phosphorylation and de-phosphorylation, which is delicately regulated by protein kinases and phosphatases, is critical for nearly all biological processes. The Apicomplexa are a large phylum which contains various parasitic protists, including human pathogens, such as *Plasmodium*, *Toxoplasma*, *Cryptosporidium* and *Babesia* species. The diverse life cycles of these parasites are highly complex and, not surprisingly, many of their key steps are exquisitely regulated by phosphorylation. Interestingly, many of the kinases and phosphatases, as well as the substrates involved in these events are unique to the parasites and therefore phosphorylation constitutes a viable target for antiparasitic intervention. Most progress on this realm has come from studies in *Toxoplasma* and *Plasmodium* of their respective kinomes and phosphoproteomes. Nonetheless, given their likely importance, phosphatases have recently become the focus of research within the apicomplexan parasites. In this review, we concentrate on serine/threonine phosphatases in apicomplexan parasites, with the focus on comprehensively identifying and naming protein phosphatases in available apicomplexan genomes, and summarizing the progress of their functional analyses in recent years.

Graphical abstract



*Corresponding author: garrizab@iu.edu.

Introduction

Apicomplexa are a phylum of the Alveolata, a superphylum of eukaryotic protists that harbor a set of sacs, called alveoli, underneath their plasma membrane (Cavalier-Smith, 1993). Apicomplexans have a unique apical complex structure made up of a system of spirally arranged tubulin filaments called the conoid, which make them distinct from the other alveolates. The known apicomplexans are mainly classified into five taxa, including Haemosporidia, Piroplasmorida, Coccidia, Cryptosporida and Gregarinasina, but only the first three are supported by phylogeny (Hinchliff *et al.*, 2015). Except for Gregarinasina, each of the apicomplexan taxa includes parasites that infect vertebrates, many of which can be pathogenic in humans. For example, the coccidian *Toxoplasma gondii*, can cause severe disease and even death in the immunocompromised or immunosuppressed, and in the congenitally infected; the piroplasmid *Babesia spp.*, which is tick-borne, is responsible for the hemolytic disease babesiosis; *Cryptosporidium spp.* can cause cryptosporidiosis, a gastrointestinal illness responsible for severe diarrhea in the immunocompromised and in children; the representative blood-infecting haemosporidians, *Plasmodium spp.*, cause malaria, perhaps the most serious of parasitic diseases in the world, resulting in hundreds of thousands deaths annually. However, due to the resistance, toxicity and inefficacy of current anti-apicomplexan agents, more efficient drugs are greatly needed (Antony & Parija, 2016). In pursuit of potential anti-apicomplexan targets, the top priority should be given to those proteins that are unique and essential to apicomplexan life cycles. In fact, apicomplexans possess a set of plant-like genes, which could be descended from their photosynthetic common ancestor (Janouskovec *et al.*, 2010). Given no counterparts are present in mammals, the proteins encoded by these plant-like genes are considered as potential anti-apicomplexan targets, and have been aggressively studied, especially for those plant-like protein kinases and phosphatases.

Protein kinases and phosphatases oppositely regulate the most common post-translational modification, phosphorylation. A recent phosphoproteomics study estimated that nearly half of the total proteins from any of human, mouse or yeast are phosphoproteins (Vlastaridis *et al.*, 2017). Therefore, it is reasonable to estimate that a comparable ratio of apicomplexan proteins is under the regulation of phosphorylation. In fact, studies of the last few decades have shown that phosphorylation plays essential roles in every aspect of apicomplexan parasites, such as including the parasite propagation, conversion and pathogenesis. The propagation of apicomplexan parasites starts from invasion of host cells and ends with egress when exhausting the nutrient. Within host cells, the parasites reside in parasitophorous vacuoles (PVs), which the parasites formed upon invasion, and replicate exponentially. Distinct from their host cells, apicomplexans adopt special division manner for their asexual replication, such as endodyogeny for *T. gondii* and schizogony for *P. falciparum*. Accordingly, it is reasonable to infer that the cell cycle regulation of apicomplexan parasites is somewhat different from their host cells. As we know, cell cycle in higher eukaryotic cells is exquisitely regulated by two key classes of regulatory molecules, cyclins and cyclin-dependent kinases (CDKs). Without exception, all apicomplexan genomes encode a variety of cyclins and CDKs, and studies has shown that they are critical cell cycle regulators (Gubbels *et al.*, 2008, Francia & Striepen, 2014, Khan *et al.*, 2002). Moreover,

apicomplexans such as *T. gondii* contain a number of atypical cyclins and CDK-related kinases (CRKs), which are essential for the regulation of endodyogeny (Alvarez & Suvorova, 2017). Therefore, the apicomplexan cell cycle is regulated by phosphorylation events, some of which are unique to Apicomplexa.

Events that occur during apicomplexan propagation such as parasite invasion and egress are mostly unique and essential to survival and pathogenesis, and, accordingly, have received much attention. Host cell invasion and egress of apicomplexans are largely regulated by the secretion protein from the apicomplexan specialized organelles, the micronemes, rhoptries, and dense granules (Lebrun *et al.*, 2014). Studies of the last few decades have identified phosphorylation, especially calcium dependent phosphorylation driven principally by the unique family of calcium dependent protein kinases (CDPKs), as playing essential roles in regulating the protein secretion from these unique organelles. For example, the CDPK1 of *T. gondii* has been shown to be part of a signaling pathway that leads to secretion of microneme proteins needed for *T. gondii* motility, invasion and egress (Lourido *et al.*, 2010). *P. falciparum*'s CDPK1, which is localized to the periphery of the parasite (Zhao *et al.*, 1994, Green *et al.*, 2008, Moskes *et al.*, 2004) has also been shown to play key roles in motility (Kato *et al.*, 2008), secretion (Bansal *et al.*, 2013), and development (Azevedo *et al.*, 2013) during the blood stages of the parasite. Its closest *T. gondii* homolog, TgCDPK3, is critical for rapid exit from the cell (McCoy *et al.*, 2012, Garrison *et al.*, 2012, Lourido *et al.*, 2012), a process which this kinase appears to regulate through the phosphorylation of the motor protein myosin A (MyoA) (Gaji *et al.*, 2015). MyoA is one of the components of the unique molecular machine termed the glideosome, an actin-myosin motor complex which the apicomplexans adopt for gliding motility, host cell invasion and egress (Opitz & Soldati, 2002, Keeley & Soldati, 2004). The glideosome anchors to a structure composed of alveoli known as the inner membrane complex (IMC), and lies in the space between the IMC and the plasma membrane (Boucher & Bosch, 2015). Besides MyoA, the glideosome includes short actin filaments, actin accessory proteins, myosin light chain (MLC) proteins and the glideosome associated proteins (GAPs) (Boucher & Bosch, 2015). The phosphorylation status of the glideosome components is known to regulate their activation. For example, the phosphorylation of the *P. falciparum* MyoA tail interacting protein (MTIP) is essential for its interaction with MyoA (Green *et al.*, 2008), and the phosphorylation of *T. gondii* GAP45 is important for its anchoring to the IMC (Gilk *et al.*, 2009, Ridzuan *et al.*, 2012).

Another biological context in which phosphorylation plays an important role in the apicomplexans is conversion between the different stages of their life cycles. Apicomplexans have complex life cycles that can include both insect and mammalian hosts (e.g. *Plasmodium* and *Babesia spp.*), latent tissue cysts stages (e.g. *T. gondii*) and environmentally exposed stages (e.g. *T. gondii* and *Cryptosporidium spp.*). Many studies have shown the key role of phosphorylation in different stage conversions. For example, phosphorylation of the eukaryotic initiation factor-2 α (eIF2 α) in *P. falciparum* sporozoites within the salivary gland of the mosquito vector represses translation of key transcripts that become expressed upon dephosphorylation of eIF2 α once parasites are injected into human host (Zhang *et al.*, 2010, Zhang *et al.*, 2016). Similarly, in *T. gondii*, eIF2 α phosphorylation and the ensuing translational repression accompany the transition for the rapidly replicating tachyzoite stage to the latent encysted bradyzoite form (Narasimhan *et al.*, 2008). The importance of this

phosphorylation event for the establishment and maintenance of the latent stage is evidenced by the fact that two inhibitors of eIF2 α dephosphorylation, salubrinal and guanabenz induce bradyzoite formation and inhibit reconversion of encysted parasites to tachyzoites (Konrad *et al.*, 2013, Benmerzouga *et al.*, 2015). Cyclic nucleotide dependent phosphorylation also appears to play a role in *T. gondii* stage conversion as inhibition of cAMP and cGMP kinases induces bradyzoite formation (Eaton *et al.*, 2006). Pathogenesis of apicomplexan parasites are also under the control of phosphorylation. Many apicomplexan virulence factors are rhoptry organelle proteins (ROPs) with protein kinase activities, which are secreted directly into host cells and interfere the innate immunity of the host. Studies have shown that ROP17 and 18 are the two protein kinases in *T. gondii* that phosphorylate and inactivate immunity related GTPases (IRGs), which are responsible for early immune response (Fentress *et al.*, 2010, Steinfeldt *et al.*, 2010, Etheridge *et al.*, 2014).

Through its involvement in cell division, propagation, stage conversion, and pathogenesis, phosphorylation is central to the life cycles of apicomplexan parasites. Compared with the extensive studies of protein kinases, protein phosphatases have received relatively less attention. However, many phosphorylation events key to propagation, conversion and pathogenesis occur in response to specific cues or at specific timing, and thus are transitory. Accordingly, phosphatases and their regulatory proteins are expected to play important roles in the life cycle of apicomplexans, and have gradually become the focus of functional studies in recent years. In this study, we comprehensively compare the protein serine/threonine phosphatases in the representative apicomplexans, *Cryptosporidium parvum*, *Plasmodium falciparum*, *T. gondii*, and *Babesia bovis*. These chosen species are representatives of each the four taxa that infect vertebrates (Cryptosporida, Coccidia, Haemosporidia and Piroplasmorida) and are all pathogenic in humans. Moreover, these four apicomplexan parasites are among the most studied at the molecular and cellular level and their genomes have been sequenced and annotated. Besides listing, classifying and naming the protein phosphatases present in these genomes, we summarize recent progress in the functional analysis of phosphatases in apicomplexans.

The phosphatases of eukaryotic cells

Reversible phosphorylation is an extensively used and highly ubiquitous regulatory mechanism. Protein kinases and phosphatases respectively take the responsibility to add and remove phosphate to/from amino acids, commonly the three hydroxyl-containing amino acids, serine, threonine and tyrosine, in eukaryotes and histidine in prokaryotes, as well as some others including arginine, lysine, and cysteine (Ciesla *et al.*, 2011). The presence of phosphate groups on any of these amino acids can alter the structure, stability, interactions, and function of proteins.

Protein phosphatases are typically subdivided into three major categories based on the amino acid targeted: serine/threonine phosphatases, tyrosine phosphatases, and dual specificity phosphatases, which act on all three residues. Based on a proteomic analysis of more than 6000 phosphorylation sites on more than 2,000 phosphoproteins the relative abundance of phosphoserine, phosphothreonine and phosphotyrosine in human cells was determined to be 86.4%, 11.8% and 1.8% respectively (Olsen *et al.*, 2006). Protein tyrosine kinases (PTKs)

and phosphatases (PTPs) usually have comparable numbers in each eukaryotic genome, but most organisms commonly have more protein serine/threonine kinases (PSKs) than the corresponding phosphatases. The discrepancy between the numbers of PSKs and protein serine/threonine phosphatases (PSPs) can be explained by the fact that a single catalytic subunit of PSPs can act on multiple substrates, with specificity conferred by interactions with a large number of regulatory and structural subunits.

PSPs, which are the focus of this review, are divided into three major families: phosphoprotein phosphatases (PPPs), protein phosphatases Mg^{2+}/Mn^{2+} dependent (PPMs), and aspartate based phosphatases (Shi, 2009). The PPP family is represented by protein phosphatase 1 (PP1), PP2A, PP2B (a.k.a. calcineurin), PP4, PP5, PP6, and PP7. For many members of the PPP family, the catalytic subunit interacts with a large array of regulatory subunits. By contrast, members of the PPM family, which includes the Mg^{2+}/Mn^{2+} dependent PP2C and pyruvate dehydrogenase phosphatase (PDP), do not possess regulatory subunits but instead have additional domains and conserved motifs that impart substrate specificity. Much less is known about aspartate based phosphatases, which are represented by TFIIIF-associating carboxyl-terminal domain (CTD) phosphatases (FCP) and small CTD phosphatase (SCP), accordingly known as FCP/SCP family phosphatases. In mammals and yeast FCP1 dephosphorylates the CTD of RNA polymerase II (Eick & Geyer, 2013), and SCP1 dephosphorylates the linker regions of Smad 1, 2 and 3 in human (Sapkota *et al.*, 2006).

The apicomplexan phosphatases

While many apicomplexan genomes have been sequenced and annotated (available at EuPathDB.org), the designation of proteins as phosphatases is likely imperfect and in many instances, does not account for classification. Protein phosphatases have been comprehensively annotated in *P. falciparum* based on three independent previous studies (Wilkes & Doerig, 2008, Pandey *et al.*, 2014, Guttery *et al.*, 2014). Accordingly, we performed a bioinformatics analysis to comprehensively identify putative protein phosphatases in the other three apicomplexan genomes: *T. gondii*, *C. parvum* and *B. bovis*, and investigated their identities and homolog relationships. Specifically, the PSPs of *P. falciparum* annotated in EuPathDB, and the listed PSPs of the genomes of human, budding yeast and *Arabidopsis* listed in the database EKP1.1 (Wang *et al.*, 2014) were used as queries in BLASTP searches to identify homologs in the other three apicomplexan genomes. The sequences of the identified putative phosphatases were used as queries in reciprocal BLASTP searches using the most recent genome databases (EuPathDB release 30) to confirm identification and establish most similar gene pairs between genomes. This was followed by protein phosphatase domain prediction with SMART and PFAM online service for further validation. Importantly, when we perform PFAM as the primary search method we obtain the same list of phosphatases. The number PSPs present in all the four apicomplexan genomes studied are summarized in Table 1. Below, within sections for the various protein phosphatase categories, we discuss the specific findings from the analysis of the four pathogenic apicomplexans that are the focus of our review, as well as current knowledge of the function of particular phosphatases.

