The Inner Membrane Complex Sub-compartment Proteins Critical for Replication of the Apicomplexan Parasite Toxoplasma gondii Adopt a Pleckstrin Homology Fold*

Michelle L. Tonkin,†1 Josh R. Beck,‡1 Peter J. Bradley,§1 and Martin J. Boulanger‡2

From the †Department of Biochemistry and Microbiology, University of Victoria, Victoria, British Columbia V8W 3P6, Canada and §Department of Microbiology, Immunology, and Molecular Genetics, UCLA, Los Angeles, California 90095

Background: Inner membrane complex sub-compartment Proteins (ISPs) are critical for proper cell division of apicomplexan parasites, but the mechanism is unknown.

Results: ISPs adopt a pleckstrin homology (PH)-fold yet do not retain phospholipid-binding activity.

Conclusion: ISPs appear to repurpose the phospholipid-binding site to recruit protein partners.

Significance: First structural characterization of an ISP from any organism.

Toxoplasma gondii, an apicomplexan parasite prevalent in developed nations, infects up to one-third of the human population. The success of this parasite depends on several unique structures including an inner membrane complex (IMC) that lines the interior of the plasma membrane and contains proteins important for gliding motility and replication. Of these proteins, the IMC sub-compartment proteins (ISPs) have recently been shown to play a role in asexual T. gondii daughter cell formation, yet the mechanism is unknown. Complicating mechanistic characterization of the ISPs is a lack of sequence identity with proteins of known structure or function. In support of elucidating the function of ISPs, we first determined the crystal structures of representative members TgISP1 and TgISP3 to a resolution of 2.10 and 2.32 Å, respectively. Structural analysis revealed that both ISPs adopt a pleckstrin homology fold often associated with phospholipid binding or protein–protein interactions. Substitution of basic for hydrophobic residues in the region that overlays with phospholipid binding in related pleckstrin homology domains, however, suggests that ISPs do not retain phospholipid binding activity. Consistent with this observation, biochemical assays revealed no phospholipid binding activity. Interestingly, mapping of conserved surface residues combined with crystal packing analysis indicates that TgISPs have functionally repurposed the phospholipid-binding site likely to coordinate protein partners. Recruitment of larger protein complexes may also be aided through avidity-enhanced interactions resulting from multimerization of the ISPs. Overall, we propose a model where TgISPs recruit protein partners to the IMC to ensure correct progression of daughter cell formation.

The phylum Apicomplexa contains >5000 obligate intracellular parasitic protozoans that cause devastating diseases on a global scale. Two of the major pathogens in this phylum are Plasmodium spp., the causative agents of malaria that are responsible for more than one million human deaths per annum (1), and Toxoplasma gondii, the etiological agent of toxoplasmosis, a widespread disease particularly affecting immune-compromised patients and congenitally infected neonates (2). The success of these parasites is largely due to various phylum-specific biological processes, cellular structures, and proteins that hold intriguing promise as therapeutic targets. The inner membrane complex (IMC)3 is an apicomplexan-specific structure that consists of an intricate system of flattened vesicles and associated filamentous network underlying the parasite plasma membrane (3–6). This patchwork of membrane sacs consists of rectangular plates around the center and bottom of the parasite, capped by a single conical apical plate, with openings at the apex and base of the cell (3). The IMC is a critical anchor point for gliding motility, a type of movement unique to apicomplexans that may be required for host cell invasion, and also serves as a scaffold for proper daughter cell formation (7, 8). The limited number of proteins known to associate directly with the IMC (9–13) combined with the importance of the IMC to parasite invasion and replication prompted our recent identification and characterization of a new family of IMC associated proteins (14).

Three proteins were initially identified and found to localize to distinct sub-compartments of the IMC in T. gondii: IMC sub-compartment protein (ISP) 1, ISP2, and ISP3 (14). The ISP family appears to be conserved throughout Apicomplexa, suggesting that these proteins are critical factors in parasite viability. In support of this notion, disruption of TgISP2 caused a

*This work was supported by Canadian Institutes for Health Research Grant MOP82915 (to M. J. B.), a National Institutes of Health Microbial Pathogenesis Training Grant T32 AI07323 (to J. R. B.), and Grant R01 AI064616 (to P. J. B.).

†The atomic coordinates and structure factors (codes 4CHJ and 4CHM) have been deposited in the Protein Data Bank (http://wwpdb.org/).

§Supported by a Natural Sciences and Engineering Research Council of Canada Alexander Graham Bell Canada Graduate Scholarship (CGSD3) and a Ventura Neale Trust Endowed PEO Scholar Award.

‡A Canadian Institutes for Health Research New Investigator, a Michael Smith Foundation for Health Research scholar, and a Canada Research Chair. To whom correspondence should be addressed: Dept. of Biochemistry and Microbiology, University of Victoria, P. O. Box 3055 STN CSC, Victoria, BC, Canada V8W 3P6. Tel.: 250-721-7072; Fax: 250-721-8855; E-mail: mboulang@uvic.ca.

