The Crystal Structure of (S)-3-O-Geranylgeranylglycerol Phosphate Synthase Reveals an Ancient Fold for an Ancient Enzyme
Jian Payandeh*, Masahiro Fujihashi†, Wanda Gillon*, Emil F. Pai‡
From the Departments of *Medical Biophysics, †Biochemistry, and ‡Molecular & Medical Genetics, University of Toronto, Toronto, Ontario, M5G 2M9
Running Title: Structure of Geranylgeranylglycerol Phosphate Synthase
Address correspondence to: Emil F. Pai, Division of Cancer Genomics & Proteomics, Ontario Cancer Institute, Princess Margaret Hospital, 610 University Avenue, Toronto, Ontario, M5G 2M9, Tel. 416 946 2968; Fax. 416 946 6529; E-mail: pai@hera.med.utoronto.ca.
†Present address: Department of Chemistry, Graduate School of Science, Kyoto University, Sakyo-ku, Kyoto 606-8502, Japan.

We report crystal structures of the citrate and sn-glycerol-1-phosphate complexes of (S)-3-O-geranylgeranylglycerol phosphate synthase from Archaeoglobus fulgidus (AfGGGPS) at 1.55 Å and 2.0 Å resolution, respectively. AfGGGPS is an enzyme that performs the committed step in archaeal lipid biosynthesis and it presents the first TIM-barrel structure with a prenyltransferase function. Our studies provide insight into the catalytic mechanism of AfGGGPS and demonstrate how it selects for the sn-G1P isomer. The replacement of “Helix 3” by a ‘strand’ in AfGGGPS, a novel modification to the canonical TIM-barrel fold, suggests a model of enzyme adaptation that involves a “greasy slide” and a “swinging door”. We propose functions for the homologous PcrB proteins, which are conserved in a subset of pathogenic bacteria, as either prenyltransferases or being involved in lipoteichoic acid biosynthesis. Sequence and structural comparisons lead us to postulate an early evolutionary history for AfGGGPS, which may highlight its role in the emergence of Archaea.

The membrane lipids found in Archaea are a defining characteristic of this domain of life (1). These lipids are based on a core architecture where branched-chain saturated hydrocarbons are connected to glycerol through ether linkages (2, 3). In hyperthermophiles, two dihytanylglycerol units are often linked covalently through their hydrocarbon tails to form tetraether lipids that completely span the membrane. In addition, archaeal membrane lipids have three general characteristics that distinguish them from their bacterial and eukaryotic counterparts (4). First, the phospholipid backbone is built upon the opposite glycerol stereoisomer, sn-glycerol-1-phosphate (G1P), not the sn-glycerol-3-phosphate (G3P) backbone found in bacteria and eukaryotes. Second, the hydrophobic chains are isoprenoid derivatives instead of fatty acids. Third, the isoprenoid chains are bound to G1P through ether, not ester, linkages. Of these traits, the glycerol phosphate stereochemistry is the most distinctive because ether-linked lipids are known to exist in some eukaryotes and bacteria (5, 6), and phospholipid fatty acids have recently been described in Archaea (7). To date, however, there is no known exception to the G1P backbone stereochemistry of archaeal lipids or to the G3P backbone stereochemistry found in bacterial and eukaryotic lipids.

The biosynthesis of archaeal membrane lipids is schematically illustrated in Figure 1. In brief, dimethylallyl diphosphate (DMAPP) and its isomer isopentenyl diphosphate (IPP) are synthesized by a mevalonate-like pathway (2, 8). Long isoprenoid chains are produced from these five-carbon precursors by consecutive condensations through the action of a prenyl diphosphate synthase. The committed step in archaeal lipid synthesis occurs with the formation of an ether linkage between G1P and an isoprenoid diphosphate, usually geranylgeranyl diphosphate (GGPP). Separate enzymes catalyze the sequential transfer of isoprenoid units onto the C3 and C2 hydroxyls of G1P, respectively (9). The enzyme catalyzing the first reaction, (S)-3-O-geranylgeranylglycerol phosphate synthase (GGGPS), is a cytosolic protein that has previously been purified and enzymatically characterized from Methanobacterium thermoautotrophicum (9-11) and Thermoplasma acidophilum (12). The enzyme catalyzing the
second prenyltransfer, (S)-2,3-di-O-geranylgeranylglyceryl phosphate synthase (DGGGPS), is an integral membrane protein which has recently been cloned and purified from Sulfolobus solfataricus (13). In ensuing reactions the isoprenoid double bonds are reduced and various polar headgroups may be attached to the lipid. Synthesis of the aforementioned tetraether lipids still remains largely unexplored.

GGGPS, the enzyme catalyzing the committed step in archaeal lipid synthesis, displays three major points of interest. First, it represents the branch point for membrane lipid biosynthesis in Archaea and imparts the unique G1P stereochemistry of the lipid backbone. Second, GGGPS belongs to a class of prenyltransferases that catalyzes the transfer of isoprenoid groups onto nonisoprenoid acceptors; this class includes protein prenyltransferases, DMAPP-tRNA transferase, DMAPP-AMP transferase, dimethylallyltryptophan synthase, and the recently characterized aromatic-prenyltransferases (11, 14). Finally, GGGPS is a homologue of PcrB, a protein of unknown function that is conserved in some Gram-positive bacteria, including the human pathogens Bacillus anthracis, Staphylococcus aureus, and Listeria monocytogenes (8, 12, 15).

