

## MINIREVIEW

# Forward Genetics in *Toxoplasma gondii* Reveals a Family of Rhopty Kinases That Mediates Pathogenesis<sup>∇</sup>

L. David Sibley,<sup>1\*</sup> Wei Qiu,<sup>2</sup> Sarah Fentress,<sup>1</sup> Sonya J. Taylor,<sup>1</sup> Asis Khan,<sup>1</sup> and Raymond Hui<sup>2</sup>

*Department of Molecular Microbiology, Washington University School of Medicine, St. Louis, Missouri 63130,<sup>1</sup> and Structural Genomics Consortium, University of Toronto, Toronto, Ontario M5G 1L7, Canada<sup>2</sup>*

*Toxoplasma gondii* is a widespread protozoan parasite that commonly infects domestic, wild, and companion animals (8). Like related tissue coccidians, *T. gondii* has a complex life cycle alternating between a sexual cycle, which occurs only within enterocytes of the small intestines of cats (all members of the Felidae appear to be susceptible), and asexual propagation in a variety of warm-blooded vertebrate hosts (8). The sexual phase culminates with fecal shedding of a spore-like stage called the oocyst, which is diploid and undergoes meiosis in the environment. During the asexual cycle, the parasite interconverts between a fast-growing lytic form known as the tachyzoite and a slow-growing, semidormant form called the bradyzoite. Tachyzoites propagate rapidly in virtually all types of nucleated cells, including macrophages, while differentiation to bradyzoites is favored in long-lived, terminally differentiated host cells (49). Tachyzoites are adept at direct migration across cellular barriers and also disseminate rapidly within leukocytes, thereby reaching sites of immune privilege such as the central nervous system and the developing fetus, where they are more likely to cause disease (2). The life cycle shows remarkable flexibility between lytic and dormant states, thus facilitating asexual transmission between intermediate hosts. This adaptation may account for the recent spread and emergence of a few dominant clonal groups within North America and Europe (46).

*Toxoplasma* is well adapted to mammalian hosts, being transmitted by ingestion of undercooked meat harboring tissue cysts and through food and water supplies contaminated with oocysts shed from cats (8). Humans are accidental hosts of *T. gondii*, yet seroprevalence rates indicate high rates of chronic infection in many countries of Europe and Central and South America (16). Infections are often mild or subclinical in healthy adults; however, toxoplasmosis can present as a clinically important infection in immunocompromised patients and the developing fetus (32). In many regions of the world, toxoplasmosis remains a frequent problem in patients infected with human immunodeficiency virus, due to lack of access to effective antiviral therapy (29). Additionally, in regions such as

southern Brazil, ocular toxoplasmosis often presents as a clinically severe infection in otherwise healthy adults (18).

Although *T. gondii* is primarily an opportunistic pathogen, it has emerged as a model for study of the biology of apicomplexan parasites, a group that contains *Plasmodium* spp. (malaria parasites), *Cryptosporidium* spp., and a variety of animal pathogens. Although the life cycles of these parasites differ substantially, they share common pathways for actin-myosin-based motility, calcium-dependent secretion, and active cell invasion (41). *Toxoplasma* offers excellent tools for studying molecular and cell biology, forward and reverse genetics, and animal models. Hence, exploration of the molecular basis of complex traits such as pathogenesis has been feasible in this system. This review summarizes recent advances using forward genetics to identify genes involved in pathogenesis, as well as the use of reverse genetics to validate the roles of prospective candidate genes.

### STRAIN DIVERSITY AND GENETIC ANALYSIS IN *T. GONDII*

*Toxoplasma gondii* has a highly unusual population structure comprising three predominant clonal populations in North America and Europe (42). These lineages are found in overlapping geographic regions, where they infect a variety of warm-blooded animals, including humans. Despite being maintained as clonal lineages, they are highly similar genetically and originated from several closely related parental types in the wild (6). Estimates of their last common ancestry indicate that this event occurred within the past 10,000 years, followed by rapid expansion of these three lineages (46). The cause of this unusual pattern is uncertain, but it may indicate a localized bottleneck followed by the expansion of a few lineages, or a strong genetic sweep that replaced abundant and more-diverse isolates. Intriguingly, this recent origin is also marked by the common inheritance of a single monomorphic chromosome Ia that is shared by the three clonal lineages (19). The population structure differs substantially in South America, where lineages are genetically more diverse and undergo both sexual recombination and asexual transmission (20, 28).

While the origins of the unique population structure are uncertain, it has profound consequences for population studies. For example, association studies generally are not informative for defining the contributions of genes to particular phenotypes, since linkage disequilibrium is so high. However, a

\* Corresponding author. Mailing address: Department of Molecular Microbiology, Washington University School of Medicine, St. Louis, MO 63130. Phone: (314) 362-8873. Fax: (314) 286-0060. E-mail: sibley@borcim.wustl.edu.

<sup>∇</sup> Published ahead of print on 22 May 2009.

number of strong phenotypic differences have been noted; for example, type I strains are acutely virulent in outbred mice (see below). Additionally, type I strains have been shown to exhibit enhanced migration across biological barriers *in vitro* and to disseminate more rapidly *in vivo* in the mouse model than other strains (3). It is not clear how these traits translate to human infections, which are caused primarily by type II strains, although several reports have indicated an increased frequency of type I strains (or strains bearing type I alleles) among clinical isolates (11, 14, 21). Although the laboratory mouse appears to be highly sensitive to toxoplasmosis compared to humans, it provides a useful starting point for investigating genes that contribute to pathogenesis.

Classical genetic crosses in *T. gondii* were pioneered by the studies of Elmer Pfefferkorn and his colleagues, who established procedures for crossing strains of the parasite by coinfecting cats and collecting the resulting recombinant oocysts. These early studies established several key properties, including the following. (i) A single cloned tachyzoite can give rise to the entire life cycle, indicating that there is not a predetermined mating type (34). (ii) The replicating stages are all haploid, and mating leads to a diploid stage (35). (iii) When coinfecting a single cat, strains undergo self-mating or outcrossing at approximately equal frequencies (35). (iv) Unlinked genes segregate randomly and yet can demonstrate epistatic interactions (33). These studies provided the beginnings of a working genetic system, which was later expanded by the development of molecular markers, thus allowing the construction of rudimentary linkage maps (44). Further expansion of the number of markers was used to support preliminary mapping of acute virulence in the mouse model; remarkably, virulence was linked to a single major locus on chromosome VII (since renamed VIIa) (47). This study was also notable because it demonstrated that acute virulence was a genetically heritable trait and suggested that fine mapping could likely be used to identify the specific gene(s) involved. For the current version of the genetic map, the segregation of more than 200 markers was compared among crosses between either types II and III or types I and III, to generate linkage maps for the 14 chromosomes (23). This composite map defined the order of markers on the chromosomes, established the basic parameters for recombination, and provided a framework for assembly of the *T. gondii* genome, which is approximately 65 Mb (<http://ToxoDB.org>). The composite genetic map consists of ~600 cM (1 cM = 1% recombination), with an average map unit of around 100 kb. Although the markers are fairly widely dispersed, the most limiting attribute of genetic mapping in *T. gondii* is the relatively modest recombination rate; hence, mapping precision is limited by the number of available progeny rather than by the frequency of markers. The various advantages and limitations of forward genetics in *T. gondii* have been discussed previously in more detail (1).