© 2014 by The American Society for Biochemistry and Molecular Biology, Inc. Published in the U.S.A.
significant loss in parasite fitness and a severe defect in endodyogeny (14), the form of internal cell budding in which two daughter cells are formed within the intact mother parasite. In particular, parasites lacking ISP2 generate abnormal numbers of daughter cells during endodyogeny. All three TgISP proteins contain N-terminal cysteine and glycine residues required for targeting to the IMC membranes by palmitoylation and myristoylation, respectively (14). The TgISP proteins are organized in a hierarchical manner through an unknown mechanism, with TgISP1 localizing exclusively to the conical apical plate, TgISP2 localizing to the IMC sub-compartment that begins after the apical cap and extends approximately two-thirds down the length of the parasite, and TgISP3 localizing to the same region as TgISP2 but is distinct in that it appears to only require palmitoylation for membrane association. TgISP4 has lower expression levels than the other family members and a different expression timing through the cell cycle. Disruption of TgISP4 did not show any defects in cell growth or replication, but whether or not redundancy plays a role in ISP function is currently unknown (15).

A fourth member of the ISP family in T. gondii was more recently identified; TgISP4 localizes to the same sub-compartment as TgISP2 but is distinct in that it appears to only require palmitoylation for membrane association. TgISP4 has lower expression levels than the other family members and a different expression timing through the cell cycle. Disruption of TgISP4 did not show any defects in cell growth or replication, but whether or not redundancy plays a role in ISP function is currently unknown (15).

The ISPs have also recently been characterized in Plasmodium berghei, which only contains two family members, ISP1 and ISP3, and both were found to be important for defining apical polarity (16). Although each of the ISPs can be disrupted in T. gondii, only ISP3 knockouts could be obtained in P. berghei, suggesting that ISP1 may be essential. Interestingly, the localization of ISP1 is altered in Δisp3 parasites, indicating these proteins act in a coordinated fashion similar to the TgISPs.
Structural Characterization of the ISP Family

Although the importance of ISPs is clear, the detailed function of these proteins remains elusive. Complicating their functional assignment is the lack of sequence homology to characterized domains or folds. Using a phylogenetic analysis we initially show that the apicomplexan ISPs are divided into two distinct clades. We then used x-ray crystallography to characterize the three-dimensional structures of TgISP1 and TgISP3, representatives of the two phylogenetic clades. Intriguingly, both structures adopt a pleckstrin homology (PH) fold often found to support protein-lipid or protein-protein interactions. Collectively, these first structures of an ISP offer insight into possible mechanisms by which the ISPs enable correct daughter cell formation.

EXPERIMENTAL PROCEDURES

Bioinformatics

Select high confidence apicomplexan ISP sequences were obtained from UniProt, NCBI, and OrthoMCL database searches combined with our previous identification of ISP paralogs and orthologs (14), aligned with MUSCLE (17), and used to generate a maximum likelihood phylogenetic tree tested with 500 bootstrap replicates in MEGA6.0 (18).

For conservation mapping with ConSurf (19), complete or partial apicomplexan ISP sequences were aligned with MUSCLE (17) and mapped onto the TgISP1 core (trimmed to remove ordered expression tag sequences).

Construct Design and Cloning

The genes encoding ISP1 (TGME49_260820), ISP2 (TGME49_237820), and ISP3 (TGME49_316540) were amplified from T. gondii type II cDNA and cloned NheI-NotI into a pET28a vector modified to contain an N-terminal hexahistidine tag separated from the ISP sequence by either a tobacco etch virus protease or thrombin cleavage site. Sequence analysis confirmed that no mutations were introduced during amplification procedures.

TgISP1—Two constructs of TgISP1 were further cloned for functional studies and crystallization trials as described previously (20): TgISP1-1 (post N-terminal localization residues Gly-2, Cys-8, Cys-9 to C terminus; Ala-9 to Ala-176) and TgISP1-2 (conserved core; Pro-58 to Ala-176).

TgISP2—Two constructs of TgISP2 were subcloned for functional studies and crystallization trials. TgISP2-1 extends from Gly-10 after the N-terminal residues involved in localization (Gly-2, Cys-5, Cys-8, and Cys-9) to Ala-160 at the C terminus of the protein. TgISP2-2 extends from Ser-31 to Ala-155 (conserved core).

TgISP3—Two constructs of TgISP3 were subcloned for functional studies and crystallization trials. TgISP3-1 extends from Asp-10 after the N-terminal residues involved in localization (Gly-2, Cys-6, Cys-7) to Asn-164 at the C terminus of the protein. TgISP3-2 extends from Pro-30 to Asn-155 (conserved core).

Protein Expression and Purification

All constructs of the three TgISP proteins were produced recombinantly in E. coli BL21-CodonPlus cells (Stratagene) and purified by nickel-affinity and size exclusion chromatography as described previously for TgISP1-2 (20), with the following exceptions: TgISP1-1, TgISP2-1, and TgISP3-1 contained an N-terminal thrombin cleavage site, and after nickel-affinity purification the proteins were buffer-exchanged into HEPES-buffered saline (20 mM HEPES, pH 7.0, 150 mM NaCl) containing 1 mM DTT or β-mercaptoethanol and 2.25 mM CaCl2, and cleaved with thrombin by overnight incubation at 18 °C.

A selenomethionine (SeMet) version of TgISP1-2 (TgISP1-2_SeMet) was generated by transforming the clone into the methionine auxotroph E. coli B34 strain and inoculating SeMet media (Molecular Dimensions) containing 50 μg/ml ampicillin. The culture was grown at 37 °C to an A600 of 0.9 and induced with 1 mM isopropyl-1-thio-β-D-galactopyranoside. After 12 h of growth at 30 °C, the cells were harvested by centrifugation, and the SeMet-labeled protein was purified using the same protocol as for the native protein (20).

