Plastid-associated Porphobilinogen Synthase from Toxoplasma gondii

KINETIC AND STRUCTURAL PROPERTIES VALIDATE THERAPEUTIC POTENTIAL*\[\[2\]\

Received for publication, January 25, 2010, and in revised form, April 19, 2010. Published, JBC Papers in Press, May 4, 2010, DOI 10.1074/jbc.M110.107243

Dhanasekaran Shanmugam1, Bo Wu1, Ursula Ramirez5, Eileen K. Jaffe4, and David S. Roos2,6

From the 1Department of Biology, University of Pennsylvania, Philadelphia, Pennsylvania 19104 and the 6Institute for Cancer Research, Fox Chase Cancer Center, Philadelphia, Pennsylvania 19111

Apicomplexan parasites (including Plasmodium spp. and Toxoplasma gondii) employ a four-carbon pathway for de novo heme biosynthesis, which is an important regulatory step in the pathway and is achieved differently in photosynthetic versus non-photosynthetic eukaryotes. The former use the “C5” pathway to adapt glutamyl tRNA using the enzymes glutamate tRNA reductase and glutamate semialdehyde aminomutase, whereas the latter use aminolevulinic acid synthase to condense glycine and succinyl coenzyme A via the “C4” pathway (1, 2). ALA is then converted into tetrapyrroles (heme, chlorophyll, B12, etc.) using a multistep pathway, the first three enzymes of which are common to all organisms. In photosynthetic eukaryotes, the complete pathway for tetrapyrrole biosynthesis is located in the chloroplast, whereas in non-photosynthetic eukaryotes, the enzymes are typically distributed between the mitochondria and cytosol (1).

The first step in ALA utilization is catalyzed by the enzyme porphobilinogen synthase (PBGS; also known as δ-aminolevulinate dehydratase; EC 4.2.1.24), which converts two molecules of ALA into porphobilinogen (3). PBGS enzymes are highly conserved metalloproteins, with well-documented enzymatic and structural properties (3, 4). Unusual characteristics include extensive variation in metal ion usage at the active and allosteric sites (5) and an unusually dynamic quaternary structure (6, 7). Cytosolic (animal/fungal) PBGS enzymes are Zn2+-dependent, whereas the plastidic (plant/algal) PBGS respond to Mg2+ (5). Sequence conservation and activity assays suggest that the latter contain two distinct Mg2+-binding sites: one presumed to be analogous to the catalytic Zn2+ in animal/fungal PBGS and an allosteric site present at an interdimer interface (5). Both the human (HsPBGS) (7) and Pismus sativum (PsPBGS) (6) enzymes exhibit allosteric regulation based on quaternary structure; homooligomers are highly active, whereas homohexamers are inactive; dissociation and conformational change at the dimer level is required for hexamer-octamer interconversion. Misregulation of the human PBGS quaternary structure equilibrium is associated with porphyric disease (8). Proteins exhibiting such alternative quaternary structure assemblies are called “morphheins” (9).

The protozoan parasites Plasmodium falciparum and Toxoplasma gondii (both human pathogens in the eukaryotic phylum Apicomplexa) are capable of de novo heme biosynthesis, and inhibition of this pathway results in parasite death (10, 11). Phylogenetic studies reveal that apicomplexan heme biosyn-

---

1. This work was supported, in whole or in part, by National Institutes of Health Grants AI28724 (to D. S. R.), ES003654 (to E. K. J.), T32CA0935 (to the Institute for Cancer Research), AI07757 (to E. K. J.), and CA006927 (to Fox Chase Cancer Center).
2. The on-line version of this article (available at http://www.jbc.org) contains supplemental Tables S1 and S2, Text S1, and Figs. S1 and S2.
3. Both authors contributed equally to this work.
4. To whom correspondence should be addressed. E-mail: droos@sas.upenn.edu.
5. The abbreviations used are: ALA, aminolevulinic acid; eGFP, enhanced green fluorescent protein; PBGS, porphobilinogen synthase; HsPBGS, H. sapiens PBGS; PsPBGS, P. sativum PBGS; PaPBGS, P. aeruginosa PBGS; TgPBGS, T. gondii PBGS; RACE, rapid amplification of cDNA ends; SPP, stromal processing protease (pitrilysin); TgSPP, T. gondii SPP; Bistris, 1,3-bis[tris(hydroxymethyl)methylamino]propane; GFP, green fluorescent protein; β-ME, β-mercaptoethanol.
thesis enzymes exhibit a mosaic of plant and animal origins and are distributed between the parasite mitochondrion, plastid (apicoplast), and cytosol (12, 43). The plastid-associated enzymes are of plant origin, and plastid PBGS is sufficiently distinct from human PBGS to offer potential as a therapeutic target. *P. falciparum* PBGS forms active octamers and exhibits an unusual metal-ion-independent activity (13).