#### MAPPING ACUTE VIRULENCE IN THE MOUSE MODEL

*Toxoplasma gondii* naturally infects a variety of rodents, and laboratory mice provide a reasonable model for both acute and chronic infections (38). Type I strains are acutely virulent in mice: limiting dilution indicates that a single tachyzoite administered by needle inoculation is lethal in many laboratory

mouse strains, including outbred mice (43). In contrast, type II and III strains are nonvirulent in outbred mice, although differences can be discerned in the susceptibilities of inbred mice (38). Death of the animals results from overwhelming cytokine production, an outcome that occurs following inoculation with low doses of type I strains but only after high doses of type II strains (12, 30). While needle inoculation bypasses the natural oral route of infection, it has the advantage of being easily quantifiable and reproducible and giving very discrete endpoints. For these reasons, genetic mapping studies of complex phenotypes have relied largely on this model. Additional phenotypic differences include the variable induction of cytokines following *in vitro* infection. For example, type II strains are potent inducers of interleukin 12 (IL-12) from naive macrophages, while type I and III strains are relatively quiescent (39). Differences in the activation of pathways for IL-12 induction have also been noted: type I strains rely primarily on a non-MyD88-dependent pathway, while type II strains trigger both dependent and independent pathways, leading to greater overall levels of IL-12 (24). While this seems counter to the results described above for *in vivo* infection (type I leads to greatly elevated cytokine levels later in infection), differences including growth rate (37), escape from cellular destruction (50), and ability to disseminate rapidly (3) also contribute to different outcomes *in vivo*.

Forward genetic analysis of acute virulence in the murine model was based on a genetic cross between the acutely virulent type I strain called GT-1 (marked by resistance to fluoro-deoxyuridine) and the type III strain CTG (marked by resistance to adenine arabinoside and sinesfungin) (47). Progeny clones from this cross were tested in outbred mice to define acute virulence based on cumulative mortality during the initial phase of infection following different doses of tachyzoites. The progeny were also genotyped in order to analyze the segregation of acute mortality with specific markers. Genetic associations of other traits were also analyzed based on the ability to migrate across polarized epithelia (transmigration) and the ability to migrate away from a focus of infection, traits that had previously been associated with type I strains (3). Differences in growth rate were also analyzed based on previous findings that type I strains have a shorter cell cycle (37). Finally, we examined the serological responses in surviving animals, since less-virulent strains readily give rise to chronic infections, while acutely virulent type I strains do not (43). Because these traits were likely to arise as the result of a combination of genetic effects, the phenotypes were analyzed using quantitative trait locus (QTL) mapping (26, 27), a statistical method for defining the association of complex phenotypes with multiple underlying genetic regions. Analyses of 34 genetically independent recombinant progeny were based on 175 of the genetic markers from the composite genetic map that were informative in the type I-versus-type III cross (the remaining markers being identical between these two lineages but different in type II) (23).

Genomewide association studies revealed the remarkable finding that nearly all of the phenotypes described above were strongly associated with the same major peak on chromosome VIIa (48). This was highly unexpected, since the assumption was that such complex biological phenotypes would be multigenic and controlled by a number of separate loci. Instead, a

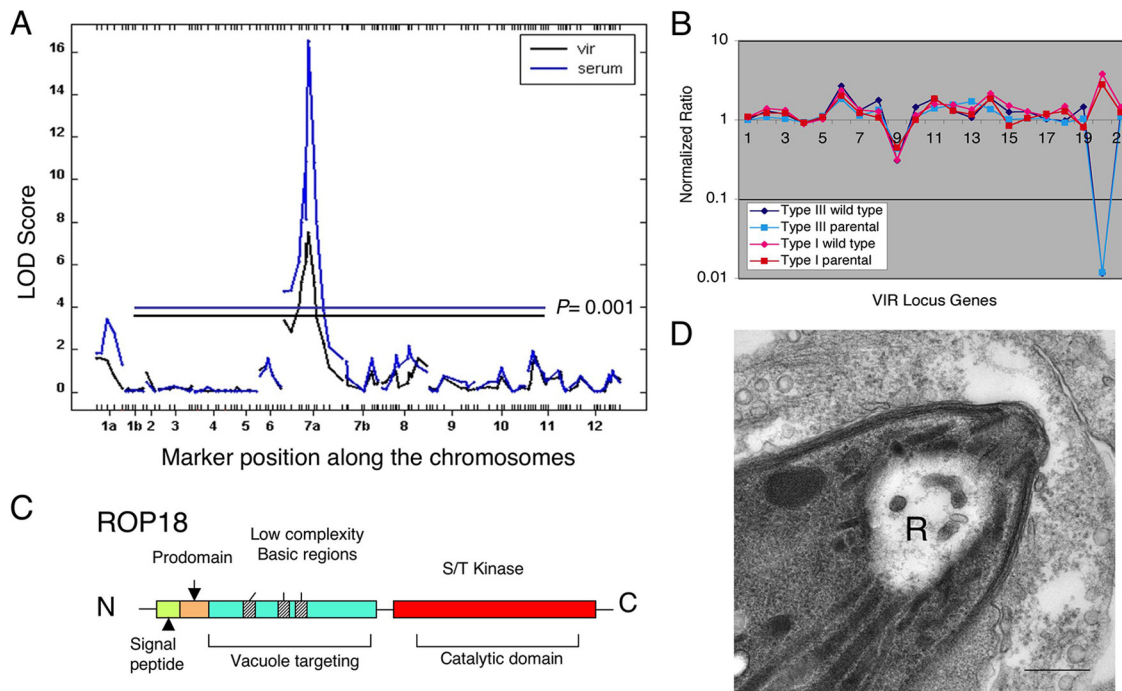


FIG. 1. Genetic mapping of ROP18 as a major virulence determinant in the mouse. (A) Whole-genome analyses of the traits of acute mortality (*vir*) and serum responsiveness (*serum*) show a strong peak on chromosome VIIa with a minor peak on chromosome Ia. The x axis shows the 175 informative markers aligned across the 14 chromosomes. LOD score, log odds ratio. (B) Microarray analysis of a cluster of 21 genes from the central region of the QTL shown in panel A. Wild-type strains are untagged strains of the type I (GT-1) and type III (CTG) backgrounds, while “parental” refers to the drug-resistant strains used in the cross. (Reproduced from reference 48.) (C) Schematic of ROP18 showing the signal peptide, processing site, N-terminal domain containing low-complexity regions, and S/T kinase domain. (D) Discharge of rhoptries accompanies cell invasion. An empty rhoptry profile of a recently invaded tachyzoite (R) is visible. Bar, 200 nm.

single locus was estimated to be responsible for 90% of the migration, transmigration, mortality, and serum response traits (a plot of two of these traits is shown in Fig. 1A). The only other significant QTL associated with virulence was a region on chromosome Ia that was positively associated with chromosome VIIa (Fig. 1A). This minor QTL on chromosome Ia has not been localized to a specific gene, but it is intriguing that it lies on a chromosome that has been shown to be recently inherited and essentially monomorphic among the clonal lineages (19). In contrast to the single major peak for virulence, growth was multigenic, with five to six separate QTLs showing statistical support, one of which coincided with the acute mortality peak on chromosome VII (48).