The purity of each protein was assessed at every stage by SDS-PAGE, and protein concentrations were determined by the Bradford assay for TgISP1 and TgISP2 constructs due to the absence of tryptophan residues and by absorbance at 280 nm for TgISP3 constructs.

Crystallization and Data Collection

TgISP1—TgISP1-1 was refractory to crystallization, whereas high quality crystals of native TgISP1-2 were identified as described previously (20). TgISP1-2_SeMet was crystallized in 2.0 M ammonium sulfate, 5% isopropyl alcohol. A single TgISP1-2_SeMet crystal was looped, cryoprotected in saturated lithium sulfate, and flash-cooled at 100 K directly in the cryostream. Diffraction data were collected on beam line 08B1-1 at the Canadian Light Source at the optimized wavelength of 0.9792 Å for the f” selenium edge.

TgISP3—Initial crystal trials for TgISP3-1 were set in 96-well plates (Emerald Biosystems), and crystals were identified in 2.0 mM sodium malonate, pH 7.0. The final sitting drops consisted of 1.5 μl of protein (14 mg/ml in HEPES-buffered saline containing 1 mM DTT) and 1.5 μl of reservoir solution and were equilibrated against 100 μl of reservoir solution. Crystals were optimized to a final condition of 2.0 mM sodium malonate, pH 7.0, 10 mM betaine hydrochloride. A single TgISP3-1 crystal was looped, stepped into a final cryoprotectant of reservoir solution supplemented with 25% glycerol, and flash-cooled to 100 K directly in the cryostream. Diffraction data were collected on beam line 11-1 at the Stanford Synchrotron Radiation Lightsource.

Data Processing, Structure Solution, and Refinement

TgISP1-2—Diffraction data for native and SeMet crystals were processed to 2.10 and 2.70 Å resolution, respectively, using Iosmos (21) and Scala (22) in the CCP4 suite of programs (23). The structure of TgISP1-2 was phased by SeMet single wavelength anomalous dispersion. A total of four selenium sites (two per monomer) were identified and refined using the SHELX C/D/E pipeline (24). High quality phases were obtained after density modification in dm (25) and enabled building and registering of ~60% of the backbone using Buccaneer (26). The remaining structure was built manually and used as a molecular
replacement model for the higher resolution native data using Phaser (27). Solvent atoms were selected using COOT (28) and refined in Refmac5 (29). Complete structural validation was performed in Molprobity (30), and stereochemical analysis of the refined TgISP-1 structure was performed with Rampage in CCP4 (23), with the Ramachandran plot showing excellent stereochemistry with 99% of the residues in the most favored conformation and no residues in disallowed orientations. Overall, 5% of the reflections were set aside for calculation of $R_{\text{free}}$. Data collection and refinement statistics are presented in Table 1.

**TABLE 1**

| Data collection and refinement statistics | TgISP1–2 native | TgISP1–2 SeMet SAD | TgISP3–1 native |
|------------------------------------------|----------------|------------------|----------------|
| Space group | P2$_1$2$_1$2 | P4$_1$2$_1$2 |  |
| $a$, $b$, $c$ (Å) | 58.11, 81.09, 120.07 | 128.15, 128.15, 128.15 |  |
| $\alpha$, $\beta$, $\gamma$ (degree) | 90, 90, 90 | 90, 90, 90 | 90, 90, 90 |
| Wavelength (Å) | 1.542 | 0.9792 |  |
| Resolution range (Å) | 29.09-2.10 (2.21-2.10) | 57.31-2.70 (2.85-2.70) |  |
| Measured reflections | 160,946 | 269,737 | 35,300 |
| Unique reflections | 33,884 | 10,430 |  |
| Redundancy | 4.7 (4.8) | 25.9 (24.0) | 5.2 (4.3) |
| Completeness (%) | 99.9 (100.0) | 100.0 (100.0) |  |
| $I/\sigma(I)$ | 16.3 (4.5) | 27.2 (7.7) |  |
| $R_{\text{merge}}$ | 0.069 (0.380) | 0.088 (0.471) | 0.041 (0.543) |

**Refinement statistics**

| Resolution (Å) | 28.15-2.32 (2.38-2.32) | 35.94-2.32 (2.41-2.32) |  |
| $R_{\text{cryst}}$/| $R_{\text{free}}$ | 0.204/0.256 (0.231/0.301) | 0.218/0.268 (0.302/0.314) |  |
| No. of atoms | 967/970/962/975 | 994 |  |
| Protein (A/B/C/D) | 31.8 | 46.1 |  |
| Solvent | 49.3 | N/A |  |
| B values (Å$^2$) | 26.7/24.5/30.7/31.8 | 64.1 |  |
| Ramachandran statistics (%) | 99.0 | 99.1 |  |
| Most favored | 99.0 | 99.1 |  |
| Allowed | 1.0 | 0.9 |  |
| Disallowed | 0.0 | 0.0 |  |

* $R_{\text{merge}} = \sum_{i,j} F_{\text{iobs}} - F_{\text{iobs}} / \sum_{i,j} F_{\text{iobs}}$, where $F_{\text{iobs}}$ is the average of symmetry related observations of a unique reflection.

* $R_{\text{cryst}} = \| F_{\text{calc}} - F_{\text{obs}} \| / \| F_{\text{obs}} \|$, where $F_{\text{obs}}$ and $F_{\text{calc}}$ are the observed and the calculated structure factors, respectively.