In this report, we describe the cloning of multiple cDNAs encoding *T. gondii* PBGS (*TgPBGS*), identification and expression of the physiologically relevant isoform, and kinetic characterization of this enzyme. Unlike *PsPBGS* and *Pseudomonas aeruginosa* PBGS (*PaoPBGS*), Mg\(^{2+}\) is not required for *TgPBGS* octamer stability, which appears to be conferred by an unusual C-terminal amino acid extension found in all apicomplexan PBGSs.

**MATERIALS AND METHODS**

*Molecular Procedures*—A Lambda Zap II *T. gondii* cDNA library (11) was used to PCR-amplify various fragments of the *TgPBGS* gene (see supplemental Table S1 for a list of primers), and full-length cDNA sequences were obtained by a combination of library screening and rapid amplification of cDNA ends (RACE). Quantitative "real-time" PCR of *TgPBGS* and *T. gondii* stromal processing peptidase (*TgSPP*) transcripts was carried out using poly(A)\(^+\) mRNA and the Power SYBR Green kit (Applied Biosystems); absolute levels were determined based on amplification of standards over a range of concentrations from 1 fg to 100 pg. Bradyzoite and sporozoite RNA was kindly provided by Drs. Florence Dzierszinski (McGill University) and Michael White (University of South Florida, respectively). A recombinant fragment of *TgPBGS* (amino acids 363–658) fused to a C-terminal His tag was expressed in *Escherichia coli* (stragatcic) and used to transform *E. coli* BLR(DE3) cells. Single colonies were inoculated into 1 liter of LB media containing 100 μg/ml ampicillin, 12.5 μg/ml tetracycline, and 0.4% glucose; incubated 12 h at 37 °C in a shaking incubator; pelleted and resuspended in 1 liter fresh LB without glucose; and equilibrated to 15 °C. Protein expression was induced by the addition of 100 μM isopropyl-1-thio-β-D-galacto- pryranoside and allowed to proceed for 15 h, followed by centrifugation at 4 °C to harvest. A 10-g pellet was resuspended in 20 ml of buffer containing 0.4 mg/ml lysozyme (Sigma), 50 mM KH\(_2\)PO\(_4\), pH 8.0, 170 mM KCl, 5 mM EDTA, 10 mM β-ME, and 0.1 mM phenylmethylsulfonyl fluoride. After stirring for 20 min, this mixture was diluted 1:1 with 20 ml of 100 mM KH\(_2\)PO\(_4\), pH 7.0, 12 mM MgCl\(_2\), 40 μM ZnCl\(_2\), 10 mM β-ME, and 0.1 mM phenylmethylsulfonyl fluoride, followed by the addition of DNase I (Sigma) to a final concentration of 25 μg/ml and a 20-min incubation at 25 °C with stirring.

Bacteria were lysed by several cycles of freezing, thawing, and sonication and sequentially precipitated using a cut between 20 and 45% ammonium sulfate. *TgPBGS* protein precipitated in the 45% pellet and was resuspended in 30 mM KH\(_2\)PO\(_4\), pH 7.5, 5% ammonium sulfate, and 0.1 mM MgCl\(_2\). The protein was applied to a 75-ml phenyl-Sepharose CL-4B column (GE Healthcare) and eluted at 4 °C using a 750-ml gradient from 30 to 2 mM KH\(_2\)PO\(_4\) (pH 7.5) and from 15 to 0% ammonium sulfate, in the presence of 0.1 mM MgCl\(_2\). 10-ml fractions were collected at a flow rate of 1 ml/min, and those containing enzyme activity were pooled and applied to a 75-ml Q-Sepharose column (GE Healthcare) equilibrated with 10 mM Tris, pH 7.0, and 1 mM MgCl\(_2\). The column was eluted at room temperature using a 1-liter Tris gradient from 10 mM to 1 M at pH 7 in the presence of 1 mM MgCl\(_2\). 10-ml aliquots were collected at 3 ml/min, and those spanning the enzyme activity peak were pooled based on purity (assessed by SDS and native gel electrophoresis) and concentrated to ~6 mg/ml. A 24-ml Superdex 200 analytical gel filtration column (GE Healthcare) equilibrated in 100 mM Tris, pH 7.0, 1 mM MgCl\(_2\), 10 mM β-ME was used to determine protein size after calibration with standard markers. All chromatography columns were run using the
T. gondii PBGS

AKTA system (GE Healthcare). A portion of the resulting TgPBGS protein was dialyzed against several changes of 10 mM Tris, pH 7, containing 10 mM β-ME to remove Mg2+ ions.

Monitoring Parasite Heme Synthesis by 14C-ALA Incorporation—Freshly isolated, extracellular tachyzoite stage parasites were washed and resuspended in minimal essential medium (Invitrogen) containing 1% diazylated fetal calf serum. 10⁸ parasites were preincubated for 2 h at 37 °C under 5% CO₂ in 1 ml of minimal essential medium containing 0.1, 0.25, 0.5, 1, or 2 mM succinylacetone (pH 7.3), followed by the addition of 1 μCi of 14C-ALA (HCl salt) and incubation for a further 6–8 h. Radiolabeling was stopped by centrifugation and washing with fresh minimal essential medium. Heme was extracted as described previously (16), and radioactivity was estimated by scintillation counting.