Initially, resolution of QTLs for acute virulence and migration was limited by the density of genetic markers, which only allowed mapping to approximately 1.5 to 2 Mbp. Finer mapping of these traits was accomplished by including additional markers on the central region of chromosome VIIa and by identifying several additional progeny that had crossovers in this region. This led to the resolution of distinct QTLs, one of which controlled both migration traits while a second controlled acute mortality and the serum response traits (48). Because the *T. gondii* genome scaffolds had been assembled using the genetic map (23), it was possible to pin these regions directly to the genome and hence to identify candidate genes. The region encompassing the mortality QTL comprised 140 kb and contained 21 predicted genes. Analysis of their expression profiles by microarray hybridization revealed a starkly different

expression level for one gene at the end of this region (Fig. 1B). This gene was named *ROP18* based on the sequential numbering of rhoptry proteins (ROPs) as they were described either from a proteomic study or through functional analyses (7, 9). Inspection of the predicted open reading frame revealed several intriguing features (Fig. 1C). First, ROP18 contained a signal sequence and was predicted to be a secretory protein. Second, it contained a conserved prodomain and processing site, typical of the ROP2 family of proteins (9). Third, the N-terminal region contained a cluster of three low-complexity regions typified by amphipathic alpha-helical regions. Finally, it contained a serine threonine (S/T) kinase domain that was predicted to be catalytically active. ROPs were known to be secreted from the rhoptries into the host cell at the time of invasion and to accumulate in the parasite-containing vacuole (15) (Fig. 1D). Hence, it was reasonable to suspect that this candidate might behave similarly and that its vital function in contributing to virulence might be related to its kinase activity.

**IDENTIFICATION OF ROP KINASES AS MEDIATORS OF PATHOGENESIS**

The ~100-fold-lower expression of ROP18 in type III strains provided a convenient background for testing activity by gain of function. This difference proved fortuitous, since allelic replacements are relatively difficult to perform in *T. gondii*, while expression of exogenous genes is straightforward. Transgenic lines expressing epitope-tagged ROP18 from the type I lineage

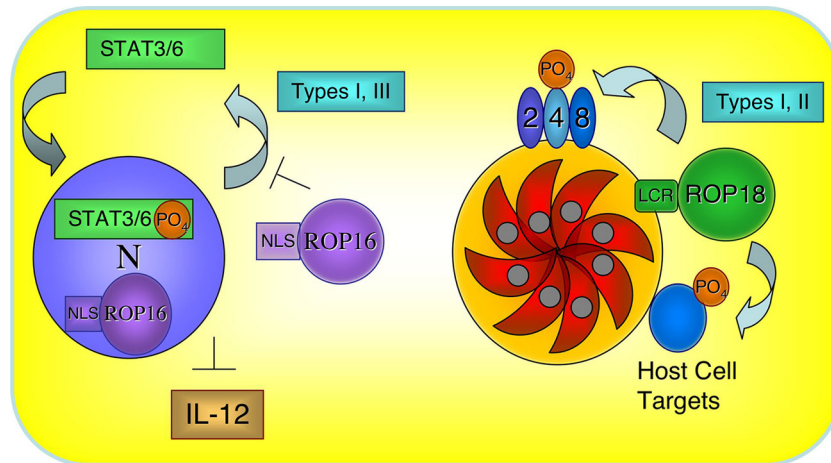


FIG. 2. Model for the proposed roles of ROP16 and ROP18. (Right) ROP18 is highly expressed in type I and II strains, and following secretion it is targeted to the PVM by a series of low-complexity helical regions (LCR) in the N terminus. ROP18 has been shown to phosphorylate other ROPs in vitro, suggesting that it may also perform this function in infected cells (i.e., ROP2, -4, and -8, as shown). Additionally, host cell proteins may be targets of ROP18. The virulence-enhancing potential of ROP18 requires kinase activity and has been associated with slightly enhanced growth. (Left) ROP16 is targeted to the host cell nucleus (N) by a nuclear localization sequence (NLS), although it may also be active in the cytosol. ROP16 from type I and III strains induces prolonged activation of STAT3/6, thus activating gene expression in the host cell nucleus but blocking the production of IL-12.

(ROP18I) were created in the type III CTG line. Correct targeting to rhoptries and secretion into the host cell during invasion were demonstrated for the transgenic line (48). Testing of the acute virulence phenotype of these transgenic lines established that ROP18 was able to recapitulate the high-mortality phenotype characteristic of type I strains and that this ability required the catalytic activity of the kinase (48). This finding established that the major virulence determinant distinguishing between type I and III strains in terms of acute mortality in the mouse model is a single gene encoding ROP18. In a separate study, it was shown that overexpression of this type I allele in the type I RH strain slightly enhanced parasite replication in vitro and that this activity was also dependent on the kinase activity (10). Following the secretion of ROP18 into the host cell, it is targeted back to the parasite-containing vacuole (48) (Fig. 2). Ectopic expression of ROP18 in BHK cells also leads to targeting to the parasite-containing vacuole membrane (PVM) (10), and this effect is mediated by the low-complexity region in the N terminus of the protein (Fig. 1C and 2) (25). This strongly implies that the critical targets of ROP18 phosphorylation are found on the surface of the PVM (Fig. 2). In vitro studies discussed below implicate other ROPs as potential targets, although the identities of substrates in vivo have remained elusive.

Differences in pathogenicity have also been recognized between other lineages of *T. gondii*, including differences among the progeny of a cross between type II and III strains (13). While both of these parental strains are avirulent in outbred mice, differences can be recognized in various inbred lines, which are less resistant to infection. Mapping of QTLs that controlled differences in pathogenesis among progeny from this cross identified five distinct QTLs, although most of these had only modest statistical support (40). Within a relatively minor QTL seen on chromosome VIIa, ROP18 proved to be the functionally important difference. The presence of an insert of DNA upstream of *ROP18* in the type III lineage was

associated with underexpression. In a result similar to that described above for the type I allele of ROP18, expression of transgenes of ROP18II in the type III background led to a gain of virulence, although this time in inbred BALB/c mice (40). A second QTL on chromosome VIIb was tracked down to a related rhoptry kinase called ROP16. In this case, introduction of the type I or III allele into the type II background attenuated virulence, consistent with the prediction of the QTL (40). Further characterization of this protein revealed that it acts by prolonging the phosphorylation of STAT3 and STAT6 (STAT3/6), hence activating some host genes but suppressing immune responses such as the induction of IL-12 (40). Type I and III strains share a highly similar allele that shows this activity, while the type II allele differs substantially and leads to only transient STAT3/6 phosphorylation (40) (Fig. 2). ROP16 appears to partially account for the previously described difference in the production of IL-12, which is induced by type II strains but not by type I or III strains (39). ROP16 is predicted to be an active kinase, and although its direct targets are unknown, it is translocated to the host cell nucleus following injection into the host cell (40). While the other QTLs identified in the type II  $\times$  type III cross have not been fully resolved, preliminary mapping of the QTL on chromosome XII identified a cluster of several nearly identical genes encoding ROP5 paralogues, and these are currently candidates for this locus (J. C. Boothroyd, personal communication).

