Comprehensive Proteomic Analysis of Membrane Proteins in Toxoplasma gondii

Fa-Yun Che‡§, Carlos Madrid-Aliste§¶, Berta Burd§, Hongshan Zhang§, Edward Nieves§**, Kami Kim§‡‡§§, Andras Fiser§¶**, Ruth Hogue Angeletti§||**, and Louis M. Weiss‡§¶¶

Toxoplasma gondii (T. gondii) is an obligate intracellular protozoan parasite that is an important human and animal pathogen. Experimental information on T. gondii membrane proteins is limited, and the majority of gene predictions with predicted transmembrane motifs are of unknown function. A systematic analysis of the membrane proteome of T. gondii is important not only for understanding this parasite’s invasion mechanism(s), but also for the discovery of potential drug targets and new preventative and therapeutic strategies. Here we report a comprehensive analysis of the membrane proteome of T. gondii, employing three proteomics strategies: one-dimensional gel liquid chromatography-tandem MS analysis (one-dimensional gel electrophoresis LC-MS/MS), biotin labeling in conjunction with one-dimensional gel LC-MS/MS analysis, and a novel strategy that combines three-layer “sandwich” gel electrophoresis with multidimensional protein identification technology. A total of 2241 T. gondii proteins with at least one predicted transmembrane segment were identified and grouped into 841 sequentially nonredundant protein clusters, which account for 21.8% of the predicted transmembrane protein clusters in the T. gondii genome. A large portion (42%) of the identified T. gondii membrane proteins are hypothetical proteins. Furthermore, many of the membrane proteins validated by mass spectrometry are unique to T. gondii or to the Apicomplexa, providing a set of gene predictions ripe for experimental investigation, and potentially suitable targets for the development of therapeutic strategies. Molecular & Cellular Proteomics 10: 10.1074/mcp.M110.000745, 1–14, 2011.

Toxoplasma gondii (T. gondii), an obligate intracellular parasitic protozoan, is a common causative agent of disease in immune competent and immune compromised humans (1, 2). It is capable of causing severe congenital neurological impairment if acquired in utero (1, 2). Infection can reactivate from latent tissue cysts in patients with immune suppression (AIDS or chemotherapy) resulting in acute infections most often manifesting as encephalitis. Toxoplasmosis is a major cause of death among patients with AIDS (3, 4). The development of new preventative and therapeutic strategies relies on an improved understanding of the interaction between pathogens and their hosts (5). Reported T. gondii infection rates can be as high as 70%, depending on the populations or geographic area studied. T. gondii can infect all warm-blooded mammals although the definitive hosts are members of the cat family.

There are two phases in the parasite’s life cycle, the sexual phase, which takes place only in felines, and the asexual phase, which takes place in any warm-blooded animal. T. gondii has three functionally distinct pathogenic forms, sporozoites (in oocysts), tachyzoites, and bradyzoites (in tissue cysts) (6). When tissue cysts (e.g. from an infected mouse) are ingested by a cat, the cysts survive the passage through the stomach to infect the epithelial cells of the small intestine, where they differentiate and reproduce sexually eventually forming oocysts, which are then shed with the feces. The oocysts contain sporozoites that develop into tachyzoites upon ingestion by mammals. In a similar fashion, if tissue cysts (which contain bradyzoites) are ingested by other mammals they can differentiate into tachyzoites. Tachyzoites, the invasive form of T. gondii, rapidly multiply asexually, invade host cells, and are distributed via the blood stream and lymphatic system throughout the body including the brain. Tachyzoites are responsible for acute infection and disease. Tachyzoites differentiate in response to stress, such as the inflammatory response, into bradyzoites, which are contained within tissue cysts. Humans, as intermediate hosts, can be exposed to and infected with T. gondii via food-borne, water-borne, or maternofetal routes. Maternofetal transmission causes congenital infection which can result in miscarriage, mental retardation, learning disabilities, blindness, microcephaly, and seizures, or death (3, 7, 8, 9). Upon invasion of host cells T. gondii can commandeer host functions affecting the secretion of cytokines (2) and other cellular processes ultimately bypassing host cell defenses (10, 11). The invasion process involves several steps including the recognition of
surface receptors by surface antigens and the secretion of proteins from specialized secretory organelles, such as rhoptries, micronemes, and dense granules. Invasion results in the formation of the parasitophorous vacuole membrane (PVM) by invaginating host plasma membrane and selective exclusion of many host integral membrane proteins (10, 11). The PVM envelops and shelters the parasite inside the parasitophorous vacuole and prevents acidification of this compartment, in which parasites can grow and replicate (12). Parasite membrane proteins play critical roles in invasion and interaction with host cells. For example, *T. gondii* apical membrane antigen-1 (TgAMA1), a type I transmembrane protein, localizes to the parasite’s micronemes, secretory organelles. TgAMA1 depletion inhibits secretion of the rhoptries whose discharge is coupled to active host cell penetration (13). Besides cell-cell interactions, other functions of membrane proteins include ion and solute transport (14), cell signaling, and catalysis. It is likely that, as in other eukaryotes, targets of successful therapeutic drugs will be identified among these parasite membrane proteins. In addition, analysis of the membrane proteome of *T. gondii* should further improve our understanding of the molecular mechanisms of parasite invasion and pathogenesis and assist in identifying new chemotherapy targets. Several thousand proteins are predicted to be membrane bound in the *T. gondii* genome according to transmembrane segment prediction algorithms. Presently, however, only a limited number of membrane proteins have been reported and characterized in this parasite (15, 16), reflecting the challenges in studying membrane proteins because of their high hydrophobicity and low abundance.

Membrane proteins are insoluble in aqueous solution, putting researchers at a dilemma when handling membrane protein samples and carrying out downstream proteomics. Solubilization of membrane proteins in aqueous solution requires the use of detergents (such as SDS or Triton X-100), nonpolar solvents, or sometimes denaturing agents. The presence of relatively high concentrations of detergents or denaturing agents in a sample generally results in inefficient enzymatic digestion (because of enzyme inactivation) and reduces the downstream performance of peptide separation on reverse phase chromatography. Furthermore, a major disadvantage of the presence of detergent in the sample is the suppression of peptide ionization and detrimental interference during electrospray ionization. Background ions raised from detergents can obscure the signal in mass spectra. In addition, detergents can cause the formation of adducts and a shift of the charge envelope (17), resulting in the interpretation of mass spectra becoming more difficult. Researchers have developed several variants of acid cleavable detergents, such as sodium 3-(4-(1, 1-bis (hexyloxy)ethyl)pyridinium-1-yl)propane-1-sulfonate, ProteaseMAX, and RapiGest SF, which can be used to help solubilize and unfold membrane proteins to improve enzymatic digestion. Following digestion, the detergent undergoes hydrolysis under acidic conditions facilitating proteomic studies, and then is removed from the sample solution.

To overcome the obstacles from using detergents to study membrane proteins, one-dimensional gel electrophoresis coupled liquid chromatography-tandem mass spectrometry analysis (one-dimensional gel LC-MS/MS) has proven to be a relatively simple and effective approach (18, 19). The high concentration of detergents can be removed after membrane proteins are resolved on SDS-PAGE gel.

