Ether-lipid membrane engineering of *Escherichia coli*

Antonella Caforio

Cover design and layout of the book: Antonella Caforio

Cover front and back: artistic interpretation of lipid rafts in a sea landscape. The lipid rafts are composed by two different lipid species representing archaeal and bacterial membrane lipds with the CarS structure embedded.

The research described in this thesis was carried out in the Department of Molecular Microbiology of the Groningen Biomolecular Sciences and Biotechnology Institute (GBB), University of Groningen, The Netherlands. It was financially supported by the biobased ecologically balanced sustainable industrial chemistry (BE-BASIC).

ISBN: 978-90-367-9379-7
ISBN: 978-90-367-9378-0 (electronic version)

Printed by Ipskamp printing, Enschede
Printing of this thesis was supported by generous contribution from the University of Groningen and the Groningen Biomolecular Sciences and Biotechnology Institute (GBB).

university of groningen

Ether-lipid membrane engineering of *Escherichia coli*

PhD thesis

to obtain the degree of PhD at the
University of Groningen
on the authority of the
Rector Magnificus Prof. E. Sterken
and in accordance with
the decision by the College of Deans.

This thesis will be defended in public on

Friday 13 January 2017 at 14.30 hours

by

Antonella Caforio

born on 14 April 1988
in Francavilla Fontana, Italy

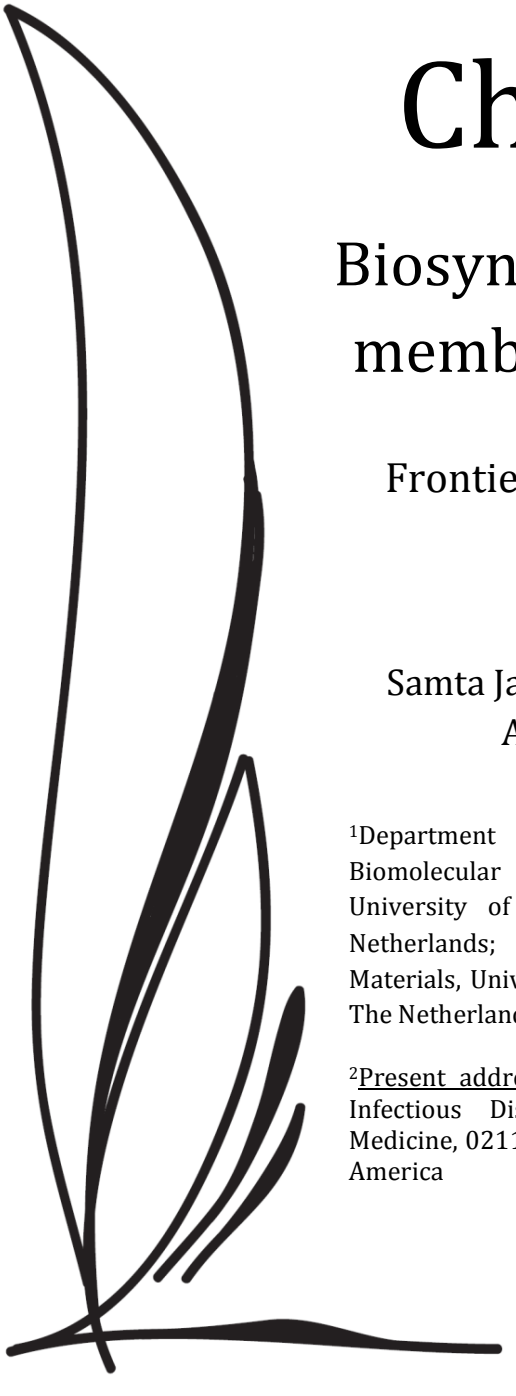
Supervisor
Prof. A.J.M. Driessen

Assessment Committee
Prof. D.J. Slotboom
Prof. O.P. Kuipers
Prof. S. Schouten

A mio padre, mia madre e mia sorella
per avermi sempre sostenuto,
incoraggiato e creduto in me
dall'inizio alla fine.

Al mio amore,
per avermi donato amore e felicitá.

Table of Contents



Chapter 1

Biosynthesis of archaeal membrane ether lipids

Frontiers in Microbiology 2014, 5:641

Samta Jain[1,2], Antonella Caforio[1] and Arnold J. M. Driessen[1]

[1]Department of Molecular Microbiology, Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, 9747 AG Groningen, The Netherlands; The Zernike Institute for Advanced Materials, University of Groningen, 9747 AG Groningen, The Netherlands

[2]Present address: Department of Medicine, Section of Infectious Diseases, Boston University School of Medicine, 02118 Boston, Massachusetts, United States of America

Abstract

A vital function of the cell membrane in all living organism is to maintain the membrane permeability barrier and fluidity. The composition of the phospholipid bilayer is distinct in archaea when compared to bacteria and eukarya. In archaea, isoprenoid hydrocarbon side chains are linked via an ether bond to the sn-glycerol-1-phosphate backbone. In bacteria and eukarya on the other hand, fatty acid side chains are linked via an ester bond to the sn-glycerol-3-phosphate backbone. The polar head groups are globally shared in the three domains of life. The unique membrane lipids of archaea have been implicated not only in the survival and adaptation of the organisms to extreme environments but also to form the basis of the membrane composition of the last universal common ancestor (LUCA). In nature, a diverse range of archaeal lipids is found, the most common are the diether (or archaeol) and tetraether (or caldarchaeol) lipids. Variations in chain length, cyclization and other modifications lead to diversification of these lipids. The biosynthesis of these lipids is not yet well understood however progress in the last decade has led to a comprehensive understanding of the biosynthesis of archaeol. This review describes the current knowledge of the biosynthetic pathway of archaeal ether lipids; insights on the stability and robustness of archaeal lipid membranes; and evolutionary aspects of the lipid divide and the last universal common ancestor LUCA. It examines recent advances made in the field of pathway reconstruction in bacteria.