Collectively these studies revealed that members of the ROP kinase family contribute to differences in pathogenesis between different strains of *T. gondii* in the mouse model. Comparison of these two genetic crosses has shown that the roles of individual genes can range from small to large quantitative effects, and interactions between the loci are also important in controlling the overall level of pathogenicity. For example, QTL mapping indicates that ROP18 is the major locus controlling differences in acute virulence between the type I and III lineages and that it also contributes to differences between

types II and III. In contrast, ROP18 contributes only partially to growth differences between type I and type III lineages, and multiple other loci also contribute to this trait (48). This result indicates that enhanced virulence is not simply the result of faster replication, although this is certainly one factor that influences pathogenesis in the mouse model. Additionally, as predicted from the genetic mapping studies, ROP18 plays no role in the enhanced migration phenotype, which also maps to chromosome VIIa (supplemental data in reference 48). Within this migration locus are a number of genes that might contribute to actin-myosin-based motility, and the precise genetic control of this trait is still under investigation. It is conceivable that traits controlled by this migration QTL may also affect the severity of disease, for example, by advancing the onset of infection.

Following the discovery that ROP18 mediates differences in acute virulence between different laboratory strains, we also investigated to what extent these properties were conserved among natural isolates. Not surprisingly, given the extremely clonal population structure in North America and Europe, other isolates of the canonical type I, II, and III lineages share the same alleles and expression profiles as the laboratory strains used in the genetic crosses (22). Hence, it is highly likely that ROP18 contributes to the pathogenesis of these isolates in a manner similar to that described for the common laboratory isolates. However, in other regions, such as South America, the population consists of a number of unique lineages, some of which have undergone greater recombination in the wild (20, 28). To analyze the diversity of ROP18 among these populations, we compared genetic diversity to a set of selectively neutral introns as well as housekeeping genes and other antigens (22). Remarkably, although genetic diversity among these isolates was high, ROP18 displayed only three distinct alleles, corresponding to those already described for the North American lineages (22). Furthermore, these alleles have coexisted over a long period of evolutionary time and show very different profiles in terms of diversifying selection. While the type III allele is the oldest, it shows little evidence of positive selection. In contrast, the type I and II alleles show high levels of diversifying selection. This is correlated with major differences in expression levels that are associated with the presence of an upstream region in the type III allele. Comparison to the outgroup *Neospora caninum* revealed that this upstream region is ancestral (22). Taken together, these findings suggest that deletion or rearrangement of this upstream region led to up-regulation of ROP18 and promoted diversifying selection (22). These findings also suggest that the stable coexistence of three subtypes results from adaptations to different niches in the environment. The long-term maintenance of the type III allele, which is associated with low levels of virulence, suggests adaptation for hosts where this trait is advantageous (or where high virulence is disadvantageous). In contrast, the more recent expansion and widespread nature of the type II and I alleles, which show moderate and high levels of virulence, respectively, suggest that these traits are adaptive in a different niche. While the phenotypes of ROP18 have been studied primarily in the laboratory mouse, this host is highly susceptible to toxoplasmosis relative to wild rodents. Thus, it may be that enhanced expression of ROP18 is an adaptive trait for the infection of

natural hosts, where rather than contributing to pathology, it enhances transmission.

### DIVERSITY AMONG THE ROP KINASES

The kinase domains of S/T kinases are characterized by the conservation of 12 subdomains that have been defined by the comparison of a large number of kinases in eukaryotic cells (17). Within these subdomains, specific residues have been implicated in functional roles based on structural information, sequence conservation, and mutational data. Highly conserved residues are shown in Fig. 3A, including a glycine-rich loop that binds the  $\alpha\beta$  phosphates of the nucleotide. As well, a critical lysine in subdomain II of protein kinase A (PKA) (K72), combined with an aspartate in the catalytic domain (D166) and a second aspartate (D184) involved in  $\gamma$  phosphate transfer, defines a key catalytic triad. Mutation of these residues abrogates activity, and their strong conservation among active kinases argues for their essentiality (17). However, there are exceptions to this pattern; for example, WNK (with-no-lysine kinase) lacks the K72 residue in subdomain II yet is active, relying on a separate K residue in subdomain I to substitute in this critical role (5). In the ROP2 family, most of the members are predicted to be catalytically inactive due to divergence at residues that are normally required for activity (Fig. 3A). Such pseudokinases have been shown in other systems to have regulatory roles, acting either to form complexes with active kinases and control their functions, to recruit substrates, or to form molecular scaffolds that regulate active kinases (5). Remarkably, the changes that are predicted to disrupt catalytic activity in ROPs are conserved between the different parasite lineages, indicating that they have been around for some time. Unlike ROP16 and ROP18, which are highly divergent between the lineages, ROP2, ROP4, and ROP8 are extremely conserved. ROP2 and ROP8 are also highly similar to each other and likely arose by a gene duplication event, since they lie in tandem in the genome upstream of a third member that is predicted to be a pseudogene (ROP2B). In contrast, ROP16 and ROP18, discussed above, as well as ROP17, retain the catalytic triad and show conservation of other residues implicated in function (Fig. 3A). In addition to those ROP members highlighted here, there are a large number of other paralogues in the genome, some of which are predicted to be pseudokinases while others are likely catalytically active.

The ROP family of S/T kinases is highly divergent compared with eukaryotic kinases, which have previously been grouped into four major classes (Fig. 3B). The cyclic-AMP-dependent (PKA), cGMP-dependent (PKG), and PKC kinases make up the AGC family. The calmodulin-regulated kinase family (CAMK) includes kinases regulated by calcium-calmodulin. Cyclin-dependent kinases, mitogen-activated protein kinases, and casein kinases make up the GMGC family. The final group, the TK family, includes the tyrosine kinases and tyrosine-like kinases. The phylogenetic divergence of the ROPs suggests that they are likely to have unique substrates and mechanisms of activation. While ROP18 is closely related to ROP5 and is a member of the ROP2 family, ROP16 and ROP17 are only distantly related to the ROP2 family (Fig. 3B). The ROP kinases also are not found in *Plasmodium falciparum*, which instead has a family of secretory S/T kinases referred to as FIKK kinases, named for a stretch of amino acids