Serine/Threonine Phosphatases: The PPP Family

As mentioned above, most eukaryotes have seven subfamilies of PPPs (PP1, PP2A, PP2B, PP4, PP5, PP6, and PP7). Of the four apicomplexan parasites we analyzed, both *T. gondii* and *P. falciparum* contain a whole set of the above-mentioned PPP subfamilies, with one member in each subfamily, except for PP2A of which there are two in *T. gondii* (Table 1). Interestingly, the other two genomes both lack members of two subfamilies: *C. parvum* lacks PP6 and PP7, and *B. bovis* lacks PP2B and PP6 (Table 1). Nonetheless, another *Babesia* species *B. microti* appears to encode a putative PP2B homolog. The lack of PP2B and PP6 is also seen in the *Theileria*, which is one of the closest related genera to *B. bovis*. This would suggest that gene loss is the likely cause for lack of these two PP1 family members in these parasites rather than incomplete genomic sequencing.

Unsurprisingly, the PPP family members of the apicomplexans are extraordinarily conserved, sharing the consensus core motifs that are conserved throughout eukaryotes, including “GDxHG”, “GDxVDRG”, “RG”, “GNHE”, “HGG” and “H” (x represents any amino acid) (Fig. 1). The subfamilies are clustered into monophyletic branches with high support values (Fig. 1). Besides having homologs of phosphatases from each of the seven subfamily members that are common across eukaryotic phyla, the apicomplexans analyzed possess members of the Kelch-like domain containing protein phosphatase (PKKL) subfamily, present only in plants and other alveolates. In addition, all the genomes except that of *B. bovis* possess a bacterial-like phosphatase subfamily that is known in the apicomplexans as the Shewanella-like phosphatases (SLP). Moreover, an EF-hand motif containing phosphatase subfamily, previously named EFPP (Kutuzov & Andreeva, 2008), is present in all the four apicomplexan genomes.

PP1

PP1 was the first identified phosphatase, characterized initially as the enzyme that dephosphorylates glycogen phosphorylase (Cori & Cori, 1945). PP1 is now known to be involved in a wide range of biological processes, such as glycogen metabolism, protein synthesis, transcription, cellular division and apoptosis (Ceulemans & Bollen, 2004). A holoenzyme of PP1 contains both a catalytic and a regulatory subunit. The catalytic subunit of PP1 is highly conserved among all eukaryotes and contains three histidines, two aspartic acids, and one asparagine that coordinate two metal ions, Mn^{2+} and Fe^{2+} (Shi, 2009). These two ions are hypothesized to interact with a water molecule to initiate a nucleophilic attack on the phosphate (Goldberg *et al.*, 1995). Although humans have only three PP1 catalytic subunits, which are nearly identical to each other in sequence, it has been estimated that around one third of protein dephosphorylation events are conducted by PP1 (Rebello *et al.*, 2015). Thus, substrate specificity is likely conferred by the about 200 proteins identified as potential PP1 regulatory subunits (Rebello *et al.*, 2015). Each of the four apicomplexans encodes for one PP1 catalytic subunit with comparable protein sequence length and high similarity to human homologs. These four PP1s are CpPP1 (cgd7_2670), TgPP1 (TGME49_310700), PfPP1 (PF3D7_1414400) and BbPP1 (BBOV_III006130), respectively.

In *T. gondii*, the presence of PP1 activity was first deduced from experiments showing that exposure to PP1 inhibitors significantly impaired parasite invasiveness, suggesting a role for

this phosphatase in host cell invasion (Delorme *et al.*, 2002). Genome sequencing allowed the identification of the *T. gondii* PP1 homolog, and transient transfection of its coding sequence fused to a cMyc epitope tag showed that is present in both the cytoplasm and nucleus (Daher *et al.*, 2007). The PP1 homolog from *P. falciparum*, PfPP1, was determined through subcellular fractionation to also have nucleocytoplasmic localization (Daher *et al.*, 2006). Moreover, when expressed in *Saccharomyces cerevisiae* lacking PP1, PfPP1 can complement the low glycogen phenotype observed with the lack of yeast PP1 (Bhattacharyya *et al.*, 2002). Abrogation of PfPP1 with short interfering RNA (siRNA) led to inhibition of DNA synthesis and cell cycle progression (Kumar *et al.*, 2002). Interestingly, PfPP1 has also been detected in the Maurer's clefts, membranous structures in the cytoplasm of erythrocytes infected with *P. falciparum* that are associated with parasite protein sorting and export. Within the Maurer's clefts PfPP1 appears to modulate the phosphorylation of the N-terminal domain of Skeleton Binding Protein 1 (PfSBP1), a Maurer's cleft transmembrane protein, whose C-terminal interacts with host cell membrane. Inhibition of PP1 with calyculin A leads to the hyperphosphorylation of PfSBP1, which impacts the release of parasites from the host cells (Blisnick *et al.*, 2006). Thus, as in other systems, apicomplexan PP1s are involved in many types of processes including glycogen metabolism and protein sorting and export and are likely to have a diversity of substrates.

Given the functional diversity of apicomplexan PP1s and the high conservation of PP1 across eukaryotic taxa, it is not surprising that many of the PP1 regulators found in mammals are also present in apicomplexans. Two independent studies reported that the leucine rich repeat protein 1 (LRR1) in both *T. gondii* and *P. falciparum*, similarly to their homologs LRR1 in mammals and Sds22 in yeast, function as negative regulators of PP1 (Daher *et al.*, 2006, Daher *et al.*, 2007). The direct interaction between PfPP1 and PfLRR1 has been confirmed through several interaction-trap and biochemical approaches (Daher *et al.*, 2006). The same study also confirmed that PfLRR1 indeed functions as an inhibitor of PfPP1 as shown through inhibition of the phosphatase activity of recombinant PfPP1 upon pre-incubation with recombinant PfLRR1 (Daher *et al.*, 2006). In *T. gondii* the physical interaction between TgPP1 and TgLRR1, and the function of TgLRR1 as a TgPP1 inhibitor have also been confirmed (Daher *et al.*, 2007).

PP1 specific inhibitory proteins I-2 and I-3 are both negative regulators of PP1 in human cells, and homologs of both proteins are present in all the four representative apicomplexans (Table 2). Studies with *P. falciparum* Pfi2 showed that the phosphatase activity of PfPP1 was severely decreased when it was preincubated with wild-type recombinant Pfi2, but not with mutant versions of the inhibitor lacking either of the two main PP1 interacting motifs (Freville *et al.*, 2013, Freville *et al.*, 2014). Surprisingly, by contrast to other eukaryotes, Pfi3 is a PP1 activator in *P. falciparum* and it has an essential role in the growth and survival of blood stage parasites (Freville *et al.*, 2012). A more recent study combined co-affinity purification followed by mass spectrometry (MS), yeast two hybrid (Y2H) screening and in silico analysis for *P. falciparum* proteins with RVxF motif, which is present in most PP1 regulators, to define the PfPP1 interactome (Hollin *et al.*, 2016). This effort identified 186 potential PP1 interactors in *P. falciparum*, of which 35 have been validated as PfPP1 interactors in an ELISA based assay (Hollin *et al.*, 2016). However, further experimentation is needed to confirm their roles as PfPP1 regulators.

PP2A, PP4 and PP6

The catalytic subunits of PP2A, PP4 and PP6 cluster as a monophyletic branch, suggesting that they are derived from a common ancestor (Uhrig *et al.*, 2013). Not only do these three enzymes share sequence similarity, they also have common functional mechanisms by forming heterodimeric and heterotrimeric complexes that combine catalytic, structural and regulatory subunits.

PP2A—Like PP1, PP2A is also involved in numerous biological processes, such as cell proliferation, cell death, cell motility and cell cycle control (Lillo *et al.*, 2014). The human genome contains two catalytic (PPP2CA and B), two structural (PPP2R1A and B) and 15 regulatory subunits. The human PP2A core enzyme is a heterodimeric complex formed by the combination of a catalytic and a structural subunit, and the PP2A holoenzyme is a heterotrimeric complex with the addition of a regulatory subunit to the PP2A core dimeric complex. The structural subunit, which acts as a scaffold, contains 15 tandem HEAT repeats that form a horseshoe-shaped structure (Groves *et al.*, 1999). The 15 PP2A regulatory subunits are subdivided into four families: B, B', B'' and B''', which share high sequence similarities within families, but are distinct across families (Table 2). The B subfamily is also known as B55 or PR55 and is comprised of PPP2R2A, B, C and D; the B' subfamily, also known as B56 or PR61, includes PPP2R4, PPP2R5A, B, C, D and E; the B'' subfamily contains three members, PPP2R3A (also known as PR48), PPP2R3B (PR72/PR130) and PPP2R3C (G5PR); the B''' subfamily includes STRN (PR110), STRN3 (PR93) and STRN4 (Table 2).

Like in the case of human cells, *C. parvum* and *T. gondii* both have two genes encoding catalytic subunits referred to as PP2A1 (cgd7_810, TGME49_224220) and PP2A2 (cgd5_4070, TGME49_215170), but *P. falciparum* and *B. bovis* both have only one each (PF3D7_0925400, BBOV_I000740) (Table 2). Detailed phylogenetic analysis of all PP2As from the four apicomplexans showed that PfPP2A shows higher similarity to CpPP2A2 and TgPP2A2, while BbPP2A shows higher similarity to CpPP2A1 and TgPP2A1 (Fig. 1), suggesting that two copies of PP2A were present in ancient apicomplexan ancestors, and that either one of them was lost in some taxa. This scenario is supported by the fact that CpPP2A1, TgPP2A1 and BbPP2A all have the motif “HAG” instead of the consensus “HGG”, which is found in PfPP2A, CpPP2A2 and TgPP2A2 (Fig. 1). *T. gondii*, *P. falciparum*, and *B. bovis*, but not *C. parvum*, have one putative structural subunit (TGME49_315670, PF3D7_1319700 and BBOV_II001110), and all four have a variable number of regulatory factors all belonging to the B' and B'' subfamilies, four in *C. parvum* (cgd7_4970, cgd6_4380, cgd1_2470 and cgd4_290), three in *T. gondii* (TGME49_200400, TGME49_283720 and TGME49_246510), and two in both *P. falciparum* (PF3D7_1356400 and PF3D7_1430100) and *B. bovis* (BBOV_IV010340 and BBOV_IV000610) (Table 2). The apparent lack of a structural subunit in *C. parvum* is possibly due to incomplete sequencing because its close relative *Cryptosporidium muris* contains one (CMU_017370) and the corresponding region coding this gene is absent in the sequenced genome of *C. parvum*. Hence, theoretically *C. parvum* could have eight possible combinations of PP2A holoenzymes if the structural subunit is present, six for *T. gondii*, and two for both *P. falciparum* and *B. bovis*.

Within the apicomplexans, PP2A has only been studied in *Plasmodium spp.* The first PfPP2A gene was isolated with multiple degenerate PCR primers designed based on the conserved amino acids among PP1, PP2A and PP2B phosphatases (Li & Baker, 1997). Investigation of the expression pattern of the PfPP2A transcript showed that PfPP2A is specifically expressed in gametocytes (Li & Baker, 1997). Biochemical characterization of PfPP2A was performed in a later study, which showed that it required Mn^{2+} for activity and was sensitive to nanomolar concentrations of okadaic acid (OA) (Dobson *et al.*, 1999). A member of SET/TAF-family, PfARP (asparagine (N)-rich protein, PF3D7_0909300) has been shown to be a specific inhibitor of PfPP2A with an IC_{50} of approximately 8 nM (Dobson *et al.*, 2003). The interaction between PP2A and Phosphotyrosyl Phosphatase Activator (PTPA), an activator of PP1A, has been reported to be conserved in *P. falciparum*, as evidenced by co-immunoprecipitation experiments performed in *Xenopus* oocytes following injection of recombinant His-tagged PfPTPA and capped mRNA encoding Myc-tagged PfPP2A (Vandomme *et al.*, 2014). Although no study has been performed in the other three species, PTPA is conserved in these genomes (cgd1_2470, TGME49_283720 and BBOV_IV010340), and the *T. gondii* PTPA appears to be important for propagation based on a low relative fitness score determined through a genome-wide CRISPR knock-out of all *T. gondii* genes (Sidik *et al.*, 2016). PP2A is also regulated by reversible methylation in mammals and it involves two proteins: leucine carboxyl methyltransferase (LCMT1) and PP2A methyltransferase (PPME1) (Shi, 2009). Homologs of LCMT1 are present in all the apicomplexan genomes discussed here but *C. parvum* (TGME49_237570, PF3D7_1439700 and BBOV_III003160). In contrast, PPME1 is only present in *T. gondii* (TGME49_262140), suggesting that the methylation-based regulation of PP2A might only be retained in *T. gondii*.