To gain more insight into the unique stereochemistry of the unusual reaction catalyzed by GGGPS, and to shed light onto the function of PcrB, we have determined crystal structures of GGGPS from the sulfate-reducing hyperthermophile Archaeoglobus fulgidus in complex with G1P and citrate. The present work reveals 1) an ancient fold for this deeply rooted enzyme and 2) establishes, for the first time, a prenyltransferase function on a TIM-barrel scaffold. Our structures further unveil 3) how the “standard phosphate-binding motif” is utilized in concert with conserved active site residues to select for the G1P substrate, 4) a long apolar cavity that bears the hallmarks of a hydrocarbon ruler, and 5) a unique change to the classical (βα)₈-TIM-barrel fold which prompts us to present a hypothesis for the functional consequences of this adaptation.

MATERIALS & METHODS

Cloning, Purification, and Crystallization - The coding sequence for AfGGGPS was amplified from genomic DNA (ATCC 49558D) and cloned into the NdeI and BamHI sites of the pET-15b expression vector. Native protein was expressed in E. coli BL21(DE3), grown in LB media at 37°C and induced at OD₆₀₀ ~0.8 with 1 mM IPTG for 8-12 hrs. Selenomethionine (SeMet) protein was expressed similarly, but in E. coli B834(DE3) and in M9 minimal media supplemented with 40 mg/L SeMet. Cells were broken by sonication in Buffer A (50 mM Tris pH 8.0, 100 mM NaCl, 5 mM imidazole, 10 mM β-mercaptoethanol) supplemented with 16 μg/mL benzamidine, 10 μg/mL aprotinin, 10 μg/mL leupeptin, and 1 mM phenylmethyl sulfonyl fluoride. Clarified lysate was passed over Ni-NTA agarose (Qiagen), the resin was washed extensively with Buffer A, and proteins were eluted in Buffer A containing 500 mM imidazole. The His₆-tag was removed by thrombin digestion (~12 hrs) in the presence of 3 mM CaCl₂. Proteins were passed over a Superdex 200 column (Amersham) in Buffer A without imidazole, and peak fractions containing AfGGGPS were pooled. Proteins were further passed over a second Ni-column and the flow-through containing AfGGGPS was concentrated using a centrifugal filter device (Millipore, 10K MWCO). Finally, the buffer was exchanged to 20 mM Tris pH 8.0, 100 mM NaCl, 1 mM Tris(2-carboxyethyl)phosphine hydrochloride by three rounds of concentration and dilution. This procedure typically yields ~30 mg of AfGGGPS per litre of culture.

AfGGGPS crystallizes readily in the R₃/H₃ space group. Crystals were grown at 22°C by the hanging drop method and micro-seeding was employed for consistent crystal quality. Equal volumes of protein (20-30 mg/mL) and seed solution (100 mM citrate pH 6.0, 10 mM MgCl₂, 16% polyethylene glycol (PEG) 4000) were mixed over a reservoir solution of 100 mM citrate pH 6.0, 10 mM MgCl₂, 4-10% PEG4000. For G1P-AfGGGPS soaks, 5-10 mM rac-G1P (Sigma-Aldrich) was added to solutions during cryoprotection. In attempts to co-crystallize the enzyme with GGPP, the addition of 80 mM n-octyl-β-D-glucopyranoside or 0.8 mM n-dodecyl-β-D-maltopyranoside to the protein solution was
found to produce crystals in the $P3_1$ space group, in otherwise identical crystallization set-ups. Cryoprotection was achieved in mother liquor supplemented with glycerol, MPD, or ethylene glycol and crystals were flash-frozen in a cold nitrogen stream.

Data Collection and Structure Determination – Typically 180° of data were collected from randomly oriented crystals maintained at 100K in a cold nitrogen stream. A 3-wavelength multiple anomalous dispersion (MAD) data set was collected at the NSLS beamline X6A on a SeMet-substituted crystal. Native data sets were subsequently collected at the APS BioCARS 14BM-D. All X-ray data were processed using the DENZO/SCALEPACK software (16) and reduced, when required, using the program TRUNCATE (17). The AfGGGPS structure was initially determined using SOLVE (18). After applying RESOLVE (19), ARP/wARP (20) built the majority of the protein structure. All subsequent structures were solved by molecular replacement methods using CNS (21) and standard refinement protocols were implemented with CNS. Guided by composite omit maps, water and substrate molecules were added at late stages of refinement. The positions of all atoms were checked manually in O (22). The quality of the structural models was assessed by PROCHECK (23).

Although discussed in the main text, the AfGGGPS structure in the $P3_1$ space group has not been reported herein due to its relatively limited resolution (~3.0 Å). Therefore, only the data collection and refinement statistics for the citrate-bound and G1P-soaked AfGGGPS structures are presented in Table 1.

