Published in final edited form as:

Biochemistry. 2012 September 25; 51(38): 7533–7540. doi:10.1021/bi3005335.

The N-terminus and the Chain-length Determination (CLD) Domain Play a Role in the Length of the Isoprenoid Product of the Bifunctional *Toxoplasma gondii* Farnesyl-diphosphate Synthase

Zhu-Hong Li, Roxana Cintrón, Noah A. Koon, and Silvia N.J. Moreno

Center for Tropical and Emerging Global Diseases and Department of Cellular Biology, University of Georgia, Athens, GA 30602

Abstract

Toxoplasma gondii possesses a bifunctional farnesyl diphosphate (FPP)/geranylgeranyl diphosphate (GGPP) synthase (TgFPPS) that synthesizes C₁₅ and C₂₀ isoprenoid diphosphates from isopentenyl diphosphate (IPP) and dimethylallyl diphosphate (DMAPP). This enzyme has a unique arrangement of the 4th and 5th amino acid upstream to the First Aspartic Rich Domain (FARM) where the 4th amino acid is aromatic and the 5th is a cysteine. We mutated these amino acids converting the enzyme to an absolute FPPS by changing the cysteine to a tyrosine. The enzyme could be converted to an absolute GGPPS by changing both the 4th and 5th amino acids to alanines. We also constructed four mutated TgFPPSs whose regions around the first aspartate-rich motif were replaced with the corresponding regions of FPP synthases from *Arabidopsis thaliana* or *Saccharomyces cerevisiae* or with the corresponding regions of GGPP synthases from *Homo sapiens* or *S. cerevisiae*. We determined that the presence of a cysteine at the 4th position is essential for the TgFPPS bifunctionality. We also found that the length of the N-terminal domain has a role in determining the specificity and the length of the isoprenoid product. Phylogenetic analysis supports the grouping of this enzyme with other Type I FPPSs but the biochemical data indicates that TgFPPS has unique characteristics that differentiate it from mammalian FPPSs and GGPPSs and is therefore an important drug target.

Isoprenoids are an extensive group of natural products with diverse structures consisting of various numbers of five carbon isoprene units. The major building reaction in the pathway is the sequential condensation of isopentenyl diphosphate (IPP) with growing allylic isoprenoid diphosphates. The enzyme farnesyl diphosphate synthase (FPPS) plays a central role in this pathway in eukaryotes by producing farnesyl diphosphate (FPP), an important precursor of sterols, dolichols, ubiquinones, heme a, and prenylated proteins. FPPS forms

Corresponding Author: Silvia N.J. Moreno, Center for Tropical and Emerging Global Diseases, Paul D. Coverdell Center, 500 D.W. Brooks Dr., Athens, GA, 30602. Tel: 706-542-4736. smoreno@uga.edu.

The authors declare no competing financial interest

ASSOCIATED CONTENT

Supporting Information

Supporting materials, including Figures S1 and S2 and Tables S1 and S2. This material is available free of charge via the Internet at <http://pubs.acs.org>

FPP by the sequential condensation of dimethylallyl diphosphate (DMAPP) with two molecules of IPP. The 15-C isoprenoid units, FPP, could further form GGPP, the 20-C isoprenoid units, in a reaction catalyzed by the geranylgeranyl diphosphate synthase (GGPPS).

In contrast to mammalian cells, *Toxoplasma gondii*, a major opportunistic pathogen of fetuses from recently infected mothers and of patients with AIDS, possesses a bifunctional FPPS/GGPPS (TgFPPS) that is able to catalyze the formation of both FPP and GGPP (1). This peculiarity was also recently reported for the *Plasmodium vivax* enzyme (2). TgFPPS has a long N-terminal extension of about 164 amino acids, 42 of which (26%) are serine, and is localized to the mitochondrion (1). TgFPPS is also a valid target for drugs since bisphosphonates, diphosphate analogues that are specific FPPS inhibitors, inhibit parasite growth *in vitro* and *in vivo* (3–8). Lipophilic bisphosphonates that block prenyl synthases also have potent activity against *Plasmodium* liver stages (9).

FPPS sequences contain seven conserved regions including two aspartate-rich domains named as FARM, for First Aspartic Rich Motif, and SARM, for Second Aspartic Rich Motif, which are critical for the catalytic action and most likely act as the sites for allylic substrates (10). The 4th and 5th amino acids upstream of the FARM have been shown to have a critical role in the determination of the length of the isoprenoid products in the reactions catalyzed by FPPSs (11). These two amino acids have aromatic side chains, like Phe and Tyr, that form the floor of the hydrophobic pocket in the interior of the enzyme that binds the growing hydrocarbon chain (11). Replacement of these aromatic rings with smaller side chains leads to the formation of longer chain products (12–14). This domain (from the FARM to the 5th amino acid upstream) was named as the chain-length determination (CLD) domain (Figure S1A). Some reports indicate that amino acids located in regions other than those defined as CLD also play an important role in product chain-length determination mechanism of prenyl synthases (15) (16). It is also important to note that the conserved aspartate residues of the two aspartate rich motif DDXXD (FARM and SARM) bind three divalent cations (Mg^{2+}) that are in turn coordinated by the phosphate backbone of bisphosphonates, which are potent inhibitors of these enzymes (17).