PP4—PP4 plays roles in many processes such as protein complex assembly (e.g. centrosome maturation and spliceosome assembly) and cell signaling (e.g. *NF- κ B* pathway, *Jnk* pathway and apoptotic signaling) (Cohen *et al.*, 2005). Like PP2A, the catalytic subunit of PP4 can also form dimers and trimers with regulatory subunits. The human genome contains one catalytic subunit gene (PPP4C) and five regulatory subunit genes (PPP4R1, 2, 3A, 3B and 4). PPP4R1 and PPP4R4 both contain HEAT repeats similar to those found in the structural subunits of PP2A. Both PPP4R1 and PPP4R4 can form heterodimeric complexes with the catalytic subunit. PPP4R2 and either of the PPP4R3 isoforms interact with the catalytic subunit to form the heterotrimeric complex PP4-PP4R2-PP4R3, which is conserved between humans and budding yeast (PPH3-PSY4-PSY2), and functions in DNA damage responses in both organisms (Lillo *et al.*, 2014). All the four apicomplexans described here contain one catalytic subunit coding gene in their genomes (Table 2). No homologs of PPP4R2 and PPP4R4 have been identified in the apicomplexan genomes analyzed, but *C. parvum*, *T. gondii* and *P. falciparum* all seem to have a potential homolog of PPP4R1 (cgd6_5000, TGME49_220910 and PF3D7_1034500), and *C. parvum*, *T. gondii* and *B. bovis* seem to have one potential homolog of PPP4R3 each (cgd3_240, TGME49_265440, and BBOV_II006300) (Table 2). As they are relatively less similar to their human homologs than the proposed regulators of PP2A, their identities as PP4 regulators need closer inspection for confirmation. Unfortunately, no PP4 functional studies have been conducted in apicomplexans to date, although it appears to be essential in *T.*

gondii based on its relative fitness score in the genome wide CRISPR screen described above (Sidik *et al.*, 2016).

PP6—This multifunctional phosphatase has roles in DNA damage responses, initiating repair of double stranded DNA breaks and ER-to-Golgi traffic (Lillo *et al.*, 2014). The human genome encodes for one catalytic subunit of PP6 and three regulatory subunits, which contain Sit4-associated protein (SAP) domains. Another three proteins with ankyrin repeat domains that have been reported to serve as structural subunits (Stefansson *et al.*, 2008). It is thought that the catalytic, regulatory and structural subunits form heterotrimeric holoenzymes in a manner analogous to PP2A (Lillo *et al.*, 2014). No PP6 catalytic subunit coding genes have been identified in *C. parvum* and *B. bovis*. Nonetheless, one copy is present in both *T. gondii* (TGME49_301010) and *P. falciparum* (PF3D7_0314400) (Table 2). Additionally, *T. gondii* and *P. falciparum* each have a potential regulatory and a structural subunit (Table 2). As is the case of PP4, no functional studies have been conducted to elucidate the roles of PP6 in apicomplexans.

PP2B, PP5 and PP7

Another monophyletic group within the PPP family consists of PP2B, PP5 and PP7, indicating a common ancestor. Unlike most PPP family phosphatases, which function as multimers of catalytic and regulatory subunits, both PP5 and PP7 are monomeric enzymes without any regulatory subunits. Instead of using regulators for the binding to target substrates, PP5 and PP7 have evolved with extra domains, which helps convey substrate specificity. By contrast PP2B, which is commonly referred to as calcineurin, is made up of a catalytic and a regulatory subunit.

PP2B—Also known as calcineurin or protein phosphatase 3 (PP3), PP2B is a calcium-dependent serine-threonine phosphatase, and participates in a diversity of calcium-dependent biological processes, including Ca²⁺-dependent T-cell signaling (Musson & Smit, 2011). The holoenzyme of PP2B consists of a catalytic subunit (CnA) and a regulatory subunit (CnB), which binds calcium through a number of EF hand domains. CnA contains four functional domains: a phosphatase domain in its N-terminal region, a helical CnB-binding domain, a Ca²⁺-calmodulin-binding domain, and an autoinhibitory domain in its C-terminus (Fig. 1) (Shi, 2009). The CnB-binding domain forms a helix which is physically distant from the catalytic domain, facilitating the binding with CnB (Shi, 2009). Binding of CnA to Ca²⁺-calmodulin causes a conformational change that displaces the autoinhibitory domain from the active site resulting in activation of the phosphatase. The autoinhibitory domain interacts with the catalytic domain when the phosphatase is inactive, blocking the interaction between the catalytic domain with substrates (Shi, 2009). The human genome encodes for three PP2B catalytic subunits and two regulatory subunits. Of the apicomplexan genomes studied, all except *B. bovis* contain one copy of the catalytic subunit and one of the regulatory subunit (Table 2). Nonetheless, another *Babesia* species *B. microti* contains one potential PP2B gene in its genome (BmR1_III07610), although is not present in all other Piroplasmid genomes.

PP2B first garnered attention in apicomplexans because of their sensitivity to Cyclosporin A (CsA) and FK506 (Dobson *et al.*, 1999), two immunosuppressants known to act through

calcineurin inhibition in mammalian cells (Jun Liu *et al.*, 1991). Both FK506 and CsA inhibit the secretion of microneme proteins in *P. falciparum*, suggesting a relationship between calcineurin with the effector(s) of Ca²⁺-dependent microneme exocytosis (Singh *et al.*, 2014). Similarly, CsA was shown to block *T. gondii* egress, which is also dependent on calcium dependent secretion and motility (Moudy *et al.*, 2001). Nonetheless, more recent studies using genetic manipulation have refined our understanding of calcineurin in these parasites. Calcineurin in both parasites is predominantly cytoplasmic in the schizont stage (Kumar *et al.*, 2005, Paul *et al.*, 2015), but both TgCnA and TgCnB become enriched at the apical end when the parasites are released into extracellular medium (Paul *et al.*, 2015). Direct knockout of either subunit of calcineurin in both *P. falciparum* and *T. gondii* proved unsuccessful, suggesting that both proteins are essential within their respective parasites (Paul *et al.*, 2015). Nonetheless, conditional knockdown of either PfCnA and TgCnB was accomplished by adding a destabilizing domain (DD) to the endogenous genes, which leads to degradation of the protein unless the cell permeable small molecule Shield-1 (Shld1) is present. Removal of Shld1 in both the PfCnA-DD and PfCnB-DD expressing strains led to strong reduction in parasite proliferation (Paul *et al.*, 2015). Detailed analysis revealed that degradation of calcineurin did not affect *P. falciparum* egress or intra-erythrocytic development but significantly disrupted re-invasion, due to a defect in attachment (Paul *et al.*, 2015). Similar results were observed in *T. gondii* when expression of TgCnA was knocked down using a tetracycline repressible promoter system. Interestingly, knockdown of calcineurin in *T. gondii* did not impact egress, suggesting that the previous blockage of egress by CsA is likely due to off target effects. In both *T. gondii* and *P. falciparum*, attachment and the subsequent invasion event are primarily dependent on proteins that are secreted from the micronemes upon calcium fluxes (Lourido *et al.*, 2012). Interestingly, while genetic disruption of calcineurin in both parasites significantly impairs attachment to the host cell, neither microneme nor rhoptry secretion is affected. Thus, calcineurin appears to regulate microneme independent attachment of both extracellular *P. falciparum* and *T. gondii* parasites to host cells before intracellular entry (Paul *et al.*, 2015). Identification of the substrates of calcineurin in either of these parasites would shed light into this poorly understood form of attachment.

PP5—Unlike most PPP family members, which contain separate catalytic and regulatory subunits, PP5 is a monomeric phosphatase with both the catalytic and regulatory domains in a single polypeptide. PP5 is involved in cellular proliferation, differentiation, apoptosis, survival and DNA damage repair (Hinds & Sanchez, 2008). The presence of N-terminal tetratricopeptide repeat (TPR) domains, which usually serve as protein-protein interaction motifs, is one of the unique characteristics of PP5. Like in other eukaryotes, all the four apicomplexans analyzed have a single PP5, all with three TPR domains (Fig. 1). Identification of the PP5 gene in *P. falciparum* was based on PCR reactions with degenerate primers designed on the basis of conserved peptide sequences of PPP family phosphatases (Dobson *et al.*, 2001b, Lindenthal & Klinkert, 2002). Stage-specific expression analysis at the transcript level showed that PfPP5 is present in the entire asexual erythrocytic life cycle and gametocytes of *P. falciparum*, and immunofluorescence assays using polyclonal antiserum showed that PfPP5 is concentrated in the nucleus but also present in the cytoplasm (Lindenthal & Klinkert, 2002). Heat shock protein 90 (hsp90) is a binding partner of PP5 in

mammalian cells (Hinds & Sanchez, 2008) and co-immunoprecipitation studies have confirmed that this interaction is conserved in *P. falciparum* (Dobson *et al.*, 2001b).

PP7—This phosphatase is also known as PPEF in mammalian cells or PPJ in *P. falciparum*. Like PP5, PP7 is also a monomeric phosphatase with both the catalytic and regulator domains in a single polypeptide. PP7 has both an N-terminal IQ calmodulin-binding motif and a C-terminal set of EF-hands, which make PP7 unique among PPP family members. Contrary to what is observed with calcineurin, the Ca²⁺-calmodulin binding domain of PP7 inhibits its phosphatase activity based on *in vitro* assays with the *Arabidopsis* phosphatase (Kutuzov *et al.*, 2001). The functions of PP7 involve various cellular processes, including the pathways of light and stress signaling in plants, and stress protective responses, cell survival, and growth in mammalian systems (Andreeva & Kutuzov, 2009). Of the four parasite species analyzed all but *C. parvum* possess a PP7 gene each. Compared with PP7 from other eukaryotes including those from *P. falciparum* and *B. bovis*, *T. gondii* PP7 is unique in that it has a cytoskeletal-associated protein glycine rich (CAP-Gly) domain in its N-terminus before the calmodulin-binding motif. Consistent with the presence of this domain, TgPP7 localizes to the apical cap compartment of the IMC (Peter Bradley's comments in ToxoDB), where it might be associated with the tubulin filaments of the conoid. By contrast the *P. falciparum* PfPP7, which lacks the cytoskeleton association domain, is predominantly cytoplasmic (Dobson *et al.*, 2001a). Given that the function of PP7s is associated with calcium signaling and that TgPP7 is localized to the apical end where calcium-dependent secretion and conoid protrusion occurs, it is plausible that TgPP7 plays a potential role in host cell invasion, egress, and/or motility. Nonetheless, functional studies of PP7 in any apicomplexan are currently lacking.

PPKL, EFPP and SLP

Besides the phosphatases also found in animals, apicomplexans possess three additional phosphatase classes within the PPP family: Kelch-like domain containing protein phosphatases (PPKL) and Shewanella-like protein phosphatases, which are both also present in plants, and a parasite specific PPP phosphatase with EF-hand motifs within their N-terminus.

Kelch-like domain containing protein phosphatases (PPKL)—PPKL is a unique PPP family member, which is shared by Viridiplantae and Apicomplexa. The Kelch domain usually has a set of five to seven Kelch repeats that form a β -propeller tertiary structure generally involved in protein-protein interactions. Four PPKL members in *Arabidopsis* have been identified, known as BSU1, BSL1, BSL2 and BSL3. PPKLs in *Arabidopsis* have been shown to positively regulate brassinosteroid signaling (Maselli *et al.*, 2014). A single PPKL with three to four Kelch-like domains is present in each apicomplexan genome discussed here (Fig. 1). Two independent studies showed that PfPPKL is highly expressed in female gametocytes and in ookinetes, and that the deletion of PfPPKL resulted in abnormal ookinete morphology, which caused deficiency in forward gliding motility and in mosquito transmission (Guttery *et al.*, 2012, Philip *et al.*, 2012). No information about the function of this unique phosphatase within *T. gondii* is currently known.

EFPP—Besides calcineurin and PP7, apicomplexans have another PPP family phosphatase that contains calcium binding EF-hand motifs. A previous review (Kutuzov & Andreeva, 2008) named these protein phosphatases as EFPP to distinguish from PP7/PPEF. It is worthwhile to note that the catalytic domains of these EFPPs have varied degrees of mutations in the core catalytic domains of PPP family members, which puts their phosphatase activity into question. All the four apicomplexans each possess one EFPP, each with one or two EF-hand motifs (Fig. 1, Table 2). Despite their uniqueness, no studies have been conducted to investigate their functions.

Shewanella-like phosphatases (SLP)—Another PPP family phosphatase member that is absent in animals but present in both plants and apicomplexan parasites is a “bacterial-like” phosphatase, which is more closely related to bacterial PPP phosphatases than eukaryotic ones. Based on the similarities to different bacterial sources, the “bacterial-like” phosphatases have been classified into three categories: SLPs (Shewanella-like phosphatases), RLPs (Rhizobiales-like phosphatases), and ALPHs (ApaH-like phosphatases) (Kerk *et al.*, 2013). The category with homologs within certain apicomplexans is the SLPs. While no SLPs have been identified in *B. bovis*, *P. falciparum* contains two SLP coding genes, and both *T. gondii* and *C. parvum* have one (Table 1). Despite their divergence from other PPP family members, the SLPs contain most of the six core catalytic motifs of PPP family phosphatases (Fig. 1). Studies of apicomplexan SLPs have been limited to *P. falciparum* where PfSLP2 was shown to possess tyrosine phosphatase activities like other SLPs and to localize to the apical end of invasive merozoites within some vesicles distinct from dense granule, rhoptry and micronemes (Fernandez-Pol *et al.*, 2013). Prior to parasite invasion, the interaction between Band3, an erythrocyte membrane protein, and the erythrocyte cytoskeleton is disrupted by the phosphorylation of Band3 to facilitate invasion (Fernandez-Pol *et al.*, 2013). It has been proposed that, after the completion of parasite invasion, secreted PfSLP2 dephosphorylates Band3 and restores the interaction between erythrocyte Band3 and cytoskeleton (Fernandez-Pol *et al.*, 2013). In a separate study PfSLP1 was shown to be expressed in all stages of parasite life cycle, and its deletion affected microneme formation, and ookinete development and resulted in complete inhibition of oocyst formation (Patzewitz *et al.*, 2013).