Figures and Modeling – All structural figures were prepared from the G1P-soaked AfGGGPS structure using the programs MOLSCRIPT (24) and RASTER3D (25). The electrostatic surface representations shown in Fig. 3A-B were calculated using SPock (26).

Prior to modeling (Fig. 3D), G1P, MPD, and two water molecules were removed from the G1P-soaked structure. A Mg$^{2+}$-GGPP substrate was manually placed in the active site cleft. Annealing and energy minimization protocols were run in CNS (21), keeping the positions of all experimentally determined atoms fixed. G1P was added back to the structure with its initial, i.e. experimentally determined coordinates.

GGGPS Assay – To avoid the use of radioactivity (9-12), we implemented an enzyme-coupled assay to detect the free pyrophosphate liberated by the GGGPS reaction. The P$_{Per}$ Pyrophosphate Assay (Molecular Probes) was performed as described by the manufacturer in the “Background Management” setting. Absorbance was monitored at 565 nm and all reactions were incubated at 37°C in between time points. Different AfGGGPS concentrations were prepared in 50 mM Tris pH 7.5 at fixed substrate concentrations: 25 μM rac-G1P, 25 μM GGPP (Sigma-Aldrich), and 1 mM MgCl$_2$. The protein used as a negative control, orotidine 5′-monophosphate decarboxylase from M. thermoautotrophicum (MtODCase) (27), was purified as described above for AfGGGPS.

RESULTS AND DISCUSSION

Overall Structure – Unless otherwise stated, we will be describing the G1P-soaked AfGGGPS crystal structure. The GGGPS enzymes can be divided into two groups based on phylogenetic analysis (8, 12). The Group I GGGPS from A. fulgidus (AfGGGPS) presents a dimeric structure (Fig. 2B-C). This is consistent with our results from gel filtration and the recent characterization of a Group II GGGPS from T. acidophilum (TaGGGPS) (12). The structure of AfGGGPS in the $P3_1$ space group confirms its dimeric nature, as does the crystal structure of its PcrB homologue from B. subtilis (BsPcrB; PDB 1VIZ). Considering these results, it is not immediately obvious from sequence alignments (Fig. 2A) how the Group II GGGPS from M. thermoautotrophicum (MtGGGPS) would form its reported higher-order oligomer (10, 11). We conclude that AfGGGPS and most of its sequence homologues are dimeric proteins (Fig. 6A). The dimensions of an AfGGGPS dimer are approximately 85 x 50 x 40 Å (Fig. 2B-C), and each monomer buries ~1,400 Å$^2$, or ~13% of its total accessible surface area.

Based on the unique three-dimensional structures of the cis-, trans-, protein-, and aromatic-prenyltransferase families (14, 28), it is surprising to see AfGGGPS exhibit a TIM-barrel
AfgGGGPS displays a central 8-stranded parallel β-barrel (β1-β8) with a tightly packed hydrophobic core, which is decorated on its periphery by α-helices (α1-α8). A striking feature of the AfgGGGPS structure is that “Helix 3” has been replaced by a ‘strand’. Although it is without any regular secondary structure, we will refer to this structural element as α3* (Fig. 2B). To the best of our knowledge, the location of this type of modification has not been previously observed in any TIM-barrel protein.

The residues preceding β1 form a “plug” (α0) at the bottom of the AfgGGGPS barrel (Fig. 2B). Most archaeal homologues have an additional N-terminal extension of 10-20 residues which tends to be enriched in basic and/or hydrophobic amino acids (Fig. 2A). In light of the AfgGGGPS structure, we suggest that these N-terminal extensions may promote an association with the membrane by directly contacting membrane lipids, or through an interaction with other proteins involved in lipid biosynthesis. Given the overall acidic nature of AfgGGGPS (theoretical pl ~ 5.0), it seems significant that the only appreciable clustering of basic residues on the electrostatic surface is also found within this region (Fig. 3A); this makes an interaction with the negatively charged phosphate groups of the membrane lipids plausible.

The AfgGGGPS active site lies at the C-terminal end of its β-barrel. This is evident from sequence conservation and the structures of AfgGGGPS in complex with various ligands (Fig. 2A-C). The presence of citrate-ions dramatically increased the diffraction limit of the AfgGGGPS crystals, and a molecule of citrate was subsequently identified within the active site. Coincidentally, soaking these crystals in G1P-containing solutions will replace citrate at this location.

The two active sites within the AfgGGGPS dimer do not behave in an equivalent fashion, irrespective of the space group or ligand(s) bound. For example, inspection of the electron density maps in the G1P-soaked structure clearly shows a molecule of G1P bound in one active site, with a ‘mixture’ of G1P and citrate found in the other. The relative active site occupancy also correlates with the rigidity of different parts of the AfgGGGPS TIM-barrel. As indicated by the average temperature factors (Fig. 2C and Table 1), a well-bound substrate molecule induces stabilization within the region surrounding α3* (i.e. α2, α3*, α4a, and α4b) but it increases the mobility elsewhere (i.e. α8 and α1). Because these observations hold true in all of our structures, and since a similar trend in B-factors is found in BsPerB, it seems unlikely that these effects are caused by crystal packing or lattice contact forces.