Enzyme Activity Assays—PBGS assays were run in 100 mM Bistris propane-HCl buffer containing 10 mM β-ME, as previously reported (17), using 10 μg/reaction protein except when testing protein concentration dependence. Reaction mixtures were preincubated at 37 °C for 10 min before starting the assay with the addition of ALA (10 mM unless otherwise stated). Assays were typically run for 20 min, followed by the addition of 0.1 mM HgCl₂ in 20% trichloroacetic acid, development for 8 min in Ehrlich’s reagent, and measurement of absorbance at 555 nm using the quartz dip probe on a Cary50 spectrometer. Enzyme activity was measured in parasite lysates using 200 μg of total protein (~10⁸ parasites) per assay and a 1-h reaction time.

Homology-based Structural Model of TgPBGS—The PaPBGS crystal structure (Protein Data Bank entry 1GZG) was used as a template for constructing a structural homology model of TgPBGS amino acids 324–643. Alignments were constructed in MolLIDE (18) using PSI-BLAST (~47% identity). Gaps in the alignment map to loops and were modeled independently for each monomer using Loopy (19). Side chain coordinates were retained from the template structure for identical residues; non-identical side chains were modeled using SCWRL 3.0 (20). Side chains for each subunit of the octamer were modeled in the presence of the other subunits to impart appropriate conformational constraints. PaPBGS forms an octamer in solution, and the crystal structure contains two monomers in its asymmetric dimeric unit. The TgPBGS octamer was therefore modeled using one template monomer for subunits A, C, E, and G and the other for subunits B, D, F, and H.

Because no template crystal structure is available for the C-terminal extension of apicomplexan PBGS, TgPBGS (644–658) (VEDMKGTQKFTEPCY) was modeled based on similar sequences in other proteins. The Protein Data Bank provides at least nine sequences (32 structures) containing ≥5 conserved and ≥4 identical amino acids. Five proteins include the sequence VEDMK, which is always helical. The sequence GTQKFT forms a loop and helix in several structures. In three different protein families containing the sequence FTEPCY, PCY was always a non-helical, non-strand loop region, and FTE was variable. Secondary structures for each subsequence were combined using ϕ-ψ angles from available crystal structures, and examination of the resulting possibilities produced only two sterically feasible combinations, one of which was clearly superior based on its interface with the rest of the TgPBGS model. This structure was added to each subunit of the dimer model, with manual adjustment to remove steric hindrance and improve the interaction surface with the two monomeric units. The resulting model was then overlaid with each of the dimers of the octameric 1GZG to form an octamer model of TgPBGS.

The protein preparation workflow in Maestro (Schrödinger, LLC) was used to add hydrogen atoms and identify overlapping atoms in the octameric structure, and the entire model was manually refined to accommodate all interfacing residues. The quality of the resulting model was estimated using the Protein Interfaces, Surfaces, and Assemblies Service (21). See supplemental Text T1 for PDB coordinates of the TgPBGS structure model.

RESULTS

Genomic Organization and Expression of T. gondii PBGS Gene—Analysis of the T. gondii ME49 reference genome (see the ToxoDB Web site) identifies a single candidate PBGS gene, on chromosome III (TGME49.053900). Few expressed sequence tags map to this locus, and microarray evidence suggests low transcript abundance, but several independent proteomic studies have identified TgPBGS peptides (see the ToxoDB Web site). A combination of library screening and reverse transcription-PCR revealed four alternatively spliced transcripts defining 17 exons (Fig. 1A). TgPBGS1 (DQ029337, 3863 nucleotides) fuses exons 1-2-4-5-6-7; TgPBGS2 (DQ029338, 3641 nucleotides) fuses exons 1-3-4-5-6-7, using an alternative splice donor and acceptor that truncate exons 1 and 3; TgPBGS3 (DQ029339, 4579 bp) transcripts appear to initiate 129 nucleotides upstream of TgPBGS1 and -2 and fuse exons 1-2-3-4-5-6-7. All three of these transcripts include conserved PBGS sequences. The fourth cDNA includes no mature PBGS sequence, fusing exons 1 and 2 to exons 8–17 (confirmed by reverse transcription-PCR using primer pairs, as indicated in Fig. 1A, and by RACE; not shown) and is predicted to encode a plastid stromal processing peptidase (TgSPP). Production of these two unrelated proteins from the same genetic locus is also observed in P. falciparum (22). Quantitative real-time PCR was used to calculate steady-state abundance of TgPBGS in RH strain tachyzoites as 155.2 ± 31.3 fg/μg of poly(A)⁺ mRNA, versus 44.7 ± 2.2 fg/μg for TgSPP (i.e. TgPBGS is 3.45 ± 0.53-fold more abundant than TgSPP) (Fig. 1B).