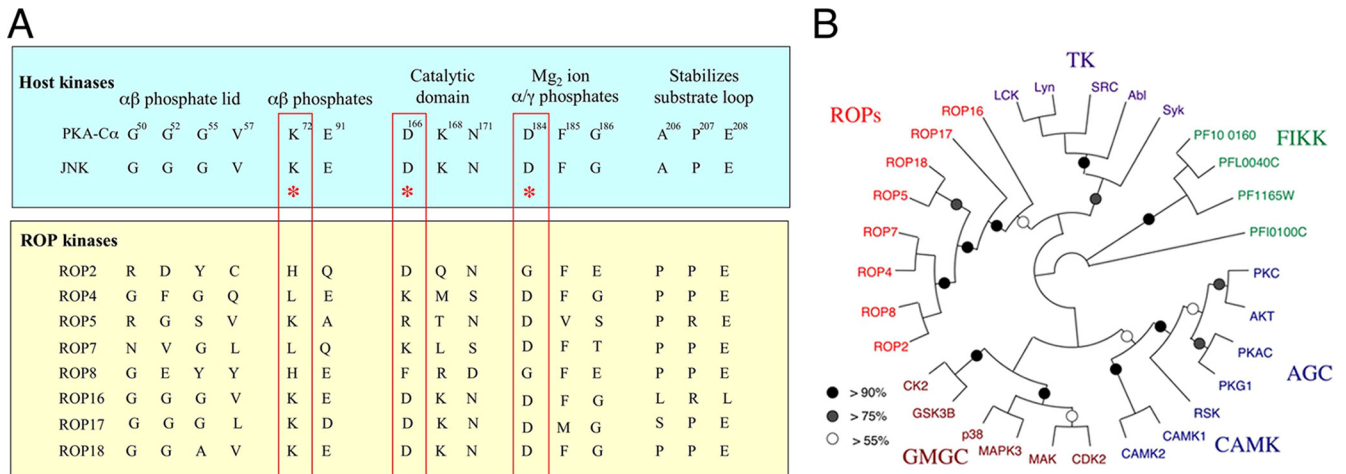


FIG. 3. Divergence of ROP2 family members and degeneracy in the conserved kinase domains. (A) Alignment of several mammalian kinases (PKA-Cα and Jun N-terminal protein kinase [JNK]) showing conservation of key residues implicated in binding to ATP and in catalysis. The catalytic triad is boxed in red. While a majority of ROP2 family members are divergent and are predicted to be inactive, ROP16, ROP17, and ROP18 conserve the key residues associated with activity. (B) Phylogenetic tree showing the relationship between ROP kinases, FIKK kinases from *P. falciparum*, and major families of human kinases. TK, tyrosine kinases; AGC, cAMP-regulated kinase, c-GMP regulated kinase, and PKC; CAMK, calcium-regulated kinases; GMCC, cyclin-dependent, mitogen-activated, and casein kinases. (Reproduced from reference 36.)

that occurs in their N termini (31). FIKK kinases are exported via a PEXEL (*Plasmodium* export element) motif and are found in the infected red blood cell in association with a membranous system, although their role there is not understood (31). ROP kinases are found in related tissue coccidia, notably *Neospora caninum*, for which there is a genome project (<http://www.sanger.ac.uk/sequencing/Neospora/caninum/>).

Early studies of ROP2 revealed that it was targeted to the PVM and suggested, based on differential staining with antibodies to the N and C termini, that it adopted a transmembrane topology (4). The N-terminal region of ROP2 was also implicated in binding to the host endoplasmic reticulum and mitochondria, which are recruited into close proximity with the PVM (45). More-recent analysis of members of the ROP2 family suggests that they adopt a kinase fold and that the originally proposed transmembrane domain is in fact buried in the core of the C lobe of the kinase domain (9). This suggests an alternative explanation for the detergent accessibility experiments described above based on protein folding rather than membrane topology. Furthermore, while the N-terminal amphipathic alpha-helical regions are involved in recruitment to the PVM (9), they may also participate in organelle recruitment, especially if ROPs form complexes at this interface. In this regard, ROP18 is more similar to the ROP2 family in that it conserves this N-terminal low-complexity region, which is important in binding to the PVM (Fig. 1C), while ROP16 does not.

#### UNIQUE STRUCTURAL AND REGULATORY FEATURES OF ROP KINASES

Initial attempts to model the kinase domain of ROPs provided only low-homology models based on comparison to known kinases (9, 48), in part due to the sequence divergence within this family. This problem was resolved by the recent solution of X-ray crystal structures for the kinase domains of ROP8 (Protein Data Bank [PDB] accession no. 3BYV) and

ROP2 (PDB accession no. 3DZO and 2W1Z) (25, 36). Importantly, these structures do not have the N-terminal domains that contain amphipathic regions in ROP2-related proteins. Rather, they are based on the conserved kinase domain plus a short N-terminal extension of ~50 amino acids that was found to be necessary for the stability and solubility of the kinase domain when expressed in *Escherichia coli* (36). From these studies it is clear that ROPs share a common structural fold with S/T kinases, composed of an N lobe, a narrow hinge, and a C lobe (Fig. 4A). In addition, ROPs contain several insertions not seen in other kinases. The most unusual of these is the short N-terminal extension that forms a β-sheet completing the N lobe and also includes two helices running down the face of the kinase and wrapping around the hinge region (Fig. 4A). Although the 12 subdomains that define S/T kinases are conserved, a number of structural features indicate that the nucleotide-binding pockets of ROP2 and ROP8 are unable to bind ATP (25, 36). Additionally, a highly conserved cysteine pair in subdomains IX and X forms a disulfide bond that is likely conserved within the family (36). Within the ROP2 family, the conserved PPE motif on the substrate binding loop strongly kinks the helix in domain VII and makes a hydrophobic interaction with the indole ring of a tryptophan from subdomain IX (36). This interaction is also stabilized by an interdomain salt bridge, providing a rigid platform for the substrate-binding region. It is noteworthy that the primary differences between ROP2 and ROP8 lie along the surfaces of this substrate-binding region, which differ in charge and hydrophobicity (36). This conservation predicts that the ROPs are involved in binding to different substrates and that they may be biologically, even if not catalytically, active.