Another approach to membrane protein analysis that specifically targets plasma membrane proteins is biotinylation of cell surface proteins and affinity purification. Cells are labeled with sulfo-NHS-SS-biotin and the labeled proteins are collected with streptavidin beads, resolved on SDS-PAGE gel, followed by the typical protein identification procedure using LC-MS/MS (20, 21). Recently, Zhu and colleagues developed a “Tube-Gel” digestion protocol for membrane proteins (22). Membrane proteins in the presence of high concentrations of detergents are incorporated into a polyacrylamide gel matrix prior to casting the gel using a miniaturized tube. The tube gel is removed without electrophoresis and washed to remove detergents before in-gel digestion. The Tube-Gel digestion protocol can be used to solubilize membrane proteins with minimal interference in downstream LC-MS/MS analysis (22). A drawback, however, with Tube-Gel protocol is that acrylamide can chemically modify some amino acid residues in proteins. A novel gel electrophoresis system, three-layer “sandwich” gel electrophoresis (TLSGE), was recently introduced by Lynn and colleagues for efficient salt removal and concentrating protein (23). A powerful feature of TLSGE is its ability to handle relatively large volumes of sample, even several milliliters, concentrating proteins on a small piece of protein gel and removing detergents and salts during electrophoresis. It is a very useful enrichment strategy for low abundance membrane proteins. In this paper we report the benefits of combining TLSGE with MudPIT (TLSGE MudPIT) and RapiGest SF for the proteomic analysis of membrane proteins in *T. gondii*.

To comprehensively decipher the membrane proteome of *T. gondii*, we implemented three approaches: one-dimensional gel LC-MS/MS, TLSGE MudPIT, and biotin-directed affinity purification (BDAP) (20, 21). The MS/MS data were analyzed in the context of a variety of alternative gene predictions from *T. gondii*, and 841 sequentially nonredundant *T. gondii* membrane protein clusters were validated from thousands of predicted membrane proteins. Approximately, 42% of these identified membrane proteins are hypothetical proteins, 50% of which are unique to *T. gondii*.

---

1 The abbreviations used are: LC-MS/MS, liquid chromatography-tandem MS; TLSGE, TLSGE, three-layer sandwich gel electrophoresis; BDAP, biotin directed affinity purification; TLSG, three-layer sandwich gel; GPI, glycosylphosphatidylinositol.
EXPERIMENTAL PROCEDURES

Reagents—Thioglycolopyranoside, Percoll, urea, thiourea, octyl-β-glucoside, dithiothreitol, iodoacetamide, tris(2-carboxyethyl)phosphine, ammonium bicarbonate, Coomassie Brilliant Blue R-250, KCl, KH2PO4, H2PO4 and Na2CO3 were purchased from SIGMA (St. Louis, MO). Acetoneitrile was from Fisher Scientific (Pittsburgh, PA). Complete™ Proteinase inhibitor mixture was purchased from Santa Cruz Biotechnology (Santa Cruz, CA). Trifluoroacetic acid, formic acid and EZ-Link™ sulfo-NHS-SS-biotin were purchased from Pierce (Rockford, IL). Sequence grade trypsin was purchased from Promega (Madison, WI). Streptavidin magnetic beads were purchased from BioLabs (Ipswich, MA). RapiGest™SF was purchased from Waters (Milford, MA). Dubbelco’s modified eagle medium, phosphate buffered saline (PBS), fetal calf serum, glutamine, and penicillin streptomycin were purchased from Invitrogen (Carlsbad, CA).

Growth of T. gondii in vitro and Isolation of Tachyzoites—RH strain T. gondii parasites were maintained by serial passage in confluent monolayers of human foreskin fibroblasts in Dulbecco’s modified eagle medium supplemented with 10% fetal calf serum, 2 mA glutamine, and 5 mA penicillin-streptomycin in a 5% CO2 incubator at 37 °C. The absence of mycoplasma contamination was monitored monthly using a PCR method (invitrogen) and only mycoplasma negative cells were used for culture. Following 4 to 5 days in culture, tachyzoites were counted in a Neubauer hemocytometer, washed three times with PBS and the freshly lysed parasites were scraped off the culture flasks. T. gondii was isolated from human foreskin fibroblasts by passage through with a 25-gauge needle followed by filtration through a Nucleopore filter (3 μm pore size) as previously published (24). Parasites were then pelleted by centrifugation. Purification of Membrane Proteins from T. gondii—About 4.0 × 109 purified T. gondii RH strain tachyzoites from culture were resuspended in 20 ml of SMD buffer (250 mM Sucrose, 10 mM MOPS-KOH, pH 7.2, 2 mM dithiothreitol, 1× protease inhibitor mixture) and disrupted by a French press at a pressure of 1000 PSI, medium setting. The lysate was centrifuged at 756 × g at 4 °C for 10 min to pellet unbroken cells. Intact parasites and large debris were resuspended in 10 ml of SMD buffer and disrupted once more by French press at a pressure of 1000 PSI, medium setting. The pooled supernatant was centrifuged at 25,000 × g at 4 °C for 20 min. The supernatant was saved for analysis as the cytosolic fraction. The pellet was resuspended in 10 ml of 30% Percoll in SMD buffer. Following centrifugation at 75,000 × g in an ultracentrifuge (Rotor TLA 100.3; 30,000 rpm) at 4 °C for 25 min, the top band was collected from the self-generated fraction. The band was diluted in SMD buffer and spun at 100,000 × g at 4 °C for 90 min (Rotor TLA 100.3; 40,000 rpm). A band collected between the buffer and resultant Percoll cushion contained the T. gondii ghost fraction consisting of membranes and cytoskeleton (24). To isolate the membrane fraction, T. gondii ghosts were resuspended in an equal volume of 2% thioglycolopyranoside in 40 mM Tris, pH 7.6, by pipetting the mixture up and down 10 times (suspension kept on ice) followed by a brief vortex. Following centrifugation at 20,000 × g in an Eppendorf centrifuge at 4 °C for 20 min, the supernatant was saved as the membrane fraction (extraction 1). This extraction was repeated twice with 300 μl of 2% thioglycolopyranoside (extraction 2 and 3, respectively). These fractions were pooled and frozen in liquid nitrogen and stored at −80 °C until protein analysis. In the parallel experiment, prior to being resuspended in thioglycolopyranoside solution, T. gondii ghosts were first washed sequentially with 1 M KCl and 0.1 M Na2CO3 twice.