Sequences upstream of the first four ATGs differ from the consensus translational start site for T. gondii (caaaATG) 23), are not conserved in the related parasite Neospora caninum (see the ToxoDB Web site), and fail to initiate a long open reading frame, but the fifth ATG permits synthesis of proteins of 462, 658, or 1715 amino acids (for TgPBGS1, TgPBGS3, and TgSPP, respectively). Use of the fourth ATG would correct for the frameshift introduced by TgPBGS2 intron 1, yielding an open reading frame of 407 amino acids. Only TgPBGS3 is predicted to encode all conserved PBGS sequences (exons 3–7), and Northern blot hybridization reveals a 4.6-kb TgPBGS3 mRNA (not shown). Similarly, Western blotting using polyclonal mouse antiserum raised against TgPBGS (see “Materials and Methods”) detected a ~56-kDa protein (Fig. 1C), larger than the predicted size of TgPBGS1 or -2 but consistent with
PBGS organization, expression, and localization. A, 23 kb spanning the T. gondii PBGS/SPP locus on chromosome III. Exons 1 and 2 encode a bipartite apicoplast-targeting signal (signal sequence + plastid-targeting domain; hatched boxes), fused to mature coding sequences for TgPBGS (exons 3–7; black boxes) or TgSPP (exons 8–17; gray boxes). The white boxes indicate 5′- and 3′-untranslated regions (precise polyadenylation sites have not been mapped for TgPBGS1 and -2). The open arrowheads indicate primers used to define the 5′-end of TgPBGS and TgSPP transcripts by reverse transcription-PCR or RACE, bars indicate fragments amplified for quantitative PCR analysis of mRNA abundance, and closed arrowheads indicate primers used for assessing TgPBGS expression in different parasite stages. Three distinct transcripts including PBGS sequences were identified; B, quantitation of TgPBGS and TgSPP transcripts by quantitative real-time PCR using gene-specific primer pairs (supplemental Table S1); TgPBGS was amplified above the Ct threshold – 2 cycles earlier than TgSPP, corresponding to a 3.5-fold difference in mRNA abundance (replicate samples at three different dilutions). C, Western blotting of parasite cell lysates detects the 56-kDa PBGS3 protein. H, primary human foreskin fibrobast cell lysate; T1, 5 × 10⁶ T. gondii tachyzoites; T2, 10⁶ tachyzoites; B, 400 T. gondii bradyzoite cysts (purified from a chronically infected mouse brain). D, a 442-bp PCR product specific for TgPBGS mRNA was detected in tachyzoites (T), bradyzoites (B), and sporozoites (S). G, amplification of genomic DNA using the same primers yields a 1.1-kb product. E, subcellular localization of TgPBGS in transgenic parasites expressing the N-terminal leader (394 amino acids) fused to eGFP (green). The apicoplast (red) was detected by immunostaining for T. gondii acyl carrier protein (ACP). F, schematic diagram of TgPBGS protein features showing the bipartite plastid-targeting region (hatched), predicted signal peptide, and experimentally determined transit peptide cleavage sites and the portions expressed as recombinant proteins. aa, amino acids; nt, nucleotides.

The predicted size of TgPBGS3 after cleavage of the apicoplast-targeting signal. TgPBGS transcripts were detected in cDNA derived from three major parasite life cycle stages: tachyzoites, bradyzoites, and sporozoites (Fig. 1D).
the partially purified protein displayed a broad elution profile from Q-Sepharose (supplemental Fig. S2A, bottom). The size and purity of recombinant proteins were confirmed by SDS-PAGE (supplemental Fig. S2B); because TgPBGS-ΔC exhibits very low activity, positive fractions were pooled and concentrated. Purified TgPBGS-ΔC protein formed a fluffy precipitate upon extended storage. No contaminating E. coli PBGS was detected (25), based on assays of TgPBGS-ΔC phenyl-Sepharose peak without metals, with Zn²⁺ alone, Mg²⁺ alone, or Zn²⁺ plus Mg²⁺ (data not shown).

**Kinetic Properties and Metal Ion Dependence of Recombinant TgPBGS**—Under optimal conditions (see “Materials and Methods”), the wild type TgPBGS enzyme exhibits specific activity of 23–25 μmol of porphobilinogen/mg of protein/h and a Kₘ of 0.149 mM (Fig. 2A), comparable with previous reports for PBGS from other species (7, 17, 26, 27). TgPBGS activity is also dependent on protein concentration, reaching maximal activity at ~10 μg/ml (~0.25 μM) and an apparent Kₜ of ~0.04 μM (Fig. 2B), comparable with values reported for PsPBGS (17) and Bradyrhizobium japonicum PBGS (26). In other systems, concentration dependence is associated with regulation based on quaternary structure because high protein concentration favors the formation of high activity octamers over low activity hexamers. The specific activity of TgPBGS does not approach zero at low protein concentrations, however, suggesting a different quaternary structure equilibrium. The excellent fit to standard Michaelis-Menten kinetics over substrate concentrations from 1 μM to 10 mM (Fig. 2A) also argues against alternative quaternary structure assemblies with differing Kₘ values (6, 29).