The unique N-terminal extension of the ROP kinase domains also suggests a unique mechanism of regulation (36). The two alpha-helical regions wrap the kinase, fitting into a hydrophobic cleft formed by subdomains V and VIa. As a result of this close proximity, the side chain of Arg228 in ROP8

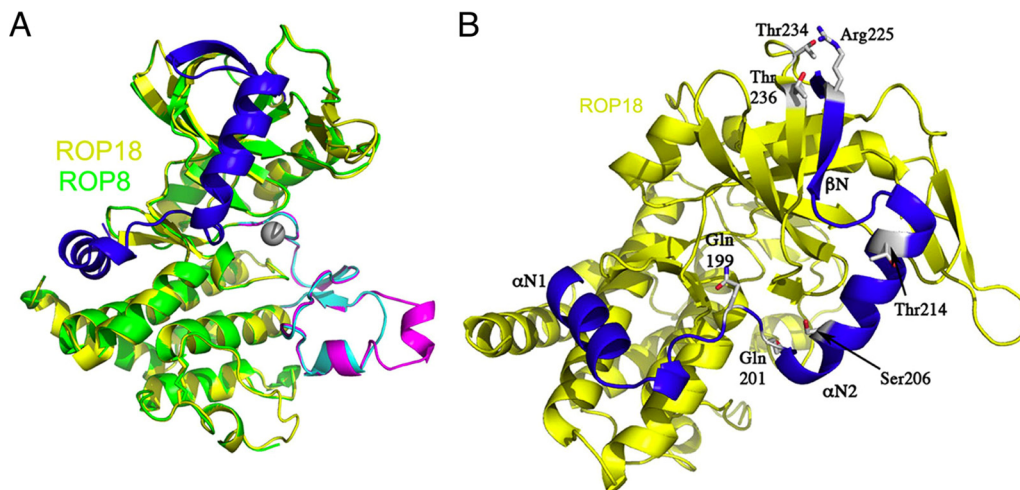


FIG. 4. Ribbon models of the ROP8 and ROP18 kinase domains. (A) Model of ROP18 derived by homology modeling (yellow), superimposed on a model derived from the X-ray crystal structure of ROP8 (green). The N-terminal extension of the kinase domain is found in both proteins and is shown in blue. This extension consists of a  $\beta$ -sheet that completes the N lobe and two alpha helices that wrap the main kinase domain. Substantial differences also occur in the substrate-binding domain, where ROP18 is shown in cyan and ROP8 in magenta. (B) Model of ROP18 showing residues that are phosphorylated and implicated in the regulation of activity. (Reproduced from reference 36.)

(Arg220 in ROP2) protrudes into the nucleotide-binding pocket, precluding ATP binding (36). Although this is inconsequential for ROP2 and ROP8, because they are inactive pseudokinases, homology studies predict a similar N-terminal subdomain in ROP18, with Gln199 instead of an arginine occluding the conserved ATP binding pocket (25, 36). The positioning of these N-terminal helices is stabilized by a number of other interactions. These sequence motifs are conserved in other ROP2 family members, including ROP4, ROP5, ROP7, and ROP18. Collectively, these findings indicate that the helical regions of the N-terminal extension wrap the kinase domain in order to stabilize it and to regulate activation, as discussed below.

Immunoprecipitation of ROP18 from parasites revealed that it is capable of phosphorylating other ROPs in vitro (36). ROP2 was weakly labeled, while ROP8 and a truncated, inactive version of ROP18 were more strongly labeled (36). These studies also revealed that ROP18 undergoes autophosphorylation and that this activity is dependent on a conserved Asp394 in the catalytic triad, indicating that the observed activity is not likely to be due to coprecipitation of another active kinase (36). In keeping with the predictions that ROP2 and ROP8 are inactive, they showed no ability to undergo autophosphorylation in vitro. In contrast, expression of a highly soluble form of ROP18 in *E. coli* also supported evaluation of its in vitro kinase activity. Mass spectrometry mapping identified several Thr residues in the N lobe and several residues in the first alpha-helix of the N-terminal extension of the ROP18 kinase domain that were autophosphorylated (36). Mutation of these residues reduced phosphorylation activity toward both ROP18 itself and the heterologous substrate myelin basic protein (36). Analogously to the interaction of the Arg220 residue seen in ROP8, ROP18 has a pair of Gln residues at 199 and 201, and the side chains of these residues extend into the nucleotide-binding pocket (Fig. 4B). Phosphorylation of key residues on the N lobe or upstream alpha-helices may alter the

position of these side chains by extending the N-terminal extension away from the kinase domain, thus relaxing inhibition. Consistent with this, mutation of these key Gln residues to Ala results in partial activation of kinase activity (36).

Collectively, these features suggest a novel means of regulation. When the N-terminal helical extension is closely apposed to the kinase domain, it stabilizes the inactive conformation. However, when the N-terminal extension becomes phosphorylated, it swings out from the kinase domain and relieves inhibition. While these studies have thus far been demonstrated in vitro, it remains to be investigated whether similar regulation is important in vivo. In this regard, the clustering of ROP18 on the PVM may be important for transphosphorylation and regulating activation. As well, the ability of ROP18 to phosphorylate other ROPs in vivo may be important in modulating its activity. Testing these and other modifications of ROP18 will be facilitated by reverse genetic approaches as described above.

## FUTURE PERSPECTIVES

Although the development of forward genetic mapping required a substantial investment, it has proven to be a powerful method for investigating complex biological traits, such as acute mortality in the mouse model and the influence of the parasite on host cell gene expression. One of the advantages of genetic mapping in *T. gondii* is that the progeny can be cryopreserved and tested repeatedly for a variety of different phenotypes. At present, only a few of the potential phenotypes or experimental models have been explored by this approach. Additional traits that would be of interest to map include differences in pathology following natural routes of infection, dissemination to deep tissues, such as the central nervous system, and escape from innate immunity. Recent advances in whole-genome analysis of gene expression in the parasite using Affymetrix arrays have provided an enhanced genetic map,

thus facilitating fine positional mapping. As well, additional genetic crosses, including a recently completed cross between type I and II strains (unpublished data), will make it possible to compare the roles of specific genes in different genetic backgrounds. To facilitate wider application of genetic analysis, the progeny from the existing crosses have been made available through the Biodefense and Emerging Infectious Disease Research Resources Repository ([www.beiresources.org/](http://www.beiresources.org/)), and the genetic maps, coordinates, and markers are freely available at <http://toxomap.wustl.edu/policy.htm>.

Thus far, the role of ROPs has been investigated only in the murine system, although some of the *in vitro* phenotypes are also expressed in human cells. It will be important to expand these findings to human infection in order to assess their potential to contribute to the variable clinical outcomes seen in toxoplasmosis. The availability of reagents for various ROPs will be useful for establishing whether they are recognized by the human immune system and whether variants of the various ROPs correlate with the clinical presentation of different isolates in human disease. Structural information combined with biochemical assays will also be useful for screening small-molecule libraries to define chemical inhibitors of the ROPs. Compounds that selectively inhibit ROP kinases might be expected to block the virulence potential of the parasite, thus allowing the immune system to control the infection and prevent pathology.

#### ACKNOWLEDGMENTS

This work was supported in part by grants from the NIH (AI036629 and AI059176, to L.D.S.).

We are grateful to John Boothroyd, David Roos, Michael White, John Wootton, Jim Ajioka, and Jean Francois Dubremetz for helpful comments and to members of our laboratories for their contributions to the work summarized here.