PPM (PP2C) family of Serine/Threonine Phosphatase in Apicomplexa

PPM family phosphatases, which require Mg^{2+}/Mn^{2+} for their phosphatase activity, include PP2Cs and PDPs. Unlike most PPP family phosphatases, which have separate catalytic and regulatory subunits, PPM family phosphatases are monomeric enzymes with both catalytic and regulatory domains in a single polypeptide. Although having similar three-dimensional structure in terms of the catalytic domains, and similar mechanism of catalysis (Shi, 2009), PPM family phosphatases do not share much similarity at the amino acid level with PPP family phosphatases. Moreover, in contrast to the significant conservation among members of the PPP family of phosphatases across eukaryotic phyla, the PP2C phosphatase domains are much more divergent from each other. Nonetheless, residues that are involved in metal coordination and phosphate binding, such as RxxxED, DGxxG, DGxWD and DN, are all maintained in most PP2Cs (Shi, 2009). In addition, the number of PP2C members among

eukaryotic taxa vary widely, with a notable expansion in plants. For example, *Arabidopsis* contains about 80 PP2C members, in contrast to only 7 in budding yeast and 18 in human. The main role of PP2Cs appears to be regulation of stress signaling, although they also are involved in other biological processes, such as cell differentiation, survival and apoptosis (Lu & Wang, 2008).

All putative PP2Cs found in the four apicomplexan genomes used for our bioinformatics analysis are listed in Table 3. It is worth noting the variation in the number of PP2C members present in each of these four apicomplexans: 4 in *B. bovis*, 13 in *P. falciparum*, 14 in *C. parvum*, and 33 in *T. gondii* (Table 3, numbers include potential PDPs). To investigate the evolution of PP2Cs in these apicomplexans, we performed a phylogenetic analysis based on the PP2C phosphatase domain amino acid sequences (Fig. 2). We included the human and yeast PPM family members in the analysis to investigate possible correlation with phosphatases with known functions. As shown in Figure 2, the apicomplexan PP2Cs are mainly clustered into ten groups. For the convenience of citation, we arbitrarily numbered these groups from I to X. Also, because most *P. falciparum* PP2Cs have been designated names in a previous study (Guttery *et al.*, 2014), we designated names for the rest of the *P. falciparum* PP2Cs and for those from *T. gondii* ones based on their homology relationships with those from *P. falciparum*. The designated names are shown in Table 3.

Among the ten clusters, groups II and X are of note. Group II includes the yeast PDP ScPTC5 and the two human PDPs, which suggests that the apicomplexan PP2Cs in this group are possibly PDPs. Group X contains a single *P. falciparum* PP2C, PfPPM3, and eight *T. gondii* homologs TgPPM3A to H, indicating that an expansion of this PP2C likely occurred during the evolution of *T. gondii*. Duplication appears to be one of the mechanisms behind this expansion based on the fact that TgPPM3H and TgPPM3G are two nearly identical PP2Cs with tandem localization. Interestingly, of the eight *T. gondii* PP2Cs in group X, six contain signal peptides, suggesting that they might be secreted phosphatases functioning in either the parasitophorous vacuole (PV) or the host cell.

A number of studies have been conducted to investigate the functions of several PP2Cs in *P. falciparum* and *T. gondii*. The first identified PP2C in *P. falciparum* was PfPPM2 (PF3D7_1138500), which contains two PP2C catalytic domains both with validated phosphatase activity (Mamoun *et al.*, 1998). Substrates of PfPPM2 include the translation elongation factor 1-beta (Pfef-1 β), suggesting a role in the regulation of translation (Mamoun & Goldberg, 2001). Another *P. falciparum* phosphatase implicated in translation regulation is UIS2 (PF3D7_1464600) (Zhang *et al.*, 2016). This phosphatase was shown to be responsible for the dephosphorylation of eIF2 α -P, a translational factor whose phosphorylation status is critical for host infection (Zhang *et al.*, 2016). UIS2 was validated as a PP2C based on the fact that it can be inhibited by EDTA and Cd²⁺, and that it shows preference for Mn²⁺ and Mg²⁺, which are required for the activity of PP2Cs (Zhang *et al.*, 2016). However, this protein lacks a canonical PP2C catalytic domain and shows high similarity to purple acid phosphatases, which have been reported to possess protein phosphatase activity (Kaida *et al.*, 2010). Interestingly, two orthologs of this phosphatase are present in *T. gondii* (*TGME49_228160* and *TGME49_228170*) tandemly localized within chromosome X, suggesting a duplication event.

In a genome-wide functional analysis of *P. falciparum* protein phosphatases complete knockout was achieved for 7 out of 9 PP2Cs targeted (Guttery *et al.*, 2014). Knockout of PfPPM2 (group VII) and PfPPM5 (group III) resulted in the most significant phenotypes. Parasites lacking PfPPM2 parasites exhibited reduced macrogamete numbers and ookinete conversion rate, and PfPPM5 parasites had a defect in both number and size of oocysts formed (Guttery *et al.*, 2014). Further analysis revealed that PfPPM2 is essential for gametocyte sex allocation and ookinete differentiation, and PfPPM5 is involved in the modulation of oocyst development (Guttery *et al.*, 2014). There are two homologs of PfPPM2 and three of PfPPM5 in *T. gondii*, although the function of these expanded PP2Cs are not currently known.

One of the few studied PP2Cs from *T. gondii* (TGME49_231850, PPM13) was shown to act in conjunction with casein kinase II to regulate the phosphorylation status of Toxofilin, an actin binding protein known to be secreted into the host (Delorme *et al.*, 2003, Jan *et al.*, 2007). Another *T. gondii* PP2C (TGME49_282055, TgPPM20 akaPP2Chn) was shown to be localized in the rhoptries and to be secreted into the host nucleus during invasion (Gilbert *et al.*, 2007). The disruption of this PP2C results in a slight deficiency in parasite growth, but the particular substrate in either the host or parasite are not known (Gilbert *et al.*, 2007).

FCP/SCP family phosphatases

FCP/SCP family phosphatases use an aspartate-based catalysis mechanism, which distinguishes them from PPP and PPM family phosphatases. Their most conserved catalytic domain motif is DxDT/V, which is present in nearly all phosphatases within this family. Humans have eight members of this family of phosphatases while yeasts have five. For the four apicomplexans discussed here, five putative FCP/SCP family phosphatases are present in *C. parvum*, eight in *T. gondii*, four in *P. falciparum* and six in *B. bovis* (Table 4). FCP1 (also named as CTDP1) is the only member of this family that has been extensively investigated in humans and yeasts because it serves as a phosphatase of the CTD of the large subunits of RNA polymerase II (Corden, 2013). This region of the polymerase contains multiple heptapeptide repeats with a consensus sequence YSPTSPS in humans and yeasts, in which all positions except the two prolines can be phosphorylated (Yang & Stiller, 2014). Both human and yeast genomes have only one FCP gene, but interestingly, two putative orthologs are present in all the four apicomplexan genomes (Table 4). Also of remark is the fact that, compared with human and yeast which contain 52 and 26 tandemly repeated heptapeptides respectively, some apicomplexans have highly degenerate CTDs. For example, although *T. gondii* has ten heptapeptides, only two of them are tandemly repeated while all the others are dispersed throughout the domain. Previous studies in yeasts showed that the smallest CTD functional unit lies in heptapeptide pairs (Stiller & Cook, 2004), thus the dispersed single heptapeptides of *T. gondii* might have lost the conventional CTD functions. In contrast, the other three apicomplexans focused in this study all have more than ten tandemly repeated heptapeptides in their CTDs (Yang & Stiller, 2014). A particularly dynamic CTD evolution might have occurred in the genus *Plasmodium*, in which two independent expansions of tandem repeats happened in primate-infecting *Plasmodium* species, including *P. falciparum* (Kishore *et al.*, 2009). As the CTD plays a role as a docking platform to interact with various transcription and processing factors in transcription cycles

(Corden, 2013), the fluid evolution of this domain in Apicomplexa suggests diverse features for these processes in these taxa. Additionally, the presence of two copies of FCP phosphatases, if validated, could be a consequence of the life cycles and multiple developmental stages observed in these parasites. Unfortunately, almost no studies have been performed to investigate the role of differential phosphorylation of the CTD in regulating the activity of the RNA polymerase II in apicomplexans. A genome-wide analysis of the occupancy of differential phosphorylated RNA polymerase II CTD in *P. falciparum* revealed that the most genes with serine 2 and 5 double phosphorylated CTD are divided into two phases, early and late, during the infectious intraerythrocytic developmental cycle (Rai *et al.*, 2014). In depth study of the RNA polymerase II CTD and the FCP/SCP family phosphatases in these parasites is likely to reveal valuable information as to the evolution of these processes and factors.

Conclusion and future perspective

The precise dynamic phosphorylation status of proteins is critical for their functions. Protein kinases and phosphatases play equal important roles in the regulation of phosphorylation. However, like ugly ducklings, protein phosphatases have not received equivalent attention in most systems including in parasites. This is likely due to the low level of specificity shown by the phosphatases themselves and the high level of complexity that arises when one starts considering regulatory and structural domains. Many studies of apicomplexan phosphatases were performed without the advantage of sequenced genomes or the modern molecular tools that now make these parasites more tractable. Therefore, the knowledge of the phosphatases in this taxon is very limited (phosphatases for which function is known are listed in table 5). As reversible phosphorylation is now recognized as a key regulator of many essential and unique processes of the Apicomplexa, phosphatases will inevitably come into focus. A particular aspect of the biology of these parasites in which the study of phosphatases is likely to yield important functional information is parasite motility. The motility machinery of the so-called “glidesome” is extensively regulated by calcium dependent phosphorylation, and while the theoretical roles for phosphatases in motility dependent events have been alluded to the identity of these remain unknown.

Identification of new drug targets is always a priority in the study of the biology of these parasites. The activity of Cyclosporin A and FK506 would support the concept of phosphatases as valid targets for antiparasitic drugs. Importantly, most PPP family members, a few PPM and FCP/SCP family members appear to be essential according to the recent genome-wide CRISPR screen performed in *T. gondii* (Sidik *et al.*, 2016). Not only does this study underscore the important role of these proteins and their potential as drug targets, but it also provides a roadmap to prioritize their study. Moreover, apicomplexans contain three families of phosphatases that are not present in mammals and, thus, likely to regulate parasite specific processes. The last decade has brought great technical advances to the study of these parasites, including gene editing and phosphor-proteomics. Thus, the field is well poised to explore the role of phosphatases in the biology and pathogenesis of these important human parasites. Given the unique place of these organisms in the tree of life, such studies would also provide valuable information as to the evolution of protein phosphatases.

Acknowledgments

Research in the Arrizabalaga's laboratory is supported by grants from NIH (R21AI119516, R01AI123457, and R03AI101624). We thank Kaice A. LaFavers for her helpful suggestions during the preparation of this manuscript.

References

- Alvarez CA, Suvorova ES. A checkpoint roadmap for the complex cell division of Apicomplexa parasites. *bioRxiv*. 2017
- Andreeva AV, Kutuzov MA. PPEF/PP7 protein Ser/Thr phosphatases. *Cell Mol Life Sci*. 2009; 66:3103–3110. [PubMed: 19662497]
- Antony HA, Parija SC. Antimalarial drug resistance: An overview. *Trop Parasitol*. 2016; 6:30–41. [PubMed: 26998432]
- Azevedo MF, Sanders PR, Krejany E, Nie CQ, Fu P, Bach LA, Wunderlich G, Crabb BS, Gilson PR. Inhibition of *Plasmodium falciparum* CDPK1 by conditional expression of its J-domain demonstrates a key role in schizont development. *Biochem J*. 2013; 452:433–441. doi: 410.1042/BJ20130124. [PubMed: 23548171]
- Bansal A, Singh S, More KR, Hans D, Nangalia K, Yogavel M, Sharma A, Chitnis CE. Characterization of *Plasmodium falciparum* calcium-dependent protein kinase 1 (PfCDPK1) and its role in microneme secretion during erythrocyte invasion. *J Biol Chem*. 2013; 288:1590–1602. doi: 1510.1074/jbc.M1112.411934. Epub 412012 Nov 411930. [PubMed: 23204525]
- Benmerzouga I, Checkley LA, Ferdig MT, Arrizabalaga G, Wek RC, Sullivan WJ Jr. Guanabenz repurposed as an antiparasitic with activity against acute and latent toxoplasmosis. *Antimicrob Agents Chemother*. 2015; 59:69396945.
- Bhattacharyya MK, Hong Z, Kongkasuriyachai D, Kumar N. *Plasmodium falciparum* protein phosphatase type 1 functionally complements a *glc7* mutant in *Saccharomyces cerevisiae*. *International journal for parasitology*. 2002; 32:739–747. [PubMed: 12062492]
- Blisnick T, Vincensini L, Fall G, Braun-Breton C. Protein phosphatase 1, a *Plasmodium falciparum* essential enzyme, is exported to the host cell and implicated in the release of infectious merozoites. *Cellular microbiology*. 2006; 8:591–601. [PubMed: 16548885]
- Boucher LE, Bosch J. The apicomplexan glideosome and adhesins - Structures and function. *J Struct Biol*. 2015; 190:93–114. [PubMed: 25764948]
- Cavalier-Smith T. Kingdom protozoa and its 18 phyla. *Microbiol Rev*. 1993; 57:953–994. [PubMed: 8302218]
- Ceulemans H, Bollen M. Functional diversity of protein phosphatase-1, a cellular economizer and reset button. *Physiol Rev*. 2004; 84:1–39. [PubMed: 14715909]
- Ciesla J, Fraczyk T, Rode W. Phosphorylation of basic amino acid residues in proteins: important but easily missed. *Acta Biochim Pol*. 2011; 58:137–148. [PubMed: 21623415]
- Cohen PT, Philp A, Vazquez-Martin C. Protein phosphatase 4—from obscurity to vital functions. *FEBS letters*. 2005; 579:3278–3286. [PubMed: 15913612]
- Corden JL. RNA polymerase II C-terminal domain: Tethering transcription to transcript and template. *Chemical reviews*. 2013; 113:8423–8455. [PubMed: 24040939]
- Cori GT, Cori CF. The enzymatic conversion of phosphorylase a to b. *The Journal of biological chemistry*. 1945; 158:321–332.
- Daher W, Browaeyns E, Pierrot C, Jouin H, Dive D, Meurice E, Dissous C, Capron M, Tomavo S, Doerig C, Cailliau K, Khalife J. Regulation of protein phosphatase type 1 and cell cycle progression by PflRR1, a novel leucine-rich repeat protein of the human malaria parasite *Plasmodium falciparum*. *Molecular microbiology*. 2006; 60:578–590. [PubMed: 16629662]
- Daher W, Oria G, Fauquenoy S, Cailliau K, Browaeyns E, Tomavo S, Khalife J. A *Toxoplasma gondii* leucine-rich repeat protein binds phosphatase type 1 protein and negatively regulates its activity. *Eukaryotic cell*. 2007; 6:1606–1617. [PubMed: 17660360]
- Delorme V, Cayla X, Faure G, Garcia A, Tardieux I. Actin dynamics is controlled by a casein kinase II and phosphatase 2C interplay on *Toxoplasma gondii* toxofilin. *Molecular biology of the cell*. 2003; 14:1900–1912. [PubMed: 12802063]