When comparing the CLD region of TgFPPS with other GGPPSs and FPPSs, we found some interesting differences (1) (Figure S1A). The most interesting and of relevance for the present work is that the 4th amino acid upstream to the FARM region is phenylalanine (F), an amino acid with a bulky aromatic side chain, but the 5th amino acid upstream to the FARM region is a cysteine (C). This arrangement is peculiar when compared to other short chain prenyl synthases because, for example Type I FPPS have two aromatic amino acids at the 4th and 5th position upstream to the FARM (Figure S1B) and Type II FPPSs and Type I GGPPS an aromatic amino acid in the 5th position upstream of the FARM and a small side chain-containing amino acid in the 4th position (Figure S1B). Interestingly, this peculiar amino acid organization is also found in the *Plasmodium vivax* (2), and *P. falciparum* FPPS orthologs: a phenylalanine in the 4th position and an alanine or a serine in the 5th position upstream of the FARM region, respectively. This uncommon combination of 4th and 5th amino acids may provide the enzymes with the unusual characteristic of being bifunctional producing both FPP and GGPP as products (1). However, no mutational study that supports

this hypothesis has been made with any of these enzymes. This feature may be convenient for the parasites because there is no evidence for the presence of genes with homology to GGPPSs of other organisms in the genomes of either *T. gondii* or *Plasmodium* spp. (<http://eupathdb.org/eupathdb/>).

In the present work mutational studies of TgFPPS were performed to investigate the mechanism responsible for the bifunctionality of TgFPPS. We found that the CLD arrangement of amino acids does have a role in the bifunctionality of the enzyme. Unexpectedly, we found that the length of the N-terminal domain has a role in determining the length of the isoprenoid product of the TgFPPS reaction. In spite of its bifunctionality, our phylogenetic analysis shows that TgFPPS groups with other Type I FPPSs and not with GGPPSs.

MATERIALS AND METHODS

Materials

Oligonucleotide primers were obtained from Integrated DNA Technologies (Coralville, IA). TaqDNA polymerase, the site directed mutagenesis kit, and restriction enzymes were from Invitrogen. ECL chemiluminescence detection kit, and nylon membranes were from Amersham Biosciences. Plasmid miniprep kit, gel extraction kit and plasmid maxiprep kit were from Qiagen Inc. (Chatsworth, CA). IPP, DMAPP, GPP, FPP, GGPP were from Isoprenoids, LC (FL, USA). Anti-FLAG M2 affinity resin was from Sigma. [4-¹⁴C]Isopentenyl diphosphate triammonium salt (55.0 mCi/mmol) was from PerkinElmer Life Sciences. All other reagents were analytical grade.

Cell Cultures

Tachyzoites of *T. gondii* RH strain were cultivated in human fibroblasts and purified as described before (18). Host cells were cultivated in Dulbecco's modified minimal essential medium supplemented with 10% fetal bovine serum. Host cells were infected with tachyzoites at a ratio of 5:1 (parasites:host). Tachyzoites were collected 2–3 days post-infection. Cell cultures were maintained at 37 °C at 5% CO₂.

Construction of TgFPPS Mutants

The construction of TgFPPS-171 has been described previously (1). The *TgFPPS-141*, and *TgFPPS-103* expression cassettes were generated by PCR (primers 15 and 16 in Table S1). The truncated *TgFPPS*s were cloned into the NcoI and XhoI sites of pET-28a using routine protocols. Site directed mutagenesis was performed following the manufacturer's instructions (Invitrogen). The pET-28a-*TgFPPS-141* was used as template and the oligonucleotides used are included in Table S1 (primers 1 to 14). All the mutants were confirmed by sequencing.

Expression and Purification of Recombinant TgFPPS in *E. coli*

The *TgFPPS-103*, *TgFPPS-141* and *TgFPPS-171* expressing plasmids were introduced into *E. coli* BL21(DE3) strain. The expression of recombinant TgFPPSs was induced by addition of 0.4 mM isopropyl β-thiogalactopyranoside at 18 °C overnight. The recombinant proteins

were purified using His-Bind 900 cartridges (Novagen) following the manufacturer's instructions. The polyhistidine tag was not removed because this treatment resulted in an almost complete loss of the catalytic activity of FPPSs from other parasites (19) while the His-tagged enzyme we previously purified from *T. gondii* had excellent activity (1).

Purification of TgFPPS-FLAG with an Anti-FLAG resin

Approximately 8×10^9 of TgFO cells (1) were purified, washed and resuspended with lysis buffer (20 mM Hepes, 50 mM KCl, 125 mM sucrose, 0.5 mM EDTA, pH 7.2). A freeze and thaw method was used to break the parasites. The supernatant was collected by centrifugation at $12,000 \times g$ for 10 min at 4 °C and then mixed with the anti-FLAG M2 affinity resin for 3 h at 4 °C. After washing with TBS buffer (50 mM Tris-HCl, 150 mM NaCl, pH 7.4), the FLAG-tagged protein was eluted with 3× FLAG peptide in TBS buffer at 200 ng/μl.

SDS-Gel Electrophoresis and Western Blot Analyses

Western blot analyses were performed as described previously (1). We used an affinity purified rabbit anti-TgFPPS antibody generated by Cocalico Biologicals (Reamstown, PA) against the recombinant TgFPPS-171 protein.