Many PBGS enzymes are completely dependent on metal ions, but some Mg²⁺-binding enzymes exhibit partial metal ion independence (13, 30). Because protein sequence alignments suggested that TgPBGS can bind Mg²⁺ at both the active site and an allosteric site, the enzyme was purified in 1 mM Mg²⁺ (see “Materials and Methods”). Close to maximal activity was observed using undialyzed TgPBGS-wt in assays containing ~2 μM Mg²⁺ (Fig. 2C, squares). Thus, if Mg²⁺ is required for activ-

**FIGURE 2. Kinetic properties of wild type TgPBGS.** A, specific activity as a function of ALA concentration. Fitting to the Michaelis-Menten equation yields a Kₘ value of 0.149 mM and Vₘ₉₉₉ of 23.2 μmol of porphobilinogen/mg of protein/h. B, TgPBGS specific activity is dependent on enzyme concentration, reaching maximal activity at ~10 μg/ml (~0.25 μM). The data are fitted to a simple hyperbolic equation with a y (offset). C, specific activity of TgPBGS as a function of pH and Mg²⁺. Undialyzed protein (squares) contains ~2 μM Mg²⁺, and dialyzed protein (circles) contains <0.01 μM Mg²⁺; ALA contributed an additional ~0.18 μM Mg²⁺ to the assay. D, effect of EDTA on TgPBGS enzyme activity (using undialyzed TgPBGS without added Mg²⁺). E, MgCl₂ titration of TgPBGS specific activity. Two Kₛ values of ~0.7 and 84 μM were determined, based on fitting a double hyperbolic equation. Unless otherwise specified, standard assay conditions included 10 μg/ml TgPBGS (except in B), 10 mM ALA (except in A), pH 8.14 (except in C and D), 1 mM MgCl₂ (except in C–E), and 10 mM β-ME; assays were initiated by the addition of ALA and incubated for 20 min at 37 °C.
ity, it has a very high affinity for the protein because the binding site seems to be significantly populated even at low Mg\(^{2+}\)/H\(_{1001}\) concentrations. \(\sim 40\%\) maximal activity was detected even after dialysis depleted Mg\(^{2+}\)/H\(_{1001}\) to \(<0.3\) ions/octamer, as determined by atomic absorption spectroscopy (Fig. 2C, circles). EDTA treatment eliminated the activity of undialyzed protein (Fig. 2D), consistent with the hypothesis of a tight binding catalytic Mg\(^{2+}\). Varying the concentration of Mg\(^{2+}\) added to the reaction suggests two affinities, with \(K_a\) values of \(~0.7\) and \(84\) \(\mu\)M (Fig. 2E), which we attribute to the active and allosteric sites, respectively. These affinities are significantly tighter than the \(K_a\) values of \(35\) \(\mu\)M and \(2\) mM reported for \(PsPBGS\) (17).

**Evaluation of the \(TgPBGS\) Quaternary Structure**—The molecular mass of recombinant \(TgPBGS\) was determined by gel filtration under optimal conditions (see “Materials and Methods”). Comparison with standard marker proteins indicates a molecular mass of \(~323\)-kDa \(TgPBGS\), consistent with an octameric structure (Fig. 3A), as observed in other PBGS enzymes (7, 31). Unlike \(HsPBGS\) or \(PsPBGS\), however, hexamers were not evident in \(TgPBGS\), even as a shoulder of the main peak.
T. gondii PBGS

peak. The allosteric Mg\(^{2+}\) ions of octameric PBGS reside at a subunit-subunit interface, and their removal from *Pa*PBGS destabilizes the octamer, resulting in hexamer accumulation (6). Purification of *Pa*PBGS without Mg\(^{2+}\) yields a dimeric complex, and dialysis against millimolar levels of Mg\(^{2+}\) produces an octameric assembly. Substrate has also been shown to stabilize octameric assembly in PBGS proteins prone to oligomer dissociation (6, 25, 31, 32, 33). Native gel electrophoresis of *Tg*PBGS reveals that this protein is primarily present as an octamer (Fig. 3B), and even dialysis to reduce Mg\(^{2+}\) to <0.3 Mg\(^{2+}\)/octamer does not affect the quaternary structure. The absence of substrate or the presence of EDTA also did not affect the stability of *Tg*PBGS octamer (data not shown).