- Delorme V, Garcia A, Cayla X, Tardieux I. A role for *Toxoplasma gondii* type 1 ser/thr protein phosphatase in host cell invasion. *Microbes Infect.* 2002; 4:271–278. [PubMed: 11909736]
- Dobson S, Bracchi V, Chakrabarti D, Barik S. Characterization of a novel serine/threonine protein phosphatase (PfPPJ) from the malaria parasite, *Plasmodium falciparum*. *Molecular and biochemical parasitology.* 2001a; 115:29–39. [PubMed: 11377737]
- Dobson S, Kar B, Kumar R, Adams B, Barik S. A novel tetratricopeptide repeat (TPR) containing PP5 serine/threonine protein phosphatase in the malaria parasite, *Plasmodium falciparum*. *BMC microbiology.* 2001b; 1:31. [PubMed: 11737864]
- Dobson S, Kumar R, Bracchi-Ricard V, Freeman S, Al-Murrani SW, Johnson C, Damuni Z, Chakrabarti D, Barik S. Characterization of a unique aspartate-rich protein of the SET/TAF-family in the human malaria parasite, *Plasmodium falciparum*, which inhibits protein phosphatase 2A. *Molecular and biochemical parasitology.* 2003; 126:239–250. [PubMed: 12615323]
- Dobson S, May T, Berriman M, Del Vecchio C, Fairlamb AH, Chakrabarti D, Barik S. Characterization of protein Ser/Thr phosphatases of the malaria parasite, *Plasmodium falciparum*: inhibition of the parasitic calcineurin by cyclophilin-cyclosporin complex. *Molecular and biochemical parasitology.* 1999; 99:167–181. [PubMed: 10340482]
- Eaton MS, Weiss LM, Kim K. Cyclic nucleotide kinases and tachyzoite-bradyzoite transition in *Toxoplasma gondii*. *International journal for parasitology.* 2006; 36:107–114. [PubMed: 16216248]
- Eick D, Geyer M. The RNA polymerase II carboxy-terminal domain (CTD) code. *Chemical reviews.* 2013; 113:8456–8490. [PubMed: 23952966]
- Etheridge RD, Alaganan A, Tang K, Lou HJ, Turk BE, Sibley LD. The *Toxoplasma* pseudokinase ROP5 forms complexes with ROP18 and ROP17 kinases that synergize to control acute virulence in mice. *Cell host & microbe.* 2014; 15:537–550. [PubMed: 24832449]
- Fentress SJ, Behnke MS, Dunay IR, Mashayekhi M, Rommereim LM, Fox BA, Bzik DJ, Taylor GA, Turk BE, Lichti CF, Townsend RR, Qiu W, Hui R, Beatty WL, Sibley LD. Phosphorylation of immunity-related GTPases by a *Toxoplasma gondii*-secreted kinase promotes macrophage survival and virulence. *Cell host & microbe.* 2010; 8:484–495. [PubMed: 21147463]
- Fernandez-Pol S, Slouka Z, Bhattacharjee S, Fedotova Y, Freed S, An X, Holder AA, Campanella E, Low PS, Mohandas N, Haldar K. A bacterial phosphatase-like enzyme of the malaria parasite *Plasmodium falciparum* possesses tyrosine phosphatase activity and is implicated in the regulation of band 3 dynamics during parasite invasion. *Eukaryotic cell.* 2013; 12:1179–1191. [PubMed: 23825180]
- Francia ME, Striepen B. Cell division in apicomplexan parasites. *Nat Rev Microbiol.* 2014; 12:125–136. [PubMed: 24384598]
- Freville A, Cailliau-Maggio K, Pierrot C, Tellier G, Kalamou H, Lafitte S, Martoriati A, Pierce RJ, Bodart JF, Khalife J. *Plasmodium falciparum* encodes a conserved active inhibitor-2 for Protein Phosphatase type 1: perspectives for novel anti-plasmodial therapy. *Bmc Biol.* 2013; 11:80. [PubMed: 23837822]
- Freville A, Landrieu I, Garcia-Gimeno MA, Vicogne J, Montbarbon M, Bertin B, Verger A, Kalamou H, Sanz P, Werkmeister E, Pierrot C, Khalife J. *Plasmodium falciparum* inhibitor-3 homolog increases protein phosphatase type 1 activity and is essential for parasitic survival. *The Journal of biological chemistry.* 2012; 287:1306–1321. [PubMed: 22128182]
- Freville A, Tellier G, Vandomme A, Pierrot C, Vicogne J, Cantrelle FX, Martoriati A, Cailliau-Maggio K, Khalife J, Landrieu I. Identification of a *Plasmodium falciparum* inhibitor-2 motif involved in the binding and regulation activity of protein phosphatase type 1. *FEBS J.* 2014; 281:4519–4534. [PubMed: 25132288]
- Gaji RY, Johnson DE, Treeck M, Wang M, Hudmon A, Arrizabalaga G. Phosphorylation of a Myosin Motor by TgCDPK3 Facilitates Rapid Initiation of Motility during *Toxoplasma gondii* egress. *PLoS pathogens.* 2015; 11:e1005268. [PubMed: 26544049]
- Garrison E, Treeck M, Ehret E, Butz H, Garbuz T, Oswald BP, Settles M, Boothroyd J, Arrizabalaga G. A Forward Genetic Screen Reveals that Calcium-dependent Protein Kinase 3 Regulates Egress in *Toxoplasma*. *PLoS pathogens.* 2012; 8:e1003049. [PubMed: 23209419]

- Gilbert LA, Ravindran S, Turetzky JM, Boothroyd JC, Bradley PJ. Toxoplasma gondii targets a protein phosphatase 2C to the nuclei of infected host cells. *Eukaryotic cell*. 2007; 6:73–83. [PubMed: 17085638]
- Gilk SD, Gaskins E, Ward GE, Beckers CJ. GAP45 phosphorylation controls assembly of the Toxoplasma myosin XIV complex. *Eukaryotic cell*. 2009; 8:190–196. [PubMed: 19047362]
- Goldberg J, Huang HB, Kwon YG, Greengard P, Nairn AC, Kuriyan J. Three-dimensional structure of the catalytic subunit of protein serine/threonine phosphatase-1. *Nature*. 1995; 376:745–753. [PubMed: 7651533]
- Green JL, Rees-Channer RR, Howell SA, Martin SR, Knuepfer E, Taylor HM, Grainger M, Holder AA. The motor complex of Plasmodium falciparum: phosphorylation by a calcium-dependent protein kinase. *The Journal of biological chemistry*. 2008; 283:30980–30989. [PubMed: 18768477]
- Groves MR, Hanlon N, Turowski P, Hemmings BA, Barford D. The structure of the protein phosphatase 2A PR65/A subunit reveals the conformation of its 15 tandemly repeated HEAT motifs. *Cell*. 1999; 96:99–110. [PubMed: 9989501]
- Gubbels MJ, White M, Szatanek T. The cell cycle and Toxoplasma gondii cell division: tightly knit or loosely stitched? *International journal for parasitology*. 2008; 38:1343–1358. [PubMed: 18703066]
- Guttery DS, Poulin B, Ferguson DJ, Szoor B, Wickstead B, Carroll PL, Ramakrishnan C, Brady D, Patzewitz EM, Straschil U, Solyakov L, Green JL, Sinden RE, Tobin AB, Holder AA, Tewari R. A unique protein phosphatase with kelch-like domains (PPKL) in Plasmodium modulates ookinete differentiation, motility and invasion. *PLoS pathogens*. 2012; 8:e1002948. [PubMed: 23028336]
- Guttery DS, Poulin B, Ramaprasad A, Wall RJ, Ferguson DJ, Brady D, Patzewitz EM, Whipple S, Straschil U, Wright MH, Mohamed AM, Radhakrishnan A, Arold ST, Tate EW, Holder AA, Wickstead B, Pain A, Tewari R. Genome-wide functional analysis of Plasmodium protein phosphatases reveals key regulators of parasite development and differentiation. *Cell host & microbe*. 2014; 16:128–140. [PubMed: 25011111]
- Hinchliff CE, Smith SA, Allman JF, Burleigh JG, Chaudhary R, Coghill LM, Crandall KA, Deng J, Drew BT, Gazis R, Gude K, Hibbett DS, Katz LA, Laughinghouse HDt, McTavish EJ, Midford PE, Owen CL, Ree RH, Rees JA, Soltis DE, Williams T, Cranston KA. Synthesis of phylogeny and taxonomy into a comprehensive tree of life. *Proceedings of the National Academy of Sciences of the United States of America*. 2015; 112:12764–12769. [PubMed: 26385966]
- Hinds TD Jr, Sanchez ER. Protein phosphatase 5. *Int J Biochem Cell Biol*. 2008; 40:2358–2362. [PubMed: 17951098]
- Hollin T, De Witte C, Lenne A, Pierrot C, Khalife J. Analysis of the interactome of the Ser/Thr Protein Phosphatase type 1 in Plasmodium falciparum. *BMC genomics*. 2016; 17:246. [PubMed: 26988354]
- Jan G, Delorme V, David V, Revenu C, Rebollo A, Cayla X, Tardieux I. The toxofilin-actin-PP2C complex of Toxoplasma: identification of interacting domains. *Biochemical Journal*. 2007; 401:711–719. [PubMed: 17014426]
- Janouskovec J, Horak A, Obornik M, Lukes J, Keeling PJ. A common red algal origin of the apicomplexan, dinoflagellate, and heterokont plastids. *Proceedings of the National Academy of Sciences of the United States of America*. 2010; 107:10949–10954. [PubMed: 20534454]
- Liu, Jun, Farmer, Jesse D., Jr, Lane, Willam S., Friedman, Jeff, Weissman, Irving, Schreiber, SL. Calcineurin is a common target of cyclophilin-cyclosporin A and FKBP-FK506 complexes. *Cell*. 1991; 66:807–815. [PubMed: 1715244]
- Kaida R, Serada S, Norioka N, Norioka S, Neumetzler L, Pauly M, Sampedro J, Zarra I, Hayashi T, Kaneko TS. Potential role for purple acid phosphatase in the dephosphorylation of wall proteins in tobacco cells. *Plant physiology*. 2010; 153:603–610. [PubMed: 20357138]
- Kato N, Sakata T, Breton G, Le Roch KG, Nagle A, Andersen C, Bursulaya B, Henson K, Johnson J, Kumar KA, Marr F, Mason D, McNamara C, Plouffe D, Ramachandran V, Spooner M, Tuntland T, Zhou Y, Peters EC, Chatterjee A, Schultz PG, Ward GE, Gray N, Harper J, Winzeler EA. Gene expression signatures and small-molecule compounds link a protein kinase to Plasmodium falciparum motility. *Nat Chem Biol*. 2008; 4:347–356. [PubMed: 18454143]