Enzymatic Activity Assay and Kinetic Studies

The enzymatic assays were performed in a total volume of 100 μl of a mixture containing 10 mM Hepes buffer, pH 7.4, 1 mM MgCl₂, 2 mM dithiothreitol, 100 μM [4-¹⁴C]IPP (10 μCi/μmol), and allylic substrate (100 μM DMAPP, 27 μM GPP, or 13 μM FPP), and 4–40 ng of protein. The reactions were carried out at 37 °C for 30 min and terminated by the addition of 10 μl of 6 M HCl. The mixture was made alkaline by addition of 15 μl of 6 M NaOH, extracted with 1 ml of hexane, washed with water, and then transferred to a scintillation vial for counting.

For kinetic studies, the concentration of DMAPP, GPP, FPP, or IPP was varied, whereas the corresponding counter-substrate was kept at saturating concentration. A nonlinear regression analysis in Sigma Plot 10.0 was used to estimate the kinetic parameters.

Reverse Phase TLC

The radioactive prenyl products synthesized during the enzymatic reactions were hydrolyzed to the corresponding alcohols by incubating them overnight with an alkaline phosphatase (Life sciences, 18011-015) at room temperature. The alcohols were subsequently extracted with hexane and separated by thin layer chromatography using HP-TLC-RP18 plates by reversed phase unibond octadecyl modified silica gel (AnalTech, DE) using acetone:H₂O (6:1; v/v) as running solvent. The position of the standard prenyl alcohol was visualized using iodine vapor. Radioactivity was visualized by autoradiography or using a phosphorImager.

Phylogenetic analysis

The prenyl synthase protein sequences were retrieved from the National Center for Biotechnology Information (NCBI) and OrthoMCL (version 1.0) databases. The phylogenetic tree from Wang and Ohnuma (11) was used as a guide to search for known farnesyl diphosphate (FPPS), geranylgeranyl diphosphate (GGPPS) and C₃₀-C₅₀ diphosphate synthase protein sequences. The *E. coli* FPPS sequence was used to query the NCBI database using the genomic Basic Local Alignment Search Protein (BLASTP) tool by selecting for organisms ranging from prokaryotes to eukaryotes. The default BLAST parameters (expected threshold: 10, low-complexity filter, BLOSSUM 62 substitution matrix) were used.

Protein sequences were used to create multiple sequence alignments to construct phylogenetic trees of prenyl synthases. Multiple sequence alignments were done using the default (slow/accurate) multiple alignment parameters of the Clustal W program (20, 21). In order to optimize the alignments, the Multiple EM for Motif Elicitation (MEME) program (22) was used to search for protein motifs and enable deletion of non-conserved regions. The default expected motif distribution was applied, which specifies the ZOOPS (zero or one per sequence – zero or one instance of the pattern per sequence). The GeneDoc (version 2.6.002) Multiple Sequence Alignment Editor and Shading Utility (23) was used to visualize the motifs identified by MEME. The phylogenetic analyses were conducted in the MEGA5 software (24) using the Dayhoff matrix based model for amino acid substitutions (25). The evolutionary divergence between sequences was estimated to construct one-hundred phylogenetic trees by the Neighbor-joining method (26).

RESULTS

Heterologous Expression and Activity of Truncated TgFPPS

With the aim of increasing the solubility of the recombinant TgFPPS expressed in bacteria we constructed expression vectors with truncated versions of the long N-terminal extension (absent in other FPPSs). Using this approach we expressed TgFPPS proteins lacking 103, 141 or 171 amino acids of the N-terminal extension (Figure S1C). This approach improved the solubility of the recombinant proteins. The protein lacking the first 171 amino acids was the most soluble and was expressed at higher levels in *E. coli*. We also overexpressed the Flag-tagged native enzyme in *T. gondii* and purified it using anti-Flag affinity resins as described before (1). Figure 1 shows the western blot analyses of the different proteins obtained using affinity-purified polyclonal antibodies against TgFPPS. It is interesting to note that the endogenous protein (Figure 1B) has a similar molecular mass (~58 kDa) than the FLAG-tagged native and recombinant truncated versions of the protein (Figure 1A, and (1)) while expression of the full length recombinant TgFPPS in bacteria has a molecular mass of ~75 kDa (data not shown), which is closer to the calculated mass of 69.6 kDa (646 amino acids). This result agrees with the prediction using Mitoprot (<http://ihg.gsf.de/ihg/mitoprot.html>) that the mitochondrial targeting signal is processed at approximately 150 amino acids. These results suggest that the endogenous TgFPPS is processed in *T. gondii* losing at least part of its N-terminal extension.

We tested the activity of these proteins (native, -171, -141, -103) and analyzed the products formed using TLC, and the data are shown in Figure 2A-D, respectively. When native TgFPPS (Figure 2A) was used a mixture of GPP, FPP, and GGPP was produced. However, the low amount of protein obtained precluded further analysis of this enzyme. We investigated the activity of the TgFPPS-103 or TgFPPS-141 recombinant enzymes with DMAPP as substrate and the reaction produced a mixture of GPP, FPP, and GGPP (Figure 2B and C). Using GPP as substrate yielded FPP and GGPP, and FPP yielded only GGPP. These results are in agreement with a stepwise or partially processive mechanism of chain elongation. Unexpectedly, TgFPPS-171 showed mostly FPPS activity and it formed GPP and FPP with DMAPP as substrate or FPP with GPP as substrate (Figure 2D). This mutant was not efficient at elongating FPP (Figure 2D), suggesting that the extra amino acids that are missing in this sequence have a role in determining the length of the isoprenoid product. Interestingly, the lack of this N-terminal extension did not affect the overall enzyme activity with DMAPP as substrate while greatly decreased its activity with FPP as substrate (Figure 2D and Table 1). Since TgFPPS-141 produced the same isoprenoid products as the native enzyme and the yield of the native enzyme (likely processed at 150 amino acids) was too low for further characterization we used this truncated version for the subsequent studies.