As noted above, apicomplexan PBGS enzymes are distinguished by their unusual termini, including the N-terminal extension required for targeting to the apicoplast (which is cleaved post-translationally) and C-terminal extension unique to these parasites (*Vibrio cholerae* also displays a shorter, unrelated C-terminal extension; supplemental Fig. S1). *Tg*PBGS-ΔC, a mutant form of the enzyme lacking the C-terminal 13 amino acids, purified as dimers according to gel filtration analysis (Fig. 3A). The absence of higher order *Tg*PBGS-ΔC oligomers was also confirmed by native gel electrophoresis (Fig. 3B) and Western blotting (Fig. 3C). This striking difference in the native structure of *Tg*PBGS-wt versus *Tg*PBGS-ΔC suggests that the C-terminal 13-amino acid extension may be required for octamer stability. Specific activity of *Tg*PBGS-ΔC was ~25-fold lower than for *Tg*PBGS-wt (~1 μmol/mg protein/h; Fig. 3D), but it is not clear if this activity is intrinsic to the dimer or due to transient octamer formation in the presence of substrate, which is known to stabilize the octamer form of PBGS in other species.

In order to further explore the role of the *Tg*PBGS C-terminal extension, we constructed a homology model based on the known structure of *Pa*PBGS (see “Materials and Methods”). Individual PBGS subunits form a ~300-amino acid α₂β₅-barrel, with an extended N-terminal arm of variable length that interacts with other subunits in oligomeric forms of the enzyme (4). Available crystal structures also show that the C terminus, which is often disordered in the crystal structure, points toward the N terminus and might therefore influence higher order oligomer formation as well. The structural model of *Tg*PBGS (Fig. 3E) predicts that the C-terminal extension increases the interaction surface area from 1604 to 2023 Å\(^2\). In addition, the solvation free energy of the protein interface (ΔG; calculated by the Protein Interfaces, Surfaces, and Assemblies Service, EBI) was ~14.6 kcal/mol for *Tg*PBGS-wt versus ~4.4 kcal/mol in the *Tg*PBGS-ΔC mutant. For comparison, the average ΔG for the equivalent interface in crystal structures of human, mouse, yeast, *Chlorobium*, and *E. coli* PBGS are ~15.4 ± 5.4 kcal/mol, strongly arguing that the C-terminal extension of *Tg*PBGS is required for the formation of an energetically favorable interface. This contrasts with *Pa*PBGS and *Pa*PBGS but is consistent with the observation that *Tg*PBGS octamers remain stable in the absence of Mg\(^{2+}\) (see above), presumably due to stability provided by the C-terminal extension.

**FIGURE 4. Inhibition of PBGS activity by active site-directed suicide inhibitors.** Inhibition of *Tg*PBGS and *Hs*PBGS enzyme activity (under optimal conditions) by succinylacetone (A, preincubated 30 min) or 4,7-dioxosebacic acid (4,7-DOSA) (B, preincubated 100 min).

**Inhibition of *Tg*PBGS and Heme Biosynthesis in *T. gondii*—** Previous studies have described various substrate mimics that specifically inhibit PBGS (28, 34–37, 41), including succinylacetone, which mimics ALA and is a potent irreversible inhibitor (38). As shown in Fig. 4A, the dose-response curve for succinylacetone inhibition of *Tg*PBGS is comparable with that of *Hs*PBGS, with IC\(_{50}\) values of ~1 μM. This is to be expected because the inhibitor binds covalently to a conserved active site lysine, causing suicide inhibition (39). In contrast, the active site-directed, species-specific irreversible inhibitor 4,7-dioxosebacic acid, which is most effective against Zn\(^{2+}\)-requiring PBGS (28, 37, 40), shows ~40-fold selectivity for *Hs*PBGS relative to *Tg*PBGS (Fig. 4B). The specificity of 4-oxosebacic acid for *E. coli* PBGS is also understood at a mechanistic level (37, 40, 41); this compound failed to inhibit either *Tg*PBGS or *Hs*PBGS (not shown), even at concentrations as high as 3 mM and a preincubation time up to 24 h.

Robust heme synthesis is observed during active stages of the *T. gondii* life cycle and can be monitored by following the incorporation of 14C-ALA. PBGS enzyme activity in parasite cell lysates was inhibited by succinylacetone in a dose-dependent manner (Fig. 5A). Succinylacetone treatment also markedly decreased parasite heme biosynthesis (Fig. 5B).
The effect of succinylacetone on the survival of *T. gondii* parasites was monitored using transgenic lines expressing a luciferase reporter gene (15) (Fig. 5 C). Once again, succinylacetone decreased *T. gondii* growth in a dose-dependent manner, ultimately killing the parasites. Similar dose-response curves were obtained for heme incorporation and parasite survival. Comparable results have been reported previously for *Plasmodium* species (10), suggesting that PBGS activity and heme biosynthesis may be essential for apicomplexan parasite survival.