- Keeley A, Soldati D. The glideosome: a molecular machine powering motility and host-cell invasion by Apicomplexa. *Trends in cell biology*. 2004; 14:528–532. [PubMed: 15450974]
- Kerk D, Uhrig RG, Moorhead GB. Bacterial-like PPP protein phosphatases: novel sequence alterations in pathogenic eukaryotes and peculiar features of bacterial sequence similarity. *Plant signaling & behavior*. 2013; 8:e27365. [PubMed: 24675170]
- Khan F, Tang J, Qin CL, Kim K. Cyclin-dependent kinase TPK2 is a critical cell cycle regulator in *Toxoplasma gondii*. *Molecular microbiology*. 2002; 45:321–332. [PubMed: 12123447]
- Kishore SP, Perkins SL, Templeton TJ, Deitsch KW. An unusual recent expansion of the C-terminal domain of RNA polymerase II in primate malaria parasites features a motif otherwise found only in mammalian polymerases. *Journal of molecular evolution*. 2009; 68:706–714. [PubMed: 19449052]
- Konrad C, Queener SF, Wek RC, Sullivan WJ Jr. Inhibitors of eIF2alpha dephosphorylation slow replication and stabilize latency in *Toxoplasma gondii*. *Antimicrob Agents Chemother*. 2013; 57:1815–1822. [PubMed: 23380722]
- Kumar R, Adams B, Musiyenko A, Shulyayeva O, Barik S. The FK506-binding protein of the malaria parasite, *Plasmodium falciparum*, is a FK506-sensitive chaperone with FK506-independent calcineurin-inhibitory activity. *Molecular and biochemical parasitology*. 2005; 141:163–173. [PubMed: 15850699]
- Kumar R, Adams B, Oldenburg A, Musiyenko A, Barik S. Characterisation and expression of a PP1 serine/threonine protein phosphatase (PfPP1) from the malaria parasite, *Plasmodium falciparum*: demonstration of its essential role using RNA interference. *Malar J*. 2002; 1:5. [PubMed: 12057017]
- Kutuzov MA, Andreeva AV. Protein Ser/Thr phosphatases of parasitic protozoa. *Molecular and biochemical parasitology*. 2008; 161:81–90. [PubMed: 18619495]
- Kutuzov MA, Bennett N, Andreeva AV. Interaction of plant protein Ser/Thr phosphatase PP7 with calmodulin. *Biochemical and biophysical research communications*. 2001; 289:634–640. [PubMed: 11716523]
- Lebrun M, Carruthers VB, Cesbron-Delauw MF. *Toxoplasma* Secretory Proteins and Their Roles in Cell Invasion and Intracellular Survival. *Toxoplasma Gondii: The Model Apicomplexan - Perspectives and Methods* (2nd). 2014:389–453.
- Li JL, Baker DA. Protein phosphatase beta, a putative type-2A protein phosphatase from the human malaria parasite *Plasmodium falciparum*. *European journal of biochemistry / FEBS*. 1997; 249:98–106.
- Lillo C, Kataya ARA, Heidari B, Creighton MT, Nemie-Feyissa D, Ginbot Z, Jonassen EM. Protein phosphatases PP2A, PP4 and PP6: mediators and regulators in development and responses to environmental cues. *Plant Cell Environ*. 2014; 37:2631–2648. [PubMed: 24810976]
- Lindenthal C, Klinkert MQ. Identification and biochemical characterisation of a protein phosphatase 5 homologue from *Plasmodium falciparum*. *Molecular and biochemical parasitology*. 2002; 120:257–268. [PubMed: 11897131]
- Lourido S, Shuman J, Zhang C, Shokat KM, Hui R, Sibley LD. Calcium-dependent protein kinase 1 is an essential regulator of exocytosis in *Toxoplasma*. *Nature*. 2010; 465:359–362. [PubMed: 20485436]
- Lourido S, Tang K, Sibley LD. Distinct signalling pathways control *Toxoplasma* egress and host-cell invasion. *The EMBO journal*. 2012; 31:4524–4534. [PubMed: 23149386]
- Lu G, Wang YB. Functional diversity of mammalian Type 2C protein phosphatase isoforms: New tales from an old family. *Clin Exp Pharmacol P*. 2008; 35:107–112.
- Mamoun CB, Goldberg DE. *Plasmodium* protein phosphatase 2C dephosphorylates translation elongation factor 1beta and inhibits its PKC-mediated nucleotide exchange activity in vitro. *Molecular microbiology*. 2001; 39:973–981. [PubMed: 11251817]
- Mamoun CB, Sullivan DJ Jr, Banerjee R, Goldberg DE. Identification and characterization of an unusual double serine/threonine protein phosphatase 2C in the malaria parasite *Plasmodium falciparum*. *The Journal of biological chemistry*. 1998; 273:11241–11247. [PubMed: 9556615]

- Maselli GA, Slamovits CH, Bianchi JJ, Vilarrasa-Blasi J, Cano-Delgado AI, Mora-Garcia S. Revisiting the evolutionary history and roles of protein phosphatases with Kelch-like domains in plants. *Plant physiology*. 2014; 164:1527–1541. [PubMed: 24492333]
- McCoy JM, Whitehead L, van Dooren GG, Tonkin CJ. TgCDPK3 Regulates Calcium-Dependent Egress of *Toxoplasma gondii* from Host Cells. *PLoS pathogens*. 2012; 8:e1003066. [PubMed: 23226109]
- Moskes C, Burghaus PA, Wernli B, Sauder U, Durrenberger M, Kappes B. Export of *Plasmodium falciparum* calcium-dependent protein kinase 1 to the parasitophorous vacuole is dependent on three N-terminal membrane anchor motifs. *Molecular microbiology*. 2004; 54:676–691. [PubMed: 15491359]
- Moudy R, Manning TJ, Beckers CJ. The loss of cytoplasmic potassium upon host cell breakdown triggers egress of *Toxoplasma gondii*. *Journal of Biological Chemistry*. 2001; 276:41492–41501. [PubMed: 11526113]
- Musson RE, Smit NP. Regulatory mechanisms of calcineurin phosphatase activity. *Curr Med Chem*. 2011; 18:301–315. [PubMed: 21110798]
- Narasimhan J, Joyce BR, Naguleswaran A, Smith AT, Livingston MR, Dixon SE, Coppens I, Wek RC, Sullivan WJ Jr. Translation regulation by eukaryotic initiation factor-2 kinases in the development of latent cysts in *Toxoplasma gondii*. *The Journal of biological chemistry*. 2008; 283:16591–16601. [PubMed: 18420584]
- Olsen JV, Blagoev B, Gnäd F, Macek B, Kumar C, Mortensen P, Mann M. Global, in vivo, and site-specific phosphorylation dynamics in signaling networks. *Cell*. 2006; 127:635–648. [PubMed: 17081983]
- Opitz C, Soldati D. ‘The glideosome’: a dynamic complex powering gliding motion and host cell invasion by *Toxoplasma gondii*. *Molecular microbiology*. 2002; 45:597–604. [PubMed: 12139608]
- Pandey R, Mohammed A, Pierrot C, Khalife J, Malhotra P, Gupta D. Genome wide in silico analysis of *Plasmodium falciparum* phosphatome. *BMC genomics*. 2014; 15:1024. [PubMed: 25425018]
- Patzewitz EM, Guttery DS, Poulin B, Ramakrishnan C, Ferguson DJP, Wall RJ, Brady D, Holder AA, Szoor B, Tewari R. An Ancient Protein Phosphatase, SHLP1, Is Critical to Microneme Development in *Plasmodium* Ookinetes and Parasite Transmission. *Cell Rep*. 2013; 3:622–629. [PubMed: 23434509]
- Paul AS, Saha S, Engelberg K, Jiang RH, Coleman BI, Kosber AL, Chen CT, Ganter M, Espy N, Gilberger TW, Gubbels MJ, Duraisingh MT. Parasite Calcineurin Regulates Host Cell Recognition and Attachment by Apicomplexans. *Cell host & microbe*. 2015
- Philip N, Vaikkinen HJ, Tetley L, Waters AP. A unique Kelch domain phosphatase in *Plasmodium* regulates ookinete morphology, motility and invasion. *PloS one*. 2012; 7:e44617. [PubMed: 22957089]
- Rai R, Zhu L, Chen H, Gupta AP, Sze SK, Zheng J, Ruedl C, Bozdech Z, Featherstone M. Genome-wide analysis in *Plasmodium falciparum* reveals early and late phases of RNA polymerase II occupancy during the infectious cycle. *BMC genomics*. 2014; 15:959. [PubMed: 25373614]
- Rebello S, Santos M, Martins F, da Cruz e Silva EF, da Cruz e Silva OA. Protein phosphatase 1 is a key player in nuclear events. *Cellular signalling*. 2015; 27:2589–2598. [PubMed: 26275498]
- Ridzuan MA, Moon RW, Knuepfer E, Black S, Holder AA, Green JL. Subcellular location, phosphorylation and assembly into the motor complex of GAP45 during *Plasmodium falciparum* schizont development. *PloS one*. 2012; 7:e33845. [PubMed: 22479457]
- Sapkota G, Knockaert M, Alarcon C, Montalvo E, Brivanlou AH, Massague J. Dephosphorylation of the linker regions of Smad1 and Smad2/3 by small C-terminal domain phosphatases has distinct outcomes for bone morphogenetic protein and transforming growth factor-beta pathways. *The Journal of biological chemistry*. 2006; 281:40412–40419. [PubMed: 17085434]
- Shi YG. Serine/Threonine Phosphatases: Mechanism through Structure. *Cell*. 2009; 139:468–484. [PubMed: 19879837]
- Sidik SM, Huet D, Ganesan SM, Huynh MH, Wang T, Nasamu AS, Thiru P, Saeij JP, Carruthers VB, Niles JC, Lourido S. A Genome-wide CRISPR Screen in *Toxoplasma* Identifies Essential Apicomplexan Genes. *Cell*. 2016; 166:1423–1435 e1412. [PubMed: 27594426]

- Singh S, More KR, Chitnis CE. Role of calcineurin and actin dynamics in regulated secretion of microneme proteins in *Plasmodium falciparum* merozoites during erythrocyte invasion. *Cell Microbiol.* 2014; 16:50–63. [PubMed: 23910910]
- Stefansson B, Ohama T, Daugherty AE, Brautigan DL. Protein phosphatase 6 regulatory subunits composed of ankyrin repeat domains. *Biochemistry.* 2008; 47:1442–1451. [PubMed: 18186651]
- Steinfeldt T, Konen-Waisman S, Tong L, Pawlowski N, Lamkemeyer T, Sibley LD, Hunn JP, Howard JC. Phosphorylation of mouse immunity-related GTPase (IRG) resistance proteins is an evasion strategy for virulent *Toxoplasma gondii*. *PLoS biology.* 2010; 8:e1000576. [PubMed: 21203588]
- Stiller JW, Cook MS. Functional unit of the RNA polymerase II C-terminal domain lies within heptapeptide pairs. *Eukaryotic cell.* 2004; 3:735–740. [PubMed: 15189994]
- Uhrig RG, Labandera AM, Moorhead GB. Arabidopsis PPP family of serine/threonine protein phosphatases: many targets but few engines. *Trends in plant science.* 2013; 18:505–513. [PubMed: 23790269]
- Vandomme A, Freville A, Cailliau K, Kalamou H, Bodart JF, Khalife J, Pierrot C. PhosphoTyrosyl phosphatase activator of *Plasmodium falciparum*: identification of its residues involved in binding to and activation of PP2A. *Int J Mol Sci.* 2014; 15:2431–2453. [PubMed: 24521882]
- Vlastaridis P, Kyriakidou P, Chaliotis A, Van de Peer Y, Oliver SG, Amoutzias GD. Estimating the total number of phosphoproteins and phosphorylation sites in eukaryotic proteomes. *Gigascience.* 2017; 6:1–11.
- Wang Y, Liu Z, Cheng H, Gao T, Pan Z, Yang Q, Guo A, Xue Y. EKPD: a hierarchical database of eukaryotic protein kinases and protein phosphatases. *Nucleic acids research.* 2014; 42:D496–502. [PubMed: 24214991]
- Wilkes JM, Doerig C. The protein-phosphatome of the human malaria parasite *Plasmodium falciparum*. *BMC genomics.* 2008; 9:412. [PubMed: 18793411]
- Yang C, Stiller JW. Evolutionary diversity and taxon-specific modifications of the RNA polymerase II C-terminal domain. *Proceedings of the National Academy of Sciences of the United States of America.* 2014; 111:5920–5925. [PubMed: 24711388]
- Zhang M, Fennell C, Ranford-Cartwright L, Sakthivel R, Gueirard P, Meister S, Caspi A, Doerig C, Nussenzweig RS, Tuteja R, Sullivan WJ Jr, Roos DS, Fontoura BM, Menard R, Winzeler EA, Nussenzweig V. The *Plasmodium* eukaryotic initiation factor-2 α kinase IK2 controls the latency of sporozoites in the mosquito salivary glands. *The Journal of experimental medicine.* 2010; 207:1465–1474. [PubMed: 20584882]
- Zhang M, Mishra S, Sakthivel R, Fontoura BM, Nussenzweig V. UIS2: A Unique Phosphatase Required for the Development of *Plasmodium* Liver Stages. *PLoS pathogens.* 2016; 12:e1005370. [PubMed: 26735921]
- Zhao Y, Franklin RM, Kappes B. *Plasmodium falciparum* calcium-dependent protein kinase phosphorylates proteins of the host erythrocytic membrane. *Mol Biochem Parasitol.* 1994; 66:329–343. [PubMed: 7808482]