**Stereoselective Binding of G1P** - G1P is found near the top inner rim of the β-barrel, running at a slight crossing angle against the C-terminal end of β7, surrounded by the neighboring regions of β6 and β8 (Fig. 3C and 4A). The interactions of G1P with AfgGGGPS can most easily be divided into those formed by water molecules and the protein backbone, or those involving direct side-chain contacts (Fig. 4A). As suggested (12), the phosphate moiety of G1P binds to the TIM-barrel “standard phosphate binding motif” (29, 30). It forms 10 hydrogen bonds in the AfgGGGPS active site: 4 directly to the protein backbone, 1 to the Asn216 side-chain, and 5 to four tightly-bound water molecules (Fig. 4A). The hydroxyl substituent of the chiral carbon center that defines G1P’s stereochemistry forms hydrogen bonds with two conserved side-chains: the hydroxyl of Tyr165, and the Oe2 of Glu167. The Oe2 of Glu167 also interacts with the amide of Ser169, while its Oe1 H-bonds to the C3 hydroxyl of G1P (Fig. 4A). Therefore, the side-chain of Glu167 is positioned in a way that allows it to hydrogen bond to the C2 and C3 hydroxyls of G1P at the same time. In the Group I GGGPS proteins, Tyr165 will also assist the side-chain of Lys11 to interact with the C3 hydroxyl of G1P (Fig. 4A).

In total, G1P forms 14 hydrogen bonds within the AfgGGGPS active site: 4 directly to the backbone, 5 through side-chains, and 5 mediated by water. While the C1 carbon and O1P oxygen of G1P are ‘recognized’ indirectly in this binding mode, AfgGGGPS conforms to the “four-location model” of stereospecificity in order to discriminate between the enantiomers G1P and G3P (31). The same conclusion is reached in the context of the stereocenter-recognition model for stereoselectivity (32); where binding of G1P is seen to involve two locations, one direction, and steric hindrance (i.e. a third “location”). With respect to the stereospecificity for G1P, it may be surprising that only two protein side-chains are
absolutely conserved between the GGGPS phylogenetic groups. Perhaps this is reconciled by the fact that Tyr165 and Glu167 simultaneously select for the stereochemistry that defines G1P (Fig. 4A).

For completeness, we detail the substitutions that occur in MtGGGPS and TaGGGPS as they relate to the AfGGGPS active site. Ser169 and Asn216 are replaced by glycines, resulting in the loss of two hydrogen bonds to G1P (Fig. 4A). Additionally, Lys11 is changed to leucine, Glu64 to phenylalanine, and Glu122 to methionine (Fig. 2A); these substitutions will drastically alter the charge character of the active site in the Group II proteins. How these amino acid differences affect substrate binding, catalysis, and the release of product are issues that remain to be addressed.

GGPP: Fit for a Groove - One prominent feature of the AfGGGPS structure is a deep cleft that runs across the top of its β-barrel. This starts within the G1P binding site at the β7-β8 junction and transverses ~20 Å across the barrel; it passes under α5’ and reaches Trp91 on α4a. The cleft becomes a tunnel as it moves underneath the side-chain of Trp91, dives down and to the right ~45°, then ends ~10 Å later at Trp99. We believe this is the binding site for the geranylgeranyl moiety of GGPP for the following reasons (Fig. 3D). First, it is conserved and has a distinctly apolar character, with the length of the cavity being particularly void of water molecules. Second, it is continuous and leads directly to the C3 hydroxyl of G1P. Third, it can accommodate a geranylgeranyl chain and it bears hallmarks of the “hydrocarbon rulers” found in other acyl- and prenyltransferases (28, 33, 34). For instance, a bulky hydrophobic residue usually marks the end of these ‘channels’, and Trp99 would presumably help select for the chain length of the substrate through steric hindrance. Finally, extra density was found in the G1P-soaked structure past Trp91 and in front of Trp99. The density can be fit by a model of 2-methyl-2,4-pentanediol (MPD), the cryoprotectant used in the experiment (Fig. 3C). This finding confirms the binding potential of the cavity for hydrophobic molecules. The position modeled for MPD probably reflects the one occupied by the terminal isoprenoid unit of the GGPP substrate (Fig. 3B-D). Although glycine or alanine replaces Trp99 in MtGGGPS and TaGGGPS respectively, these substitutions are likely compensated by the phenylalanine found at the position equivalent to Val80 (Fig. 2A). As Val80 is juxtaposed to the bulky side-chain of Trp99, the Group II enzymes also seem to contain the steric hindrance-based chain-length selecting feature characteristic of a hydrocarbon ruler.