One-Dimensional SDS-PAGE and In-gel Digestion—For SDS-PAGE, ~60 μg of proteins was mixed with 50 μl of 2× SDS loading buffer and heated at 85 °C for 10 min prior to loading on a precast gel (Invitrogen, Carlsbad, CA, USA). Proteins were afterward stained with Coomassie Brilliant Blue R-250. Usually, 40–45 gel bands were cut across the whole sample lane. Each band was excised further into small pieces (1 cm × 2 mm) and placed into a 0.65 ml micro-centrifuge siliconized tube (PGC Scientific) and washed with 250 μl of Millipore water twice for 5 min. In-gel trypsin digestion was performed. Gel pieces were washed three times each with an appropriate volume of 25 mM NH4HCO3 in 50% acetonitrile. The gel pieces were dried completely in a SpeedVac. The reduction reaction was allowed to proceed at 56 °C for 1 h in an appropriate volume of 10 mM dithiothreitol in 25 mM NH4HCO3 and the alkylation was allowed to proceed in the dark for 45 min at room temperature in an appropriate volume of 55 mM iodoacetamide in 25 mM NH4HCO3. The gel pieces were washed with 100 μl of 25 mM NH4HCO3 for 10 min and dehydrated with 100 μl of 25 mM NH4HCO3 in 50% acetonitrile for 5 min. The wash and dehydration step was repeated once. Following drying in a SpeedVac, the gel pieces were mixed with an appropriate volume of 12.5 ng/μl of trypsin and incubated on ice for 40 min and 25 mM NH4HCO3 was added as needed to cover the gel pieces. Digestion was then carried out at 37 °C overnight. To extract the tryptic peptides from the gel pieces, an appropriate volume of 60% acetonitrile, 0.2% TFA, was added. Following 20 min of vortex and 5 min of sonication, the supernatant was taken and saved. The extraction was repeated once and all the supernatants were combined (in-gel digestion protocol obtained from http://ms-facility.ucsf.edu/ingel.html). Following the evaporation of acetonitrile in a SpeedVac, the sample was desalted with a C18 ZipTip (Millipore), and half of the eluate was analyzed with nanoLC-MS/MS.

Preparation of Integral Plasma Membrane Proteins Using Biotin Labeling—T. gondii RH strain tachyzoites (~4.0 × 109 cells) collected from cell culture were washed five times in 20 ml PBS (prewarmed at 37 °C) and resuspended in 20 ml PBS. About 25 mg EZ-Link™ sulfo-NHS-SS-biotin was added to the mixture and incubated for 30 min at room temperature. The mixture was centrifuged for 20 min at 4150 rpm and 15 °C. The supernatant was removed, and the tachyzoites resuspended in 20 ml PBS containing 1 mg/ml lysine and incubated for 15 min at room temperature. The mixture was centrifuged for 20 min at 4150 rpm at 15 °C and, the supernatant removed. T. gondii tachyzoites were washed twice with 20 ml of PBS. T. gondii ghost fraction was prepared as described (see above “Purification of Membrane Proteins from Toxoplasma gondii”). The ghost fraction was resuspended in 500 μl of 2% thioglycolopyranoside in 40 mM Tris, pH 7.6, 1× protease inhibitor, containing around 9–10 mg prewashed streptavidin magnetic beads. Following 2 h incubation at 4 °C, magnetic beads were collected using a magnetic rack (Bio-Labs). The beads were sequentially washed three times with 1 ml of ice-cold 1 M KCl, three times with 1.2 ml of 0.1 M Na2CO3, then once with 1.2 ml of 2% thioglycolopyranoside in 40 mM Tris, pH 7.6, 1× protease inhibitor. Finally, the beads were suspended in 85 μl of 2× SDS-PAGE sample buffer containing 100 μl of dithiothreitol, 30 mM biotin, 6 μl urea, and 2 μl thiourea. Incubation was carried out for 10 min at room temperature followed by 10 min at 95 °C. The supernatant was collected following centrifugation for 3 min at 11,000 × g. Aliquots of the supernatant were separated with SDS-PAGE. 41 gel bands were cut across the whole sample lane. In-gel digestion was performed for each band as described above.

Three-Layer Sandwich Gel Electrophoresis (TLSGE), In-gel Digestion and Strong Cation Exchange Chromatography of Tryptic Peptides—The preparation of the three-layer sandwich gel (TLSG) was performed as described previously (23). The three-layer “sandwich” gel electrophoresis (TLSGE) was carried out with an electro-eluter device (Bio-Rad, model 422). The bottom layer or sealing layer was first formed by casting 1 ml of 40% polyacrylamide gel in a glass tube (1 cm I.D. × 6 cm, Bio-Rad) (Fig. 1). Immediately on top of the sealing layer is cast the concentrating layer with ~200–250 μl of 10% acrylamide gel. About 500 μl of 2× SDS-PAGE sample loading buffer was

Molecular & Cellular Proteomics 10.1
10.1074/mcp.M110.000745–3
mixed with ~500–1000 μl of membrane protein fraction. The resulting solution was then mixed with 0.5–1.0 ml of preheated 1% agarose in the elution buffer containing 25 mM Tris base, 192 mM glycine, and 0.2% SDS, and immediately added on top of the concentrating layer forming the upper layer of the TLSG. After the glass tube cooled down, the tube was placed into the electro-eluter with the upper and lower elution buffers, a 150 constant voltage was applied and the electrophoresis was allowed to run for 4–5 h. When a potential is applied between the upper and lower layers of TLSG, proteins migrate from the agarose gel into the middle layer and are concentrated in this layer; the bottom layer prevents protein leakage from the middle layer because of the high density acrylamide gel (40%) (Fig. 1). Consequently, TLSGE concentrates almost all proteins from a solution into the middle layer gel slice through electrophoresis whereas salts and detergents are removed from this layer gel. Following electrophoresis, the concentrating layer of gel was taken out, excised into small pieces (~1 × 2 mm), and placed into a 1.5 ml microcentrifuge siliconized tube (PGC Scientific). The gel pieces were washed with 600 μl of Millipore water for 10 min, and this wash step was repeated at least four or five times. In-gel tryptic digestion was performed according to the procedure described above with some modifications: prior to the addition of trypsin, 200–300 μl of 0.1% RapiGest™SF was added to the dried gel pieces before incubating in a 37 °C water bath for 30 min, followed by drying the gel pieces completely in a SpeedVac. The extracted tryptic peptide solution was concentrated in a SpeedVac to ~300–400 μl and mixed with 600 μl of buffer A used during strong cation exchange chromatography. The resulting sample was vortexed briefly and centrifuged at 13,000 rpm for 10 min. The supernatant was taken, mixed with 9 ml of strong cation exchange chromatography buffer A and loaded on a 10 ml sample loop of the AKTA purifier (Amersham Biosciences). The strong cation exchange chromatography separation column was a Polysulfoethyl A column (2.1 mm I.D. × 100 mm, 5 μm, 300Å, PolyLC Inc.). Buffer A was 10 mM KH₂PO₄, 25% acetonitrile, pH 2.8; buffer B was 10 mM KH₂PO₄, 700 mM KCl, 25% acetonitrile, pH 2.8. The gradient was from 0% to 50% buffer B in 120 min and from 50% to 100% B in 20 min with a flow rate of 0.1 ml/min. One fraction was collected every minute. Each fraction was concentrated to 20 μl in a SpeedVac, desalted with a C18 ZipTip, and half of the eluate was analyzed by nanoLC-MS/MS.