Construction of Mutated TgFPPS with Partial Sequences of FPP Synthases

FPPSs have been divided into two types: type I (eukaryotic) and type II (eubacterial), while GGPPS have been divided into three types: type I GGPPS, which includes archaeal GGPPS, type II GGPPS, from eubacteria and plants, and type III GGPPS that include yeast and mammalian GPPS (Figure S1B). The CLD region, especially the amino acids occupying the 4th and 5th position upstream of the FARM as well as two amino acid insertions within the FARM region (Figure S1B), have been reported to determine the final products of the reaction they catalyze (11). TgFPPS, as well as *P. vivax* FPPS (2), have just one aromatic amino acid (phenylalanine) at the 4th position and a cysteine or alanine (or serine in case of *P. falciparum*), respectively, at the 5th position upstream of the FARM (Figure S1B) (1). In order to investigate the role of the region around the FARM in the product specificity, we replaced the six residues sequence between position 325 and 330 of TgFPPS-141 (³²⁵SCFLVM³³⁰), with the corresponding sequences of two eukaryotic FPPS, from *Arabidopsis thaliana* (³²⁵TYFLVL³³⁰), or *Saccharomyces cerevisiae* (³²⁵TYFLVA³³⁰), respectively (27). The mutant enzymes were overexpressed in *E. coli*, purified, and assayed for enzyme activity. Their *K_m* for DMAPP was not affected as compared to TgFPPS-141, while only the enzyme with the *S. cerevisiae* sequence had lower *V_{max}* (Table 1). Both mutants behave as FPPSs generating GPP and FPP when DMAPP was used as co-substrate (Figure 3) but were unable to elongate FPP to GGPP (Figure 4). This mutagenesis analysis indicates that the bifunctional role of TgFPPS is also determined in this CLD.

In order to determine whether the 4th or the 5th amino acids upstream of the FARM played a significant role in the activity or in the determination of the chain length of the products of TgFPPS, the normal (TgFPPS-141) and mutant enzymes, in which an alanine, threonine or tyrosine replaced the cysteine at the 5th position, or a glycine replaced the phenylalanine at the 4th position were also overexpressed in *E. coli*, purified and assayed for activity. While substitution of the cysteine (³²⁶Cys) by either alanine or threonine did not prevent the

formation of GGPP using either DMAPP (Figure 3) or FPP (Figure 4) as substrate, substitution by the bulky amino acid tyrosine prevented the generation of GGPP in agreement with the postulated role of the amino acid in the 5th position upstream of the FARM region in limiting the product length of FPPSs (13). In either case the V_{max} was decreased by about 30% while only the K_m for DMAPP was significantly decreased when tyrosine replaced ³²⁶Cys, in agreement with its conversion into an exclusive FPPS. Conversion to an exclusive FPPS in the case of TgFPPS-171 also resulted in a similar reduction in the K_m for DMAPP (Table 1). Substitution of the bulky phenylalanine at the 4th position by a glycine greatly decreased the activity with either DMAPP (Figure 3) or FPP (Figure 4) as co-substrate. Presumably, this mutation causes additional structural changes that affect the affinity of the enzyme for the co-substrate.

Conversion of TgFPPS Into a Synthase Capable of Producing GGPP

To further confirm the role of the region around the FARM in the product specificity, we replaced the sequence between positions 325 and 330 of TgFPPS-141 (³²⁵SFLVM³³⁰), with the corresponding sequences of two eukaryotic GGPPS, from *Homo sapiens* (³²⁵NASLLI³³⁰), or *Saccharomyces cerevisiae* (³²⁵NSSLLI³³⁰), respectively (Figure S1B). The mutant enzymes were overexpressed in *E. coli*, purified, and assayed for enzyme activity. The mutants were only able to produce GPP when DMAPP was used as co-substrate (Figure 3) with very low activity (Table 1) while they have good activity with FPP as co-substrate (Table 1) and were able to generate GGPP (Figure 4). A double mutation of the 4th and 5th amino acid upstream of the FARM (³²⁶CF³²⁷ to ³²⁶AA³²⁷) also converted TgFPPS into a straight GGPPS (Figures 3 and 4). The analysis of these mutants further supports that the unique combination of 4th and 5th amino acid upstream of the FARM is responsible for the bifunctionality of TgFPPS.

Phylogenetic analysis of TgFPPS

Taking into account the bifunctionality of the TgFPPS it was of interest to analyze its phylogenetic profile to see if the sequence grouped with other FPPSs or with GGPPSs. We used the *E. coli* FPPS protein sequence to search the NCBI database for the E-isoprenyl diphosphate synthases, FPPS, GGPPS and C₃₀-C₅₀ PP (HEPPS, OPPS, DPPS) (Table S2). We also searched for TgFPPS orthologs using the OrthoMCL database. We recovered orthologs from *P. falciparum*, *C. hominis* and *C. parvum*. We selected a set of sequences which included archaeobacteria, eubacteria and eukaryotic species.