None of our structures showed electron density that could be interpreted as a Mg²⁺-ion. To identify a potential binding site for the Mg²⁺-diphosphate moiety of GGPP, we considered the strictly conserved ‘active site’ residues in Group I and Group II GGGPSs. In this context, Tyr124 is too far away from the C3 hydroxyl of G1P to promote substrate binding and catalysis. However, the side-chain of Asp13 lies in a pocket adjacent to G1P, where it forms hydrogen bonds to nearby waters. In fact, a small network of water molecules exists within this pocket. We also noticed that the position equivalent to Thr39 is always either a serine or threonine (Fig. 2A). This is significant because Thr39 is found juxtaposed to Asp13 (Fig. 4B). Together Asp13, Thr39, the surrounding backbone carbonyls and water molecules are excellent candidates to ligate the Mg²⁺-diphosphate group of GGPP. This scenario is supported by the fact that the amino acids around Asp13 and Thr39 are two of the most highly conserved regions within the GGGPS proteins (Fig. 2A).

In the area surrounding α3*, binding of G1P also induces rigidification of the β2-α2 loop, i.e. residues Gly38-Glu45, including Thr39 (Fig. 2C). While this stretch of the polypeptide chain is more mobile in the citrate-bound structure, soaking with G1P results in much stronger electron density. It is tempting to speculate that binding of one substrate (G1P) would generate the proper “mold” for binding of the second substrate (GGPP), implying an ordered reaction mechanism.

Components of Catalysis - Poulter and co-workers have interpreted their results on MtGGGPS in favour of an electrophilic alkylation reaction over a nucleophilic attack (9-11), which is consistent with the proposals for most other prenyltransferases (28, 35). The reaction mechanisms presented for the well-characterized protein farnesyl- and type I geranylgeranyltransferases however, contain both
electrophilic and nucleophilic components (36, 37). As we have only obtained crystals of AfGGGPS in complex with G1P, a discussion of the interactions between the enzyme and GGPP is based on our model of a ternary complex. In this model the C3 hydroxyl of G1P is directed toward the postulated Mg\textsuperscript{2+}-diphosphate binding site, where the C1 of GGPP is appropriately situated to attack the C3 oxygen of G1P (Fig. 3D). As such, we feel that the limited conservation between Group I and Group II GGGPSs considerably narrows the choice when pinpointing essential catalytic residues.

A general feature of prenyltransferase catalysis is thought to be rupture of the carbon-oxygen bond in the isoprenoid diphosphate substrate, which generates a highly electrophilic allylic carbocation that subsequently alkylates a prenyl acceptor (35). Formation of the electrophilic intermediate is aided by the presence of an allylic double bond in the prenyl donor (9, 35). In AfGGGPS, several acidic residues within the active site generate a negatively charged environment (Fig. 3B), lowering the activation energy for creation of a carbocation. At the same time, a magnesium counterion will support the pyrophosphate leaving group. Asp13, an absolutely conserved residue in all GGGPS proteins (Fig. 2A), is well positioned to act as a ligand for a Mg\textsuperscript{2+}-ion.

Glu167 is another residue with a potential role in catalysis. Its Oc2 forms hydrogen bonds to the amide of Ser169 and to the C2 hydroxyl of G1P, while its Oc1 interacts with the C3 hydroxyl of G1P (Fig. 4A). This network of hydrogen bonds leads to a precise positioning of G1P; it also insures, by polarizing the C3 hydroxyl bond, that this oxygen atom becomes an excellent prenyl acceptor. The side chain of Glu167 can further act as an intermediary acceptor of the proton released upon formation of the ether bond.

The Missing Helix: Implications of \(\alpha3^*\) - To the best of our knowledge, the \(\alpha3^*\) modification of AfGGGPS is a previously unobserved feature in any TIM-barrel structure (29, 30). However, members of the phosphatidylinositol phospholipase C (PI-PLC) family have long been known to replace “Helix 4” and “Helix 5” with ‘strands’ (38, 39). The functional significance of the PI-PLC alterations may involve the steric requirements to dock a TIM-barrel domain onto a membrane; in order to allow access of phospholipid head groups into the active site, and release product following catalysis (39). In AfGGGPS the \(\alpha3^*\) alteration stands out, in part because of the dramatic changes in B-factors observed within this region depending upon the relative active site occupancy (see above and Fig. 2C). Since the AfGGGPS structure could easily accommodate a helix at this position, there is no obvious steric reason for the \(\alpha3^*\) modification. On the contrary, the replacement of “Helix 3” creates a large gap in the protein structure. This essentially leaves a void behind Trp99, the terminal residue of the hydrophobic channel that presumably harbors the isoprenoid chain of the GGPP substrate, or that of the GG1P product.