* $R_{\text{free}}$ is $R$ using 5% of reflections randomly chosen and omitted from refinement.

**Accession Numbers**

The atomic coordinates and structure factors have been deposited in the Protein Data Bank under the codes 4CHM (TgISP1–2) and 4CHJ (TgISP3–1).

**Phospholipid Binding Assay**

TgISP1–1, TgISP2–1, and TgISP3–1 were tested for their ability to bind a variety of membrane-associated lipids using Echelon Biosciences PIP Strip assay (P-6001), containing eight phosphoinositides and seven other biologically important lipids. Phosphatidylinositol 4,5-diphosphate-specific binding protein was used as a positive control, and the strips were probed according to the manufacturer’s instructions.

**Homology Modeling**

Secondary structure predictions were performed using PsiPred to determine the core conserved elements for TgISP2 (32). A homology model for TgISP2core (Gly-40 to Ala-152) was generated using Modeler (33) embedded in Chimera (34) with chain A of TgISP1–2 as an input model. The highest confidence model (indicated by the lowest zDOPE score) was chosen and inspected manually for quality.

**RESULTS**

Apicomplexan ISPs Are Separated into Two Clades Represented by TgISP1 and TgISP3—In our previous study we showed that apicomplexan ISPs are broadly conserved and separated into two ortholog groups (14). Here we have performed a phylogenetic analysis on selected ISPs, showing that TgISP1, TgISP2, and TgISP4 cluster in one clade, whereas TgISP3 clusters in a distinct, second clade (Fig. 1B) in agreement with recent observations (16). It is noteworthy that parasites such as T. gondii, Neospora caninum, and the Plasmodium spp. have representatives of both the first and second ortholog groups, whereas parasites such as Cryptosporidium spp. have only a TgISP1/2 ortholog. Based on currently available sequence data, the tissue dwelling coccidia (Toxoplasma, Neospora, and Eimeria spp.) appear to be the only parasites with an ISP1 paralog (i.e. ISP2) (Fig. 1B). Although the ISPs share relatively low sequence identity, we previously found that all three TgISP3s are predicted to have the same core secondary structure elements.
Structure Determination of TgISP1-2 and TgISP3-1—Crystallization trials were performed using all six TgISP constructs. After several rounds of optimization, diffraction quality crystals were obtained for TgISP1-2 and TgISP3-1. Consistent with the lack of sequence identity with reported structures, no suitable molecular replacement model was identified, and phases for TgISP1-2 were ultimately obtained from a SeMet single wavelength anomalous dispersion experiment. The 118-residue structure of TgISP1-2 contains a single methionine, with the addition of a second methionine provided by the residual sequence of length anomalous dispersion experiment. The 118-residue TgISP molecular replacement model was identified, and phases for lack of sequence identity with reported structures, no suitable crystallization trials were performed using all six (18.3 kDa) corresponds to a molecular mass of 27 kDa.

Two constructs were designed for TgISP1 (Fig. 1C, top): nearly full-length (lacking the N-terminal sites for myristoylation and palmitoylation, TgISP1-1) and conserved core (TgISP1-2). TgISP1-1 (19.5 kDa) migrated during gel filtration as a single major peak corresponding to 36.4 kDa, suggesting the formation of a dimer (Fig. 1C, black). The truncated TgISP1-2 (13.9 kDa) eluted in one major peak corresponding to 16.4 kDa with two significant shoulders (Fig. 1C, purple); this elution profile suggests the presence of a conformationally impure monomer population, possibly due to redox instability given the high proportion of cysteines in this small domain (20). The gel filtration traces for both TgISP2 constructs indicate a nearly equal mixture of monomer and dimer (Fig. 1D), with TgISP2-2 (14.5 kDa), for example, eluting as two major peaks corresponding to molecular masses of 35.3 and 18.3 kDa (Fig. 1D, blue). The two constructs of TgISP3 showed similar elution profiles with only one major peak, likely corresponding to a semi-globular monomer (Fig. 1E); the major peak for TgISP3-1 (18.3 kDa) corresponds to a molecular mass of 27 kDa.

Structure of TgISP1-2 and TgISP3-1 crystallized as a monomer with four molecules in the asymmetric unit (AU) of the primitive cubic unit cell. In contrast, native TgISP1-2 crystallized as a monomer with four molecules in the AU of the primitive orthorhombic unit cell (35). The four monomers of native TgISP1-2 overlay well, with root mean square deviations relative to chain A of 0.75 Å over 124 Ca (chain B), 0.95 Å over 124 Ca (chain C), and 0.79 Å over 123 Ca (chain D). All structural analyses were performed with chain A unless otherwise noted. The 2.10 Å resolution structure of TgISP1-2 is modeled completely from Ser-54 to Ala-178 (Fig. 2A). At both termini, the residual sequence from the expression vector was well ordered; the first TgISP1 residue modeled is Pro-58, and the last is Ala-176. The structure of TgISP1-2 emerged as a suitable molecular replacement model for TgISP3-1, which crystallized as a monomer with one molecule in the AU of the body-centered tetragonal unit cell (35). TgISP3-1 was refined to a resolution of 2.32 Å with the final model extending from Ala-15 through Glu-153 with two regions of disorder (Fig. 2B). Although a significant portion of the N-terminal coil is ordered, a five-residue region between Val-25 and Ile-31 could not be modeled without ambiguity. In addition, a portion of the loop between Ser-97 and Lys-108 was disordered, although some intervening density was present.