#### REFERENCES

- Ajioka, J. W., and L. D. Sibley. 2007. Development and application of classical genetics in *Toxoplasma gondii*, p. 367–389. In L. M. Weiss and K. Kim (ed.), *Toxoplasma gondii* the model Apicomplexan: perspectives and methods. Elsevier/Academic Press, New York, NY.
- Barragan, A., and L. D. Sibley. 2003. Migration of *Toxoplasma gondii* across biological barriers. *Trends Microbiol.* **11**:426–430.
- Barragan, A., and L. D. Sibley. 2002. Transepithelial migration of *Toxoplasma gondii* is linked to parasite motility and virulence. *J. Exp. Med.* **195**:1625–1633.
- Beckers, C. J. M., J. F. Dubremetz, O. Mercereau-Puijalon, and K. A. Joiner. 1994. The *Toxoplasma gondii* rhoptry protein ROP2 is inserted into the parasitophorous vacuole membrane, surrounding the intracellular parasite, and is exposed to the host cell cytoplasm. *J. Cell Biol.* **127**:947–961.
- Boudeau, J., D. Miranda-Saavedra, G. J. Barton, and D. R. Alessi. 2006. Emerging roles of pseudokinases. *Trends Cell Biol.* **16**:443–452.
- Boyle, J. P., B. Rajasekar, J. P. J. Saeij, J. W. Ajioka, M. Berriman, I. Paulsen, L. D. Sibley, M. White, and J. C. Boothroyd. 2006. Just one cross appears capable of dramatically altering the population biology of a eukaryotic pathogen like *Toxoplasma gondii*. *Proc. Natl. Acad. Sci. USA* **103**:10514–10519.
- Bradley, P. J., C. Ward, S. J. Cheng, D. L. Alexander, S. Coller, G. H. Coombs, J. D. Dunn, D. J. Ferguson, S. J. Sanderson, J. M. Wastling, and J. C. Boothroyd. 2005. Proteomic analysis of rhoptry organelles reveals many novel constituents for host-parasite interactions in *T. gondii*. *J. Biol. Chem.* **280**:34245–34258.
- Dubey, J. P. 2007. The history and life cycle of *Toxoplasma gondii*, p. 1–17. In L. M. Weiss and K. Kim (ed.), *Toxoplasma gondii* the model Apicomplexan: perspectives and methods. Elsevier/Academic Press, New York, NY.
- El Hajji, H., E. Demey, J. Poncet, M. Lebrun, B. Wu, N. Galeotti, M. N. Fourmaux, O. Mercereau-Puijalon, H. Vial, and J. F. Dubremetz. 2006. The ROP2 family of *Toxoplasma gondii* rhoptry proteins: proteomic and genomic characterization and molecular modeling. *Proteomics* **6**:5773–5784.
- El Hajji, H., M. Lebrun, S. T. Arold, H. Vial, G. Labesse, and J. F. Dubremetz. 2007. ROP18 is a rhoptry kinase controlling the intracellular proliferation of *Toxoplasma gondii*. *PLoS Pathog.* **3**:e14.
- Fuentes, L., J. M. Rubio, C. Ramirez, and J. Alvar. 2001. Genotypic characterization of *Toxoplasma gondii* strains associated with human toxoplasmosis in Spain: direct analysis from clinical samples. *J. Clin. Microbiol.* **39**:1566–1570.
- Gavrilescu, L. C., and E. Y. Denkers. 2001. IFN- $\gamma$  overproduction and high level apoptosis are associated with high but not low virulence *Toxoplasma gondii* infection. *J. Immunol.* **167**:902–909.
- Grigg, M. E., S. Bonnefoy, A. B. Hehl, Y. Suzuki, and J. C. Boothroyd. 2001. Success and virulence in *Toxoplasma* as the result of sexual recombination between two distinct ancestries. *Science* **294**:161–165.
- Grigg, M. E., J. Ganatra, J. C. Boothroyd, and T. P. Margolis. 2001. Unusual abundance of atypical strains associated with human ocular toxoplasmosis. *J. Infect. Dis.* **184**:633–639.
- Häkansson, S., A. J. Charron, and L. D. Sibley. 2001. *Toxoplasma* vacuoles: a two-step process of secretion and fusion forms the parasitophorous vacuole. *EMBO J.* **20**:3132–3144.
- Hall, S., K. A. Ryan, and D. Buxton. 2001. The epidemiology of toxoplasma infection, p. 58–124. In D. H. Joynton and T. G. Wreghitt (ed.), *Toxoplasmosis: a comprehensive clinical guide*. Cambridge University Press, Cambridge, England.
- Hanks, S. K., and T. Hunter. 1995. Protein kinases 6. The eukaryotic protein kinase superfamily: kinase (catalytic) domain structure and classification. *FASEB J.* **9**:576–596.
- Jones, J. L., C. Muccioli, R. Belfort, Jr., G. N. Holland, J. M. Roberts, and C. Silveira. 2006. Recently acquired *Toxoplasma gondii* infection, Brazil. *Emerg. Infect. Dis.* **12**:582–587.
- Khan, A., U. Bohme, K. A. Kelly, E. Adlem, K. Brooks, M. Simmonds, K. Mungall, M. A. Quail, C. Arrowsmith, T. Chillingworth, C. Churcher, D. Harris, M. Collins, N. Fosker, A. Fraser, Z. Hance, K. Jagels, S. Moule, L. Murphy, S. O'Neil, M. A. Rajandream, D. Saunders, K. Seeger, S. Whitehead, T. Mayr, X. Xuan, J. Watanabe, Y. Suzuki, H. Wakaguri, S. Sugano, C. Sugimoto, I. Paulsen, A. J. Mackey, D. S. Roos, N. Hall, M. Berriman, B. Borell, L. D. Sibley, and J. W. Ajioka. 2006. Common inheritance of chromosome Ia associated with clonal expansion of *Toxoplasma gondii*. *Genome Res.* **16**:1119–1125.
- Khan, A., B. Fux, C. Su, J. P. Dubey, M. L. Darde, J. W. Ajioka, B. M. Rosenthal, and L. D. Sibley. 2007. Recent transcontinental sweep of *Toxoplasma gondii* driven by a single monomorphic chromosome. *Proc. Natl. Acad. Sci. USA* **104**:14872–14877.
- Khan, A., C. Su, M. German, G. A. Storch, D. Clifford, and L. D. Sibley. 2005. Genotyping of *Toxoplasma gondii* strains from immunocompromised patients reveals high prevalence of type I strains. *J. Clin. Microbiol.* **43**:5881–5887.
- Khan, A., S. Taylor, J. W. Ajioka, B. M. Rosenthal, and L. D. Sibley. 2009. Selection at a single locus leads to widespread expansion of *Toxoplasma gondii* lineages that are virulent in mice. *PLoS Genet.* **5**:e1000404.
- Khan, A., S. Taylor, C. Su, A. J. Mackey, J. Boyle, R. H. Cole, D. Glover, K. Tang, I. Paulsen, M. Berriman, J. C. Boothroyd, E. R. Pfefferkorn, J. P. Dubey, D. S. Roos, J. W. Ajioka, J. C. Wootton, and L. D. Sibley. 2005. Composite genome map and recombination parameters derived from three archetypal lineages of *Toxoplasma gondii*. *Nucleic Acids Res.* **33**:2980–2992.
- Kim, L., B. A. Butcher, C. W. Lee, S. Uematsu, S. Akira, and E. Y. Denkers. 2006. *Toxoplasma gondii* genotype determines MyD88-dependent signaling in infected macrophages. *J. Immunol.* **177**:2584–2591.
- Labesse, G., M. Gelin, Y. Bessin, M. Lebrun, J. Papoin, R. Cerdan, S. T. Arold, and J. F. Dubremetz. 2009. ROP2 from *Toxoplasma gondii*: a virulence factor with a protein-kinase fold and no enzymatic activity. *Structure* **17**:139–146.
- Lander, E., and L. Kruglyak. 1995. Genetic dissection of complex traits: guidelines for interpreting and reporting linkage results. *Nat. Genet.* **11**:241–247.
- Lander, E. S., and D. Botstein. 1989. Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics* **121**:185–199.
- Lehmann, T., P. L. Marcet, D. H. Graham, E. R. Dahl, and J. P. Dubey. 2006. Globalization and the population structure of *Toxoplasma gondii*. *Proc. Natl. Acad. Sci. USA* **103**:11423–11428.
- Mariuz, P., and R. T. Steigbigel. 2007. *Toxoplasma* infection in HIV-infected patients, p. 147–177. In D. H. M. Joynton and T. G. Wreghitt (ed.), *Toxoplasmosis: a comprehensive clinical guide*. Cambridge University Press, Cambridge, England.
- Mordue, D. G., F. Monroy, M. La Regina, C. A. Dinarello, and L. D. Sibley. 2001. Acute toxoplasmosis leads to lethal overproduction of Th1 cytokines. *J. Immunol.* **167**:4574–4584.
- Nunes, M., J. P. D. Goldring, C. Doerig, and A. Scherf. 2007. A novel protein kinase family in *Plasmodium falciparum* is differentially transcribed and secreted to various cellular compartments of the host cell. *Mol. Microbiol.* **63**:391–403.
- Petersen, E., and O. Liesenfeld. 2007. Clinical disease and diagnostics, p. 81–100. In L. M. Weiss and K. Kim (ed.), *Toxoplasma gondii* the model Apicomplexan: perspectives and methods. Elsevier/Academic Press, New York, NY.
- Pfefferkorn, E. R., and L. H. Kasper. 1983. *Toxoplasma gondii*: genetic