MEME analysis of the selected sequences identified the seven conserved domains of FPPS and GGPPS and also the FARM and SARM. Some peculiarities are also highlighted after analysis of the alignments (Figure S2). Amino acid residues of domains IV and V are mostly conserved in type I FPPSs. Type I FPPS and most type II GGPPSs share domain VII, but different amino acids are found in the sequences of most type II FPPSs, type I GGPPSs, type III GGPPS, C₃₀-C₅₀ PP, and bifunctional enzymes.

The phylogeny of the prenyl synthase was constructed using the MEGA5 software and it is shown in Figure 5. We built 100 phylogenetic trees using 47 sequences, which included FPPS, GGPPSs and C₃₀-C₅₀ PP synthases. The final tree displays bootstrap values 50% or

higher. Most evolutionary relationships are supported by significant bootstrap values (above 75%). Lower bootstrap values can be attributed to an incomplete phylogenetic inventory (missing species) and sequence divergence among species. Our phylogenetic tree shows that the Apicomplexan putative FPPSs cluster with type I FPPS enzymes. Other putative isoprenoid biosynthetic enzymes found in Apicomplexa cluster with long chain prenyl synthases.

DISCUSSION

In most organisms the short chain prenyl synthases GPP, FPP and GGPP synthases elongate DMAPP to products of 10-C, 15-C, and 20-C, respectively. FPPS usually synthesizes FPP (the 15-C product) from DMAPP (5-C product) or GPP (C₁₀ product) and GGPPS uses FPP as co-substrate to synthesize GGPP (20-C product). It has been demonstrated that the chain length of the final product does not exceed the limit determined by the specificity of the enzyme (11). Previous studies by Ohnuma et al. (13), using random chemical mutagenesis aimed at altering the chain-length selectivity of prenyl synthases, have shown that *Bacillus stearothermophilus* FPP synthase could be converted into a GGPPS by replacing only one amino acid, tyrosine 81 of the FPP synthase, situated at the 5th position upstream of the first aspartic rich domain (FARM). Similar results were reported by Tarshis et al. (28), who mutated the phenylalanine 112 of avian FPP synthase located at the 5th position upstream of the FARM to alanine and detected GGPP formation. Ohnuma et al. (13) studies showed that the region around the FARM determines the product specificity of FPP and GGPP synthases. Specifically when both the 4th and 5th amino acids upstream of the FARM have large bulky aromatic residues like phenylalanine and tyrosine, the final product that FPPSs produce is FPP. This is apparently because these bulky amino acids would block further condensation of compounds greater than 15-C. It was proposed that the region between the 4th and 5th amino acid upstream of the FARM all the way to the end of the FARM domain determines the final products of all prenyl diphosphate synthases and it was named as the chain-length determination (CLD) region (11). Interestingly, *T. gondii* FPPS (1), as the *P. vivax* FPPS (*Pv*GGPPS) (2), is able to synthesize both FPP and GGPP. *T. gondii* apparently relies on TgFPPS for making both products because no other gene coding for an apparent GGPPS has been detected in the genome of this parasite. When comparing the TgFPPS amino acid sequence around the FARM with other FPPS and GGPPS several interesting features were evident. The first was that the 4th amino acid upstream of the FARM is phenylalanine and the 5th is cysteine (1). This is different from any of the patterns previously described: (1) both are aromatic amino acids (type I FPPSs); (2) neither of them is aromatic amino acid (type II and type III GGPPSs); (3) only the fifth amino acid is aromatic (type II FPPSs and type I GGPPs). In contrast, TgFPPS (and also *Pv*GGPPS) has just one aromatic amino acid at the fourth position upstream of the FARM instead of at the fifth position. This organization is unusual although similar to the organization of other Apicomplexan enzymes (Figure S1B). The *P. vivax* enzyme has alanine at the 5th position and phenylalanine at the 4th position upstream to the FARM region (2) (Figure S1B). This different combination of fourth and fifth amino acids may provide this enzyme its bifunctionality.

In this work, we describe that the replacement of the six-residue sequence upstream of the FARM region of TgFPPS by the amino acid sequences present in *A. thaliana* or *S. cerevisiae* FPPSs or the replacement of the 5th amino acid (cysteine) upstream of the FARM region by a tyrosine caused a change of the enzyme activity so that it is converted into an exclusive FPP synthase. On the other hand, the conversion of *T. gondii* FPPS into a GGPPS was possible when the six-amino acid residues upstream of the FARM region were replaced with the residues found in the same position of *Homo sapiens* or *S. cerevisiae* GGPPSs. The replacement of amino acids at positions 4th and 5th upstream of the FARM region by alanine also determined the conversion of TgFPPS into a GGPPS. Unexpectedly, TgFPPS produced more FPP and very little GGPP when the N-terminal 171 amino acids were deleted.

Taken together, our results indicate that the composition of the sequence upstream of the FARM region of TgFPPS and the N-terminal region of the enzyme determine the length of the products of this enzymatic reaction. Although previous work with the *S. cerevisiae* type III GGPPS has shown that the N-terminal of this GGPP protrude from the helix core into the other subunit and contribute to the tight dimer formation, and that deletion of some of these amino acids caused the dissociation of dimer into monomers and abolished the activity of the enzyme (29), to our knowledge, this is the first report to describe that the N-terminal region of the enzyme has a role in isoprenoid elongation in prenyl synthases. It is interesting to note that the predicted mitochondrial targeting sequence suggests that the targeting signal is processed at approximately 150 amino acids. However, we found the most important differences when we used a truncated version lacking 171 amino acids, thus suggesting that the changes observed are probably due to this extra region of 20 amino acids that is maintained after processing.

