Comparative Genomic and Phylogenetic Analyses of Calcium ATPases and Calcium-Regulated Proteins in the Apicomplexa

Kisaburo Nagamune and L. David Sibley
Department of Molecular Microbiology, Washington University School of Medicine

The phylum Apicomplexa comprises a large group of early branching eukaryotes that includes a number of human and animal parasites. Calcium controls a number of vital processes in apicomplexans including protein secretion, motility, and differentiation. Despite the importance of calcium as a second messenger, very little is known about the systems that control hoemostasis or that regulate calcium signaling in parasites. The recent completion of many apicomplexan genomes provides new opportunity to define calcium response pathways in this group of parasites in comparison to model organisms.

Whole-genome comparison between the apicomplexans Plasmodium spp., Cryptosporidium spp., and Toxoplasma gondii revealed the presence of several P-Type \( \text{Ca}^{2+} \) transporting ATPases including a single endoplasmic reticulum (ER)-type sarcoplasmic–endoplasmic reticulum \( \text{Ca}^{2+} \) ATPase, several Golgi-like \( \text{Ca}^{2+} \)/\( \text{H}^{+} \) exchangers, and a single \( \text{Ca}^{2+} \)/\( \text{H}^{+} \) exchanger. Only \( T. \text{gondii} \) showed evidence of plasma membrane-type \( \text{Ca}^{2+} \) ATPases or voltage-gated calcium channels. Despite pharmacological evidence for IP3 and ryanodine-mediated calcium release, animal-type calcium channels were not readily identified in parasites, indicating they are more similar to plants. Downstream of calcium release, a variety of EF-hand–containing proteins regulate calcium responses. Our analyses detected a single conserved calmodulin (CaM) homologue, 3 distinct centrins (CETN)-caltractin–like proteins, one of which is shared with ciliates, and a variety of deep-branching, CaM-CETN–like proteins. Apicomplexans were also found to contain a wide array of calcium-dependent protein kinases (CDPKs), which are commonly found in plants. \( T. \text{gondii} \) contains more than 20 CDPKs.

Genomic and phylogenetic comparisons revealed that apicomplexans contain a variety of unusual calcium response pathways that are distinct from those seen in vertebrates. Notably, plant-like pathways for calcium release channels and calcium-dependent kinases are found in apicomplexans. The experimental flexibility of \( T. \text{gondii} \) should allow direct experimental manipulation of these pathways to validate their biological roles. The central importance of calcium in signaling and development, and the novel characteristics of many of these systems, indicates that parasite calcium pathways may be exploited as new therapeutic targets for intervention.

Introduction

Apicomplexans are early branching eukaryotes related to ciliates and dinoflagellates (Baldauf 2003). Apicomplexans are not closely related to plants, fungi, or metazoan animals; hence, their basic biology is often distinct from that of model organisms. Included in the phylum Apicomplexa are several agents of human disease such as \( P. \text{falciparum} \) (malaria), \( C. \text{parvum} \), and \( T. \text{gondii} \). The recent completion of whole-genome sequences for members of this genus has generated a large volume of data for comparative analyses (Aravind et al. 2003; Huang et al. 2004; Templeton et al. 2004). However, despite extensive efforts at genome annotation, our understanding of core regulatory systems such as calcium signaling remains limited.

Calcium is an important second messenger in eukaryotic cells (Tsien 1990; Clapham 1995; Berridge et al. 2000b). Calcium is maintained at 10 000-fold lower levels in the cytosol of the cell compared with the extracellular environment, and its rapid release or influx is coupled to a number of key physiological responses. Calcium enters cells through a variety of calcium channels (Miller and Fox 1990; Berridge 1995), and conversely, it can be sequestered or pumped back out of the cell through the action of calcium ATPases or exchangers at the plasma membrane or in the secretory pathway (Wuytack et al. 2002; Nagata et al. 2004). The endoplasmic reticulum (ER) contains a major calcium pump for sequestering calcium called the sarcoplasmic–endoplasmic reticulum \( \text{Ca}^{2+} \) ATPase (SERCA), and in animal cells, this store can be released through the action of IP3-responsive calcium channels (Berridge 1993). A second system for rapid calcium release is mediated in certain cell types by ryanodine receptors, which respond to the ligands cADPR and nicotinic acid adenine dinucleotide phosphate (Chini and De Toledo 2002). Specialized receptors for IP3 (IP3R) and ryanodine (RYR)-mediated calcium release have evolved during differentiation of vertebrates (Sorrentino et al. 2000). Plants (Wu et al. 1997; White 2000) and protozoa (Thiel et al. 1990; Masuda et al. 1997) also respond to agonists of these intracellular calcium channels, although the calcium release channels involved in these responses are not conserved with animals.

Calcium homeostasis in protozoan parasites is mediated by several organellar systems that sequester intracellular calcium including the ER, mitochondria, and acidicalcisomes (Moreno and Docampo 2003). Acidicalcisomes are acidic calcium storage organelles that are conserved across a wide range of organisms, and they play an important role in polyphosphate metabolism (Docampo et al. 2005). Calcium controls a number of important events in apicomplexan parasites including motility (Wetzel et al. 2004), secretion (Carruthers, Moreno, Sibley 1999), and differentiation (Bilker et al. 1998, 2004).