**DISCUSSION**

**The Unusual PBGS Gene Structure of Apicomplexan Parasites**—Heme biosynthesis in *T. gondii* and *P. falciparum* is distinct from their animal hosts in that three enzymes of the common pathway, including PBGS, are located in the apicoplast (12). Motifs required for plastid targeting (42) can be identified in the sequence of TgPBGS and *P. falciparum* PBGS, and apicoplast localization has been confirmed in both species (12, 13, 43). In both species, PBGS sequences (TGME49_053900 and PF14_0381) lie immediately upstream of a stromal processing peptidase gene (pitrilysin; SPP; TGME49_053890 and PF14_0382). These two genes form a single genetic locus, exploiting the same upstream exons to encode an apicoplast-targeting signal (Fig. 1 E), which is post-translationally removed by signal peptidase and pitrilysin to yield two distinct mature proteins (Fig. 1 A). At the transcript level, TgPBGS is ~3–5-fold more abundant than TgSPP (Fig. 1B); the analogous transcripts in *Plasmodium* are equally abundant (22). Mapping the mature N terminus of TgPBGS confirms that processing occurs near the 3'-end of exon 2, removing most of the sequences shared by TgPBGS and TgSPP (Fig. 1F). Three alternatively spliced cDNAs encoding PBGS sequence were identified in *T. gondii* cDNA libraries (Fig. 1A), but bioinformatic, RNA, protein, and biochemical analyses indicate that TgPBGS3 is by far the most abundant (Fig. 1C) and likely to be the only functional species.

**TgPBGS Structure and Function**—The activity of recombinant TgPBGS (synthesized without the N-terminal plastid-targeting signals and low complexity sequence; Fig. 1F), is comparable with *Hs*PBGS but lower than that reported for *Ps*PBGS. Efforts to express the full-length mature wild type TgPBGS protein (amino acids 148–658) were unsuccessful, probably because of low complexity sequence in the N-terminal region (amino acids 148–301). This domain is not well conserved between *Plasmodium* and *Toxoplasma* or even between *T. gondii* and its close relative *N. caninum*. Several insertions and deletions in the low complexity region are observed in different *T. gondii* isolates. We therefore expect that the TgPBGS-wt protein used for these studies (amino acids 302–658) accurately reflects the enzyme in vivo.

Dialysis, EDTA treatment, and Mg<sup>2+</sup> titration experiments lead us to conclude that TgPBGS requires Mg<sup>2+</sup>, which is bound very tightly, but that the importance of Mg<sup>2+</sup> to enzyme quaternary structure differs from that observed in plant/algal PBGS. In PsPBGS, Mg<sup>2+</sup> is required for activity, binding to both the active site and an allosteric site located at a subunit interface. The latter is required for the assembly of active octamers rather than inactive hexamers (6, 31). Similarly, the removal of
Mg$^{2+}$ from PaPBGS promotes the formation of inactive dimers. Although titration demonstrates the importance of Mg$^{2+}$ for optimal TgPBGS activity (Fig. 2E), which was abolished by complete Mg$^{2+}$ removal (Fig. 2D), TgPBGS purified in 1 mM MgCl$_2$ retained its octameric structure and ~40% maximal enzyme activity even after removal of Mg$^{2+}$ by dialysis (Fig. 2C). Atomic absorption spectroscopy confirmed very low levels of Mg$^{2+}$ ions in dialyzed TgPBGS, but the addition of 10 mM ALA to the assay introduces trace Mg$^{2+}$ that may be sufficient for activity. PsPBGS shows no activity under such conditions because its affinity for Mg$^{2+}$ is much lower (see the legend to Fig. 2C). Mg$^{2+}$ concentration in the chloroplast varies between ~1 and 10 mM to support chlorophyll synthesis; although Mg$^{2+}$ concentration in the apicoplast is unknown, it is likely to be far lower because parasites do not synthesize chlorophyll. Tight binding of Mg$^{2+}$ may therefore be necessary for TgPBGS function.

Quaternary structure has been shown to dramatically affect the activity of PBGS from many species (29). The plant and human enzymes are now known to exist as a mixture of low activity hexamers and high activity octamers whose interconversion occurs through a conformational change in the dissociated dimeric assembly. These non-additive quaternary isoforms of the enzyme are called “morpheein forms,” and proteins that behave in this fashion are called “morpheins” (9). Interconversion of morphein forms requires oligomer disassembly, conformational change in the dissociated state, and reassembly as a structurally and functionally distinct oligomer. The equilibrium of quaternary structure assemblies of PBGS is affected by the presence of metal ions and substrate molecules, both of which favor formation of high activity octamers (7, 25). For PsPBGS, the equilibrium between octamers and hexamers results in a protein concentration dependence of the specific activity. Under optimal conditions, the specific activity of TgPBGS also exhibits protein concentration dependence (Fig. 2B), although there is no evidence for hexamers (Fig. 3, A–C). Thus, the quaternary structure equilibrium for TgPBGS is unlike that of the prototype plant PBGS.