In its rigid state, a striking network of ordered water molecules surrounds the \(\alpha3^*\) element. This water network is generally two molecules wide and can be described as an upside down “L” (Fig. 4B). The L begins near the ‘bottom’ end of \(\alpha3^*\) and tracks up along it, beside \(\alpha4b\) and \(\alpha4a\). The ‘short arm’ of waters actually threads under \(\alpha3^*\), and moves into a crevice found just beneath the \(\beta2\)-\(\alpha2\) loop region. The ‘long arm’ of waters continues along \(\alpha3^*\) and passes by Trp99, landing along the MDP molecule that occupies the end of the aforementioned hydrophobic cavity. The track of water molecules closest to \(\alpha3^*\) makes numerous hydrogen bonds to the protein backbone. By contrast, the waters next to the \(\alpha4a\)-\(\alpha4b\) region are found in a relatively hydrophobic environment. These characteristics probably impose a greasy quality along the length of this peculiar area. Because an analogous feature is known in other protein structures, we refer to this attribute by an established term, a “greasy slide” (40). It is significant that this greasy slide starts right where the hydrophobic binding site of the geranylgeranyl chain ends; then it runs past a region of the protein that is known to be highly mobile at times, and ends near a surface that may be associated with the membrane bilayer. In addition, we suggest that the entire \(\alpha3^*\) modification can move; in essence, that it may swing back and forth like a door on hinges.

A case can be made for the functional significance of a “greasy slide” and “swinging door” in the AfGGGPS structure. As with many acyl- and prenyltransferase systems, GGGPSs may
require a significant amount of detergent for optimal enzymatic activity (12). Moreover, the release of product is the rate-limiting step in some of these systems (36, 37, 41). Thus, because AfGGGPS turns an already hydrophobic substrate (GGPP) into an even more hydrophobic product (GGG1P), there are obvious implications for the release of GGG1P. As hydrogen bonds to the C3 hydroxyl of G1P are abolished and anchoring of the geranylgeranyl chain by the Mg$^{2+}$-diphosphate moiety is released during the reaction cycle, these events are likely propagated along the AfGGGPS structure. Probably linked through the $\beta_2$-$\alpha_2$ loop region and the underlying L-shaped water network, $\alpha_{3*}$ might become mobile and further open access to the greasy slide beyond Trp99 as an exit for the GGG1P product. A back and forth swinging motion of $\alpha_{3*}$ could also act like a molecular ratchet. In this way, the AfGGGPS TIM-barrel modification might promote the delivery of its hydrophobic product in proximity to the membrane bilayer; where the next enzyme in the pathway lies in wait (13).

A Possible Function for PcrB - PcrB is a homologue of AfGGGPS (~35% sequence identity, Fig. 2A) and a protein of unknown function that is conserved in some Gram-positive bacteria, including the human pathogens B. anthracis, S. aureus, and L. monocytogenes (8, 12, 15). Based on the similarity between AfGGGPS and BsPcrB crystal structures (1.3 Å RMSD over 209 C$\alpha$ atoms, Fig. 5), it is clear that the PcrB proteins may function as prenyltransferases. Although the AfGGGPS “greasy-slide” has been filled in, and the $\alpha_{3*}$ “swinging door” stabilized, the PcrB proteins could still accommodate a geranyl chain within their hydrophobic surface clefts. The remainder of the PcrB active site is virtually identical to AfGGGPS, where all PcrB homologues could maintain the 14 hydrogen bonds to G1P (Fig. 2A, Fig 4A and Fig. 5).

A DALI search (42) shows that the top structural alignments against AfGGGPS and BsPcrB are TIM-barrel proteins which are known to bind aromatic substrates. It is therefore tempting to speculate that the PcrB proteins might perform a function similar to the recently characterized class of aromatic-prenyltransferases (14). Alternatively, because G1P is a component of the lipoteichoic acid polymers found in Gram-positive bacteria (43), it is possible that the PcrB proteins could play a role in these biosynthetic pathways.

**PERSPECTIVES**

**GGGPS: An Ancient TIM-barrel Prenyltransferase** - Nature has clearly exploited the TIM-barrel structural scaffold with a plethora of enzymatic functions (29, 30). We now add a long-chain prenyltransferase to the list (Fig. 6B). This result is particularly fascinating when considering the highly specialized structures and unique folds that are known for other prenyltransferases (14, 28). Given the complete restriction of GGGPS homologues to Archaea, the extremely narrow distribution of the GGGPS TIM-barrel prenyltransferase in Nature has two immediate implications. The generation of a GGGPS function on a TIM-barrel fold seems to be a uniquely archael invention (8); and the limited occurrence of PcrB proteins within bacteria points to a horizontal gene transfer event from Archaea (8). In addition to representing the committed step in lipid synthesis, GGGPS function alone can provide all three distinguishing characteristics of archael phospholipids: the unique backbone stereochemistry, an isoprenoid hydrophobic chain, and an ether linkage. Since biological membranes presumably display a vertical line of inheritance, we believe that the evolutionary history of GGGPS reflects the emergence of Archaea.