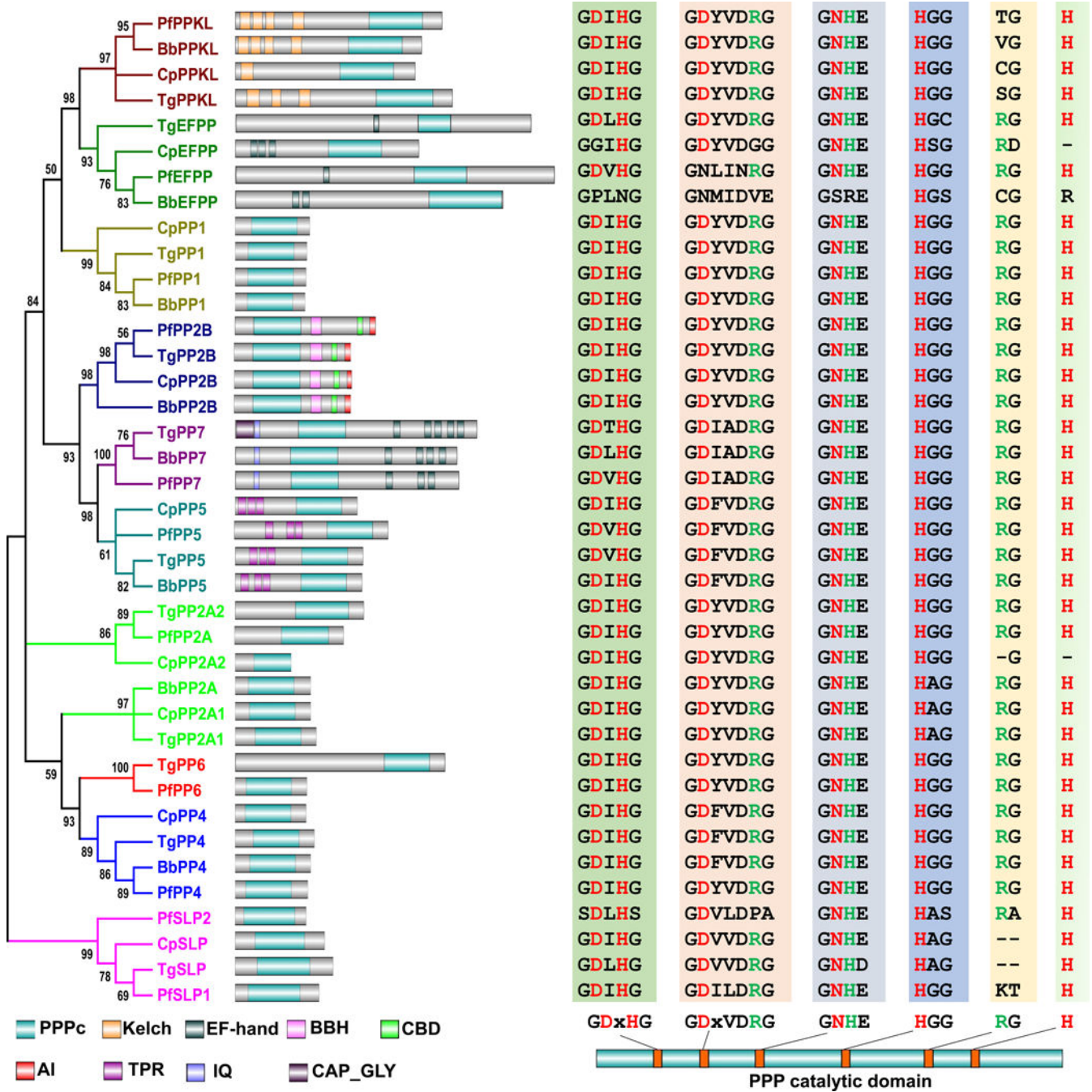


Figure 1. The PPP family phosphatases of *T. gondii*, *P. falciparum*, *B. bovis*, and *C. parvum*
 The phylogenetic tree shown was assembled based on the alignments of the catalytic domains of the PPP family phosphatase members found in the genomes of *T. gondii* (Tg), *P. falciparum* (Pf), *C. parvum* (Cp), and *B. bovis* (Bb). Protein alignment was performed with MUSCLE, and Gblocks 0.91b was used to select conserved blocks for tree-building. The phylogenetic analysis was performed using PHYML 3.0 with a maximum likelihood method under the LG model of amino acids substitution. Branch support values were estimated by Approximate Likelihood-Ratio Test (aLRT). The final tree was condensed with a cut-off

value 50%. The branches of each subfamily are displayed in an individual color. The domain architecture of each phosphatase is shown to the right of its branch. Domains shown include PPPc, PPP catalytic domain; Kelch, kelch domain; EF-hand, EF-hand motif; BBH, CNB-binding helix; CBD, Ca²⁺-calmodulin-binding motif; AI, autoinhibitory sequence; TPR, Tetratricopeptide repeat domain; IQ, Calmodulin-binding motif; CAP_GLY, Cytoskeleton-associated protein glycine-rich domain. The alignments of the core catalytic motif sequences are displayed to the right of each corresponding domain architecture. The relative location and consensus sequences of each core catalytic motif are shown underneath the alignments. The residues shown in red or green are those contribute to metal coordination or phosphate binding.

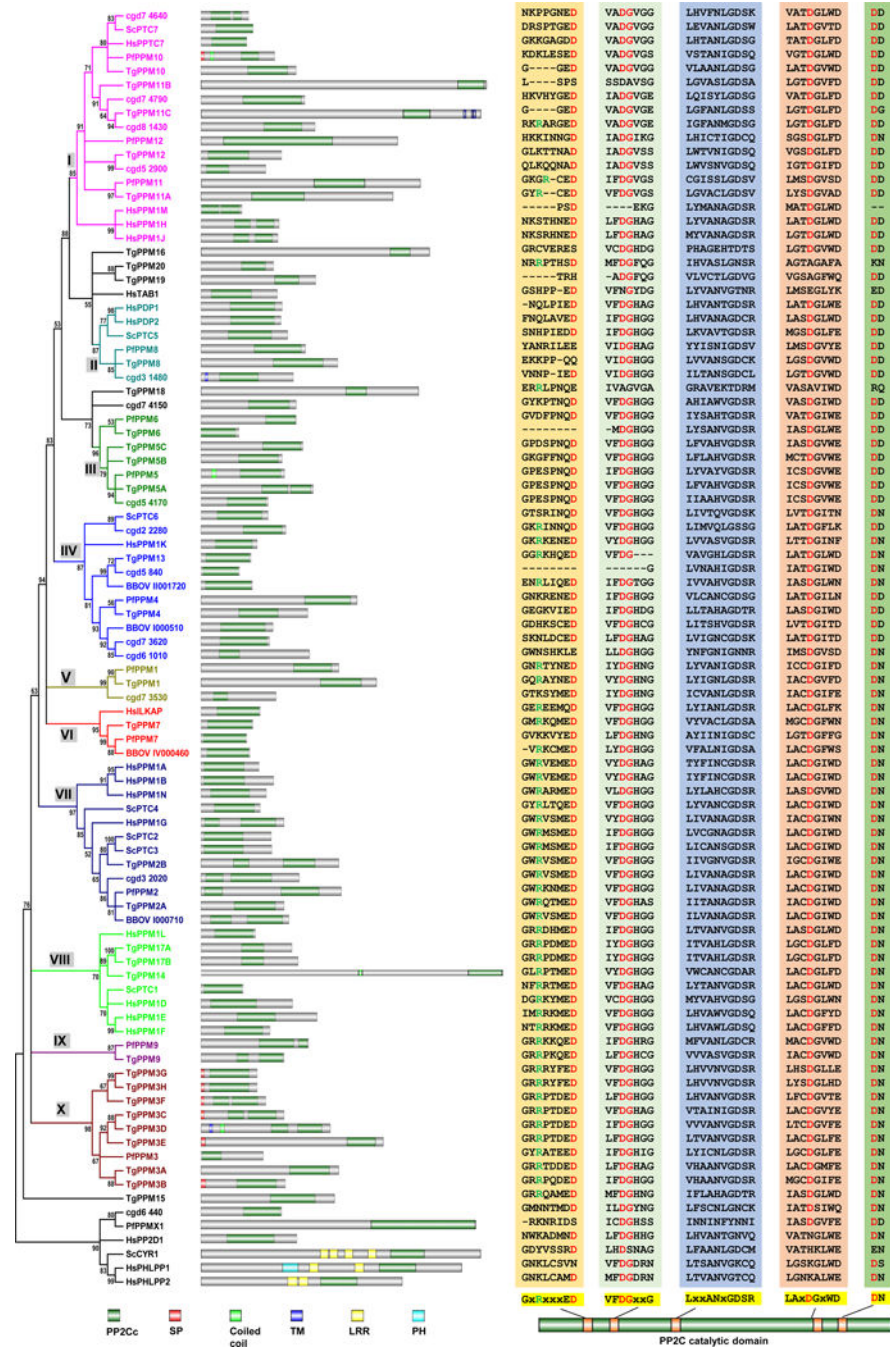


Figure 2. The PPM family phosphatases of *T. gondii*, *P. falciparum*, *B. bovis*, and *C. parvum*. The phylogenetic tree was constructed based on the alignments of the PP2C catalytic domains of the PPM family members from the four apicomplexans and the 20 PPM members (18 PP2Cs and 2 PDPs) of human and 8 (7 PP2Cs and 1 PDP) of budding yeast. Alignments and tree assembly were performed as in figure 1. Domain architecture and alignments of the core catalytic motif sequences are also shown as in figure 1. Domains include PP2Cc, PP2C catalytic domain; SP, Signal peptide; Coiled coil, Coiled coil motif; TM, Transmembrane domain; LRR, Leucine-rich repeats; PH, Pleckstrin homology domain.

The PPM family phosphatase are clustered into 10 groups, each annotated with a different color and a group name from I to X.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

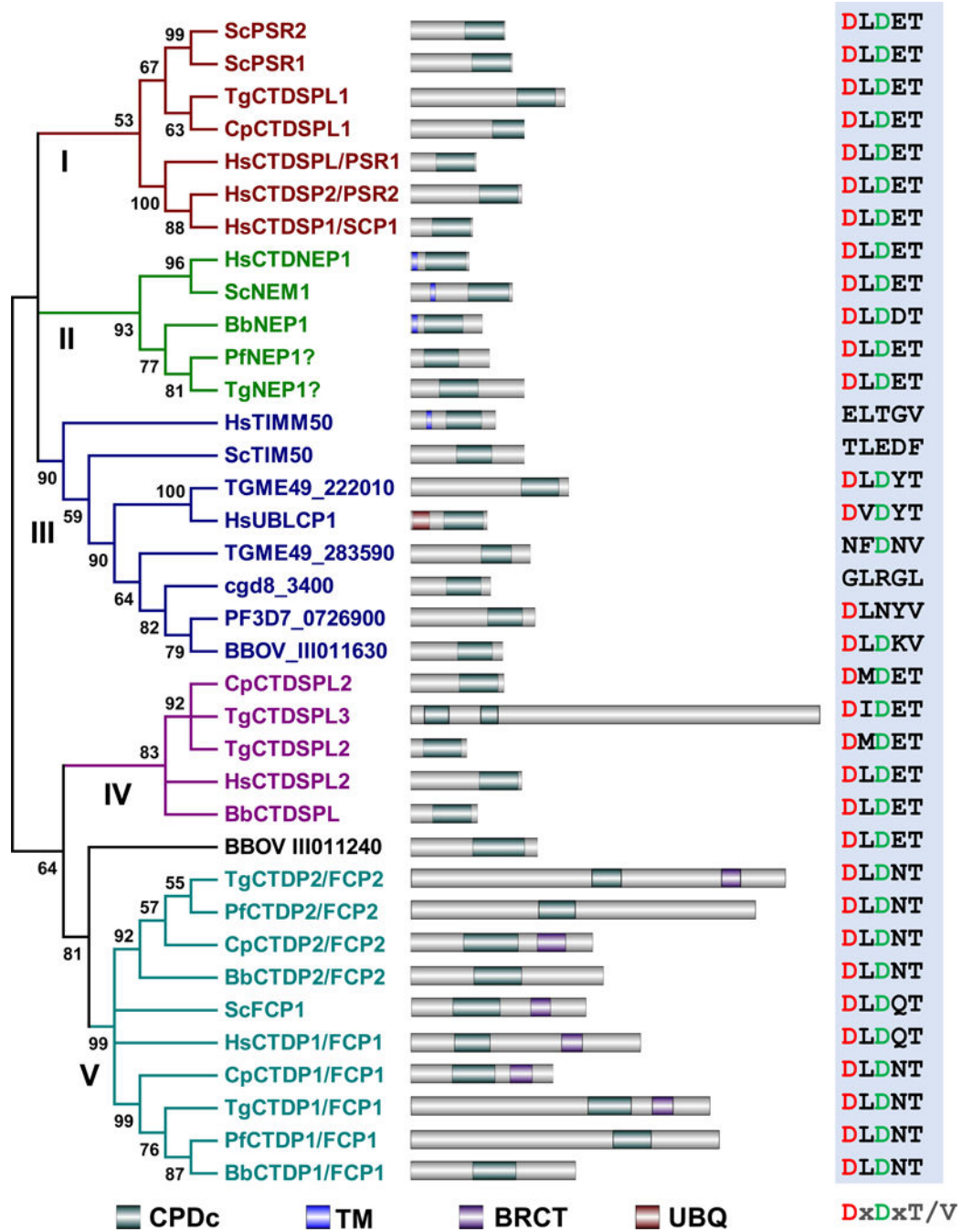


Figure 3. The FCP/SCP family phosphatases of *T. gondii*, *P. falciparum*, *B. bovis*, and *C. parvum*
 The phylogenetic tree shown was built based on the alignments of the catalytic domains of the FCP/SCP family phosphatases from humans, budding yeast, and the four apicomplexans studied. Alignments and tree assembly were done as for figure 1. The FCP/SCP family phosphatase are clustered into five groups, each annotated with a different color. Domain architecture and alignments of the core catalytic motif sequences are shown as in figure 1. Domains include CPDc, CTD-like phosphatase catalytic domain; TM, Transmembrane

domain; BRCT, Breast cancer carboxy-terminal domain; UBQ, Ubiquitin homologue domain.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Ser/Thr protein phosphatases of Human, *Arabidopsis*, budding yeast and four pathogenic Apicomplexa parasites

Table 1

The number of members within each phosphatase family/subfamily members are shown for human (*Homo sapiens*), a plant (*Arabidopsis thaliana*), yeast (*Saccharomyces cerevisiae*) and four apicomplexan parasites (*C. parvum*, *B. bovis*, *P. falciparum*, and *T. gondii*). PP1, protein phosphatase 1; PP2A, protein phosphatase 2A; PP2B, Protein phosphatase 2B aka calcineurin; PP4, protein phosphatase 4; PP5, protein phosphatase 5; PP6, protein phosphatase 6; PP7, protein phosphatase 7; PPKL, Kelch-like domain containing protein phosphatase; EFPP, EF-hand motif containing protein phosphatase; SLP, Shewanella-like protein phosphatases; PPM, protein phosphatases Mg²⁺/Mn²⁺ dependent; FCP/SCP, FCP/SCP family phosphatases.