**NanoLC-MS/MS Analysis**—Nanoelectrospray LC-MS/MS analysis was performed on a linear ion trap mass spectrometer (LTQ, Thermo, San Jose, CA) interfaced with a TriVersa NanoMate nanoelectrospray ion source (Advion BioSciences, Ithaca, NY). An Ultimate Plus nano-HPLC system with a FAMOS autosampler (Dionex Corporation, Sunnyvale, CA), was coupled with the TriVersa NanoMate. Peptides, 25 μl, were loaded on a C18 μ-Precolumn™ Cartridge (5 μm, 300 μm i.d. × 5 mm) by the autosampler with a 25 μl sample loop at a flow rate of 15 μl/min. Mobile phase A was 2% acetonitrile and 0.1% formic acid in water and mobile phase B was 80% acetonitrile and 0.1% formic acid in water. The high-pressure liquid chromatography flow rate used was 250 nL/min. Following injection of 25 μl sample, and washing for 20 min with 100% mobile phase A, the precolumn was switched in line with the analytical column, C18 PepMap100, 3 μm, 100 Å, 75 μm i.D. × 150 mm (Dionex Corporation, Sunnyvale, CA). Then mobile phase B was increased from 2% to 55% over 70 min, held for 5 min, increased to 95% over 20 min and held at 95% B for 5 min. The four most intense ions having a charge state between +2 to +4, determined from an initial survey scan from 300–1800 m/z, were selected for zoom scan and MS/MS. In some nanoLC-MS/MS experiments, the 10 most intense ions from each initial MS survey were selected for MS/MS only. MS/MS was performed using an isolation width of 2 m/z; normalized collision energy of 35% and a minimum signal intensity of 1000 counts. Dynamic exclusion was enabled, such that once a certain ion is selected twice over 2 to 4 min, this ion is excluded from being selected again for MS/MS during the next 120 s.

**Data and Bioinformatics Analysis**

Protein Database Search of MS/MS Spectra—DTA files were created from the raw LC-MS/MS data, and searched with Mascot Daemon software (Version 2.2.2) against a database comprised of the human proteome (~40,000 sequences) and predicted T. gondii protein sequence sets (38,184 sequences) organized in our Experimental
Proteomics DataBase for T. gondii and Cryptosporidium parvum EPIC-DB (24). EPIC-DB is a comprehensive collection of experimental and computational data that contains all available protein sequences for T. gondii and C. parvum, which represents the theoretical proteome of the respective organisms. The available T. gondii sequences were compiled from five different gene prediction datasets: TigrScan, TwinScan, GlimmerHMM, ToxoDB Release 4.30, ME49 (Release 6.0, http://ToxoDB.org), and experimentally characterized sequences obtained from the NCBI protein database. The combination of alternative gene predictions was necessary because there was a surprisingly small overlap between any two alternative sets (24). In addition, MS/MS analysis showed that all gene prediction sets contain experimentally validated unique predictions. A false discovery rate for peptide identification was assessed by decoy database searching. The following parameters were used for all searches: trypsin, two missed cleavages; variable modifications of carbamidomethylation (Cys), deamidation (Asn and Gln) and oxidation (Met); monoisotopic masses; peptide mass tolerance of 3.5 Da; and product ion mass tolerance of 0.6 Da. Proteins were considered identified having at least two bold red peptides (BR, the most logical assignment of a peptide to a proteins and prevents duplicate homologous proteins to be reported) and having a peptide ion score cut-off of 49 or greater (corresponding to \( p < 0.05 \)). For those proteins identified by only one BR peptide and having a peptide ion score cut-off of 49 or greater (corresponding to \( p < 0.05 \)), the mass spectrometry raw data was checked manually to meet the following criteria: the difference between the measured and theoretical peptide parent mass divided by theoretical one was within 100 ppm; MS/MS product ions within 0.5 Da of the predicted b and y ions; and 80% of the high intensity product ions matched to either b or y ions; and both b and y ions matched. A BLAST search against other genomes was also performed with the peptide hits to rule out possible protein hits from other contaminating species. A list of all significant T. gondii protein hits are provided in Supplemental Tables 1, 2, and 3. All references to subsequent protein hits refer to those of T. gondii only, unless specified as human.

Clustering Protein Sequences, Transmembrane Predictions—Redundant T. gondii protein sequences were clustered using a greedy clustering algorithm (CD-HIT) using a 90% sequence identity threshold (24, 26–28). Transmembrane segments were predicted using Phobius, a combined trans-membrane protein topology and signal peptide predictor (29).

GPI Prediction and Plotting Transmembrane Domains—GPI-anchor proteins were predicted using PredGPI (http://www.biomedcentral.com/1471-2105/9/392) (30). Two-dimensional topology images of transmembrane proteins were created using TOPO2 software (http://www.sacs.ucsf.edu/TOPO2/).

GO Annotations and Functional Annotation—Homologous sequences in apicomplexan genomes, human genome and the NCBI NR database were identified by a BLAST search (without low complexity filtering and e-value of 0.001). The homologous sequences found by BLAST in the three dataset were mapped to the GO database by their GI number. The following protocol was used: GI identifiers of the homologous sequences identified by the BLAST search were used to retrieve UniProt IDs making use of a mapping file from iProClass database (http://pir.georgetown.edu/pirww/dbinfo/iproclass.shtml). The iProClass is an integrated database for protein functional analysis, provides comprehensive descriptions of all proteins, with links to over 50 databases of proteins family, function, pathway, interaction, modification, structure, genome, ontology, literature, and taxonomy.

RESULTS

Membrane proteins include integral and peripheral membrane proteins. The integral membrane proteins are a constitutive part of the membrane and can be classified into two categories: transmembrane proteins that span the entire membrane and integral monopatic proteins, which attach to the membrane from one side only. Peripheral membrane proteins are temporarily bound either to the lipid bilayer or to integral proteins by a combination of noncovalent interactions, such as hydrophobic and electrostatic interactions. T. gondii genome data is available at http://ToxoDB.org (http://EuPathDB.org). In addition, a predicted proteome based on several gene predictions for T. gondii is available at EPIC-DB (http://toro.aecom.yu.edu/biodefense/). Theoretically, membrane proteins of T. gondii can be identified in these sequences using membrane segment prediction algorithms. However, the analysis is complicated by the fact that different protein predictions in T. gondii have about a 30% error rate in identifying coding regions (24, 31–33) and the predicted membrane proteins may be expressed differently in different developmental stages. To experimentally validate gene predictions in T. gondii and comprehensively analyze the membrane proteome of T. gondii, we implemented three proteomics strategies as shown in Fig. 1. Strategy 1 and 2 target the whole membrane proteome, whereas strategy 3 targets a particular subset of the membrane proteome: the integral plasma membrane proteome.