The TgISP1-2 adopts a conserved fold—despite sharing only 25% sequence identity, structural superposition reveals that both TgISP1 and TgISP3 adopt the same structural scaffold, as evidenced by a root mean square deviation of just 1.45 Å over 105 Ca. Overall, an N-terminal helix is positioned on the periphery of a twisted, seven-stranded antiparallel β-sandwich bordered on one end by interstrand loops (open end) and capped at the other end by an amphipathic C-terminal helix (closed end) (Fig. 2C). One side of the sandwich is formed by β1-β2-β3-β4, whereas the other side is formed by β5-β6-β7; for both structures this differs from the predicted secondary structure where β5 was predicted to be α-helical (20). Although the majority of connecting segments between the secondary structure elements are short turns, an extended variable loop is present between β5 and β6 (Fig. 2C, dotted arrow). In TgISP1-2, this loop contains an α-helix (Fig. 2, A and C) and along with the ordered coil packs against the core domain; in contrast, in TgISP3-1, this loop is completely coiled and extends away from the core domain, and in an unusual organization, the N-terminal coil threads through the partially disordered β5-β6 loop (Fig. 2, B and C). Additionally, a disulfide bond is present in the structure of TgISP1-2, linking the middle of β2 to the tip of β3 (Fig. 2C, solid arrow). A structure-based sequence alignment constructed from the structures of TgISP1 and TgISP3 and a high confidence homology model of TgISP2 (data not shown) support the conclusion of a conserved fold across the ISP family (Fig. 2D).

The ISP family adopts a pleckstrin homology fold—having established the fold conservation of the TgISP, we next sought to compare the ISP fold to those of known structures. A DaliLite search (36) revealed that the overall topology and organization of secondary structure elements for TgISP1-2 and TgISP3-1 (Fig. 3, A and B) displays significant similarity with PH domain-containing proteins: phospholipase Cζ1 (PLCζ1_PH; PDB ID 1MAI-A) (Fig. 3C) was identified as the top hit for both TgISP1-2 (Z-score of 12.4) and TgISP3-1 (Z-score of 10.8) (37). Despite showing negligible sequence identity (≤15%), the clear structural homology of both TgISP1 and TgISP3 with multiple PH domains has prompted our classification of the ISP proteins as adopting a PH fold. It is noteworthy that low sequence identity is commonly observed between characterized PH domains and explains why this domain was not readily apparent in earlier bioinformatic analyses. The N-terminal helix of both TgISP1-2 and TgISP3-1 (Fig. 3) is not part of the canonical PH domain; however, a short helix in this position is observed in some other PH domain-containing proteins such as PLCζ1 (37).

PH domains typically mediate protein-lipid and protein-protein interactions in diverse protein families. Their function is
best characterized in higher eukaryote proteins with little known about the role of proteins containing this fold in single celled organisms. In fact, PH-like domains were only recently identified in bacteria (38). The structures of $TgISP_{1-2}$ and $TgISP_{3-1}$ presented here are, to our knowledge, the first structural characterization of PH domains from a eukaryotic pathogen.

$TgISP$ PH Domains Are Cysteine-rich with a Unique Disulfide in the ISP1 Ortholog Group—All three of the major $TgISP$ proteins have notably high cysteine content in the core domain. The 176-residue $TgISP_{1}$ contains 7 cysteines (5 in the core domain), and the 160-residue $TgISP_{2}$ contains 9 cysteines (6 in the core domain), whereas $TgISP_{3}$ is 164 residues and contains 6 cysteines (4 in the core domain). Despite being intracellular proteins, the structure of $TgISP_{1-2}$ shows one disulfide bond linking the middle of $\beta_9$ to the tip of $\beta_{10}$ (Cys-90 to Cys-105; Fig. 3A, black arrow), which generates a more compact core domain. Of the three free cysteines, two (Cys-142 and Cys-166) are somewhat buried, whereas one (Cys-153) is surface-exposed and displays two conformations (Fig. 3A). A disulfide bond between $\beta_2$ and $\beta_3$ is likely a structural feature of the $TgISP1$ ortholog group, as both cysteines appear broadly conserved, with the exception of $TgISP1$ orthologs in Cryptosporidium (14). A disulfide is also predicted in a similar position for $TgISP2$ (Fig. 2D). $TgISP3$ does not cluster with the $TgISP1$ ortholog group (14) (Fig. 1A). Cys-105 of $TgISP1$ is not conserved in $TgISP3$ (Fig. 2D), and no disulfide bonds are present in the structure of $TgISP3$ (Fig. 3B). However, Cys-114 at the base of $\beta_6$ is surface-exposed and capped with a $\beta$-mercaptoethanol molecule (Fig. 3B). In addition, Cys-104 is not modeled but would be surface-exposed and part of the extended $\beta_5$-$\beta_6$ loop. The cysteine-rich nature of the $TgISP$s and the number of surface-exposed cysteines may result in redox instability and explain the multiple conformations observed during purification (Fig. 1C) (20) and may also facilitate higher order multimerization (Fig. 1D).