\( T. \text{gondii} \) has been particularly useful for studying calcium responses in parasites due to the ease of culture and facility for cell and molecular biology studies in this organism. Fura-2 measurements indicate that resting calcium levels in \( T. \text{gondii} \) are maintained at a low level of \( \approx 90–100 \) nM (Moreno and Zhong 1996), similar to other eukaryotic cells. Monitoring of calcium using fluo-4–loaded...
cells reveal that cytosolic Ca$^{2+}$ undergoes oscillations during gliding motility of parasites (Lovett and Sibley 2003). Coincident with cell invasion, cytosolic Ca$^{2+}$ levels are dramatically dampened in the parasite (Lovett and Sibley 2003), consistent with the cessation of motility. Calcium also controls protein secretion in T. gondii and elevation of intracellular Ca$^{2+}$ results in discharge of micronemes, whereas chelation prevents this response and thus compromises entry (Carruthers, Giddings, Sibley 1999; Carruthers, Moreno, Sibley 1999). Although less extensively studied, similar experiments in P. berghei (Gantt et al. 2000) and Cryptosporidium parvum (Chen et al. 2004) indicate that this calcium-dependent secretion pathway is conserved in the Apicomplexa. Pharmacological studies indicate that treatment of T. gondii with IP$_3$, caffeine, or ryanodine results in increases in intracellular Ca$^{2+}$, which stimulates microneme secretion (Lovett et al. 2002). Recent evidence indicates that T. gondii contains both a specific cyclase and hydrolase for generation and turnover of the second messenger cyclic ADP ribose (cADPR) (Chini et al. 2005). Inhibitor studies demonstrate that both IP$_3$ and cADPR pathways are important for governing calcium-mediated secretion in T. gondii (Lovett et al. 2002; Chini et al. 2005).

The systems that control calcium responses in parasites are incompletely understood. A plasma membrane-type Ca$^{2+}$ ATPase has previously been described in T. gondii (TgA1), and this protein is found both on the plasma membrane and acidicolcisome (Luo et al. 2001). TgA1 is able to complement yeast cells deficient in the vacuolar Ca$^{2+}$ ATPase PMC1 (Luo et al. 2001, 2005). Mutants that do not express TgA1 have reduced polyphosphate levels, show elevated cytoplasmic levels of calcium, are less able to respond to agonists of secretion, and show impaired motility and cell invasion (Luo et al. 2005). Several P-type Ca$^{2+}$ ATPases have been described in malaria including PfATPase2 and PfATPase4, which contain both unique features and motifs conserved with other organisms (Trottein and Corman 1995; Trottein et al. 1995; Krishna et al. 2001). A Ca$^{2+}$ ATPase that is similar to PfATPase4 has also been described from C. parvum (Zhu and Keithly 1997). Additionally, Plasmodium contains a SERCA-type pump (PfATPase6) that is implicated as the target of the drug artemisinin (Eckstein-Ludwig et al. 2003).

Despite compelling evidence that calcium regulates important events in parasites, no systematic study has been made of calcium transporters or calcium-regulated proteins in the Apicomplexa. Thus, it remains unclear to what extent systems described in plants, fungi, or animals are also conserved in these early branching eukaryotes.

**Materials and Methods**

The recently annotated whole-genome sequence of T. gondii (http://ToxoDB.org. Release v3.0) was analyzed for genes related to calcium metabolism by key word searches of the BlastP results using the search terms “calcium,” “centrin (CETN),” and “EF-hand” (cutoff $\leq 1 \times 10^{-10}$). Text word searches were also done against the annotated Plasmodium falciparum genome (http://www.plasmodb.org/plasmo/home.jsp, Release 4.4) using the key words “calcium,” “ATPase,” and “exchanger” (cutoff $\leq 1 \times 10^{-10}$). BlastP comparisons using T. gondii calmodulin (CaM) (GenBank id: Y08373) were used to identify additional orthologues containing EF-hands by comparison to the T. gondii genome database (cutoff $\leq 1\times10^{-10}$). Calcium-related proteins from T. gondii and P. falciparum were then used for BlastP comparison against the SwissPro database (cutoff $\leq 1 \times 10^{-5}$) and analyzed for domains using the InterProScan database (http://www.ebi.ac.uk/interpro/) (Quevillon et al. 2005). Calcium-related proteins identified in T. gondii were also used to search using BlastP against the GenBank Protozoa genome database (http://www.ncbi.nlm.nih.gov/sutils/blast_table.cgi?taxid=Protozoa) to identify P. falciparum and C. parvum orthologues (cutoff $\leq 1 \times 10^{-10}$), Tetrahymena orthologues were identified by BlastP searches against the ciliate genome database (http://www.tigr.org). Results of these searches were compared using BlastP to identify reciprocal best matches among the Apicomplexa. In cases where orthologues were not identified in C. parvum, we also analyzed the complete genome of Cryptosporidium hominis. A complete list of all taxa and accession/contig numbers used in these studies is provided in Supplemental Tables 1–5 (Supplemental Material online).

For phylogenetic analyses, a collection of orthologues was chosen for each T. gondii protein including the corresponding proteins from P. falciparum and C. parvum plus the top 4–5 taxonomically distinct and non-apicomplexan hits from the search of GenBank. Protein sequences were aligned in ClustalX ( Higgins et al. 1996) using default parameters (gap opening = 10, gap extension = 0.1, matrix = Gonnet250, slow–accurate alignment). A complete list of the alignment used here is provided at: http://www.sibleylab.wustl.edu/Publications.htm/. Phylogenetic analyses were conducted using PAUP* 4.0 (Swofford 1998). Only informative regions as defined by PAUP* were used in the analyses (excluded characters were defined by choosing the “Informative only” option under the setting to include–exclude characters). Phylogenetic comparisons were made using distance criteria analyzed by neighbor joining. Alternatively, a full heuristic search was conducted using parsimony. Consensus trees were drawn according to the 50% majority rule after bootstrapping 1000 times, and phylograms were presented as unrooted trees. Parsimony and neighbor-joining trees were highly similar in all cases, and therefore, only neighbor-joining trees are shown here.

**Results and Discussion**

We analyzed the recently completed genomes of T. gondii, Cryptosporidium spp., and Plasmodium spp. with the goal of identifying all the major calcium pathways in these apicomplexan parasites. We searched the apicomplexan genome databases with text word queries to identify Ca$^{2+}$ ATPases, calcium, CaM, EF-hand, and CETN. BlastP analyses and InterProScan domain analyses were then used to authenticate all the calcium pathway proteins from these organisms. Whole-genome annotations have typically been done with a combination of similar approaches, which are generally sufficient to broadly classify the functional classes of putative genes. However, the degree of similarity...
either within gene families (paralogues) or between different taxa (orthologues) is not directly inferable from these tools. Molecular phylogeny is valuable for predicting possible evolutionary histories either due to inheritance from a common ancestor or due to more abrupt patterns like horizontal gene transfer. Phylogenetic reconstruction of early branching eukaryotes is particularly useful for defining plant-like, protist-like, or animal-like attributes. Such traits are useful for predicting conserved function and for identifying proteins that may have distinctly different properties in parasites. Consequently, clustal alignments and phylogenetic analyses were further used to classify families of calcium-related proteins and to identify orthologues and paralogues within the Apicomplexa in relation to plants, animals, fungi, and other protists. This combined homology-based searching and phylogenetic analyses allowed accurate assignment of orthologues and classification of plant versus animal-like pathways as well as identifying a number of parasite-specific genes that are likely involved in calcium regulation. Our results reveal several important features about calcium regulation in apicomplexans and provide a preliminary cataloging of calcium-related proteins in these organisms that will aid in future functional studies.