- crosses reveal phenotypic suppression of hydroxyurea resistance by fluoro-deoxyuridine resistance. *Exp. Parasitol.* **55**:207–218.
34. **Pfefferkorn, E. R., L. C. Pfefferkorn, and E. D. Colby.** 1977. Development of gametes and oocysts in cats fed cysts derived from cloned trophozoites of *Toxoplasma gondii*. *J. Parasitol.* **63**:158–159.
  35. **Pfefferkorn, L. C., and E. R. Pfefferkorn.** 1980. *Toxoplasma gondii*: genetic recombination between drug resistant mutants. *Exp. Parasitol.* **50**:305–316.
  36. **Qiu, W., A. Wernimont, K. Tang, S. Taylor, V. Lunin, M. Schapira, S. J. Fentress, R. Hui, and L. D. Sibley.** 2009. Novel structural and regulatory features of rhoptyry secretory kinases in *Toxoplasma gondii*. *EMBO J.* **28**: 969–979.
  37. **Radke, J. R., B. Striepen, M. N. Guerini, M. E. Jerome, D. S. Roos, and M. W. White.** 2001. Defining the cell cycle for the tachyzoite stage of *Toxoplasma gondii*. *Mol. Biochem. Parasitol.* **115**:165–175.
  38. **Reichard, U., and U. Gross.** 2007. *Toxoplasma* animal models and therapeutics, p. 153–184. *In* L. M. Weiss and K. Kim (ed.), *Toxoplasma gondii* the model Apicomplexan: perspectives and methods. Elsevier/Academic Press, New York, NY.
  39. **Robben, P. M., D. G. Mordue, S. M. Truscott, K. Takeda, S. Akira, and L. D. Sibley.** 2004. Production of IL-12 by macrophages infected with *Toxoplasma gondii* depends on the parasite genotype. *J. Immunol.* **172**:3686–3694.
  40. **Saeij, J. P. J., J. P. Boyle, S. Coller, S. Taylor, L. D. Sibley, E. T. Brooke-Powell, J. W. Ajioka, and J. C. Boothroyd.** 2006. Polymorphic secreted kinases are key virulence factors in toxoplasmosis. *Science* **314**:1780–1783.
  41. **Sibley, L. D.** 2004. Invasion strategies of intracellular parasites. *Science* **304**:248–253.
  42. **Sibley, L. D., and J. W. Ajioka.** 2008. Population structure of *Toxoplasma gondii*: clonal expansion driven by infrequent recombination and selective sweeps. *Annu. Rev. Microbiol.* **62**:329–351.
  43. **Sibley, L. D., and J. C. Boothroyd.** 1992. Virulent strains of *Toxoplasma gondii* comprise a single clonal lineage. *Nature (London)* **359**:82–85.
  44. **Sibley, L. D., A. J. LeBlanc, E. R. Pfefferkorn, and J. C. Boothroyd.** 1992. Generation of a restriction fragment length polymorphism linkage map for *Toxoplasma gondii*. *Genetics* **132**:1003–1015.
  45. **Sinai, A. P., and K. A. Joiner.** 2001. The *Toxoplasma gondii* protein ROP2 mediates host organelle association with the parasitophorous vacuole membrane. *J. Cell Biol.* **154**:95–108.
  46. **Su, C., D. Evans, R. H. Cole, J. C. Kissinger, J. W. Ajioka, and L. D. Sibley.** 2003. Recent expansion of *Toxoplasma* through enhanced oral transmission. *Science* **299**:414–416.
  47. **Su, C., D. K. Howe, J. P. Dubey, J. W. Ajioka, and L. D. Sibley.** 2002. Identification of quantitative trait loci controlling acute virulence in *Toxoplasma gondii*. *Proc. Natl. Acad. Sci. USA* **99**:10753–10758.
  48. **Taylor, S., A. Barragan, C. Su, B. Fux, S. J. Fentress, K. Tang, W. L. Beatty, E. L. Haij, M. Jerome, M. S. Behnke, M. White, J. C. Wootton, and L. D. Sibley.** 2006. A secreted serine-threonine kinase determines virulence in the eukaryotic pathogen *Toxoplasma gondii*. *Science* **314**:1776–1780.
  49. **Weiss, L. M., and K. Kim.** 2007. Bradyzoite development, p. 341–366. *In* L. M. Weiss and K. Kim (ed.), *Toxoplasma gondii* the model Apicomplexan: perspectives and methods. Elsevier/Academic Press, New York, NY.
  50. **Zhao, Y., D. J. Ferguson, D. C. Wilson, J. C. Howard, L. D. Sibley, and G. S. Yap.** 2009. Virulent *Toxoplasma gondii* evade immunity-related GTPase-mediated parasite vacuole disruption within primed macrophages. *J. Immunol.* **182**:3775–3781.