Besides BsPcrB, a DALI search (42) reveals that AfGGGPS is most closely related to the FMN-dependent oxidoreductase and phosphate binding (FMOP) family of TIM-barrel proteins (29). Specifically, HisF and HisA are listed at positions 1 and 17, respectively. This becomes intriguing because a large body of evidence indicates that HisA and HisF have evolved by 2-fold duplication and fusion of a ($\beta_\alpha$)-half-barrel ancestry (44-46). Based on the facts stated above, accumulating corroborating evidence of a ($\beta_\alpha$)-half-barrel ancestry (47-49), the presumed time of HisA and HisF appearance (50), and AfGGGPS’s profound sequence and structural relatedness to HisA and HisF (not shown), a host of evolutionary arguments can be used to implicate GGGPS as the founding feature which drove the eventual divergence of Archaea from the ‘last universal ancestor’ (51).
REFERENCES

ACKNOWLEDGEMENTS

J.P. is grateful for the continued support of family, friends, and Ms. Emily Cowan. We acknowledge suggestions provided by Drs. Ning Wu and Alex Ghetu. We thank staff at the BioCARS beamlines of the Advanced Photon Source (APS) and the X6A beamline at the National Synchrotron Light Source (NSLS) for their time commitments and expert help. Use of the APS was supported by the United States Department of Energy (DOE) under contract W-31-109-Eng-38 and BioCARS was supported by the National Institutes of Health (NIH) under grant RR07707. Use of the NSLS was supported by the DOE under contract DE-AC02-98CH10886 and X6A was funded by the NIH under agreement Y1 GM-0080-03. M.F. received an Overseas Fellowship and a Research Fellowship for Young Scientists from the Japan Society for the Promotion of Science. This work was supported by the Canada Research Chairs Program and the National Sciences and Engineering Research Council of Canada (E.F.P.).

FOOTNOTES

a The abbreviations used are: G1P, sn-glycerol-1-phosphate; GGGPS, (S)-3-O-geranylgeranylglyceryl phosphate synthase; Af, Archaeoglobus fulgidus; Bs, Bacillus subtilis; Mt, Methanobacterium thermoautotrophicum; Ta, Thermoplasma acidophilum; TIM, triose phosphate isomerase; GGPP, geranylgeranyl diphosphate; MPD, 2-methyl-2,4-pentanediol.

b J.P. Unpublished data
c J.P. and E.F.P. Unpublished observations

Data Deposition: Atomic coordinates and structure factors have been deposited in the Protein Data Bank, www.rcsb.org, with the PDB ID 2F6U and 2F6X for the citrate- and G1P-complex of AfGGGPS, respectively.

FIGURE LEGENDS

Table 1. Crystallographic data collection and refinement statistics.

Fig. 1. Schematic diagram of lipid synthesis in Archaea.

Fig. 2. Sequence alignment and structure of AfGGGPS. A. Group I proteins from A. fulgidus and Halobacterium NRC-1 (HbGGGPS) are shown, PcrB homologues from B. subtilis and B. anthracis are included (BsPcrB and BaPcrB), and the Group II proteins from M. thermoautotrophicum and T. acidophilum are presented in the alignment. Residues in α0 and the N-terminal extensions are coloured: red - acidic, blue - basic, and green - hydrophobic. Beyond these, residues with side-chains thought to interact directly with the Mg2+-diphosphate moiety of GGPP are coloured red. Residues that interact with G1P are coloured purple. Key hydrophobic residues are coloured green: a bulky residue at the beginning of α4a, and the steric “floor” at the end of the hydrocarbon ruler. B. A ribbon diagram of the AfGGGPS dimer, top and side views. Note the location of the α3* element. G1P and MPD are shown as ball-and-stick representations. C. Top and side views illustrating the average temperature factors. B-factors <30 Å2 are coloured blue, >65 Å2 are coloured red, and those ranging from 30 Å2 – 65 Å2 are shaded from blue to red. The subunit on the right contains the well-bound ligand molecules (see main text). Note the differences surrounding the α3* elements within the dimer. Also notice that the N-terminal extensions (found in homologues) may be in a position to interact with the membrane lipids and/or other proteins. α5’ and α0 indicate the ‘top’ and ‘bottom’ of the protein, respectively.

Fig 3. Electrostatic surface potential and active site of AfGGGPS. A-B. The ‘bottom’ and ‘top’ of AfGGGPS are shown, acidic regions are shaded red and basic regions are shaded blue. A. The
concentration of positive charge on the bottom surface could promote an interaction or association of AfGGGPS with the membrane. B. The acidic character of the active site surrounding G1P may play a role in catalysis. Relative to all other figures, G1P has been enlarged here for clarity. Notice that the MPD molecules are ‘obscured’ in this view. C. Stereoview of the active site cleft with G1P and MPD bound. The corresponding electron densities were taken from an Fo-Fc map and contoured at 3.5σ. Note the locations of the Trp91 and Trp99 side-chains. D. Mg^{2+}-GGPP is modeled in its proposed binding site with G1P. Mg^{2+} is represented as an aqua coloured sphere.

Fig. 4. Structural details of the AfGGGPS active site. A. Stereoview of G1P bound in the active site. Residues involved in G1P binding are shown, and water molecules that mediate an interaction with G1P are represented as green spheres. Hydrogen bonds are coloured according to their association with substrate oxygens. B. Stereoview of the network of water molecules around α3*. MPD, Trp91 and Trp99 are shown, as are Asp13 and Thr39. Asp13 and Thr39 are implicated in binding the Mg^{2+}-GGPP substrate.