Calcium Pumps and Transporters

P-Type ATPases are found in prokaryotes and eukaryotes: they use the energy from ATP hydrolysis to transport cations across biological membranes. Eukaryotic Ca\(^{2+}\) ATPases have been divided into secretory or type IIA, which includes the sarco(endo)plasmic reticulum Ca\(^{2+}\) ATPases (SERCA) and yeast PMR homologues, or type IIB in the plasma membrane (i.e., PMCA) (Kuhlbrandt 2004). Apicomplexans contain several members of the type IIA family, whereas members of the type IIB family were only detected in *T. gondii* (fig. 1, Supplemental Table 1, Supplementary Material online). InterproScan analyses revealed that these orthologues have predicted topologies for P-type ATPases including multiple transmembrane (TM) domains and a conserved Asp residue in the P-loop that becomes phosphorylated (fig. 1, Supplemental Table 1, Supplementary Material online). Ca\(^{2+}\) ATPases were further classified by molecular phylogeny using a variety of model organisms for comparison (fig. 2, Supplemental Table 1, Supplementary Material online).

SERCA pumps are responsible for refilling calcium in the ER store, which represents the most readily mobilizable source of calcium for signaling (Berridge et al. 2000a, 2000b). Vertebrates contain 3 SERCA genes, each of which shows alternative spliced forms (Wuytack et al. 2002). In the present study, apicomplexans were found to contain a single SERCA-like Ca\(^{2+}\) ATPase, previously described as ATP6 in *P. falciparum* (Kimura et al. 1993). The SERCA of malaria has been described as the target of artemisinin (Eckstein-Ludwig et al. 2003), a natural product derived from wormwood (*Artemisia annua*). Artemisinin is a potent...
antimalarial compound (Haynes and Krishna 2004; O’Neill 2004), demonstrating that disruption of calcium homeostasis provides an effective means of combating infections with apicomplexan parasites. Thapsigargin also targets the activity of SERCA by blocking the E1–E2 transition, which is required for activity (Sagara and Inesi 1991; Sagara et al. 1992). Both *T. gondii* (Moreno and Zhong 1996) and *P. falciparum* (Varotti et al. 2003) are sensitive to this plant alkaloid. Phylogenetic analysis revealed that SERCA in *P. falciparum* (PfATP6) is more divergent, whereas *T. gondii* (TgSERCA) and *C. parvum* (Cp90400) each contain a single similar orthologue (fig. 2). The apparent divergence of the *P. falciparum* SERCA may result from the strongly AT-rich genome that results in a computational bias in amino acids, as reported previously (Musto et al. 1995, 1999), a feature that can complicate phylogenetic analyses.

All 3 apicomplexans have orthologues of the yeast PMR1, which is found in the Golgi (figs. 1 and 2, Supplemental Table 1, Supplementary Material online). This gene in *P. falciparum* has previously been called ATP4, and it reportedly localized to the plasma membrane in asexual stages of malaria (Dyer et al. 1996). All 3 apicomplexans also contain orthologues of *Schizosaccharomyces pombe* ATP4, a member of the type IIA family that is localized to the ER. In fission yeast, ATP4 is essential for calcium homeostasis, microtubule stability, and cytokinesis (Facanha et al. 2002). In addition, *P. falciparum* (Pf703505, Pf703385) and *C. parvum* (Cp88954, Cp88619) each contain 2 paralogues that are distantly related to this group. Although these putative Ca^{2+} ATPases contain putative signal peptides, indicating that they are within the secretory system, there is no specific information available about their subcellular location or functions, and analysis...
of their domain architecture does not provide further insight into this important question (fig. 2, Supplemental Table 1, Supplementary Material online). Cryptosporidium parvum (CP89054, Cp90606) and T. gondii (Tg50.m03192) contain orthologues of ATP2, which was previously described in P. falciparum (Trottein and Corman 1995) (fig. 1, Supplemental Table 1, Supplementary Material online). These orthologues are most similar to phospholipid transporters (fig. 2), and hence, they may not play a major role in calcium homeostasis.

Among the apicomplexans, only T. gondii was found to contain a plasma membrane-type Ca2+ ATPases represented by a previously described TgAl (Luo et al. 2001), along with a second paralogue that is designated as Tg44.mo2812 (figs. 1 and 2, Supplemental Table 1, Supplementary Material online). TgAl is localized to the plasma membrane and acidocalcisome in T. gondii (Luo et al. 2001). TgA1 is able to complement yeast deficient in the vacuolar Ca2+ ATPase known as PMC1, demonstrating that it is active in calcium homeostasis (Luo et al. 2005). Mutants in igal in T. gondii show alterations in polyphosphate levels and disruption of basal cytosolic calcium regulation and have decreased infectivity (Luo et al. 2005).

Finally, all 3 apicomplexans contain orthologues of Ca2+/H+ exchangers that are similar to proteins found in the plant vacuole and also found in yeast but not found in metazoan animal cells (Nagata et al. 2004) (figs. 1 and 2, Supplemental Table 1, Supplementary Material online). Conversely, Na+/Ca2+ exchangers, which are commonly found in animal cells, were not found in apicomplexans.