Of the numerous structurally characterized PH domains, a high cysteine content is relatively uncommon but is present in selected PH domains such as Evcetin-2 PH (PDB ID 3VIA). Furthermore, only two disulfide bond positions have been reported in characterized PH domains. The PH domain of DAPP1/PHISH (for example PDB ID 1FB8) has a disulfide linking the base of $\beta_6$ to the tip of $\beta_7$, whereas the PH domains of PKB/AKT1 (for example PDB IDs 1UNP/1P6S/3O96) have a disulfide linking the base of $\beta_5$ to the tip of $\beta_6$. Therefore, the disulfide found in $TgISP1-2$ (Fig. 3A) is unique among structurally characterized PH domains. It is worth noting that the structure of the kindlin-2 PH domain (PDB ID 4F7H) shows two free cysteines in the same backbone positions as the disulfide found in $TgISP1-2$.
fide bound cysteines of TgISP1-2; however, the side chains are directed away from the disulfide axis.

The Phospholipid Binding Properties of Many PH Domains Are Not Conserved in the TgISPs—The vast majority of proteins containing a PH domain are multimodular, with some PH domains having a membrane anchoring role through an interaction with inositol phosphate (IP) head groups of membrane-associated phosphatidylinositides (39, 40). In fact, the PH domain of PLC\(\gamma_{1}\) displays the highest structural similarity to the PH domain of TgISP1-2 and TgISP3-1, binding inositol-1,4,5-trisphosphate with high affinity (37). We, therefore, tested the ability of the recombinant TgISPs to coordinate a variety of mono-, di-, and tri-phosphorylated phosphatidylinositols and seven other biologically relevant lipids including phosphatidylserine using commercial lipid arrays. Despite obtaining expected results with the positive control protein, no detectable interaction was observed between the nearly full-length constructs of TgISP1, TgISP2, or TgISP3 with any of the tested phospholipids (data not shown).

DISCUSSION

The ISPs are small single domain proteins conserved across phylum Apicomplexa (Fig. 1B) that are critical for proper asexual cell division in T. gondii (14) and for defining apical polarity in P. berghei (16). Our structural and biochemical characterization of the three major TgISPs revealed unexpected and intriguing layers of complexity imparted by conformational heterogeneity, multimerization, and semi-globular forms (Fig. 1) assembled around a PH core.

Comparative Structural Analysis Supports the Absence of TgISP Phospholipid Binding Properties—Classically, PH domains play important roles in mediating membrane localization through IP binding (41). Key structural features that mediate these interactions are apparent. For example, a highly basic patch on the surface of PLC\(\gamma_{1}\) to coordinate a variety of mono-, di-, and tri-phosphorylated phosphatidylinositols and seven other biologically relevant lipids including phosphatidylserine using commercial lipid arrays. Despite obtaining expected results with the positive control protein, no detectable interaction was observed between the nearly full-length constructs of TgISP1, TgISP2, or TgISP3 with any of the tested phospholipids (data not shown).

DISCUSSION

The ISPs are small single domain proteins conserved across phylum Apicomplexa (Fig. 1B) that are critical for proper asexual cell division in T. gondii (14) and for defining apical polarity in P. berghei (16). Our structural and biochemical characterization of the three major TgISPs revealed unexpected and intriguing layers of complexity imparted by conformational heterogeneity, multimerization, and semi-globular forms (Fig. 1) assembled around a PH core.

Comparative Structural Analysis Supports the Absence of TgISP Phospholipid Binding Properties—Classically, PH domains play important roles in mediating membrane localization through IP binding (41). Key structural features that mediate these interactions are apparent. For example, a highly basic patch on the surface of PLC\(\gamma_{1}\) to coordinate a variety of mono-, di-, and tri-phosphorylated phosphatidylinositols and seven other biologically relevant lipids including phosphatidylserine using commercial lipid arrays. Despite obtaining expected results with the positive control protein, no detectable interaction was observed between the nearly full-length constructs of TgISP1, TgISP2, or TgISP3 with any of the tested phospholipids (data not shown).
and their organization strongly support the biochemically observed lack of phospholipid headgroup binding. For TgISP1, Lys-108 and Arg-110 are the only positively charged residues in the region, and the high sulfate concentrations used during crystallization resulted in a single sulfate bound to Arg-110 in a position that is inconsistent with any of the phosphates of the IP bound to PLC\(\delta\)1_PH (Fig. 4B). In fact, the disulfide bond of TgISP1-2 (and likely other ISP1 orthologs) lies within the region of IP coordination of other PH domains and restricts rearrangement of the\(\beta_2\)-\(\beta_3\) loop that would likely be required for spatial accommodation of a phospholipid headgroup (Fig. 4B, bottom). For TgISP3, Lys-81 and Arg-83 are analogous to the Lys/Arg pair in TgISP1. TgISP3 additionally contains Arg-125 on \(\beta_7\), but a clustering of the residues was not observed (Fig. 4C). Thus the divergent surface chemistry of the TgISPs relative to classical PH domains correlates with the lack of observed phospholipid binding. It is not necessarily surprising that TgISPs do not require phospholipid binding functionality given that they each harbor multiple N-terminal residues predicted to be myristoylated or palmitoylated (Fig. 2D), leading to proper localization and anchoring to the IMC. Indeed, membrane anchorage by these acylated N-terminal residues may position the key IP interacting face of each ISP PH domain away from the membrane surface (see Fig. 6).