Introduction

The "Woesian Revolution" in 1977 defined the three domains of life as the Eukarya, the Bacteria and the Archaea [1]. The archaeal membrane lipid composition is one of the most remarkable feature distinguishing Archaea from Bacteria and Eukarya where the hydrocarbon chain consists of isoprenoid moieties which are ether linked to the enantiomeric glycerol backbone, glycerol-1-phosphate in comparison to glycerol-3-phosphate of bacteria and eukarya that is ester linked to the fatty acid derived hydrocarbon chain. Polar head groups on the other hand are common in all three domains of life. Other than this core archaeal diether lipid structure, a bipolar tetraether lipid structure is also prevalent in many archaea that span the entire archaeal membrane forming a monolayer [2]. It should be stressed that ether-linked lipids are not unique to Archaea *per se*, but are also found in Bacteria and Eukarya, although not ubiquitously distributed and usually only a minor component of the lipid membrane.

The stereo specificity of archaeal lipids and their unique structure was hypothesized to be chemically more stable thereby rendering the organism with the ability to resist and thrive in extreme environmental conditions [3]. However, archaea are also found in mesophilic and neutrophilic environment where such a structural role of ether lipids is still not postulated. At the same time, the distinguishing lipid structures have formed the basis to the evolutionary studies describing archaeal and bacterial differentiation. Several models hypothesizing the early evolution of archaeal and bacterial phospholipid biosynthesis were proposed to answer intriguing questions about the nature of the ancestral membrane lipid composition [4]. Understanding the archaeal lipid biosynthetic pathway is crucial to the above studies.

Decades of studies on the biosynthesis of archaeal lipids have advanced our knowledge on the major enzymatic processes but the pathway is, however still not completely understood. Several enzymes of the pathway have been studied and characterized biochemically but there are also gaps in our understanding of the archaeal lipid biosynthetic pathway and little is known about its regulation. With more genome sequences becoming available, advanced phylogenetic studies have been performed recently

[5–8] and this helped to more precisely define its evolution. This review will focus on existing knowledge and recent studies on the enzymes of the pathway, the physicochemical properties of archaeal lipids, and the theories on the lipid divide.

Biosynthesis of archaeal membrane lipids

Isoprenoid building blocks and chain elongation

Isoprenoids are ubiquitous to all three domains of life. They are structurally diverse, forming more than 30,000 different compounds in nature ranging from steroids, quinones, carotenoids and membrane lipids. The building blocks of isoprenoids are universal carbon five subunits called isopentenyl pyrophosphate (IPP) and dimethylallyl pyrophosphate (DMAPP) that are isomers. The biosynthetic pathway leading to the synthesis of IPP and DMAPP vary in different organisms (reviewed in [9,10]). To date, three pathways have been reported – 2-C-methyl-D-erythritol 4-phosphate/1-deoxy-D-xylulose 5-phosphate pathway (MEP/DOXP pathway) and two Mevalonate (MVA) pathways. The MEP pathway genes share no homology to genes of the MVA pathway where pyruvate and glyceraldehyde-3-phosphate molecules are condensed together to form 1-deoxy-D-xylulose-5-phosphate (DXP) which is subsequently converted to IPP and DMAPP by five enzymes [11] (**Figure 1**). The MEP pathway is most common in bacteria although some Firmicutes possess the MVA pathway. The MVA pathway consists of seven enzymatic reactions where two acetyl-CoA molecules are condensed to form acetoacetyl-CoA, which is further condensed to form 3-hydroxy-3-methylglutaryl CoA (HMG-CoA). HMG-CoA undergoes phosphorylation and decarboxylation to form IPP via the formation of MVA (**Figure 1**). The classical MVA pathway is common to eukaryotes while some plants and photosynthetic eukaryotes possessing the MEP pathway in addition [9].

Interestingly, homologs of the last three enzymes of the classical MVA pathway, i.e. phosphomevalonate kinase, diphosphomevalonate decarboxylase and isopentenyl diphosphate isomerase could not be found in the majority of archaea (except in Sulfolobales that have classical MVA