Calcium Channels

Notably absent from apicomplexan genomes are orthologues of the 2 families of intracellular calcium release channels found in metazoan animals, known as IP3R or RyR channels (Sorrentino et al. 2000). Drosophila and Caenorhabditis elegans each have a single IP3R and RyR calcium release channel, whereas multiple copies are present in vertebrates. IP3R and RyR contain multiple TM domains, MIR domains (mannosyltransferase, IP3R and RyR domains), and internal repeat IRH domains (IP3R and RyR homology domains). Additionally, RyR channels contain SPRY domains, also found in Sp1A, a kinase involved in differentiation in Dictyostelium (Ponting et al. 1997). Although 2 SPRY domain-containing proteins are found in T. gondii, neither contains other domains nor motifs consistent with calcium channels (Chini et al. 2005). Homologues for IP3R or RyR have not been found in fungi, protozoa (Thiel et al. 1990; Masuda et al. 1997; Lovett et al. 2002; Chini et al. 2005), or plants (Wu et al. 1997; White 2000), despite functional evidence for calcium release in response to calcium channel agonists in the later 2 groups. Collectively, these observations suggest that a more primitive calcium release channel may exist in protists and plants and that the domains that characterize IP3 and RyR channels arose during evolution of metazoan animals.

In vertebrate cells, phosphoinositid phospholipase C (PI-PLC) acts on phoshoinositid 4,5 bisphosphate (PIP2) to generate IP3 and diacyl glycerol (DAG) (Berridge 1993). Although IP3 acts to release intracellular calcium, DAG is involved in activating protein kinase C (PKC) to regulate a number of cellular responses (Berridge 1993). Apicomplexans have a delta-type PI-PLC that cleaves PIP2 under physiological conditions (Fang et al. 2005). TgPI-PLC is localized to the cytoplasmic face of the parasite cell membrane, and this enzyme is a candidate for generating IP3 that has previously been shown to play a role in calcium-mediated secretion in T. gondii (Lovett et al. 2002). Although DAG is also likely a product of this reaction, we were not able to identify a direct orthologue of PKC using rat PKC (NP_036760) or yeast PKC (NP_009445) to search ToxoDB using BlastP (data not shown). Instead, T. gondii appears to contain genes for protein kinase A, protein kinase B, and protein kinase G, similar to Plasmodium, as reported previously (Doerig et al. 2005). Whether apicomplexans lack a DAG-dependent PKC or whether they contain more divergent orthologues will require further computational and biochemical analyses.

Calcium influx channels bear overall domain similarity to the super family of ion channels that includes voltage-gated channels, cyclic nucleotide-gated channels, and transient receptor potential channels (Sorrentino et al. 2000). Excitable cells in animals and plants express ligand- and voltage-gated calcium channels (VGCC), which are most commonly localized in the plasma membrane. Based on the present analyses, apicomplexans appear to lack cyclic nucleotide-gated calcium channels and glutamate-gated channels found in animals and plants (Nagata et al. 2004). VGCC typically respond to changes in membrane polarization and allow selective influx of cations (Clapham 1995; Berridge et al. 2000b). Calcium-selective VGCC consist of an α1 subunit that comprises the major ion-transporting pore of the channel, as well as α2, β, γ, and δ subunits (Miller and Fox 1990). Typically VGCC contain 4 subdomains consisting of 6 TM segments that each forms a typical Shaker-type channel; a domain architecture that is shared with voltage-gated Na+ and K+ channels (Miller and Fox 1990; Catterall et al. 2005). VGCC in vertebrates are divided into broad groups: high-voltage–responsive VGCC consists of L type and non–L types (N, P, Q, R, and S), which are distinguished by their different sensitivities to inhibitors, whereas low-voltage–gated channels are known as T type (Miller and Fox 1990; Catterall et al. 2005). Invertebrates also contain VGCC of both the L and non–L types, and C. elegans also contains T-type channels (Jeziorski et al. 2000). Additionally, a novel calcium channel (NCA) related to VGCC has been described in C. elegans and rat, and this channel is similar to a calcium channel in yeast known as Csh1 (Jeziorski et al. 2000). Finally, plants contain a 2-pore channel (TPC) that contains 2 Shaker-type units rather than the 4 subunit architecture of conventional VGCC (White 2000). 2-pore complex 1 (TPC1) channels are found in the plant vacuole where they mediate the slow vacuolar channel that releases calcium in response to elevated intracellular calcium (Peiter et al. 2005). TPC1-like channels are also found in vertebrates, although their function there is unknown (Ishibashi et al. 2000).
Searches of the *T. gondii* genome revealed several candidate genes encoding VGCC (Supplemental Table 1, Supplementary Material online). Homologues were then further analyzed by BlastP comparison of these candidates against the genome databases. No direct orthologues were found in *P. falciparum* or *C. parvum* (Supplemental Table 1, Supplementary Material online). However, the putative *T. gondii* VGCC genes had top hits to various mammalian N-type VGCC and to plant TPC1 channels (Supplemental Table 1, Supplementary Material online). To provide a phylogenetic framework for classifying *T. gondii* VGCC, representative members of vertebrate and invertebrate classes of L, non-L VGCCs, and NCAs as well as the TPC channels were included in the analysis (Supplemental Table 3, Supplementary Material online). Previous studies have indicated that VGCC exist in ciliates based on physiological and pharmacological studies (Plattner 2002). Therefore, we searched the recently completed genome of *Tetrahymena thermophila* to identify putative VGCC (Supplemental Table 3, Supplementary Material online). Phylogenetic analyses agreed closely with previous studies of plant channels (White 2000), vertebrate channels (Catterall et al. 2005), or invertebrate channels (Jeziorski et al. 2000) and provided an excellent framework for classifying those genes found in *T. gondii* and *T. thermophila* (fig. 3).