One Dimensional SDS-PAGE and NanoLC-MS/MS Analysis for T. gondii Membrane Protein Extract—One dimensional SDS-PAGE protein separation and in-gel tryptic digestion followed by reverse phase liquid chromatography-tandem mass spectrometry (LC-MS/MS) analysis (Strategy 1 in Fig. 1) is the most common and straightforward approach for the analysis of a complicated protein sample. The main advantage of one-dimensional SDS-PAGE separation over two-dimensional gel electrophoresis for membrane proteins is that many membrane proteins precipitate under isoelectric focusing conditions, but in general, most membrane proteins are soluble under SDS-PAGE conditions (34, 35). The membrane proteins were resolved on SDS-PAGE and 40 to 45 gel bands were cut from each gel lane. Tryptic digestion was performed for each gel band following reduction and alkylation and the tryptic peptide mixture extracted from each band was subject to analysis by nanoLC-MS/MS. In the initial experiment, a wash step with a high salt and high pH buffer was not included in the membrane protein preparation procedure. A total of 1606 proteins were identified (see the green circle in Fig. 2A, 1169 + 437 = 1606), of which 734 were identified as membrane proteins with one or more predicted transmembrane segments (see the green circle in Fig. 2B, 521 + 213 = 734). These 734 proteins can be further grouped into 266 protein clusters (see the green circle in Fig. 2B, 178 + 88 = 266; note that the number within a bracket presents the number of protein clusters). The clustering approach groups together amino acid sequences even with very different lengths if the overlapping parts share more than 90% sequence identity (our clustering approach is described in detail in our recent paper,
Membrane Proteome of *T. gondii*

![Venn diagrams showing the number of proteins identified by Mascot having peptide ion scores of 49 or greater from the one-dimensional gel LC-MS/MS experiment (Strategy 1).](image)

(A) Total number of proteins identified with or without Na₂CO₃ and KCl washes. (B) Number of membrane proteins identified in A having one or more predicted transmembrane segments with or without Na₂CO₃ and KCl washes. The numbers in the square brackets refer to the number of protein clusters. The expression of “w/” means “with” and “w/o” means “without.”

Proteins are listed in Supplementary Tables.

(24)). The protein clustering strategy was designed to group together alternative gene predictions of the same genomic regions, some of which may present possible splice variants and protein post-translational modifications, such as protease cleavages and covalent modifications.

High-salt buffer and alkaline buffer can disrupt noncovalent interactions between proteins that can then be used to wash away a large proportion of plasma membrane-associated cytosolic proteins (21, 36). Therefore, in the second experiment, we applied high salt and high pH buffer (1 M KCl and 0.1 M Na₂CO₃) wash steps in the preparation of membrane proteins. Following extraction of the membrane fraction, the membrane protein extract was washed twice sequentially with 1 M KCl and 0.1 M Na₂CO₃. In this experiment, 2938 proteins were identified (see the blue circle in Fig. 2A), of which 981 proteins (corresponding to 383 protein clusters) were predicted to have at least one transmembrane segment, indicating that this wash step had led to a remarkable increase in the number of membrane proteins identified (see the blue circle in Fig. 2B). There are 521 membrane proteins (in 178 protein clusters) that are common to both of the above experiments (Fig. 2B).

Relatively high abundance plasma membrane-associated cytosolic proteins, such as 2C-methyl-D-erythritol 2, 4-cyclohexane-1, 3-diphosphate synthase, cytochrome C oxidase, glycosyl transferase, were removed from the membrane protein fraction by washing with high salt and high pH buffer. The removal of abundant cytosolic proteins resulted in a decrease of protein dynamic range in the sample, permitting many more proteins to be identified by mass spectrometry, including H⁺-translocating inorganic pyrophosphatase TVP1, P-type ATPase, seven transmembrane receptor domain-containing proteins and putative sodium hydrogen exchangers. H⁺-translocating pyrophosphatase (H⁺-Pases) has 14 transmembrane segments and is the primary H⁺ pump that uses inorganic pyrophosphate (PPi) instead of ATP as an energy source. P-type ATPase and sodium hydrogen exchanger are predicted to have 11 and 13 transmembrane segments, respectively. Fig. 3A demonstrates the peptide coverage of a membrane protein (TgGlmHMM_2178), Ca²⁺-dependent ATPase with 11 predicted transmembrane segments. In total, 14 and 6 peptides (in red and blue colors) were identified with a Mascot peptide ion score higher than 60 or within a range of 49–60, respectively. Fig. 3B shows the MS/MS spectrum of a peptide (LGGLDSLADEK) of protein TgGlmHMM_2178 (the Mascot peptide ion score is 63, p < 0.05). As shown in Fig. 3A, this peptide extends into the membrane. The zoom scan (inserted spectrum in Fig. 3B) clearly indicates that the peptide ion is doubly charged; and 83% of the b and y ions in the MS/MS spectrum match to the theoretical b and y ions, resulting in the unambiguous identification of this peptide.

**TLSGE MudPIT Analysis for *T. gondii* Membrane Proteins**—In Strategy 2 (Fig. 1), the three-layer sandwich gel electrophoresis (TLSGE) method and multidimensional chromatograph protein identification technology (MudPIT) were integrated for TLSGE MudPIT. Traditionally, MudPIT employs in-solution digestion of a complex protein mixture prior to two-dimensional liquid chromatography peptide separation. In TLSGE MudPIT a single in-gel digestion is needed before two-dimensional liquid chromatography. TLSGE MudPIT shares most of the advantages of conventional MudPIT alone, but has unique features well suited for membrane protein identification. Using this technique on the membrane protein preparation for *T. gondii*, using the same amount of organism as used for one-dimensional gel LC-MS/MS, without the high salt and high pH washes, 2938 proteins corresponding to 923 protein clusters were identified (see the green circle in Fig. 4A). Of these proteins, 1461 proteins were identified as membrane proteins with one or more predicted transmembrane segments, which grouped into 525 membrane protein clusters (green circle in Fig. 4B). When high salt and high pH wash steps were included in the sample preparation, 3116 proteins were identified (blue circle in Fig. 4A), of which 1428 were identified as membrane proteins (blue circle in Fig. 4B). These membrane proteins classified into 546 protein clusters (blue circle in Fig. 4B). There were 1825 proteins common to the two experiments (region of overlap in Fig. 4A), of which 893 proteins in 312 protein clusters are predicted with one or more transmembrane segments (region of overlap in Fig. 4B). In total, the two experiments led to identification of 4229 proteins from 1403 proteins clusters, of which 1996 proteins from 759 protein clusters were found to be transmembrane proteins. In summary, the TLSGE MudPIT method identified...
proteins identified with or without Na$_2$CO$_3$ and KCl washes. The numbers in the square brackets refer to the number of protein clusters. The expression of “w/” means “with” and “w/o” means “without”. Proteins are listed in Supplementary Tables.

376 more transmembrane protein clusters than the one-dimensional gel LC-MS/MS method.