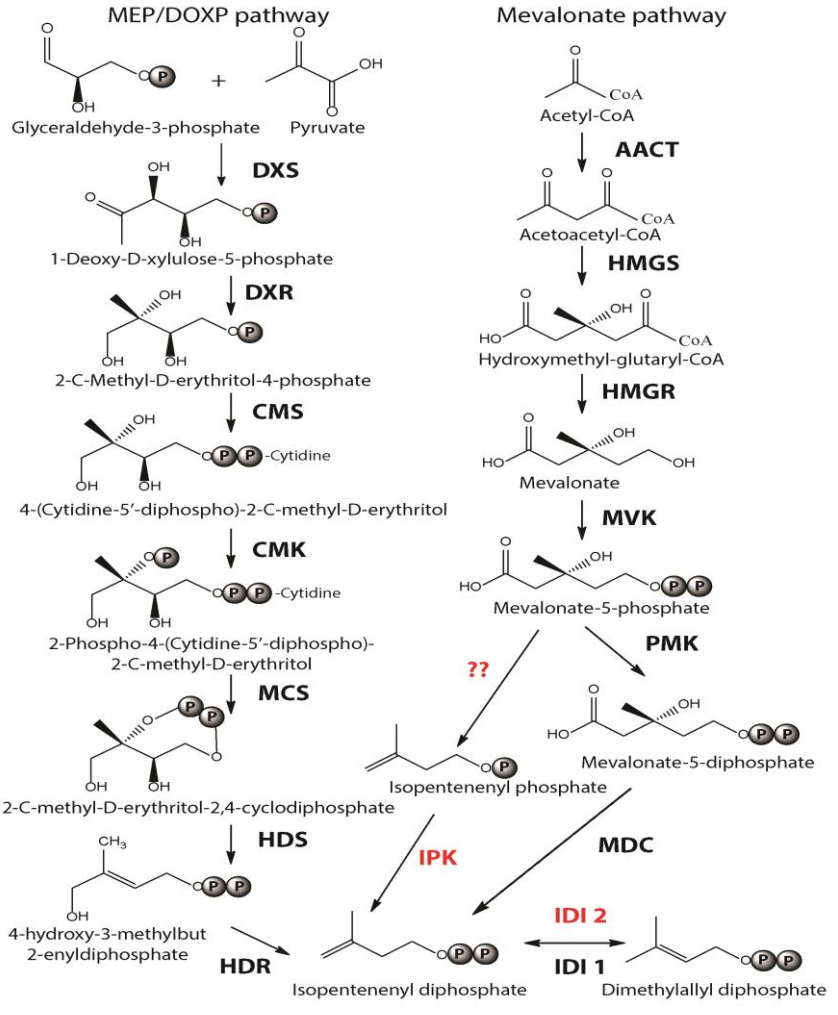


Figure 1| Biosynthesis of Isoprenoid building blocks. The three pathways leading to the synthesis of IPP and DMAPP- MEP/DOXP pathway, Mevalonate and the alternate mevalonate pathway (in red), which is prevalent in archaea are shown. Mevalonate and the alternate mevalonate pathways share four of their seven enzymatic steps. The enzyme acronyms are written in bold. DXS, 1-Deoxy-d-xylulose 5-phosphate synthase; DXR, 1-Deoxy-d-xylulose 5-phosphate reductase; CMS, 2-C-methyl-d-erythritol 4-phosphate sytidyltransferase; CMK, 4-diphosphocytidyl-c-C-methyl-d-erythritol kinase; MCS, 2-C-methyl-d-erythritol 2,4-cyclodiphosphate synthase; HDS, (E)-4-hydroxy-3-methyl-but-2-enyl pyrophosphate synthase; HDR, (E)-4-hydroxy-3-methyl-but-2-enyl pyrophosphate reductase; IDI, isopentenyl diphosphate isomerase; AACT, acetocetyl-CoA thiolase; HMGS, 3-hydroxy-3-methylglutaryl-CoA synthase; HMGR, 3-hydroxy-3-methylglutaryl-CoA reductase; MVK, mevalonate kinase; PMK, phosphomevalonate kinase; MDC, mevalonate-5-decarboxylase; IPK, isopentenyl phosphate kinase.

pathway) [5]. This search led to the discovery of the alternate MVA pathway that differs from the classical one in the last three steps (**Figure 1**). The enzyme isopentenyl kinase (IPK) was first discovered in *Methanocaldococcus jannaschii* and found to be conserved in archaea [12]. Its structure was determined [13] and IPK enzymes from *Methanothermobacter thermautotrophicus* and *Thermoplasma acidophilum* were characterized biochemically [14]. In the alternate MVA pathway, phosphomelavonate is decarboxylated to isopentenyl phosphate by a decarboxylase (enzyme yet to be identified in archaea), which is subsequently phosphorylated to IPP by IPK. Furthermore, instead of the typical IDI1 isomerase that performs the last step of the classical MVA pathway, archaea have IDI2 which is not homologous to IDI1 but that performs the same reaction [10]. Interestingly, a decarboxylase enzyme that converts phosphomelavonate to isopentenyl phosphate was found in green non-sulphur bacteria *Roseiflexus castenholzii* along with the presence of IPK enzymes indicating the existence of alternate MVA pathway in organisms other than archaea [15]. In general, IPP and DMAPP are synthesized by the MEP pathway in most of the bacteria and by two MVA pathways in eukarya and archaea. The classical MVA pathway of eukaryotes and the alternate MVA pathway of archaea share four of their seven steps.

The isoprenoid building blocks IPP and DMAPP undergo sequential condensation reactions where DMAPP acts as the first allylic acceptor of IPP leading to the formation of a carbon 10 (C10) compound termed geranyl diphosphate (GPP). Further condensation reactions proceed with the addition of IPP molecules where the chain length increases each time by a C5 unit forming farnesyl (C15), geranylgeranyl(C20), farnesylgeranyl (C25) diphosphate etc. This reaction of chain elongation is catalyzed by enzymes belonging to the family of prenyl transferases that are common to all three domains of life [16,17]. Depending on the length and geometry of the final molecule, prenyl transferases can have several members in its family. The geometry of the molecule could be *cis* or *trans* and the chain length of the *trans* form generally ranges from C10 (e.g. monoterpenes) to C50 (e.g. Coenzyme Q10) and even longer for the *cis* forms. The chain length found in archaeal membrane lipids is always in the *trans* form and

composed mostly of C20 (geranylgeranyl diphosphate, GGPP) or C25 (farnesylgeranyl diphosphate). Tetraethers are composed of a C40 chain length, the synthesis of which is still unknown (discussed below). The archaeal prenyltransferase enzymes GGPP synthase and farnesylgeranyl diphosphate synthase synthesize specifically C20 or C25 product chain lengths, respectively [18–21]. They belong to short chain *trans* prenyl transferases family (that catalyze reactions ranging from C10-C25). Interestingly, a bifunctional prenyltransferase that catalyzes the synthesis of both C15 and C20 isoprenoids has been characterized from *Thermococcus kodakaraensis* [22] and *Methanobacterium thermoautotrophicum* and is considered to be an ancient enzyme [23]. Multiple sequence alignment of homologues of the family display high sequence similarity with seven conserved regions where region two and seven contain the highly conserved aspartate rich sequences called first aspartate-rich motif (FARM) and second aspartate-rich motif (SARM) domains, respectively. Numerous mutagenesis and structural studies including several members of the family show that the region within the aspartate rich domains are involved in the binding and catalysis of the substrate while the regions flanking these domains are the major determinants of the chain length as they contribute to the size of the active site hydrophobic pocket. For example, GGPP synthase from *Sulfolobus acidocaldaricus* could be mutated (at Phe-77 which is fifth amino acid upstream of FARM) to catalyze longer chain length (C30-C50) products [24] and farnesyl pyrophosphate (FPP) synthase of *Escherichia coli* could be mutated (at Tyr-79, also fifth amino acid upstream of FARM) to create GGPP synthase [25]. Physical factors have also been shown to influence the chain length of the product, e.g. the bifunctional enzyme farnesyl diphosphate/geranylgeranyl diphosphate synthase of *Thermococcus kodakaraensis* shows an increase in the FPP/GGPP ratio with the reaction temperature [22].

Glycerol-1-phosphate backbone

The glycerophosphate backbone of archaea has an opposite stereoconfiguration than that of bacteria and eukarya. The archaeal

enzyme responsible is G1P dehydrogenase that shares homology with alcohol and glycerol dehydrogenases but no homology to the bacterial/eukaryal G3P dehydrogenase. They belong to two separate families. However, both catalyze the reduction of dihydroxyacetone phosphate (DHAP) using nicotinamide adenine dinucleotide hydrogen (NADH) or nicotinamide adenine dinucleotide phosphate hydrogen (NADPH) as substrate (**Figure 2; Table 1**). G1P dehydrogenase uses Zn^{2+} for metal ion interaction in its active site [26] and transfers the pro-R hydrogen of NADH in contrast to G3P dehydrogenase that transfers the pro-S hydrogen [27]; both enzymes bind the nicotinamide ring in an opposite orientation. G1P dehydrogenase is conserved in archaea. The enzyme has been purified and characterized from *Methanothermobacter thermoautotrophicus* as an octamer [28], *Aeropyrum pernix* as a homodimer [29] and from *Sulfolobus tokodaii* [30]. Its activity has been accessed in cell free homogenates of several archaea [2,21].