Of the 3 VGCC-like genes found in *T. gondii*, one has a single Shaker-type unit with 6 TM domains (TwinScan_3277), one has a TPC-type conformation consisting of 12 TM domains (Tg583.m05406), and the third has a conventional 4-Shaker subunit conformation consisting of 24 TM domains (fig. 1). These putative VGCC in *T. gondii* group together in a clade that is not similar to conventional L-type or non–L-type channels in vertebrates and invertebrates (fig. 3). Instead, the *T. gondii* VGCC-like genes form one of several deep-branching groups within a common node that also contains the yeast Cch1 channel.
Calcium Regulation in Apicomplexans

Comparison of apicomplexan genomes reveals that each contains a single highly homologous CaM, a calcium-binding protein implicated in a variety of signaling events (Berridge et al. 2000b). Like animals and plants (Nagata et al. 2004), apicomplexans also contain a diverse array of EF-hand–containing, CaM-like genes (fig. 5, Supplemental Table 1, Supplementary Material online). CaM contains 4 EF-hands, and a variety of other calcium-responsive proteins also contain this conserved domain. EF-hands are defined by helix-loop-helix structure containing acidic residues that bind calcium, thus causing a conformational change in the protein (Yap et al. 1999). Phylogenetic analysis indicated that T. gondii, P. falciparum, and C. hominis each contain a single conventional CaM consisting of 4 EF-hands, whereas no direct orthologue of CaM was found in C. parvum (fig. 5, Supplemental Table 1, Supplementary Material online). Additional analyses of the apicomplexan genomes analyzed here contain 3 other EF-hand–containing proteins that resemble CETN or caltractin (Supplemental Table 1, Supplementary Material online). CETNs are found in the ciliates, apicomplexans, yeast, worms, and higher metazoans, suggesting that it represents a common ancestral channel present in the early eukaryotic crown group that has since been retained in these divergent taxa.

CETNs may be involved in formation of microtubule spindle bodies or the formation of the flagellar axoneme. Potentially, these CETNs may be involved in basal bodies or the formation of the flagellar axoneme. All 3 apicomplexans also contain CETN-like genes that group with the CETNs CETN3 and CETN4, which are found in the ciliates Paramecium and Tetrahymena (fig. 5, Supplemental Table 1, Supplementary Material online, perhaps reflecting the more streamlined genomes of the latter 2 organisms.

EF-Hand–Containing Proteins

NCAs, and the TPCs from animals and plants (fig. 3). The 8 VGCC-like genes in T. thermophila formed 2 separate groups that were also included in this deep-branching assemblage (fig. 3). The conservation of this unusual calcium channel in ciliates, apicomplexans, yeast, worms, and higher metazoans, suggests that it represents a common ancestral channel present in the early eukaryotic crown group that has since been retained in these divergent taxa.

VGCC contains several features that mediate gating and ion selectivity of the pore. For example, conserved positively charged residues in TM segment 4 mediate voltage sensitivity (Miller and Fox 1990; Catterall et al. 2005). Positive charges are conserved in segment 4 in the 4 separate subdomains of putative protein encoded by Tg20.m03897 and are partially conserved in the other 2 putative VGCC from T. gondii (figs. 1 and 4). Additionally, conserved Glu (E) residues in the P-loop between TM segments 5 and 6 are characteristic of calcium-selective pores (Miller and Fox 1990; Catterall et al. 2005). The putative VGCC encoded by Tg20.m03897 contains EEEE residues in the 4 respective P-loops, and E also partially conserved the other 2 putative VGCC from T. gondii (figs. 1 and 4). These features strongly suggest that these genes encode voltage-gated calcium-selective channels in T. gondii. Potentially, these VGCC could be involved in influx of calcium at the plasma membrane in response to depolarization or calcium release from an intracellular store following an influx of calcium from the extracellular medium. VGCC-like genes appear to be lacking in the genomes of Plasmodium and Cryptosporidium, perhaps reflecting the more streamlined genomes of the latter 2 organisms.

EF-Hand–Containing Proteins

Comparison of apicomplexan genomes reveals that each contains a single highly homologous CaM, a calcium-binding protein implicated in a variety of signaling events (Berridge et al. 2000b). Like animals and plants (Nagata et al. 2004), apicomplexans also contain a diverse array of EF-hand–containing, CaM-like genes (fig. 5, Supplemental Table 1, Supplementary Material online). CaM contains 4 EF-hands, and a variety of other calcium-responsive proteins also contain this conserved domain. EF-hands are defined by helix-loop-helix structure containing acidic residues that bind calcium, thus causing a conformational change in the protein (Yap et al. 1999). Phylogenetic analysis indicated that T. gondii, P. falciparum, and C. hominis each contain a single conventional CaM consisting of 4 EF-hands, whereas no direct orthologue of CaM was found in C. parvum (fig. 5, Supplemental Table 1, Supplementary Material online). Additional analyses of the apicomplexan genomes analyzed here contain 3 other EF-hand–containing proteins that resemble CETN or caltractin (Supplemental Table 1, Supplementary Material online). CETNs are found in the ciliates, apicomplexans, yeast, worms, and higher metazoans, suggesting that it represents a common ancestral channel present in the early eukaryotic crown group that has since been retained in these divergent taxa.

CETNs may be involved in formation of microtubule spindle bodies or the formation of the flagellar axoneme. Potentially, these CETNs may be involved in basal bodies or the formation of the flagellar axoneme. All 3 apicomplexans also contain CETN-like genes that group with the CETNs CETN3 and CETN4, which are found in the ciliates Paramecium and Tetrahymena (fig. 5, Supplemental Table 1, Supplementary Material online, perhaps reflecting the more streamlined genomes of the latter 2 organisms.

EF-Hand–Containing Proteins

Comparison of apicomplexan genomes reveals that each contains a single highly homologous CaM, a calcium-binding protein implicated in a variety of signaling events (Berridge et al. 2000b). Like animals and plants (Nagata et al. 2004), apicomplexans also contain a diverse array of EF-hand–containing, CaM-like genes (fig. 5, Supplemental Table 1, Supplementary Material online). CaM contains 4 EF-hands, and a variety of other calcium-responsive proteins also contain this conserved domain. EF-hands are defined by helix-loop-helix structure containing acidic residues that bind calcium, thus causing a conformational change in the protein (Yap et al. 1999). Phylogenetic analysis indicated that T. gondii, P. falciparum, and C. hominis each contain a single conventional CaM consisting of 4 EF-hands, whereas no direct orthologue of CaM was found in C. parvum (fig. 5, Supplemental Table 1, Supplementary Material online). Additional analyses of the apicomplexan genomes analyzed here contain 3 other EF-hand–containing proteins that resemble CETN or caltractin (Supplemental Table 1, Supplementary Material online). CETNs are found in the ciliates, apicomplexans, yeast, worms, and higher metazoans, suggesting that it represents a common ancestral channel present in the early eukaryotic crown group that has since been retained in these divergent taxa.