Fig. 5. Superposition of AfGGGPS and BsPcrB. A stereoview of the structural alignment between AfGGGPS and BsPcrB showing the protein backbones of a single subunit. Using AfGGGPS numbering, some side-chains mentioned in the main text are indicated in one letter code. AfGGGPS is coloured purple and BsPcrB is coloured blue.

Fig. 6. Activity of Recombinant AfGGGPS. A. 5 μg of AfGGGPS was run on a 15% gel for SDS-PAGE. AfGGGPS migrates at the expected molecular weight (~26.4 kDa) when the sample is boiled (Lane I) but remains mostly dimeric, and migrates aberrantly (~37 kDa), when unboiled (Lane II). A molecular weight ladder is indicated. B. Absorbance at 565 nm was monitored over time (minutes) in an enzyme-coupled assay to detect liberated pyrophosphate. Red, yellow, and green plots indicate AfGGGPS at final concentrations of 2, 0.2, and 0.02 μM respectively. The “activity” of MtODCase at 2 μM, a structurally related (42) but functionally unrelated enzyme (27), is plotted in pink for comparison.
### Table 1

<table>
<thead>
<tr>
<th>Data Set</th>
<th>Citrate-Bound</th>
<th>G1P-Soaked</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Data Collection</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Source</td>
<td>BioCARS 14BM-D</td>
<td>BioCARS 14BM-D</td>
</tr>
<tr>
<td>Wavelength, Å</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>Space group</td>
<td>$H_3$</td>
<td>$H_3$</td>
</tr>
<tr>
<td>Unit cell dimensions, Å</td>
<td>$a = b = 95.24$, $c = 166.62$</td>
<td>$a = b = 95.88$, $c = 166.30$</td>
</tr>
<tr>
<td>Resolution, Å*</td>
<td>50 - 1.55 (1.61 - 1.55)</td>
<td>50 - 2.00 (2.07 - 2.00)</td>
</tr>
<tr>
<td><strong>No. of reflections</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observed</td>
<td>1 211 058</td>
<td>743 573</td>
</tr>
<tr>
<td>Unique</td>
<td>76 814</td>
<td>37 336</td>
</tr>
<tr>
<td>Completeness, %*</td>
<td>99.9 (100)</td>
<td>99.9 (100)</td>
</tr>
<tr>
<td>$d/\langle d\rangle$*</td>
<td>39.7 (5.1)</td>
<td>46.4 (6.9)</td>
</tr>
<tr>
<td>$R_{\text{merge}}$†</td>
<td>5.1 (29.4)</td>
<td>4.2 (25.8)</td>
</tr>
<tr>
<td><strong>Refinement</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$R/R_{\text{free}}$, %‡</td>
<td>19.1/21.7</td>
<td>18.8/22.7</td>
</tr>
<tr>
<td>RMSD bond length, Å/bond angles, °</td>
<td>0.008/1.36</td>
<td>0.009/1.39</td>
</tr>
<tr>
<td>No. of atoms§</td>
<td>3744/341/26</td>
<td>3766/222/20/16</td>
</tr>
<tr>
<td>Average B factors, Å²¶</td>
<td>20.3/29.3/24.8</td>
<td>34.1/36.4/43.4/51.7</td>
</tr>
<tr>
<td>Ramachandran plot, %¶¶</td>
<td>92.3/7.2/0.5/0</td>
<td>90.3/9.2/0.5/0</td>
</tr>
</tbody>
</table>

* Values in parentheses are for the highest resolution shells  
† $R_{\text{merge}} = \Sigma |I_i - \langle I_i\rangle| / \Sigma I_i$, where $I_i$ is the observed intensity and $\langle I_i\rangle$ is the average intensity over symmetry equivalent measurements  
‡ $R = \Sigma |F_{\text{obs}} - |F_{\text{calc}}| / \Sigma |F_{\text{obs}}|$, $R_{\text{free}}$ is the same as $R$, but for a 5% subset of all reflections that were not used in crystallographic refinement  
§ protein/water/citrate and protein/water/G1P/MPD, respectively  
¶ B factors for the small molecules in separate active sites of the AfGGGPS dimer are as follows: citrate: 20.1/29.5  
¶¶ most favoured/additionally allowed/generously allowed/disallowed regions
Figure 1

[Chemical pathway diagram showing the conversion of DMAPP to Archaeal Membrane Lipids via IPP, GPP, GGGPS, GGG1P, and DGGG1P.]
Figure 4
Figure 5
Figure 6
The crystal structure of (S)-3-O-geranylgeranylglyceryl phosphate synthase reveals an ancient fold for an ancient enzyme
Jian Payandeh, Masahiro Fujihashi, Wanda Gillon and Emil F. Pai

J. Biol. Chem. published online December 23, 2005

Access the most updated version of this article at doi: 10.1074/jbc.M509377200

Alerts:
• When this article is cited
• When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 0 references, 0 of which can be accessed free at http://www.jbc.org/content/early/2005/12/23/jbc.M509377200.citation.full.html#ref-list-1