Until recently it was thought that the stereo specificity is the hallmark of the 'lipid divide' where the G1P backbone is exclusively attributed to archaea. This was challenged by the discovery and characterization of the bacterial G1P dehydrogenase homolog of *Bacillus subtilis* which is annotated as 'AraM' [31]. It is also found in other related Gram positive and negative bacteria. Similar to the G1P dehydrogenase of *Aeropyrum pernix*, AraM forms a homodimer and performs G1P dehydrogenase activity. However, the two enzymes have different catalytic efficiencies and AraM is Ni^{2+} ion dependent [29,31]. Remarkably, the G1P molecule eventually becomes part of an archaea type ether lipid heptaprenylglyceryl phosphate in *B. subtilis*, the function of which is still unknown [32].

Ether linkages

The first and second ether bonds between G1P and GGPP is catalyzed by the enzyme geranylgeranylglycerly diphosphate (GGGP) synthase and di-*0*-geranylgeranylglycerly diphosphate (DGGGP) synthase respectively. GGGP synthase is a conserved enzyme found in all archaea except Nanoarchaeota, which is a symbiont and possesses no genes of the lipid biosynthesis pathway [33]. It is also found in some bacteria where the polyprenyl diphosphate substrate chain length could vary, e.g. PcrB of *B.*

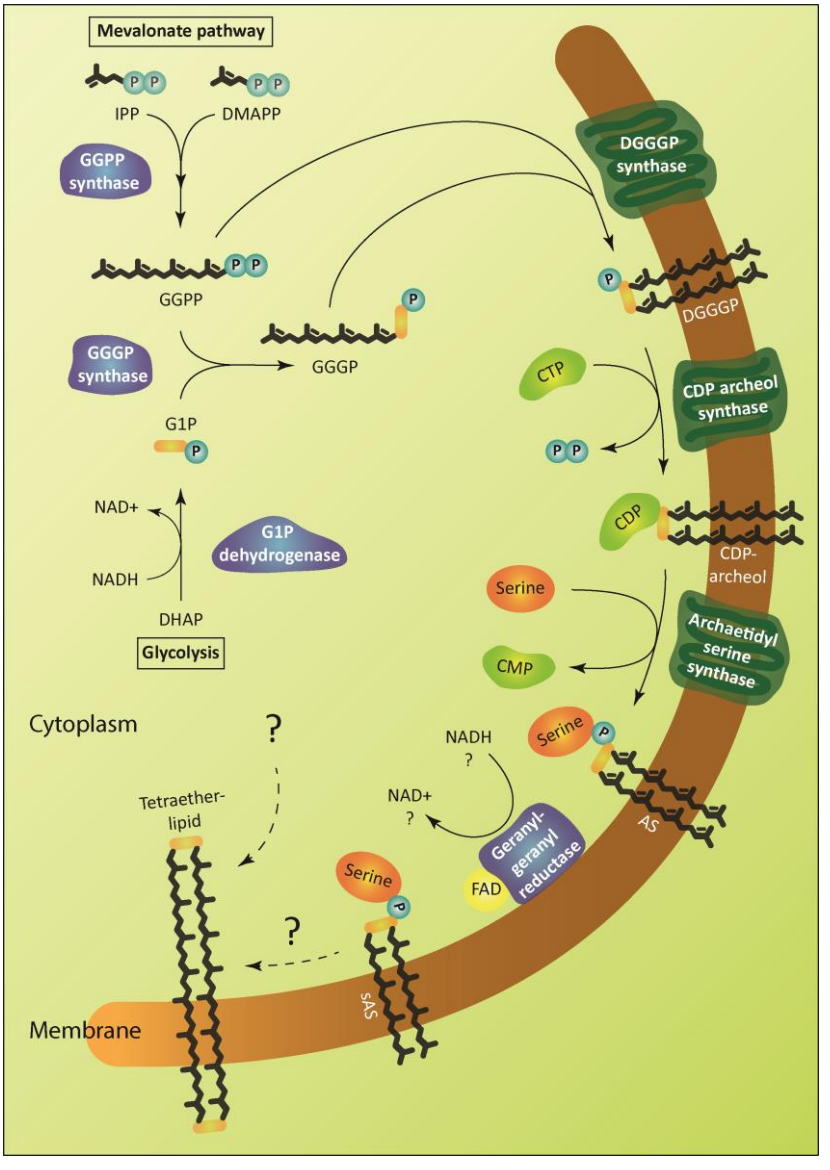


Figure 2|Enzymatic pathway of archaeal lipid biosynthesis. The soluble enzymes of the pathway are colored in blue and the membrane proteins in green. The biosynthetic steps leading to the formation of tetraethers is unknown. Archaeatidylserine (AS) and saturated AS (sAS) are depicted as an example of polar head group modification. Isopentenyl pyrophosphate (IPP) and dimethylallyl pyrophosphate (DMAPP) are isomers. GGPP, geranylgeranyl diphosphate; G1P, *sn*-glycerol-1-phosphate; GGGP, 3-*O*-geranylgeranyl-*sn*-glycerol-1-phosphate; DGGGP, 2,3-bis- *O*- geranylgeranyl-*sn*-glycerol-1-phosphate.