CETNs may be involved in formation of microtubule spindle bodies or the formation of the flagellar axoneme. Potentially, these CETNs may be involved in basal bodies or the formation of the flagellar axoneme. All 3 apicomplexans also contain CETN-like genes that group with the CETNs CETN3 and CETN4, which are found in the ciliates Paramecium and Tetrahymena (fig. 5, Supplemental Table 1, Supplementary Material online, perhaps reflecting the more streamlined genomes of the latter 2 organisms.

EF-Hand–Containing Proteins

Comparison of apicomplexan genomes reveals that each contains a single highly homologous CaM, a calcium-binding protein implicated in a variety of signaling events (Berridge et al. 2000b). Like animals and plants (Nagata et al. 2004), apicomplexans also contain a diverse array of EF-hand–containing, CaM-like genes (fig. 5, Supplemental Table 1, Supplementary Material online). CaM contains 4 EF-hands, and a variety of other calcium-responsive proteins also contain this conserved domain. EF-hands are defined by helix-loop-helix structure containing acidic residues that bind calcium, thus causing a conformational change in the protein (Yap et al. 1999). Phylogenetic analysis indicated that T. gondii, P. falciparum, and C. hominis each contain a single conventional CaM consisting of 4 EF-hands, whereas no direct orthologue of CaM was found in C. parvum (fig. 5, Supplemental Table 1, Supplementary Material online). Additional analyses of the apicomplexan genomes analyzed here contain 3 other EF-hand–containing proteins that resemble CETN or caltractin (Supplemental Table 1, Supplementary Material online). CETNs are found in the ciliates, apicomplexans, yeast, worms, and higher metazoans, suggesting that it represents a common ancestral channel present in the early eukaryotic crown group that has since been retained in these divergent taxa.

CETNs may be involved in formation of microtubule spindle bodies or the formation of the flagellar axoneme. Potentially, these CETNs may be involved in basal bodies or the formation of the flagellar axoneme. All 3 apicomplexans also contain CETN-like genes that group with the CETNs CETN3 and CETN4, which are found in the ciliates Paramecium and Tetrahymena (fig. 5, Supplemental Table 1, Supplementary Material online, perhaps reflecting the more streamlined genomes of the latter 2 organisms.
online). *Paramecium* has numerous CETNs that are localized to cilia and the infraciliary lattice, which is a contractile cytoskeletal network at the cell cortex (Kim et al. 2002). In *Tetrahymena*, 4 CETNs have been described, and they are localized to the basal bodies, ciliary rootlets, the oral apparatus, and the axoneme (Guerra et al. 2003; Stemm-Wolf et al. 2005). CETN1 and CETN2 are primarily found in basal bodies, CETN3 is localized to oral membranes,
and CETN4 shares this latter location and is also found in contractile vacuole pores (Stemm-Wolf et al. 2005). Apicomplexans lack the oral feeding apparatus of ciliates; yet they share a highly specialized apical microtubular organizing center called the conoid (Morrisette and Sibley 2002). Consistent with this, T. gondii CETN2 (Tg583.05347) has been localized to the centriole, peripheral annulus at the apical pole, and the prenoidal rings of dividing cells (Hu et al. 2006).

Finally, apicomplexans contain numerous other CETN-like genes that contain between 2 and 4 conserved EF-hands (figs. 1 and 5, Supplemental Table 1, Supplementary Material online). Our analysis finds that T. gondii contains at least 9 additional CETN-like genes, whereas C. parvum and P. falciparum each contain 5 genes (fig. 5, Supplemental Table 1, Supplementary Material online). These are all deep-branching members of the present phylogeny, and various divergent ciliate CETNs are also contained within this grouping (fig. 5). Deep-branching CETNs have diverse functions in different organisms. For example, CETN2 has been described to affect homologous DNA repair in Arabidopsis, which expresses multiple CETN-like genes (Moliner et al. 2004). CETN 1p in Paramecium modulates voltage-gated calcium channels in the cilia, thus affecting swimming behavior (Gonda et al. 2004). The function of these divergent EF-hand–containing proteins in apicomplexans is uncertain, but their retention and diversity argues for a variety of nonredundant roles in parasite biology.

EF-hands are also present in calreticulin and calnexin, which function as chaperones in the ER (Bedard et al. 2005). Calnexin is normally a type 1 membrane protein in vertebrates, whereas calreticulin is a luminal ER protein (Bedard et al. 2005). Toxoplasma gondii contains an orthologue (Tg583.05347) that groups with calnexins from various organisms (fig. 5, Supplemental Table 1, Supplementary Material online). This predicted orthologue in T. gondii has 2 TM domains that flank 4 EF-hand domains, which likely protrude into the lumen of the ER (fig. 1). A number of other calcium-binding, sequestering proteins are known from animal or plant cells including S100, troponin C, ERC55, calumenin, calsequestrin, and endoplasmin (Nagata et al. 2004). Based on the present analyses, no direct orthologues of these proteins were identified in apicomplexans.

Calcium and CaM-Dependent Kinases

Calcium- and CaM-dependent kinases comprise one subgroup of serine threonine (S/T) kinases that have evolved differently in plants and animals. Animal cells generally contain Ca\(^{2+}\)-CaM-dependent kinases (CaMK), consisting of an N-terminal S/T kinase domain, an autoinhibitory domain, and a C-terminal domain that binds to CaM (Hook and Means 2001). Binding of Ca\(^{2+}\)–CaM to the C-terminal domain relieves autoinhibition and leads to activation of the kinase (Hook and Means 2001). In contrast, plants contain calcium-dependent protein kinases (CDPKs) consisting of a S/T kinase domain fused to 4 EF-hand domains that function analogously to CaM (Harmon et al. 2000; Harper and Harmon 2005). An autoinhibitor domain lies between the kinase and EF-hand domains and binding to calcium relieves autoinhibition (Harmon et al. 2000; Harper and Harmon 2005). Arabidopsis contains more than 30 CDPKs that regulate a diversity of calcium-dependent functions including drought response, cold and salt stress, mechanical wounding, hormonal signaling, growth, and development (Harmon et al. 2000; Harper and Harmon 2005). Previous studies in malaria and T. gondii have revealed that parasites contain plant-like CDPKs. CDPK1 in T. gondii has been implicated in calcium-dependent protein secretion of microneme proteins, a process required for motility and cell invasion (Kieschnick et al. 2001). The closest orthologue of this gene in malaria is known as CDPK4, and it governs calcium-dependent development of male gametocytes (Bilker et al. 2004). Recently, it was shown that CDPK3 in Plasmodium is required for migration of ookinetes across the peritrophic membrane in the mosquito gut (Ishino et al. 2006).