Plasma Integral Membrane Proteome Analysis by Biotin-Directed Affinity Purification (BDAP) and NanoLC-MS/MS—To specifically target integral plasma membrane proteins located on the T. gondii cell surface, we implemented cell surface biotinylation and affinity purification (20, 21, 37, 38), using the same amount of organism as used for one-dimensional gel LC-MS/MS (see strategy 3 in Fig. 1). In total, 1029 proteins in 304 protein clusters were identified (the green circle in Fig. 5 A), of which 329 proteins have predicted transmembrane segments and grouped into 118 membrane protein clusters (the green circle in Fig. 5 B). Besides many hypothetical membrane proteins, these include many known or putative integral plasma membrane proteins in T. gondii with one or multiple predicted transmembrane segments, such as rhoptry proteins 4, 5, 13, 14, and 18; rhoptry neck proteins 1, 2, and 3; dense granule protein GRA3, 6, and 7; ADP/ATP carrier; calcium ATPase SERCA-like protein; Rab11; Ca$^{2+}$ ATPase; facilitative glucose transporter GT1; dihydrolipoamide acetyltransferase; ATP-binding cassette protein subfamily B member 3; H$^+$ translocating inorganic pyrophosphatase TVP1 protein; oxoglutarate/malate translocator protein; vacuolar H$^+$ translocating ATPase subunit A; vacuolar ATP synthase catalytic subunit A; oxalate formate antiporter; and NMDA receptor glutamate-binding chain.

Comparison of the Three Proteomics Strategies—Fig. 5 summarizes all the identified proteins and transmembrane proteins by the three approaches. In total, 2241 T. gondii membrane proteins from 841 protein clusters were identified (Fig. 5B). Approximately 66% were identified by at least two peptide matches, each with a Mascot score of 49 or higher. Meanwhile 15.4% were identified by a single peptide match with a score larger than 60, and 18.8% by a single peptide match with a score between 49 and 60. Fig. 6A shows the distribution of peptide Mascot scores for all the identified T. gondii transmembrane proteins by all the three methods. The majority (96%) of these peptides had a Mascot score greater than 60 and were repeatedly found in three of the different proteomics experiments. The distribution of identified membrane proteins across the number of predicted transmembrane segments is shown in Fig. 6B. Approximately 62% of all identified transmembrane proteins have two or more predicted transmembrane segments and 20.8% of all identified transmembrane proteins have seven or more transmembrane segments. One protein (52.m01648), which has 22 predicted transmembrane segments, was identified as the adenylate and guanylate cyclase catalytic domain-containing protein (Supplemental Table 1).

We investigated the overlap of the identified proteins as well as membrane proteins among the three approaches (Fig. 5). The overlap between one-dimensional gel LC-MS/MS and TLSGE MudPIT reveals that 959 membrane proteins from 358 protein clusters were common to both methods (Fig. 5B). These common membrane proteins represent 48% and 80.3% of the membrane proteins identified by TLSGE MudPIT or one-dimensional gel LC-MS/MS, respectively. The 285 identified membrane proteins from 98 protein clusters were common to all three methods, which were 23.9%, 14.3%, and 86.6% of the membrane proteins identified by one-dimensional gel LC-MS/MS, TLSGE MudPIT, and BDAP LC-MS/MS approaches, respectively. The two former methods targeted the whole membrane proteome, whereas BDAP LC-MS/MS targets a particular fraction of the membrane proteome: integral plasma membrane proteins. Only 10 membrane proteins from four protein clusters were exclusively identified by BDAP LC-MS/MS. The combination of one-dimensional gel LC-MS/MS and TLSGE MudPIT analysis was able to identify about 97% of integral plasma membrane proteins. However, BDAP LC-

Fig. 3. A, Amino acid sequence diagram of the transmembrane protein Ca$^{2+}$ dependent ATPase (TgGlmHMM_2178). Peptides in red were identified by LC-MS/MS with a Mascot score of 60 or more. Peptides in blue were identified by LC-MS/MS with a Mascot score in a range of 49–60. The transmembrane segments were predicted with Phobius program. B, MS/MS spectrum of the tryptic peptide LLGCLDSLAADEK, transmembrane region 2 in red, identified from all the three proteomic strategies, the Mascot peptide ion score is 63, $p < 0.05$. This peptide ion at m/z 702.84 (monoisotopic) is doubly charged (see inserted zoom scan). The observed peptide parent mass is 1403.66 Da and the theoretical peptide monoisotopic mass is 1403.70. The major ions in this MS/MS spectrum comprise 83% of the predicted peptide’s b and y ions. Note that the C residue is carbamidomethyl cysteine.
MS/MS analysis is important, as it allows screening specifically for integral plasma membrane proteins, many of which are potential drug targets. Supplemental Table 2A to G lists all the identified proteins in regions “a” to “g” in Fig. 5A. Supplemental Table 3A to G lists all the identified transmembrane proteins in regions “a” to “g” in Fig. 5B.

Analysis of the Proteins Identified by the Combined Strategies—Supplemental Table 1 contains all of the identified T. gondii membrane proteins with protein accession key, number of predicted transmembrane segments, number of peptides hits assigned to the protein, the experiments in which the protein was found, as well as name or function of the protein. By single clicking the web linked sequence keys in Supplemental Table 1, the detailed information of each identified transmembrane protein cluster from the EPIC-DB database (http://toro.aecom.yu.edu/cgi-bin/biodefense/main.cgi) can be retrieved, including MS experimental data and results from the bioinformatics analysis.

All proteins identified with one or more predicted transmembrane segments were subjected to Gene Ontology (GO)
As demonstrated in Fig. 7A, 23% of the proteins are classified as membrane proteins, 21% are integral to membrane, 3% are plasma membrane proteins, as well as 3% as endoplasmic reticulum membrane proteins. Biological process GO analysis indicates that a large number of the identified membrane proteins (15%) are involved in transport, including 3% and 2% in cation and ion transport, respectively (Fig. 7B). Furthermore, many of the identified membrane proteins (4%) are involved in ATP biosynthetic processes and 7% in metabolic processes. Approximately 4% of the identified membrane proteins function in proteolytic and protein phosphorylation processing (Fig. 7B). In molecular function analysis, many identified membrane proteins are characterized as ATP, nucleotide, protein or metal ion binding molecules (Fig. 7C). Many are predicted to have a membrane-related enzymatic activity, such as ATPase activity coupled to transmembrane movement of ions or other substances or phosphorylative mechanism, as well as hydrolase, transferase and oxidoreductase activities. GPI modification site prediction analysis reveals 29 proteins highly likely to be glycosylphosphatidylinositol (GPI) anchored proteins, such as surface antigen 43, a SRS domain containing protein and GPI-anchored surface BSR4-related antigen. In addition, many proteins are likely to be GPI anchored, for instance, SAG2-related antigen SAG2E, signal recognition particle 54-Kda protein, sporozoite-specific SAG protein and so on. Interestingly, many of these predicted GPI anchored proteins are hypothetical proteins.

It was striking that more than 42% of membrane proteins identified in the current study are hypothetical proteins. Furthermore, half of these hypothetical proteins appear to be unique to T. gondii. For example, a hypothetical membrane protein (TgGlmHMM_0264) having eight predicted transmembrane segments, was identified by all three proteomics approaches with eight peptide matches, four with Mascot peptide ion scores greater than 60 and four with Mascot ion scores between 49 and 60. The negative outcome of BLAST searches and GO annotations suggest that this protein is of unknown function and unique to T. gondii.