subtilis which is a heptaprenyl diphosphate synthase [34]. GGGP synthase is a crucial enzyme in the biosynthetic pathway of phospholipid metabolism in archaea as it bringstogether the three important characteristic features of the archaeal lipid structure - stereoisomeric G1P glycerol backbone and isoprenoid GGPP side chain linking them together via an ether bond (**Figure 2; Table 1**). Phylogenetic analysis of the GGGP synthase enzymes distinguishes it into two families, group I and group II, both comprising of archaeal and bacterial sequences. Several enzymes from both the groups have been characterized and a recent study performed the biochemical analysis of 17 members of GGGP synthase family [35]. The enzymes of group I form dimers (except the monomeric GGGP synthase of *Halobacterium salinarum*) and the group II enzymes are dimeric or hexameric in nature. Both the groups are further subdivided into Ia, Ib, IIa and IIb with a and b corresponding archaea and bacteria, respectively. Crystal structures of enzymes from all the four groups have been solved. The first crystal structure from the group I GGGP synthase of *Archaeoglobus fulgidus* displays a modified triose phosphate isomerase (TIM)-barrel structure [36]. It forms a dimer bound to the G1P substrate with a central eight-stranded parallel β-barrel and a hydrophobic core surrounded by α-helices (**Figure 3A**). Helix-3 is replaced by a 'strand' which is a novel TIM-barrel modification not observed previously. The substrate GGPP binds to the deep cleft traversing the top of the β-barrel. There is a 'plug' at the bottom of the barrel and the active site lies at the C-terminal end. The G1P molecule sits near the top inner rim of the barrel and the phosphate group binds to the standard phosphate-binding motif of the TIM-barrel. G1P forms 14 hydrogen bonds within the active site. The (βα)8-barrel fold is found in all the other structures of the GGGP synthases as well with the active site at the C-terminus. The crystal structure of group II archaeal hexameric GGGP synthase of *Methanothermobacter thermoautrophicus* displays a combination of three dimers that resemble the group I dimer (**Figure 3B** and **3C**). In group II, however, the plug of the barrel is longer than in group I and there are 'limiter residues' that restrict the length of hydrophobic pocket to accommodate the polyprenyl diphosphates of a specific length. Interestingly, an aromatic anchor residue

is responsible for the hexameric configuration of the enzyme, mutation of which causes it to dimerize without any loss of activity [35].

The intrinsic membrane protein DGGGP synthase catalyzes the formation of the second ether bond between the substrate GGGP and GGPP to form DGGGP (**Figure 2**). It belongs to the family of ubiquinone-biosynthetic (UbiA) prenyltransferases, the members of which are responsible for the biosynthesis of respiratory quinones, chlorophyll, heme etc. by transferring a prenyl group to the acceptors that generally

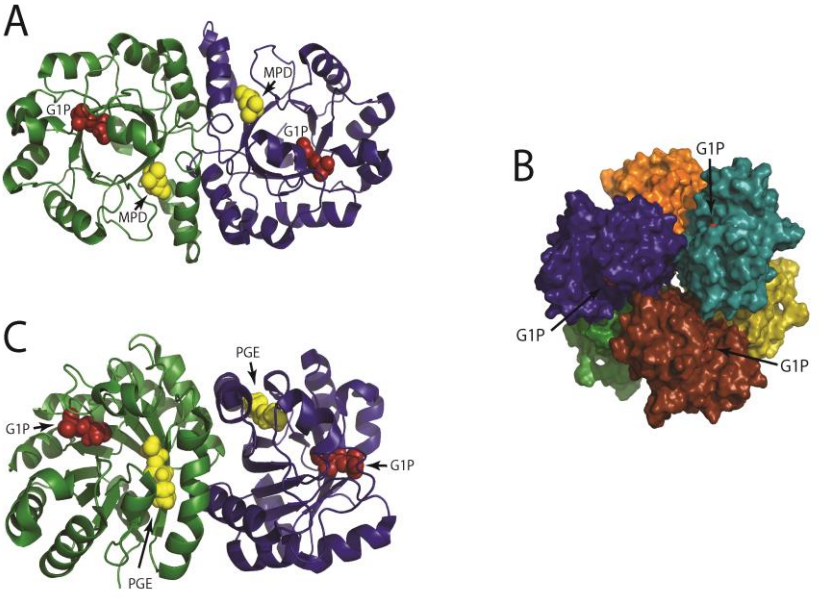


Figure 3| Structure of 3-*O*-geranylgeranyl-*sn*-glyceryl-1-phosphate synthase (GGGP synthase). (**A**) Crystal structure of group 1 dimeric GGGP synthase from *Archaeoglobus fulgidus* solved at 2.0 Å resolution (PDB: 2F6X) has a novel TM-barrel modification [36]. *Sn*-glycerol-1-phosphate (G1P) and a cryoprotectant (MPD) used in the experiment are colored as red and yellow spheres respectively. The cavity for MPD likely represent the binding site for the second substrate geranylgeranyl diphosphate (GGPP). (**B**) Crystal structure of group II GGGP synthase from *Methanothermobacter thermoautotrophicus* that forms a hexamer as a combination of three dimers was solved at 2.8 Å resolution (PDB: 4MM1) [35]. G1P is displayed as red spheres. (**C**) A dimer from (**B**) is represented to compare with the dimeric group 1 GGGP synthase of (**A**). The G1P and the triethelene glycol (PGE) are colored as red and yellow spheres respectively. The hexameric subunits are rotated to each other unlike the dimers in (**A**). The anchoring of G1P in group I and II takes place by the standard phosphate binding motif but their G1P binding pocket are different.

have hydrophobic ring structures. DGGGP synthase is divergent among archaea and could not be identified in the genomes of Thaumarchaeota [6]. Unlike other enzymes of the pathway, DGGGP synthase has not been well characterized probably due to technical limitations with overexpression of the membrane protein. The DGGGP synthase activity was first found in the membrane fraction of *Methanothermobacter marburgensis* [37]. Later the gene was identified in the genome of *Sulfolobus sulfotaricus* as UbiA-2, cloned in *E. coli* and purified to study the Mg^{2+} dependent enzymatic activity using radiolabeled substrates and mass spectrometry [38]. The ratio of the substrates utilized in the reaction was found to be 1:1.1 in a double labeling experiment using [^{3}H]GGPP and [^{14}C]GGGP, respectively. Specificity for GGPP and GGGP was also measured by substituting them with different prenyl substrates, of which none of them were used in the reaction by DGGGP synthase. In another study, DGGGP synthase was shown to accept both the S and R form of GGGP showing that unlike GGGP synthase, it is enantio unselective [39]. DGGGP synthase activity of *Archaeoglobus fulgidus* [21] and *Methanosarcina acetivorans* [40] was also observed in *E. coli* when the corresponding genes were expressed along with four previous enzymes of the pathway. However, the expression level of the enzyme was either too low to detect [21] or not investigated [40]. In a later study, a higher expression level of DGGGP synthase of *A. fulgidus* was obtained in *E. coli* by changing the ribosome-binding site and the activity of purified DGGGP synthase was monitored [41].