Our analyses reveal that apicomplexans contain a much larger number of CDPKs than previously recognized (fig. 6, Supplemental Table 1, Supplementary Material online). In total, T. gondii contains 11 CDPK-like genes, whereas P. falciparum contains 8 and C. parvum contains 7. Previous analysis of the kinome of malaria has indicated a total of 5 conventional CDPKs (Doering et al. 2005), and the greater number of putative orthologues reported here is a reflection of the unconventional domain structure of several of these (fig. 6). It has previously been suggested that apicomplexan CDPKs are plant-like (Huang et al. 2004); however, whether this reflects a common ancestry with plants or the horizontal acquisition from an algal endosymbiont is unclear. Consequently, we compared apicomplexan CDPKs with animal CaMK, and CDPKs from higher plants (Nictotinia, Zea, Arabidopsis, Oryza, etc.), algae (Chlamydomonas), and ciliates (Tetrahymena and Paramecium) (Supplemental Table 4, Supplementary Material online). Clustal alignment and phylogenetic analyses revealed that apicomplexan CDPKs fall into 4 broad groups: 2 of these are plant-like and one is a sister group of...
with CDPKs in ciliates, whereas others are divergent or similar to animal CaMK (figs. 6 and 7). The substrate specificity of plant CDPKs indicates that they recognize a general motif of K/R-X-X-S/T (Harper and Harmon 2005). Whether substrate recognized by parasite CDPKs relies on a similar motif awaits identification of the substrates of these kinases.

All 3 apicomplexans contained 5 conventional plant-like CDPKs with 4 EF-hand domains C-terminal to a S/T kinase domain. Based on the present analysis, these were found to fall into 2 separate clades: apicomplexan group 1 is a sister group to CDPKs from ciliates, whereas group 2 is unique to apicomplexans (the numbering of CDPK orthologues in apicomplexans is not identical for historical reasons) (fig. 7, Supplemental Table 1, Supplementary Material online). Both of these conventional CDPK groups are more similar to plants in their overall domain structure, suggesting that they are regulated directly by binding to calcium. The similarity of group 1 CDPKs in apicomplexans with ciliates supports the suggestion that they may have been acquired from the nuclear genome of an algal endosymbiont that was engulfed by a common ancestor of ciliates and apicomplexans (Huang et al. 2004). Additionally, it is clear that parasite CDPKs may have diverged to a greater extent than either plant or ciliate CDPKs since their split with a common ancestor (fig. 7).

Apicomplexans contain divergent calcium-dependent kinases that do not group with ciliate or plant CDPKs, and several of these are only present in T. gondii (fig. 6, Supplemental Table 1, Supplementary Material online). Three CDPKs in T. gondii have fewer than 4 EF-hands (Twin_Scan0307, Tg86.m00003, and TwinScan_5615) (fig. 6). The former 2 proteins contain 3 C-terminal EF-hands, somewhat similar to plant CCaMK, although both are deep-branching in the present phylogeny. TwinScan_5615 has an unusual domain architecture with 2 N-terminal EF-hands followed by a pleckstrin-like homology (PH) domain and then a C-terminal S/T kinase domain (fig. 6). This domain pattern is shared by the P. falciparum orthologue Pf701102, which was previously annotated as an "arginine-rich protein." InterProScan analysis of Pf701102 reveals that it contains an EF-hand domain similar to recoverin, followed by a PH domain, and C-terminal S/T kinase domain. Interestingly, these 2 genes group closer to mammalian CaMK (fig. 7); whether this reflects their unusual domain structure or core similarity in the kinase region is uncertain.

Toxoplasma gondii also contains a CDPK protein with 5 predicted EF-hands (Tg37.m00003), one of which occurs N-terminally to the kinase domain (fig. 6, Supplemental Table 1, Supplementary Material online). Tg37.m00003 branches deeply with T. gondii CDPK3, which also has an unusual domain structure consisting of 6 EF-hands, 2 of which lie within the kinase domain and 4 of which are C-terminal (figs. 6 and 7). Toxoplasma gondii contains another 6 EF-hand domain protein (Tg38.m00014) that also contains 2 N-terminal EF-hands, a S/T kinase domain, and...
A similar domain structure is conserved in P. falciparum Pf701099, a gene previously annotated as a 'hypothetical protein.' Interpro-Scan analysis of the P. falciparum protein reveals that it contains one or two N-terminal EF-hands, a S/T kinase domain, and three or four C-terminal EF-hands. Two unusual CDPKs from C. parvum (CDPK3) and C. hominis (Ch38411) also group with this divergent group of CDPKs (fig. 7). The unusual number and placement of EF-hand domains in these proteins, suggests that they have been generated by recombination from a more primitive conventional CDPK. There is no apparent relationship between the protein domains seen in several of these CDPKs (i.e., TgCDPK1, TgCDPK2, 37.m0003, see fig. 6) and intron–exon boundaries (data not shown), suggesting that these rearrangements are not due to simple domain shuffling. One of these unusual configurations is shared by proteins in T. gondii (38.m0014), P. falciparum (Pf701102), and C. hominis (Ch38411) (fig. 6), suggesting that it was acquired early during the ancestry of the phylum. Others of these more deeply branching CDPKs likely represent members of this S/T kinase family that have diverged during the evolution of apicomplexan parasites, as they are typically unique to one lineage (i.e., 86.m0003, TgCDPK3). A number of T. gondii CDPKs contain potential N-myristoylation motifs (fig. 6), and these are also
found in *Plasmodium* but not in any of the *Cryptosporidium* orthologues (Supplemental Table 1, Supplementary Material online). N-myristoylation has been observed in plant CDPKs to influence intracellular targeting to membranes, suggesting that it may play a similar role in parasites.