	PP1	PP2A	PP2B	PP4	PP5	PP6	PP7	PPKL	EFPP	SLP	PPM	FCP
<i>H. sapiens</i>	3	2	3	1	1	1	2	0	0	0	20	8
<i>A. thaliana</i>	9	5	0	2	1	2	3	4	0	2	82	27
<i>S. cerevisiae</i>	1	2	2	1	1	1	0	0	0	0	8	5
<i>C. parvum</i>	1	2	1	1	1	0	0	1	1	1	14	5
<i>B. bovis</i>	1	1	0	1	1	0	1	1	1	0	4	6
<i>P. falciparum</i>	1	1	1	1	1	1	1	1	1	2	13	4
<i>T. gondii</i>	1	2	1	1	1	1	1	1	1	1	33	8

Table 2

PPP family members of the apicomplexans

The identity numbers for the PPP family members of the four apicomplexans analyzed, *T. gondii*, *P. falciparum*, *B. bovis*, and *C. parvum*, are shown. Human homolog genes are shown as references. When relevant the catalytic, structural and regulatory subunits are included. Because of the large quantity of PP1 regulatory subunit members only three are shown as examples. Asterisk mark indicates that identity as a particular type of phosphatase might be questionable based on homology level. Fitness scores of the *T. gondii* proteins obtained from the genome-wide CRISPR screen study (Sidik *et al.*, 2016) are included in parenthesis next to the gene IDs.

	Human gene	Alias	<i>C. parvum</i>	<i>T. gondii</i>	<i>P. falciparum</i>	<i>B. bovis</i>
PP1						
Catalytic subunits	PPP1CA, B, C		egd7_2670	TGME49_310700 (-4.72)	PF3D7_1414400	BBOV_III006130
Regulatory subunits (examples)	PPP1R2	I-2	egd8_390	TGME49_232760 (-0.55)	PF3D7_0320000	BBOV_II004630
	PPP1R7	SDS22	egd6_750	TGME49_267550 (-3.78)	PF3D7_1032800	BBOV_II003040
	PPP1R11	I-3	egd4_3940	TGME49_226825 (-2.11)	PF3D7_1031700	BBOV_III008470
PP2A						
Catalytic subunits	PPP2CA, B		egd7_810	TGME49_224220 (-1.50)		BBOV_I000740
Structural subunits	PPP2R1A, B	PR65	egd5_4070	TGME49_215170 (-4.39)	PF3D7_0925400	
Regulatory subunits	PPP2R2A, B, C, D	B55 or PR55	CMU_017370	TGME49_315670 (-4.03)	PF3D7_1319700	BBOV_II001110
	PPP2R3A, B	PR48, PR72/PR130	egd7_4970	TGME49_200400 (-3.17)	PF3D7_1356400	
	PPP2R3C	G5PR	egd6_4380			
	PPP2R4	PR53	egd1_2470	TGME49_283720 (-4.71)	PF3D7_1430100	BBOV_IV010340
	PPP2R5A, B, C, D, E	B56 or PR61	egd4_290	TGME49_246510 (-2.32)		BBOV_IV000610
	STRN, STRN3, 4	PR110, PR93				
PP2B						
Catalytic subunits	PPP3CA, B, C		egd6_4200	TGME49_311310 (-3.24)	PF3D7_0802800	BmR1_III07610
Regulatory subunits	PPP3R1, 2		egd1_1370	TGME49_213800 (-3.62)	PF3D7_1451700	
PP4						
Catalytic subunits	PPP4C		egd1_2360	TGME49_286210 (-4.79)	PF3D7_0927700	BBOV_III010990
Regulatory subunits	PPP4R1		egd6_5000*	TGME49_220910* (-2.05)	PF3D7_1034500*	
	PPP4R2					

	Human gene	Alias	<i>C. parvum</i>	<i>T. gondii</i>	<i>P. falciparum</i>	<i>B. bovis</i>
	PPP4R3A, B PPP4R4		cgd3_240*	TGME49_265440* (-1.90)		BBOV_II006300*
PP5	PPP5C		cgd2_2960*	TGME49_312200* (0.71)	PF3D7_1355500*	BBOV_IV000160*
PP6	Catalytic subunits Structural subunits					
	PPP6C			TGME49_301010 (-4.05)	PF3D7_0314400	
	ANKRD28			TGME49_216680* (0.90)	PF3D7_1313900*	
	ANKRD44					
	ANKRD52					
	PPP6R1					
	PPP6R2					
	PPP6R3			TGME49_319930* (-3.76)	PF3D7_0418100*	
PP7	PPP7C	PPEF1		TGME49_251850 (0.05)	PF3D7_1423300	BBOV_IV0005220
PPKL			cgd3_250	TGME49_290170 (-5.02)	PF3D7_1466100	BBOV_II001630
EFPP			cgd2_1640	TGME49_269460 (-0.81)	PF3D7_1018200	BBOV_III0005730
SLP			cgd8_290	TGME49_254770 (0.81)	PF3D7_1469200	
					PF3D7_1206000	

Table 3

PPM family members of the apicomplexans

Proteins with predicted PPM domains are shown in this table. The designated names of *Plasmodium* and *Toxoplasma* PPM family members are shown. Horizontal borders partition the homology relationships of the PPM family members among the four apicomplexans (no borders between P1PPM5 and P1PPM6 because they are paralogs). Asterisk mark indicates that identity is not certain based on homology levels. Fitness scores of the *T. gondii* proteins (Sidik *et al.*, 2016) are included in parenthesis next to the gene IDs.

<i>P. falciparum</i>	<i>T. gondii</i>	<i>C. parvum</i>	<i>B. bovis</i>
P1PPM1 PF3D7_0410300	TgPPM1	TGME49_224240 (-3.89)	egd7_3530
P1PPM2 PF3D7_1138500	TgPPM2A TgPPM2B	TGME49_232340 (-0.25) TGME49_267100 (-1.30)	egd3_2020 BBOV_I000710
P1PPM3 PF3D7_1455000	TgPPM3A TgPPM3B TgPPM3C TgPPM3D TgPPM3E TgPPM3F TgPPM3G TgPPM3H	TGME49_237500 (1.13) TGME49_276920 (0.75) TGME49_270320 (1.24) TGME49_202610 (0.96) TGME49_244450 (0.51) TGME49_278510 (1.01) TGME49_201520 (-0.25) TGME49_201630 (-0.06)	
P1PPM4 PF3D7_1249300	TgPPM4	TGME49_208500 (0.84)	egd7_3620 BBOV_I000510
P1PPM5 PF3D7_0810300	TgPPM5A TgPPM5B TgPPM5C	TGME49_272280 (-0.55) TGME49_318660 (0.83) TGME49_281580 (-1.06)	egd5_4170 egd7_4150
P1PPM6 PF3D7_1309200	TgPPM6	TGME49_293450 (1.75)	
P1PPM7 PF3D7_0810500	TgPPM7	TGME49_310760 (0.55)	BBOV_IV000460
P1PPM8 PF3D7_1135100	TgPPM8	TGME49_218590 (0.43)	egd3_1480
P1PPM9 PF3D7_0520100	TgPPM9	TGME49_220610 (1.23)	
P1PPM10 PF3D7_1009600	TgPPM10	TGME49_254410 (-1.89)	egd7_4640

<i>P. falciparum</i>	<i>T. gondii</i>	<i>C. parvum</i>	<i>B. bovis</i>
PF3D7_1208900*	TgPPM11A	TGME49_289490* (-2.46)	
PIPPM11	TgPPM11B	TGME49_223985* (-0.27)	
	TgPPM11C	TGME49_304955* (-1.17)	
PIPPM12	TgPPM12	TGME49_207590 (0.37)	cgd5_2900
	TgPPM13	TGME49_231850 (-4.39)	cgd5_840 BBOV_II001720
	TgPPM14	TGME49_232010 (-0.65)	
	TgPPM15	TGME49_236260 (0.05)	
	TgPPM16	TGME49_237410 (-2.01)	
	TgPPM17A	TGME49_265542 (-2.32)	
	TgPPM17B	TGME49_265650	
	TgPPM18	TGME49_270190 (-0.11)	
	TgPPM19	TGME49_275840 (-0.14)	
	TgPPM20, PP2Chn	TGME49_282055 (0.90)	
PF3D7_0615900*			
			cgd2_2280
			cgd6_440
			cgd6_1010
			cgd8_1430
			cgd7_4790*

Table 4
FCP/SCP family phosphatase members of the apicomplexans

The proteins of the four apicomplexans with predicted FCP/SCP family phosphatase domains are listed. Specific names are designated for the homologs among species that are shown in same rows. Once for which identity is not as clear include asterisks. Fitness scores of the *T. gondii* proteins (Sidik *et al.*, 2016) are included in parenthesis next to the gene IDs.

Designation	<i>T. gondii</i>	<i>P. falciparum</i>	<i>C. parvum</i>	<i>B. bovis</i>
CTDSPL1	TGME49_310660 (0.77)		cgd2_3810	BBOV_III004040
CTDSPL2	TGME49_263380 (0.65)		cgd4_3240	
CTDSPL3	TGME49_202550 (-0.39)			
NEP1	TGME49_243990* (-1.94)	PF3D7_0515900*		BBOV_II003180
CTDP1/FCP1	TGME49_269700 (-4.37)	PF3D7_1355700	cgd7_4250	BBOV_IV000150
CTDP2/FCP2	TGME49_228330 (-3.54)	PF3D7_1012700	cgd8_4810	BBOV_III000220
TIM50	TGME49_222010* (-0.63)			
	TGME49_283590 (-3.76)	PF3D7_0726900	cgd8_3400	BBOV_III011630
				BBOV_III011240

Table 5

Summary of the characterized protein phosphatases.

Name	ID	Localization	Regulators	Substrates	Function	References
TgPP1	TGME49_310700	Cytoplasm and Nucleus	TgLRR1		Host cell invasion; Essential in parasite growth	Delorme, V., et al. (2002); Daher, W., et al. (2007)
PfPP1	PF3D7_1414400	Cytoplasm and Nucleus; Maurer's clefts of infected erythrocytes	PfLRR1; Pf2; Pf3	PfSBP1	DNA synthesis and cell cycle progression	Bhattacharyya, M. K., et al. (2002); Kumar, R., et al. (2002); Blisnick, T., et al. (2006); Daher, W., et al. (2006); Freville, A., et al. (2012); Freville, A., et al. (2013); Freville, A., et al. (2014); Hollin, T., et al. (2016)
PfPP2A	PF3D7_0925400		PfARP; PfPTPA		Sexual stage development	Li, J. L. and D. A. Baker (1997); Dobson, S., et al. (1999); Dobson, S., et al. (2003); Vandamme, A., et al. (2014)
TgPP2B or TgCnA	TGME49_311310	Cytoplasm and enriched in apical end for extracellular parasites	TgCnB		Stabilize attachment of extracellular parasites to host cells	High, K. P., et al. (1994); Moudy, R., et al. (2001); Paul, A. S., et al. (2015)
PfPP2B or PfCnA	PF3D7_0802800	Cytoplasm	PfCnB		Stabilize attachment of extracellular parasites to host cells	Bell, A., et al. (1994); Berrinan, M. and A. H. Fairlamb (1998); Kumar, R., et al. (2005); Singh, S., et al. (2014); Paul, A. S., et al. (2015); Philip, N. and A. P. Waters (2015)
PfPP5	PF3D7_1355500	Cytoplasm and Nucleus	PfHsp90			Dobson, S., et al. (2001); Lindenthal, C. and M. Q. Klinkert (2002)
PfPP7	PF3D7_1423300	Cytoplasm			Ookinete development	Dobson, S., et al. (2001)
PfPPKL	PF3D7_1466100					Guttery, D. S., et al. (2012); Philip, N., et al. (2012)
PfSLP1	PF3D7_1469200				Microneme Development; Ookinete development	Patzewitz, E. M., et al. (2013)
PfSLP2	PF3D7_1206000	The apical end of invasive merozoites and unknown vesicles		PfBand3	Regulation of Band 3 Dynamics during Parasite Invasion	Fernandez-Pol, S., et al. (2013)
PfPPM2	PF3D7_1138500			PfEF-1β	Gametocyte sex allocation and ookinete differentiation	Mamoun, C. B., et al. (1998); Mamoun, C. B. and D. E. Goldberg (2001); Guttery, D. S., et al. (2014)
PfPPM5	PF3D7_0810300				Oocyst development	Guttery, D. S., et al. (2014)
TgPPM13	TGME49_231850			Toxofilin	Regulation of phosphorylation status of Toxofilin	Delorme, V., et al. (2003); Jan, G., et al. (2007)
TgPPM20 or TgPP2Cmn	TGME49_282055	Rhoptries and host nucleus				Gilbert, L. A., et al. (2007)