CDP archaeol formation

The next step in the archaeal lipid biosynthetic pathway is the activation of DGGGP by cytidine triphosphate (CTP) to form the substrate for polar head group attachment called cytidine diphosphate (CDP)-archaeol (**Figure 2; Table 1**). The reaction is brought about by the enzyme CDP-archaeol synthase (CarS), the activity of which was first studied in the membrane fraction of *Methanothermobacter thermoautotrophicus* [42]. Using various synthetic substrate analogs, the activity was found to be specific for unsaturated archaetidic acid with geranylgeranyl chains and did not depend on the stereo specificity or ether/ester bond of the substrate. Minute amount of CDP-archaeol were also detected in growing

cells labeled with inorganic ^{32}P. The gene responsible for this activity was only identified in a recent study [41]. The enzyme CarS is conserved among archaea (except Nanoarchaeota). However, like the enzyme DGGGP synthase, it could not be identified in the families of Thaumarchaeota.

Interestingly, an analogous reaction is found in the bacterial phospholipid biosynthetic pathway where phosphatidic acid is activated by CTP to form CDP diacylglycerol by the enzyme CDP diacylglycerol synthase (CdsA). Although the sequence similarity between CdsA and CarS is very low, hydropathy profile alignment of the two families shows similarity in their secondary structure with overlapping transmembrane segments and cytoplasmic loop regions residing in the C-terminus half. CarS from *Archaeoglobus fulgidus* was expressed and purified from *E. coli*. Similar to CdsA, CarS activity was found to be dependent on Mg^{2+}, both accepts CTP and deoxycytidine triphosphate (dCTP) as substrates and does not utilize adenosine triphosphate (ATP), guanosine triphosphate (GTP), or thymidine triphosphate (TTP) nucleotides in the reaction using substrate DGGGP. However, the two enzymes displayed distinct activity with respect to the lipid substrate specificity where CarS only accepts unsaturated archaetidic acid with geranylgeranyl chains, while CdsA takes phosphatidic acid [41].

Polar head group attachment

The polar head groups serine, ethanolamine, glycerol and *myo*-inositol are found in the phospholipids in all three domains of life. The enzymes that catalyze the replacement of the cytidine monophosphate (CMP) entity of CDP-archaeol or CDP-diacylglycerol with a polar head group are homologous and belong to CDP-alcohol phosphatidyltransferase family [43]. Archaetidylserine (AS) synthase catalyzes the formation of AS from CDP-archeaol and L-serine (**Figure 2**) and is homologous to bacterial phosphatidylserine (PS) synthase. The enzyme can be classified into two subclasses. Subclass I includes enzymes distributed in Gram-negative bacteria, such as *E. coli* while subclass II enzymes are widespread among Gram-positive bacteria (*B. subtilis*), yeast and Archaea. Studies using cell free extracts of *Methanothermobacter thermautotrophicus*, *B. subtilis* and *E. coli* showed that both the AS and PS synthase from *M. thermautotrophicus*

Table 1: Summary of kinetic parameters derived for enzymes of archaeal lipid biosynthetic pathway.

Name	Organism	Reaction	Kinetic parameters	Ref.
GGPP synthase	*Methano-bacterium thermoformicicum*	IPP+FPP → GGPP (60°C)	K_m IPP = 51.4 μM K_m FPP = 28 μM V_{max} = 437.9 nmol/min/mg	[44]
FPP/ GGPP synthase	*Thermococcus kodakaraensis*	IPP+DMAPP → GPP	K_m IPP = 23 μM, DMAPP = 9.5 μM	[22]
		IPP+GPP → FPP	K_m IPP =22 μM, GPP = 2.2 μM K_m IPP =3 μM, FPP = 1.7 μM	
		IPP+FPP → GGPP (70°C)	K_m IPP = 79 μM, DMAPP = 13 μM	
		IPP+DMAPP → GPP	K_m IPP =31 μM, GPP = 3 μM	
		IPP+GPP → FPP	K_m IPP =16 μM, FPP = 0.81 μM	
		IPP+FPP → GGPP (90°C)		
G1P dehydro-genase	*A. pernix*	DHAP+NADH → G1P (65°C)	K_m DHAP = 0.46 mM, k_{cat} (min^{-1}) = 154.25 K_m NADH = 0.032 mM, k_{cat} (min^{-1}) = 143.96	[29]
GGGP synthase	*M. thermo-autotrophicum*	G1P+GGPP → GGGP (55°C)	K_m G1P = 13.5 μM K_m GGPP = 0.51 μM k_{cat} (s^{-1})= 0.34	[45]
	T. acidophilum	G1P+GGPP → GGGP (55°C)	K_m G1P = 21.2 μM K_m GGPP = 75 nM k_{cat} (s^{-1})= 6.1, Vmax = 13.5 μmol/min/mg	[46]
CDP-archaeol synthase	*A. fulgidus*	DGGGP+CTP → CDP archaeol (65°C)	K_m DGGGP = 0.12 mM k_{cat} (s^{-1})= 0.55	[41]

and *B. subtilis* have a broad substrate specificity and can accept lipid derivatives from archaea or bacteria. On the other hand, the *E. coli* PS synthase was specific for bacterial lipid derivatives only [47]. Archaetidylinositol phosphate (AI) synthase catalyzes the reaction where precursors L-*myo*-inositol-1-phosphate and CDP-archaeol are converted to AI phosphate as an intermediate which is further dephosphorylated to AI [48]. This reaction is similar to the bacterial phosphatidylinositol phosphate (PI) synthase. Similar to AS and PS synthase, the AI and PI synthase show a broad substrate specificity accepting both, archaeal and

bacterial lipid derivatives as substrates [49]. Enzymes homologous to PS decarboxylase and phosphatidylglycerol (PG) synthase have been identified in archaea as AS decarboxylase and archaetidylglycerol (AG) synthase but not yet characterized biochemically [7].

Saturation of double bonds

The mature phospholipids of archaea exist in their fully saturated form. The archaeal enzyme digeranylgeranylglycerophospholipid reductase catalyzes the hydrogenation or saturation of the geranylgeranyl chains of unsaturated archaetidic acid (DGGGP) in a stereospecific manner [50]. It belongs to the geranylgeranyl reductase (GGR) family that includes GGR from plant and prokaryotes that are mainly involved in photosynthesis. Prenyl reductases other than GGRs are also found in all three domains of life and these enzymes catalyze the complete or partial reduction of isoprenoid compounds like respiratory quinones, tocopherol, dolichol and other polyprenols [51].