Because T. gondii parasites were cultured in human foreskin fibroblasts, there is a possibility that the membrane protein preparation of T. gondii may contain human protein contaminates. To remove the human protein hits, we performed Mascot searches for all MS/MS data against a protein sequence database combining the human proteome and the predicted T. gondii protein sequence data sets. In total, 1161 proteins in 662 clusters were identified as human proteins. 1045 out of these 1161 human proteins do not have any

---

**Fig. 7.** Pie charts of functional distribution per cellular component (A), biological processes (B) and (C) molecular function from the membrane proteins identified by all three proteomic strategies (Fig. 1). Numbers represent the percentage from the total of identified membrane proteins. Proteins are listed in Supplementary Tables.
peptide hits that match *T. gondii* proteins, which suggests that they are unique to human. The rest of the 116 human proteins have one or more peptide hits in common with *T. gondii* proteins, however, only two (NCBI gi number 115511024 and 153945715) have predicted transmembrane segments.

All the identified *T. gondii* proteins with at least one predicted transmembrane segment are listed in Supplemental Table 1. They represent ~25.2% of the predicted membrane proteome from the *T. gondii* genome (8897 predicted membrane proteins that group into 3858 clusters). Because we investigated only the tachyzoite life stage in this study the coverage of the *T. gondii* membrane proteome by these experiments is quite high. All the data for *T. gondii* membrane proteome analysis as well as bioinformatics analysis have been deposited on the Experimental Proteomics DataBase for *Toxoplasma gondii* in the Einstein Biodefense Proteomics Research Center (http://toro.aecom.yu.edu/cgi-bin/biodefense/main.cgi) and the data provided to ToxodB (http://ToxodB.org), which is part of EuPathdB.

**DISCUSSION**

Membrane proteins are critical players in various physiological processes, such as regulation of intracellular pH; the control of cell growth and proliferation; transepithelial absorption and secretion of sodium, hydrogen, bicarbonate, chloride, and organic anions; and the metabolic response to hormones such as insulin and glucocorticoids (40). Many diseases or disorders are related to the functions of membrane proteins. A large portion of current pharmaceutical drug targets are membrane proteins. Similar to other organisms, about 30% of proteins encoded in the genome of *T. gondii* are predicted to be membrane proteins. However, a great number of them are unknown or have not been characterized although the genome sequencing of this parasite has been completed for several years. Expression profiling of *T. gondii* membrane proteins at different developmental stages and accurate membrane gene annotation is important for understanding the invasive mechanisms and the interaction between *T. gondii* and their hosts. We report here that by utilizing three proteomic strategies, one-dimensional gel LC-MS/MS, TLSGE MudPIT, and BDAP LC-MSMS, we were able to identify and validate more than 2200 membrane proteins of *T. gondii* tachyzoites, classified into 841 protein clusters with a requirement of 90% sequence identity threshold in each cluster. Because the tandem mass spectra data were searched against the hypothetical *T. gondii* proteome generated through a combination of computationally predicted proteins from TigrScan, TWINSCAN, GlimmerHMM, Release 4.3 (ToxodB) and ME49 (Release 6.0 ToxodB), and the available experimental *T. gondii* sequences from the NCBI nonredundant protein database, a protein clustering strategy is essential to remove redundancy. In addition, the protein sequence similarity clustering strategy also provides insights into potential splicing products within the genome. Each nonredundant cluster may represent a single sequence or multiple sequences. It was found that the nonredundant protein clusters generated for *T. gondii* have an average size of 2.02 sequences (24). Multiple sequences in a cluster may be truly redundant sequences or are the shorter forms generated from the longer form because of naturally occurring enzymatic processes.

Currently, MS/MS database searching is the primary approach for protein identification in large scale proteomic analysis. False positive matches are known to occur (41). Randomized decoy database search strategies are frequently used to estimate the False Discovery Rate (FDR) of peptide identification (42, 43). We implemented the Mascot search tool with an automatic decoy search feature to perform MS/MS database searches. In this study, the Mascot peptide ion score threshold for accepting a MS/MS match to a certain peptide was set at 49 (individual peptide ion scores > or = 49 indicate identity or extensive homology, p < 0.05). This threshold led the overall FDR of peptide identification to be 1.1%. For the proteins identified with a single peptide match, additional manual checks were performed, to see if the adjusted observed peptide mass after removing the systematic error of mass spectrometry analysis is within the range of the theoretical mass plus or minus 100 ppm (this mass tolerance falls within the specific internal mass error of LTQ ion trap mass spectrometer), if the product ion mass tolerance is within 0.5 Da, if the MS/MS product ions have both b and y ions, and if 80% of the most intense product ions match to either b or y ions. A recently published work investigated the effect of the “two-peptide” rule on protein identification (44). The authors demonstrated that the “two-peptide” rule decreases the number of protein identifications in the target database more significantly than in the decoy database, resulting in increased FDR, when compared with the case when proteins identified by only a single peptide match are not discarded (44). A large fraction of proteins may be missed by the “two-peptide” rule in proteomic experiments (44). In our case, about 34% of the 841 membrane protein clusters were identified by only a single peptide match. Most of these single peptide matches were repeatedly found in the experiments carried out independently by the three different membrane proteomic approaches, further supporting that these single peptide matched proteins were present in the organism. Therefore, we included those membrane proteins identified with one peptide match (Supplemental Table 1). The single-peptide-hit identified membrane proteins are most likely to be those proteins, which are very low in abundance. In addition, integral membrane proteins have a significant percentage of sequence embedded in the membrane and this may also account for the single peptide matches observed. The commonly utilized chromatography conditions for LC-MS/MS for proteomic study is generally suited to the soluble peptide segments of integral membrane proteins and cannot ef-
Membrane Proteome of *T. gondii*

...efficiently access membrane-embedded peptides (45). Therefore, the identified sequence coverage of an integral membrane protein is often lower than that of a cytosolic protein. Because of these issues and the fact that different protein predictions in *T. gondii* have been shown to each have about a 30% error rate in identifying coding regions (24), it is important to have proteomics strategies in place that focus on recognizing these low abundant proteins and identify the membrane-embedded peptides.

To achieve a global analysis of the membrane proteome of even such a small organism, one cannot rely on one approach alone. Because the proteome is much more complicated and less defined than the genome, more comprehensive proteomic strategies are essential. We applied three approaches, orthogonal to each other, in order to decipher the membrane proteome of *T. gondii*. The BDAP LC-MS/MS method specifically targets integral plasma membrane proteins; whereas one-dimensional gel LC-MS/MS and TLSGE MudPIT methods targets the whole membrane proteome. The common membrane proteins identified by using all three approaches represent 13% of all the identified membrane proteins but 87% of the membrane proteins identified by BDAP LC-MS/MS method (Fig. 7). The membrane proteins identified either by one-dimensional gel LC-MS/MS or TLSGE MudPIT share about 43% of all the identified membrane proteins from *T. gondii* tachyzoites. Both approaches uniquely identified a large number of membrane proteins, and the unique membrane proteins identified by TLSGE MudPIT were about three times more than those identified by one-dimensional gel LC-MS/MS. These data demonstrate that none of these approaches can completely replace the other, which is especially true for the one-dimensional gel LC-MS/MS and TLSGE MudPIT approaches. This also suggests that complementary proteomic strategies must be employed in different developmental stages of an organism to achieve more complete proteomic coverage.