**FIGURE 4. Surface analysis and basic side chain distribution support the lack of a phospholipid binding role for the TgISPs.** A, chimera (34) coulombic-colored surface representation of the PH domain of phospholipase C\(\delta\)1 (PLC\(\delta\)1_PH; PDB ID 1MAI) shows a concentration of basic charge around the inositol-(1,4,5)-trisphosphate (Ins(1,4,5)P\(_3\); green ball-and-stick representation) binding site at the bottom of the molecule open end (top), which is due to a clustering of basic residues forming specific interactions with the ligand (bottom). B, TgISP1-2; coulombic-colored surface shows a strong hydrophobic patch and just two basic residues forming separate small basic patches (top); a sulfate ion is seen coordinated predominately by Arg-110 near the top of the open end, and the constraining disulfide bond is indicated by a black arrow (bottom). C, TgISP3-1; coulombic-colored surface shows both basic and hydrophobic character (top), but the residues are dispersed (bottom).

**Structural Analysis of the ISPs in the Context of PH Domain Interactions Suggests a Protein Binding Functional Mechanism—**

The majority of PH-phospholipid interactions characterized thus far show coordination of the IP or phosphatidylserine by the inner face of the \(\beta_1\)-\(\beta_2\)-\(\beta_3\)-\(\beta_4\) sheet, with atypical coordination by the outer face of the \(\beta_1\)-\(\beta_2\) and \(\beta_5\)-\(\beta_6\) loops (Fig. 5A, left). However, an increasing number of studies imply a much broader functional range for PH domains with mounting evidence supporting roles in protein-protein interactions that employ several surfaces on the PH scaffold (Fig. 5A, right) (44, 45). For example, PH domains have been found to use a number of basic residues located on \(\beta_1\) and \(\beta_2\) to bind filamentous actin (46), a cleft between \(\beta_5\) and the C-terminal helix to accommodate phosphotyrosine peptides (47), or the outer surface of the \(\beta_5\)-\(\beta_6\)-\(\beta_7\) sheet to bind polyproline peptides (48). In fact, nearly every possible surface on the PH domain scaffold has been shown to bind a protein partner or facilitate interdomain interactions of multimodular proteins (44). From these analyses it is clear that PH domains have a high degree of plasticity supported by the ability to functionalize different regions of the PH fold.

Based on the current knowledge of PH domain functions combined with the lack of any observable affinity of the TgISPs for IPs, we hypothesize that the TgISPs have protein binding partners. Because the TgISPs were shown to be tightly regulated
during cell division with highly specific localization patterns (14), a protein-protein interaction functional mechanism would fit with the TgISPs recruiting protein partners to specific sub-compartments of the IMC to enable proper daughter cell formation. Although our initial attempts to identify a binding partner for TgISP1 were unsuccessful (14), the possibility of interactions that are lower affinity, transient, or specific to a certain stage in the cell cycle remains a possibility.

Conservation Analysis Suggests the TgISPs Repurpose the Phospholipid-binding Site—To hone in on which surface of the PH scaffold is most likely to be utilized by the ISPs, an analysis of the distribution of conserved and variable residues across all known apicomplexan ISP sequences was mapped onto the TgISP1 core structure using the ConSurf server (19). Conservation mapping showed a clear bias of conserved residues on the open side of the molecule composed of the β1-β2, β3-β4, and β6-β7 loops (Fig. 5B); this surface overlaps significantly with the phospholipid binding pocket of other PH domains. Given the highly polarized distribution of conserved residues, the ISPs likely maintain a conserved function utilizing the surface at the open end of the PH domain.

Crystal Packing Suggests a Possible Mechanism for TgISP Protein Binding Functionality—An indication of how the ISPs might coordinate a protein partner in this area can be observed in the deep groove present on the surface of TgISP1-2 (Fig. 5C, left), and the ability of this groove to coordinate the N-terminal tail of a neighboring molecule (Fig. 5C, bottom). Although this specific coordination event may be an artifact of crystallization, it is interesting that it is observed in both the TgISP1-2 cubic and orthorhombic crystal forms. Ultimately, the observed association does offer insight into a possible mechanism of macromolecular assembly. Overall, nearly 2200 Å² of surface area is buried at the interface between the two chains, with β7 and the β5-β6 loop of one molecule forming 14 hydrogen bonds with the N-terminal tail, β5-β6 loop, and C-terminal helix of a neighboring molecule (35). More than 50% of the total buried surface at the interface of the two molecules is due to the extended, buried N-terminal residues. In particular, a Met residue buries into a deep pocket lined by hydrophobic residues (Fig. 5C, bottom). Although the individual residues are not conserved with TgISP3, a similar hydrophobic pocket is present. Additionally, TgISP1 Arg-110, which is completely conserved among known apicomplexan ISP sequences (Fig. 5B), coordinates a sulfate ion that sits 4 Å from the N-terminal Ala residue (Fig. 5C, inset). It is tempting to speculate that the binding site on the physiological partner for each ISP has a strong hydrophobic component adjacent to a negatively charged or phosphorylated amino acid. Ultimately, the presence of the groove on TgISP1 that can accommodate a non-native peptide may highlight the fundamental plasticity of this region to promote complex formation between divergent partners.