The structures of the archaeal GGR monomer from *Thermoplasma acidophilum* [50] and *Sulfolobus acidocaldarius* [52] show that they belong to p-hydroxybenzoate hydroxylase (PHBH) superfamily of flavoproteins (**Figure 4A**). The GGR from the thermophilic archaea *T. acidophilum* was crystalized in complex with flavin adenine dinucleotide (FAD) where FAD adopts the close confirmation that possibly changes with the binding of the substrate, like in other members of the PHBH family. The reduction of FAD is brought about by either NADH or other reducing agents. Since the protein overexpressed in *E. coli*, a surrogate lipid-like ligand assigned as phosphatidylglycerol (PGX) was found in the active site forming an imperfect fit to the substrate binding pocket. The lipid binding cavity of GGR is R shaped having two tunnels where the larger tunnel B is more permissive than the smaller tunnel A which is restricted in shape (**Figure 4B**). The *S. acidocaldaricus* GGR is structurally similar to GGR from *T. acidophilum* in FAD binding and the catalytic region (**Figure 4C**, **D**) but not in the C-terminal domain which is longer in *S. acidocaldaricus* GGR. The conserved sequence motif (YxWxFPx7-8GxG) lies in the large cavity of the catalytic domain and is thought to keep the substrate in position for the reduction reaction as also indicated by mutational studies. Although the

enzymes reduce GGGP, they also reduce the double bonds of related compounds like GGGP and GGPP [52]. Another study where the *Methanosarcina acetivorans* GGR was expressed in *E. coli* along with four previous genes of the archaeal lipid biosynthetic pathway, the DGGGP derivative with a fully saturated isoprenoid chain could be obtained [53]. Interestingly, the saturation only took place when GGR was coexpressed with a ferredoxin gene found upstream of GGR in the genome of *M. acetivorans*, the ferredoxin possibly functioning as a specific electron donor. However, no ferredoxin coexpression was required when the *M. acetivorans* GGR was replaced by *S. acidocaldaricus* GGR in the same study.

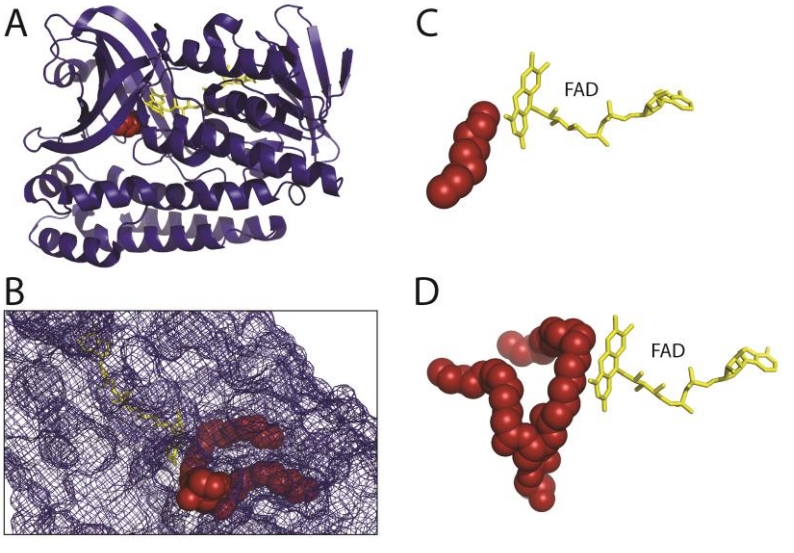


Figure 4: Structure of geranylgeranyl reductase (GGR). (**A**) The crystal structure of *S. acidocaldaricus* GGR at 1.85 Å resolution (PDB: 3ATQ) is shown [52] in complex with FAD (yellow). It is similar to the structure of GGR from *T. acidophilum* [50] sharing the FAD binding and the catalytic domain. Both the enzymes were crystallized where a lipid molecule (red) originating from the host organism was found in the substrate binding site (**B**) The lipid binding cavity of GGR is R shaped having two tunnels where the larger tunnel B is more permissive and the smaller tunnel A is restricted in shape. One edge is located at the branching point of the tunnel. (**C, D**) Comparison of the 'near perfect' placement of the hydrocarbons of lipid moieties on the re face of the FAD from *S. acidocaldaricus* GGR (**C**) and *T. acidophilum* GGR (**D**).

Also, the conservation of the ferredoxin gene upstream of GGR in other archaea was not analyzed.

It is not known at what step of the biosynthetic pathway, hydrogenation takes place. However since that CDP-archaeol synthase is specific for the unsaturated substrate, saturation probably takes place after the formation of CDP-archaeol. Although the enzyme AS synthase can accept both saturated and unsaturated substrates for catalysis, the detection of unsaturated AS in the cells of *M. thermautotrophicus* suggests that hydrogenation may already take place after the polar head groups are attached [2].

Tetraether formation

Tetraether (caldarchaeol) lipid structure with varying number (0 to 8) of cyclopentane moieties are widespread among archaea and a dominating membrane lipid structure in Crenarchaeota and Thaumarchaeota. Euryarchaeota synthesize archaeol or both archaeol and caldarchaeols. On the other hand, Thaumarchaeota have characteristic tetraether lipids with four cyclopentane moieties and a cyclohexane moiety [6]. One of the most intriguing steps in the archaeal biosynthetic pathway is the tetraether formation. *In vivo* studies suggested that tetraethers are formed from saturated diethers via head to head condensation reaction. Pulse chase and labeling experiment of *Thermoplasma acidophilum* cells with [^{14}C]-mevalonate showed that the label first incorporates into the archaeol until saturation and only then into caldarchaeol. When an inhibitor of tetraether lipid synthesis (terbinafine) is used, pulse labeling leads to the accumulation of diethers and this phenomenon can be reversed by removal of the inhibitor [54]. However in another study, radiolabelled archaeol was not incorporated into the tetraethers of *Methanospirillum hungatei* cells and the presence of double bonds was necessary for the incorporation of labeled DGGGP into the tetraethers of *M. thermoautotrophicus* cells [55,56]. The enzyme responsible for the formation of the presumed and unusual C-C bond for tetraethers has not been identified and there is no *in vitro* data to support this hypothesis [2]. Recently, an alternative pathway for tetraether and cyclopentane ring formation was hypothesized [6]. A multiple lock and key mechanism was

proposed owing to the 'greater functional plasticity' of the enzymes IPP synthase, GGGP synthase and DGGGP synthase so that they can accommodate prenyl substrates with a ring structure and chain length longer than C20. The cyclopentane rings would be formed early in the pathway before attachment of the glycerol moiety and the C20 geranyl molecules would couple together via a tail-to-tail mechanism to form the C40 phytoene chain by phytoene synthase, an enzyme that is wide spread in archaea. However, both possibilities still need to be experimentally demonstrated.