We report for the first time the phylogenetic analysis of *E*-isoprenyl diphosphate synthases by a multigenome comparison with Apicomplexa. Our biochemical evidence indicates that the TgFPPS is bifunctional, catalyzing the synthesis of both FPP and GGPP (1). However, the phylogenetic analysis of the TgFPPS classifies it as a Type I FPPS. Moreover, the *Cryptosporidium spp.* and *Plasmodium spp.* FPPSs also cluster with Type I FPPS. In addition, other putative Apicomplexan prenyl synthases, including one from *T. gondii*, cluster with long chain prenyl synthases, which suggests that these enzymes are involved in the synthesis of 30–50-C isoprenoid unit compounds. This is important to mention since the bifunctionality of the *P. vivax* FPPS has been demonstrated (2) and we provide further evidence that TgFPPS synthesizes both FPP and GGPP implying that other apicomplexan FPPSs could also be specific in the synthesis of short chain isoprenoid units.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Funding: This work was supported by National Institutes of Health grants AI068467 and AI102254 (to SNJM).

We thank Allysa Smith, Samantha Lie Tjauw and Leena Padmanabhan for technical help.

ABBREVIATIONS

FPPS	farnesyl diphosphate synthase
GGPPS	geranylgeranyl diphosphate synthase
DMAPP	dimethylallyl diphosphate
IPP	isopentenyl diphosphate
GPP	geranyl diphosphate
FPP	farnesyl diphosphate
FOH	farnesol
GGPP	geranylgeranyl diphosphate
GGOH	geranylgeraniol
FARM	first aspartic rich motif
SARM	second aspartic rich motif
CLD	chain length determination
TLC	thin layer chromatography

References

1. Ling Y, Li ZH, Miranda K, Oldfield E, Moreno SN. The farnesyl-diphosphate/geranylgeranyl-diphosphate synthase of *Toxoplasma gondii* is a bifunctional enzyme and a molecular target of bisphosphonates. *J Biol Chem.* 2007; 282:30804–30816. [PubMed: 17724033]
2. Artz JD, Wernimont AK, Dunford JE, Schapira M, Dong A, Zhao Y, Lew J, Russell RG, Ebetino FH, Oppermann U, Hui R. Molecular characterization of a novel geranylgeranyl pyrophosphate synthase from *Plasmodium* parasites. *J Biol Chem.* 2011; 286:3315–3322. [PubMed: 21084289]
3. Martin MB, Grimley JS, Lewis JC, Heath HT 3rd, Bailey BN, Kendrick H, Yardley V, Caldera A, Lira R, Urbina JA, Moreno SN, Docampo R, Croft SL, Oldfield E. Bisphosphonates inhibit the growth of *Trypanosoma brucei*, *Trypanosoma cruzi*, *Leishmania donovani*, *Toxoplasma gondii*, and *Plasmodium falciparum*: a potential route to chemotherapy. *J Med Chem.* 2001; 44:909–916. [PubMed: 11300872]
4. Yardley V, Khan AA, Martin MB, Slifer TR, Araujo FG, Moreno SN, Docampo R, Croft SL, Oldfield E. In vivo activities of farnesyl pyrophosphate synthase inhibitors against *Leishmania donovani* and *Toxoplasma gondii*. *Antimicrob Agents Chemother.* 2002; 46:929–931. [PubMed: 11850291]
5. Ling Y, Sahota G, Odeh S, Chan JM, Araujo FG, Moreno SN, Oldfield E. Bisphosphonate inhibitors of *Toxoplasma gondii* growth: in vitro, QSAR, and in vivo investigations. *J Med Chem.* 2005; 48:3130–3140. [PubMed: 15857119]
6. Rosso VS, Szajnman SH, Malayil L, Galizzi M, Moreno SN, Docampo R, Rodriguez JB. Synthesis and biological evaluation of new 2-alkylaminoethyl-1,1-bisphosphonic acids against *Trypanosoma cruzi* and *Toxoplasma gondii* targeting farnesyl diphosphate synthase. *Bioorg Med Chem.* 2011; 19:2211–2217. [PubMed: 21419634]
7. Szajnman SH, Garcia Linares GE, Li ZH, Jiang C, Galizzi M, Bontempi EJ, Ferella M, Moreno SN, Docampo R, Rodriguez JB. Synthesis and biological evaluation of 2-alkylaminoethyl-1,1-bisphosphonic acids against *Trypanosoma cruzi* and *Toxoplasma gondii* targeting farnesyl diphosphate synthase. *Bioorg Med Chem.* 2008; 16:3283–3290. [PubMed: 18096393]
8. Szajnman SH, Rosso VS, Malayil L, Smith A, Moreno SN, Docampo R, Rodriguez JB. 1-(Fluoroalkylidene)-1,1-bisphosphonic acids are potent and selective inhibitors of the enzymatic