Multimerization and Phosphorylation May Play a Role in Regulating Biologically Relevant Forms of TgISPs—Multimerization may also contribute to the specific functions of each of the TgISPs, possibly by providing composite interaction surfaces or increasing the avidity of ligand binding (Fig. 6) (49). Based on reproducible size exclusion chromatograms, TgISP1 dimerization appears to be dependent on the presence of the

![Structural Characterization of the ISP Family](image)

**FIGURE 5.** The region corresponding to lipid binding functionality observed for other PH domains is highly conserved in the apicomplexan ISPs and capable of coordinating a polypeptide structure in TgISP1. A, surfaces on PH domains identified to mediate interactions with phospholipid headgroups (left) or proteins/peptides (right) are mapped onto TgISP1. B, mapping of conserved (burgundy) and variable (teal) residues onto the TgISP1 core using ConSurf (19) shows a clear bias for conservation on one side of the protein. Viewpoint in (B, left) is aligned with RnPLC-Δ1 of Fig. 4A; C, left, the purple surface of TgISP1-2 shows a deep groove (black line) that partially overlays with the phospholipid binding region of other PH domains. Bottom, in every chain of the AU in two different space groups, the N terminus of a neighboring chain (gray schematic; SMASPQV sequence is shown as in ball and stick representation) buries into the surface groove. Inset, the positioning of the sulfate ion coordinated by the highly conserved Arg-110 is consistent with an acidic or phosphorylated amino acid at the A position (gray dotted line).
full N-terminal tail (Figs. 1C and 6), which is predicted to contain ordered regions of secondary structure (Fig. 2D). The native N-terminal tails of TgISP1 are unlikely to promote dimerization through interactions with the surface groove as no strongly hydrophobic residues are present in the coil regions with the structural flexibility to be coordinated in the groove (Fig. 2D); also, this type of N-terminal interaction with the core domain of a second molecule generates an oligomer in the crystal, not a distinct dimer as observed in gel filtration. TgISP2 has a long N-terminal tail predicted to be disordered (Fig. 2D), and the elements mediating the monomer-dimer equilibrium are fully retained in the core domain (Figs. 1D and 6). Finally, the way that the TgISP3 N-terminal tail wraps loosely around the core domain and through the β5-β6 loop helps explain its semiglobular nature (Figs. 1E and 6), and the monomeric state of TgISP3 is consistent with the crystal structure; only one molecule in the AU that does not show significant interaction interfaces with any neighboring molecules.

Phosphorylation may also play a role in ISP function, either influencing the ability to dimerize or to interact with other binding partners. A previous study on the phosphoproteomes of T. gondii and Plasmodium falciparum revealed that PfISP3, TgISP2, and TgISP3 are all phosphorylated in the parasites (50). Based on our structures and homology model, the three phosphorylation sites on TgISP2 map to a tandem pair of serine residues in the N-terminal tail and a serine at the tip of the binding partners. A previous study on the phosphoproteomes influencing the ability to dimerize or to interact with other faces with any neighboring molecules.

FIGURE 6. Multimeric state may play a role in ISP function. A schematic representation of experimentally supported organizations of the ISPs on the inner membrane complex of T. gondii is shown. ISP structures and homology models are shown as a colored schematic backbone with a semi-transparent gray surface. Purple ovals represent predicted N-terminal helices of TgISP1. Orange starbursts indicate phosphorylation sites. TgISP1 core and TgISP3 full were structurally determined, whereas TgISP2 full-length monomer and dimer are homology models.

REFERENCES
1. World Health Organization. (2010) World Malaria Report, World Health Organization, Geneva, Switzerland
2. Hill, D. E., Chirukandoth, S., and Dubey, J. P. (2005) Biology and epidemiology of Toxoplasma gondii in man and animals. Anim. Health Res. Rev. 6, 41–61
3. Pochet, E., and Torpier, G. (1977) Freeze fracture study of Toxoplasma and Sarcocystis infective stages (author’s transl.). Z. Parasitenkd. 54, 101–124
4. D’Haese, J., Mehlhorn, H., and Peters, W. (1977) Comparative electron microscope study of pellicular structures in coccidia (Sarcocystis, Besnoitia, and Eimeria). Int. J. Parasitol. 7, 505–518
5. Mann, T., and Beckers, C. (2001) Characterization of the subpellicular network, a filamentous membrane skeletal component in the parasite Toxoplasma gondii. Mol. Biochem. Parasitol 115, 257–268
6. Morrisette, N. S., Murray, I. M., and Roos, D. S. (1997) Subpellicular microtubules associate with an intramembranous particle lattice in the protozoan parasite Toxoplasma gondii. J. Cell Sci. 110, 35–42
7. Keeley, A., and Soldati, D. (2004) The glideosome: a molecular machine powering motility and host-cell invasion by Apicomplexa. Trends Cell Biol. 14, 528–532
8. Striepen, B., Jordan, C. N., Reiff, S., and van Dooren, G. G. (2007) Building

Acknowledgment—We gratefully acknowledge the staff at Stanford Synchrotron Radiation Lightsource and Canadian Light Source.
Miyazaki, M., Trüb, T., Shoelson, S. E., and Fesik, S. W. (1996) Structural basis for IL-4 receptor phosphopeptide recognition by the IRS-1 PTB domain. *Nat. Struct. Biol.* 3, 388–393

48. Prehoda, K. E., Lee, D. I., and Lim, W. A. (1999) Structure of the enabled/ VASP homology 1 domain-peptide complex: a key component in the spatial control of actin assembly. *Cell* 97, 471–480

49. Lemmon, M. A., and Ferguson, K. M. (2000) Signal-dependent membrane targeting by pleckstrin homology (PH) domains. *Biochem. J.* 350, 1–18

50. Treeck, M., Sanders, J. L., Elias, J. E., and Boothroyd, J. C. (2011) The phosphoproteomes of *Plasmodium falciparum* and *Toxoplasma gondii* reveal unusual adaptations within and beyond the parasites’ boundaries. *Cell Host Microbe.* 10, 410–419