Plants also contain a family of S/T kinase that are related to CDPK but which lack conserved C-terminal EF-hands (Harmon et al. 2000). These have been termed CDPK-related kinases (CRKs), and in the present analysis, they group phylogenetically closer to plant CDPK than animal CaMK (Harmon et al. 2000). Although these proteins have not been studied as extensively, evidence suggests that they are not regulated by calcium (Harmon et al. 2000). Apicomplexans also contain a diverse array of CRK-like kinases (Supplemental Table 1, Supplementary Material online). Similar to the case of CDPKs, *T. gondii* contains more CRK-like genes (total 10), whereas *P. falciparum* (5) and *C. parvum* (1) each contain fewer members. The roles of these kinases in parasite biology have not been investigated; however, from the large complement of S/T kinases expressed by *T. gondii*, it is likely that they are involved in a wide range of biological functions. This diversification may be related to expression at different stages of the life cycle, although confirmation of this will require genome-wide expression or proteomic studies.

**Conclusions**

Comparison of calcium pathways between apicomplexan parasites, which represent an early branching eukaryotic lineage and plant, fungal, and animal lineages provides a window on the development of complex signaling mechanisms that arose during evolution. Comparative genomic and phylogenetic analyses of calcium pathways in apicomplexans reveal both conserved pathways and several interesting differences from model organisms. We have based our analysis on conventional methods for recognizing orthologues: reciprocal BlastP comparisons and domain searching using InterProScan, which takes advantage of a number of protein motif searches. The cutoffs used here are fairly stringent (*E* values of $10^{-3}$–$10^{-10}$); hence, it is unlikely that false positives have been identified. However, we may have missed more divergent proteins that participate in calcium metabolism in apicomplexans. Previous studies have emphasized that the highly biased nucleotide composition in malaria results in an amino acid bias (Musto et al. 1995, 1999), and combined with the frequent insertions seen in malarial proteins can compromise BlastP alignments, as highlighted previously (Bastien et al. 2004). Future studies using iterative strategies based on PSI-Blast (Altschul et al. 1997), improved hidden Markov models (Eddy 1995, 1998), or hydrophobic cluster models (Cellebaut et al. 2005) may provide a more complete analysis of the calcium-related proteins in Apicomplexa. We also used fairly conventional methods for Clustal alignments and molecular phylogenies based on distance and parsimony methods. The resulting phylogenies were highly similar, suggesting that they are robust. Nonetheless, further analyses using more advanced models such as Bayesian or maximum likelihood might identify additional orthologues not recognized in the present analyses.

Despite these potential limitations, our analyses identify a large number of calcium-related proteins in the Apicomplexa and provides a preliminary cataloging of these genes for future functional analyses.

Our analyses indicate that apicomplexans contain a variety of P-type ATPases including a single SERCA, and several PMR-like proteins likely govern calcium homeostasis. Calcium channels are the targets of plant alkaloids and marine cone snail toxins, and it may be feasible to exploit the differences in parasite channels to identify selective inhibitors. The potential for this approach has been previously demonstrated based on the natural product artemisinin, which targets malarial SERCA (Eckstein-Ludwig et al. 2003). Apicomplexans also contain numerous plant-like, calcium-dependent kinases that serve essential roles in their biology. Cell cycle kinases have previously been validated as potential targets for cancer therapy (Hannah 2005; Schwartz and Shah 2005) and for malaria (Doerig et al. 2005). The diversity of calcium response proteins containing EF-hands or serine/threonine kinase domains implies that many cellular processes are controlled by calcium-responsive conformational changes (i.e., EF-hand–containing proteins) or by phosphorylation cascades (calcium-dependent S/T kinases). This diversity complicates the analysis of function based on heterologous antibodies or inhibitors, which may interact more than one isoform of these proteins. Fortunately, experimental tools for both forward (Su and Wootten 2004; Khan, Taylor et al. 2005) and reverse genetics (Roos et al. 1994; Carvalho and Menard 2005) are available in *T. gondii* and *Plasmodium* spp., and these approaches will be important for defining the specific roles of these genes in calcium metabolism and signaling.

Elucidation of calcium response pathways in parasites has the potential to identify unique pathways that are distinct from those found in their vertebrate hosts. In a general biological sense, these differences may inform us about recent adaptations that have arisen in vertebrate ancestry and that are not found in early branching eukaryotes. On a more practical level, differences in calcium signaling in parasites may identify potential targets for selective therapeutic intervention against parasites. Such targets might arise from one of several mechanisms: 1) retention of plant-like pathways that are ancestral to the acquired secondary endosymbiont, 2) divergence from vertebrates since their common ancestry (either divergence at the whole protein or within selected domains), or 3) acquisition by horizontal gene transfer. Our ability to capitalize on these differences will require appropriate functional analyses of calcium response pathways in parasites.

**Supplementary Material**

A complete list of genes used in analyses provided as supplemental material in Tables 1–5 is available at *Molecular Biology and Evolution* online (http://www.mbe.oxfordjournals.org).

**Acknowledgments**

This work was supported by grants from the National Institutes of Health (AI067051, AI34306) (L.D.S.) and the
Uehara Memorial Foundation (K.N.). Preliminary genomic and/or cDNA sequence data were provided by The Institute for Genomic Research, the Wellcome Trust Sanger Institute, Washington University, and the University of Pennsylvania. We are grateful to Con Beckers, Eduardo Chini, Kami Kim, Sylvia Moreno, Naomi Morrissette, Boris Striepen, and Gary Ward for helpful comments and discussions.

**Literature Cited**


Lovett JL, Marchesini N, Moreno SN, Sibley LD. 2002. Intracellular calcium stores in


Nagamune and Sibley


Geoffrey McFadden, Associate Editor

Accepted May 25, 2006