- activity of *Toxoplasma gondii* farnesyl pyrophosphate synthase. *Org Biomol Chem*. 2012; 10:1424–1433. [PubMed: 22215028]
9. Singh AP, Zhang Y, No JH, Docampo R, Nussenzweig V, Oldfield E. Lipophilic bisphosphonates are potent inhibitors of *Plasmodium* liver-stage growth. *Antimicrobial agents and chemotherapy*. 2010; 54:2987–2993. [PubMed: 20457823]
10. Gabelli SB, McLellan JS, Montalvetti A, Oldfield E, Docampo R, Amzel LM. Structure and mechanism of the farnesyl diphosphate synthase from *Trypanosoma cruzi*: implications for drug design. *Proteins*. 2006; 62:80–88. [PubMed: 16288456]
11. Wang K, Ohnuma S. Chain-length determination mechanism of isoprenyl diphosphate synthases and implications for molecular evolution. *Trends Biochem Sci*. 1999; 24:445–451. [PubMed: 10542413]
12. Ohnuma S, Nakazawa T, Hemmi H, Hallberg AM, Koyama T, Ogura K, Nishino T. Conversion from farnesyl diphosphate synthase to geranylgeranyl diphosphate synthase by random chemical mutagenesis. *J Biol Chem*. 1996; 271:10087–10095. [PubMed: 8626566]
13. Ohnuma S, Narita K, Nakazawa T, Ishida C, Takeuchi Y, Ohto C, Nishino T. A role of the amino acid residue located on the fifth position before the first aspartate-rich motif of farnesyl diphosphate synthase on determination of the final product. *J Biol Chem*. 1996; 271:30748–30754. [PubMed: 8940054]
14. Ohnuma S, Hirooka K, Hemmi H, Ishida C, Ohto C, Nishino T. Conversion of product specificity of archaeobacterial geranylgeranyl-diphosphate synthase. Identification of essential amino acid residues for chain length determination of prenyltransferase reaction. *J Biol Chem*. 1996; 271:18831–18837. [PubMed: 8702542]
15. Hemmi H, Noike M, Nakayama T, Nishino T. An alternative mechanism of product chain-length determination in type III geranylgeranyl diphosphate synthase. *Eur J Biochem*. 2003; 270:2186–2194. [PubMed: 12752438]
16. Noike M, Katagiri T, Nakayama T, Koyama T, Nishino T, Hemmi H. The product chain length determination mechanism of type II geranylgeranyl diphosphate synthase requires subunit interaction. *Febs J*. 2008; 275:3921–3933. [PubMed: 18616462]
17. Aripirala S, Szajnman SH, Jakoncic J, Rodriguez JB, Docampo R, Gabelli SB, Amzel LM. Design, Synthesis, Calorimetry and Crystallographic analysis of 2-Alkylaminoethyl-1,1-Bisphosphonates as inhibitors of *Trypanosoma cruzi* Farnesyl Diphosphate Synthase. *J Med Chem*. 2012; 55:6445–6454. [PubMed: 22715997]
18. Miranda K, Pace DA, Cintron R, Rodrigues JC, Fang J, Smith A, Rohloff P, Coelho E, de Haas F, de Souza W, Coppens I, Sibley LD, Moreno SN. Characterization of a novel organelle in *Toxoplasma gondii* with similar composition and function to the plant vacuole. *Mol Microbiol*. 2010; 76:1358–1375. [PubMed: 20398214]
19. Montalvetti A, Bailey BN, Martin MB, Severin GW, Oldfield E, Docampo R. Bisphosphonates are potent inhibitors of *Trypanosoma cruzi* farnesyl pyrophosphate synthase. *J Biol Chem*. 2001; 276:33930–33937. [PubMed: 11435429]
20. Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res*. 1994; 22:4673–4680. [PubMed: 7984417]
21. Thompson JD, Plewniak F, Poch O. BALiBASE: a benchmark alignment database for the evaluation of multiple alignment programs. *Bioinformatics*. 1999; 15:87–88. [PubMed: 10068696]
22. Bailey TL, Elkan C. Fitting a mixture model by expectation maximization to discover motifs in biopolymers. *Proc Int Conf Intell Syst Mol Biol*. 1994; 2:28–36. [PubMed: 7584402]
23. Nicholas KB, Nicholas HB, Deerfield DW. GeneDoc: Analysis and Visualization of genetic variation. *EMBNW NEWS*. 1997; 4:14.
24. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol*. 2011; 28:2731–2739. [PubMed: 21546353]
25. Eck RV, Dayhoff MO. Evolution of the Structure of Ferredoxin Based on Living Relics of Primitive Amino Acid Sequences. *Science*. 1966; 152:363–366. [PubMed: 17775169]

26. Saitou N, Nei M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol.* 1987; 4:406–425. [PubMed: 3447015]
27. Ohnuma S, Hirooka K, Ohto C, Nishino T. Conversion from archaeal geranylgeranyl diphosphate synthase to farnesyl diphosphate synthase. Two amino acids before the first aspartate-rich motif solely determine eukaryotic farnesyl diphosphate synthase activity. *J Biol Chem.* 1997; 272:5192–5198. [PubMed: 9030588]
28. Tarshis LC, Proteau PJ, Kellogg BA, Sacchettini JC, Poulter CD. Regulation of product chain length by isoprenyl diphosphate synthases. *Proc Natl Acad Sci U S A.* 1996; 93:15018–15023. [PubMed: 8986756]
29. Chang TH, Guo RT, Ko TP, Wang AH, Liang PH. Crystal structure of type-III geranylgeranyl pyrophosphate synthase from *Saccharomyces cerevisiae* and the mechanism of product chain length determination. *J Biol Chem.* 2006; 281:14991–15000. [PubMed: 16554305]

Figure 1.