A key structural difference between TgPBGS and PsPBGS is the C-terminal extension present in the former. Protein sequence alignments show that all apicomplexan PBGS enzymes contain such a C-terminal extension, although the precise sequence is not conserved (supplemental Fig. S1). Removal of this sequence in the mutant enzyme TgPBGS-ΔC completely eliminated its ability to assemble beyond the dimeric state (Fig. 3, A–C), pointing to the C terminus as likely to be responsible for observed differences in quaternary structure equilibria between TgPBGS and other plant-type PBGS enzymes. Fig. 3E provides a model suggesting how the C-terminal extension may hold the octamer together. Note that the C-terminal extension is predicted to lie at the edge of PBGS dimers (blue-green pairs), interacting with the C-terminal extension of adjacent dimers in an energetically favored conformation. Why TgPBGS might have evolved to use an octamer-stabilizing C-terminal extension whereas allosteric Mg$^{2+}$ is sufficient to stabilize the octameric structure in PsPBGS is unclear but provides scope for further studies on morpheein structure and function and structural differences that might be exploited for therapeutic development. Crystallographic data for TgPBGS would certainly help to expedite these goals.

**Capitalizing on Apicomplexan PBGS for Drug Discovery**—PBGS is likely to be essential for parasite survival, both by analogy to other systems and based on the observations that treatment with succinylacetone inhibits heme biosynthesis and kills *T. gondii* (albeit only at micromolar concentrations, as also observed in other eukaryotic systems). Most known inhibitors of PBGS are substrate mimics that bind the highly conserved active site, making it challenging to design small molecules that can discriminate between the human and parasite enzymes (although note that 4,7-diosebacic acid is far more active against *Hs*PBGS than TgPBGS). Allosteric inhibitors may be more readily developed to exploit phylogenetic differences in protein sequence and structure. Recent studies have demonstrated the potential of species-specific allosteric structure-based inhibitors of PBGS that function by trapping the protein in its inactive hexameric assembly (33, 44). Such inhibitors, termed “morphlocks,” have been selected through *in silico* screening for molecules that bind to a phylogenetically variable surface cavity found only in the low activity PBGS hexamer. A similar structure-based approach may be envisioned for discriminating human versus parasite PBGS, focusing on the parasite-specific C-terminal domain required for active octamer formation.

**Acknowledgments**—We acknowledge Linda Stith for technical assistance, Dr. Manami Nishi for assistance with microscopic imaging, Prof. Roland Dunbrack and the Fox Chase Cancer Center Molecular Modeling Facility for assistance with construction of the TgPBGS model, and Prof. Reinhard Neier for 4,7-diosebacic acid.

**REFERENCES**

1. Papenbrock, J., and Grimm, B. (2001) *Planta* **213**, 667–681
2. Heinemann, I. U., Jahn, M., and Jahn, D. (2008) *Arch. Biochem. Biophys.* **474**, 238–251
3. Jaffe, E. K. (2004) *Bioorg. Chem.* **32**, 316–325
4. Jaffe, E. K. (2000) *Acta Crystallogr. D Biol. Crystallogr.* **56**, 115–128
5. Jaffe, E. K. (2003) *Chem. Biol.* **10**, 25–34
6. Breining, S., Kervinen, J., Stith, L., Wasson, A. S., Fairman, R., Wlodawer, A., Zdanov, A., and Jaffe, E. K. (2003) *Nat. Struct. Biol.* **10**, 757–763
7. Tang, L., Stith, L., and Jaffe, E. K. (2005) *J. Biol. Chem.* **280**, 15786–15793
8. Jaffe, E. K., and Stith, L. (2007) *Am. J. Hum. Genet.* **80**, 329–337
9. Jaffe, E. K. (2005) *Trends Biochem. Sci.* **30**, 490–497
10. Surolia, N., and Padmanabhan, G. (1992) *Biochem. Biophys. Res. Commun.* **187**, 744–750
11. Srivastava, P., and Pandey, V. C. (1998) *Exp. Parasitol.* **86**, 60–63
12. Wu, B. (2006) *Heme Biosynthetic Pathway in Apicomplexan Parasites*, Ph.D. thesis, University of Pennsylvania, Philadelphia, PA
13. Dhanasekaran, S., Chandra, N. R., Chandrasekhar Sagar, B. K., Rangarajan, P. N., and Padmanabhan, G. (2004) *J. Biol. Chem.* **279**, 6934–6942
14. Roos, D. S., Donald, R. G., Morriseotte, N. S., and Moulton, A. L. (1994) *Methods Cell Biol.* **45**, 27–63
15. Matrajt, M., Nishi, M., Fraunholz, M. J., Peter, O., and Roos, D. S. (2002) *Mol. Biochem. Parasitol.* **120**, 285–289
16. Bonday, Z. Q., Taketani, S., Gupta, P. D., and Padmanaban, G. (1997) *J. Biol. Chem.* **272**, 21839–21846
17. Kervinen, J., Dunbrack, R. L., Jr., Litwin, S., Martins, I., Scarow, R. C., Volin, M., Yeung, A. T., Yoon, E., and Jaffe, E. K. (2000) *Biochemistry* **39**, 9018–9029
18. Canutescu, A. A., and Dunbrack, R. L., Jr. (2005) *Bioinformatics* **21**, 2914–2916