High salt and high pH buffer (sodium carbonate) can be used to strip membrane-associated proteins and enrich membrane proteins (46–52). *T. gondii* has unique structures associated with the cell membrane including the pellicle, apical rings, polar rings, conoid, rhoptries, and micronemes, as compared with other cells. Its outer covering pellicle consists of plasmalemma and two closely applied membranes that form an inner membrane complex (IMC). Between the parasite plasma membrane and IMC, there are the myosins A motor complex and the micronemal protein-host receptor complex (via aldolase/F-actin) at the parasite pellicule (53). Furthermore, beneath the IMC, 22 subpellicular microtubules run longitudinally along the length of the cell. Such complicated and tight interactions among these structures often lead to contamination of cytoskeleton or cytoplasm proteins in membrane preparations. Washing with high salt and high pH buffer substantially removed the abundant membrane-associated proteins from our *T. gondii* membrane preparation thus, making it possible to additionally identify a large number of membrane proteins. Interestingly, high salt and high pH buffer wash appeared to impact the identification of membrane proteins in one-dimensional gel LC-MS/MS more so than in the TLSGE MudPIT strategy. When combining the proteins identified by experiments with or without high salt and high pH buffer washes, the latter approach identified 288 more membrane protein clusters than the former (Fig. 2B and Fig. 4B). However, with such washes, one-dimensional gel LC-MS/MS identified 44% more membrane protein clusters than without these washes, whereas, for TLSGE MudPIT experiment, the number of identified membrane protein clusters with these wash steps is comparable to those without these washes (Fig. 2 and Fig. 4). This phenomenon probably indicates that the dynamic range of one-dimensional gel LC-MS/MS analysis is smaller than that of TLSGE MudPIT analysis. Although a high salt and high-pH buffer wash is able to significantly decrease the abundance of membrane-associated proteins, this harsh and stringent wash may also decrease the abundance of some membrane proteins. This has been observed previously by other researchers (54) and is also strongly supported by our proteomic data from both one-dimensional gel LC-MS/MS and TLSGE MudPIT experiments. Peripheral membrane proteins dissociate from transmembrane proteins following a treatment with a solution having an elevated pH or high salt concentrations. Also, the high salt and high pH conditions might denature some membrane proteins, possibly releasing some of them into the wash solution.

To our knowledge, this work represents the most comprehensive analysis of the membrane proteome of a microorganism by integrating three proteomic approaches with different sample preparation methods and computational bioinformatics investigations. This integrative analysis identified thousands of membrane proteins in *T. gondii*. Many are known membrane proteins, including proteins from plasma membrane and rhoptry, microneme, dense granule, parasitophorous vacuole organelles. Many membrane proteins putatively function as membrane enzymes, as transporters or carriers, ion exchangers, surface antigens, receptors, structural membrane-anchoring proteins, as well as chaperone proteins (e.g. disulfide isomerases and regulators of vesicular traffic like ADP-ribosylation factors). The most intriguing finding is that a large number of transmembrane proteins identified in this work are hypothetical proteins. Furthermore, ~50% of these hypothetical proteins have no GO annotations, suggesting that these sequences are most likely unique to *T. gondii* parasites or the Apicomplexa. These unique and novel membrane proteins may represent potential drug targets and further systematic analysis of the membrane proteome of *T. gondii* will provide additional fundamental insights into the biology and molecular mechanisms underlying the pathophysiology of infection of this obligate intracellular parasite.
42. Elias, J. E., and Gygi, S. P. (2007) Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry. Nat. Methods. 4, 207–214
43. Käll, L., Storey, J. D., MacCoss, M. J., and Noble, W. S. (2008) Assigning significance to peptides identified by tandem mass spectrometry using decoy databases. J. Proteome Res. 7, 29–34
44. Gupta, N., and Pevzner, P. A. (2009) False discovery rates of protein identifications: a strike against the two-peptide rule. J. Proteome Res. 8, 4173–4181
45. Speers, A. E., Blackler, A. R., and Wu, C. C. (2007) Shotgun analysis of integral membrane proteins facilitated by elevated temperature. Anal. Chem. 79, 4613–4620
46. Pedersen, S. K., Harry, J. L., Sebastian, L., Baker, J., Traini, M. D., McCarthy, J. T., Manoharan, A., Wilkins, M. R., Gooley, A. A., Righetti, P. G., Packer, N. H., Williams, K. L., and Herbert, B. R. (2003) Unseen proteome: mining below the tip of the iceberg to find low abundance and membrane proteins. J. Proteome Res. 2, 303–311
47. Zhang, L. J., Wang, X. E., Peng, X., Wei, Y. J., Cao, R., Liu, Z., Xiong, J. X., Yin, X. F., Ping, C., and Liang, S. (2008) Proteomic analysis of low-abundant integral plasma membrane proteins based on gels. Cell Mol. Life Sci. 63, 1800–1804
48. Moebius, J., Zahedi, R. P., Lewandrowski, U., Berger, C., Walter, U., and Sickmann, A. (2005) The human platelet membrane proteome reveals several new potential membrane proteins. Mol. Cell. Proteomics. 4, 1754–1761
49. Schirmer, E. C., Florens, L., Guan, T., Yates, J. R., 3rd, and Gerace L. (2003) Nuclear membrane proteins with potential disease links found by subtractive proteomics. Science. 301, 1380–1382
50. Fujiki, Y., Hubbard, A. L., Fowler, S., and Lazarow, P. B. (1982) Isolation of intracellular membranes by means of sodium carbonate treatment: application to endoplasm is reticulum. J. Cell Biol. 93, 97–102
51. Mattern, J., Siejak, F., Hagens, K., Schmidt, F., Koehler, C., Treumann, A., Schaible, U. E., and Kaufmann, S. H. E. (2007) An improved strategy for selective and efficient enrichment of integral plasma membrane proteins of mycobacteria. Proteomics. 7, 1687–1701
52. Ruth, M. C., Old, W. M., Emrick, M. A., Meyer-Arendt, K., Aveline-Wolf, L. D., Pierce, K. G., Mendoza, A. M., Sevinsky, J. R., Hamady, M., Knight, R. D., Resing, K. A., and Ahn, N. G. (2006) Analysis of membrane proteins from human chronic myelogenous leukemia cells: comparison of extraction methods for multidimensional LC-MS/MS. J. Proteome Res. 5, 709–719
53. Soldati, D., and Meissner, M. (2004) Toxoplasma as a novel system for motility. Curr. Opin. Cell Biol. 16, 32–40
54. Mastrogiacomo, A., Kohan, S. A., Whitelegge, J. P., and Gundersen, C. B. (1998) Intrinsic membrane association of Drosophila cysteine string proteins. FEBS Lett. 436, 85–91

In order to cite this article properly, please include all of the following information: Che, F.-Y., Madrid-Aliste, C., Burd, B., Zhang, H., Nieves, E., Kim, K., Fiser, A., Angeletti, R. H., and Weiss, L. M. (2011) Comprehensive Proteomic Analysis of Membrane Proteins in Toxoplasma gondii. Mol. Cell. Proteomics 10(1):M110.000745. DOI: 10.1074/mcp.M110.000745.