Western blot analysis of affinity purified FLAG-tagged native, His-tagged truncated recombinant, and endogenous TgFPPS. (A) Purified native FLAG-tagged TgFPPS expressed in *T. gondii* (*Native*) has a size similar to the two N-terminal truncated versions without 171 (TgFPPS-171) or 141 (TgFPPS-141) amino acids. TgFPPS-FLAG was expressed in *T. gondii*, and purified by affinity column chromatography. His-tagged TgFPPS-141 and TgFPPS-171 were expressed in *E. coli* and purified using Nickel columns. The purified proteins were subjected to SDS-PAGE on 10% polyacrylamide gels, transferred to nylon membranes, and probed with affinity purified polyclonal antibody against TgFPPS (1:2000). (B) Western blot analysis of the endogenous TgFPPS from total lysates of *T. gondii* RH tachyzoites. After SDS-PAGE, and transfer to nylon membranes as indicated above, the membranes were probed with the affinity purified polyclonal antibody against TgFPPS (1:2000). Arrow shows the band corresponding to the endogenous protein (~58 kDa).

Figure 2.

Reverse phase thin layer chromatography of the alcohols obtained by enzymatic hydrolysis of the products formed by the mutated TgFPPSs. Truncated versions of His-tagged TgFPPSs were expressed in *E. coli*, purified using nickel affinity chromatography and incubated (40 ng of each enzyme purified from *E. coli* or 4 ng of FLAG-TgFPPS purified from *T. gondii* lysates) with [$1\text{-}^{14}\text{C}$]IPP and the different allylic substrates (DMAPP, GFP, FPP) indicated, at 37 °C for 1 hour. Other conditions are as described under Materials and Methods. GOH, geraniol, FOH, farnesol, GGOH, geranylgeraniol. (A) Comparison of the products formed by the native TgFPPS purified from a *T. gondii* lysate as described under Materials and Methods. The *Trypanosoma cruzi* FPPS (TcFPPS) (19) was used as control. Note that there was only 4 ng of FLAG-TgFPPS used in this study while there were 40 ng of TcFPPS. (B) Products formed by the recombinant enzyme lacking 103 amino acids, expressed and purified from *E. coli*. IPP was used at saturating concentrations for all the assays and the second substrate is indicated at the top. (C) Products formed by the enzyme TgFPPS-141 in the presence of saturating concentrations of IPP plus one of the three substrates indicated at the top. (D) Products formed by the enzyme TgFPPS-171 in the presence of saturating concentrations of IPP plus one of the three substrates indicated at the top.

Figure 3.

Conversion of TgFPPS to a FPPS by site directed mutagenesis. Reverse phase thin layer chromatography of the alcohols obtained by enzymatic hydrolysis of the products formed by mutated TgFPPS-141. The His-tagged TgFPPS mutants were constructed as described under Materials and Methods. Each of the mutants was expressed in *E. coli* and purified using nickel affinity columns. The enzyme reactions were performed by incubating 40 ng of each protein with DMAPP and IPP as substrates at 37 °C for 1 hour. Other conditions are as described under Materials and Methods.

Figure 4. Conversion of TgFPPS to a GGPPS by site-directed mutagenesis. Reverse phase thin layer chromatography of the alcohols obtained by enzymatic hydrolysis of the products formed by mutated TgFPPS-141 in the presence of FPP and IPP. The TgFPPS mutants were those described in Figure 3. The enzyme reactions were performed by incubating 40 ng of proteins with FPP and IPP as substrates at 37 °C for 1 hour. Other conditions are as described under Materials and Methods.

Figure 5.

Phylogenetic tree of prenyl synthases. Apicomplexan prenyl synthases cluster in two main groups: FPPSs belong to the Type I FPPS group, whereas other Apicomplexan synthases cluster with long chain prenyl synthases. Forty-seven amino acid sequences were analyzed. The alignments are presented in Figure S2 and the sequence identification number in Table S2. Bootstrap values higher than 50% are shown in the tree. Bars, 0.5 changes per amino acid position.

Table 1

Enzymatic properties of TgFPPS

Mutation ^a	DMAPP		FPP	
	K _m (μM)	V _{max} (μmol/min × mg protein)	K _m (μM)	V _{max} (μmol/min × mg protein)
WT (-141)	18.7 ± 7.0	31.8 ± 5.5	2.3 ± 1.1	4.8 ± 0.4
WT (-171)	7.6 ± 2.0	19.6 ± 1.6	NA	NA
³²⁵ SCFLVM ³³⁰ - ³²⁵ TYFLVA ³³⁰	18.2 ± 3.1	14.4 ± 0.5	NA	NA
³²⁵ SCFLVM ³³⁰ - ³²⁵ TYFLVL ³³⁰	19.0 ± 3.8	37.1 ± 0.3	NA	NA
C326A	16.2 ± 4.7	22.1 ± 3.4	ND	2.7 ± 1.2
C326Y	9.3 ± 0.3	20.4 ± 0.3	NA	NA
³²⁵ SCFLVM ³³⁰ - ³²⁵ NASLLI ³³⁰	5.6 ± 0.3	0.3 ± 0.2	4.7 ± 0.6	10.7 ± 0.2
³²⁵ SCFLVM ³³⁰ - ³²⁵ NSSLLI ³³⁰	44.9 ± 36.9	1.5 ± 0.8	7.2 ± 0.6	6.1 ± 0.5

^aThe recombinant proteins containing these mutated amino acids were expressed and purified and their activities measured as described in Materials and Methods.

DMAPP: Dimethyl allyl diphosphate; FPP: Farnesyl diphosphate

NA: no activity

ND: not determined