Physicochemical properties of archaeal lipids

Archaeal membrane lipid composition - response to stress

Within the archaeal lipids there is a great diversity varying in length, composition and configuration of the side chains (**Figure 5**). The most common archaeal core lipid is *sn*-2,3-diphytanylglycerol diether, generally called archaeol, which can undergo several modifications including hydroxylation and condensation. Elongated hydrocarbon chains (C20-C25) are found in some Halobacteriales [57] and methanogens (**Figure 5A**) [58]. A seemingly head-to-head condensation of two diether lipid molecules is one of the most frequent and functionally important structural variations that leads to a glycerol-dialkyglycerol tetraether lipid, known as caldarcheol. It should be stressed, however, that the enzymatic mechanism resulting in this lipid species is entirely unresolved. This core lipid is the most widespread in Archaea, characterized by different modifications depending on the archaeal species. It is in particular abundant among the phyla Euryarcheaota, Creanarchaeota and Thaumarchaetoa. Up to 8 cyclopentane moieties can be found in lipids of the Thermoplasmatales (**Figure 5B**) and in the Euryarchaeota phylum, in general [59–61]. Interestingly, the presence of cyclopentane and cyclohexane ring (**Figure 5E**) is a distinct feature of Thaumarchaeota leading to a structure known as creanarcheol (**Figure 5C**) [62,63]. In some thermoacidophiles and methanogens, a polar head group called nonitol is found which is composed of nine-carbon chain. Recent studies have revealed that the 9-carbon nonitol structure is often found as a

polyhydroxylated cyclopentanic form called calditol [64]. Therefore, these structure are now known as glycerol-dialkyl-calditol-tetraether (**Figure 5D**) and are the major component of the membrane of *Sulfolobales* species [65–67]. The presence of tetraether lipids and the ratio of diether/tetraether lipids varies depending on the archaeal species and also depend upon growth conditions [68]. Likewise, there is also a wide diversity of polar lipids in archaea including phospholipids, glycolipids, phosphoglycolipids, sulpholipids and aminolipids [64]. The occurrence of different polar head groups depends on the archaeal family and can be used as unique taxonomic marker [69]. Aminolipids, for instance, are prevalent in methanogens and are completely absent in halophiles and thermophiles [68]. Bacteria and Eukarya use several mechanisms to maintain the membrane fluidity over a range of temperatures such the regulation of fatty acid composition adapting the degree of branching, saturation and chain length [70]. The homeoviscous adaption theory states that the lipid compositions in the membrane varies in response to environmental stresses in order to preserve a proper membrane fluidity [71]. However, the exact changes in fatty acid composition in the membrane upon, for instance, a temperature shift differs from species to species. In *E. coli* the degree of fatty acid unsaturation increases along with a lowering of the temperature while some Bacillus species increase the amount of iso-fatty acids with the growth temperatures. The membrane also has to maintain its permeability barrier and in general it is believed that at the growth temperature, the lipid bilayer is a liquid crystalline state. The phase transition temperature at which the membrane is transferred from the crystalline into the liquid state is considered as a very important characteristic, and depends on the length of the hydrocarbon chains, the degree of saturation and the position of methyl groups; in Bacteria, the transition temperature ranges between -20 to up to 65°C compared to archaea where the range is much wider between even up to 100°C, a temperature where some archaea grow [3].

Archaea use different mechanisms to maintain the liquid crystalline phase over the entire growth temperature range. One control mechanism has been reported in the archaeon *Methanococcus jannaschii*, in which case the membrane properties are finely tuned by varying the ratio of diether,

macrocyclic diether and tetraether lipid [72]. In contrast, at higher temperatures, hyperthermophilic archaea may incorporate a higher degree of cyclopentane rings in their isoprenoid chain that increase the transition temperature. The presence of such lipid structures in Archaea is an indication of a need to preserve the membrane function at the hostile environmental conditions. In particular, the presence of tetraether lipids and the chemically stable ether bonds are major adaptions [73]. The latter confers resistance to phospholipases attack from other organisms. Despite the general thought that the isoprenoid chains of the ether lipid are involved in thermal resistance, it has been shown that they are not absolutely required for tolerance to high temperature. Archaeol and caldarcheol were also found in the thermophilic methanoarchaeota (65°C) and in mesophilic species (37°C) [74]. This suggests that the remarkable properties of the membranes of hyperthermophiles depend not exclusively on the tetraether composition of their lipids but that other aspects are involved as well. When fully stretched, the tetraether lipids span the entire membrane thickness leading to a monolayer, which is believed to stiffen the membrane in the presence of high growth temperatures [73,75]. This may also protect the membrane from possible lysis at high temperatures [73]. The presence of cyclic structures, in particular cyclopentene rings [76], is a hallmark for high growth temperatures and causes an increased membrane packing and thus a reduction in membrane fluidity [77]. However, also the characteristics of the polar head groups may influence membrane fluidity since the proper balance of negative and positive charges at the membrane surface is essential for its functioning. Therefore, varying the proportion of different polar head groups might be another way to response to external stresses [71]. Furthermore, modification of the polar head groups with carbohydrates increases hydrogen bonding between the lipids and thus will influence the stability of the membranes. In halophiles, the presence of glycerol methylphosphate attached to the archaeol moiety, contributes to the low membrane permeability under high salt concentration [78]. A further example is cold adaption in the psychrophile archaeon *Methanococcoides burtonii*, where an increase in the degree of unsaturation in the isoprenoid chains allow growth at low temperatures as they exist in glacial environments [71,79].

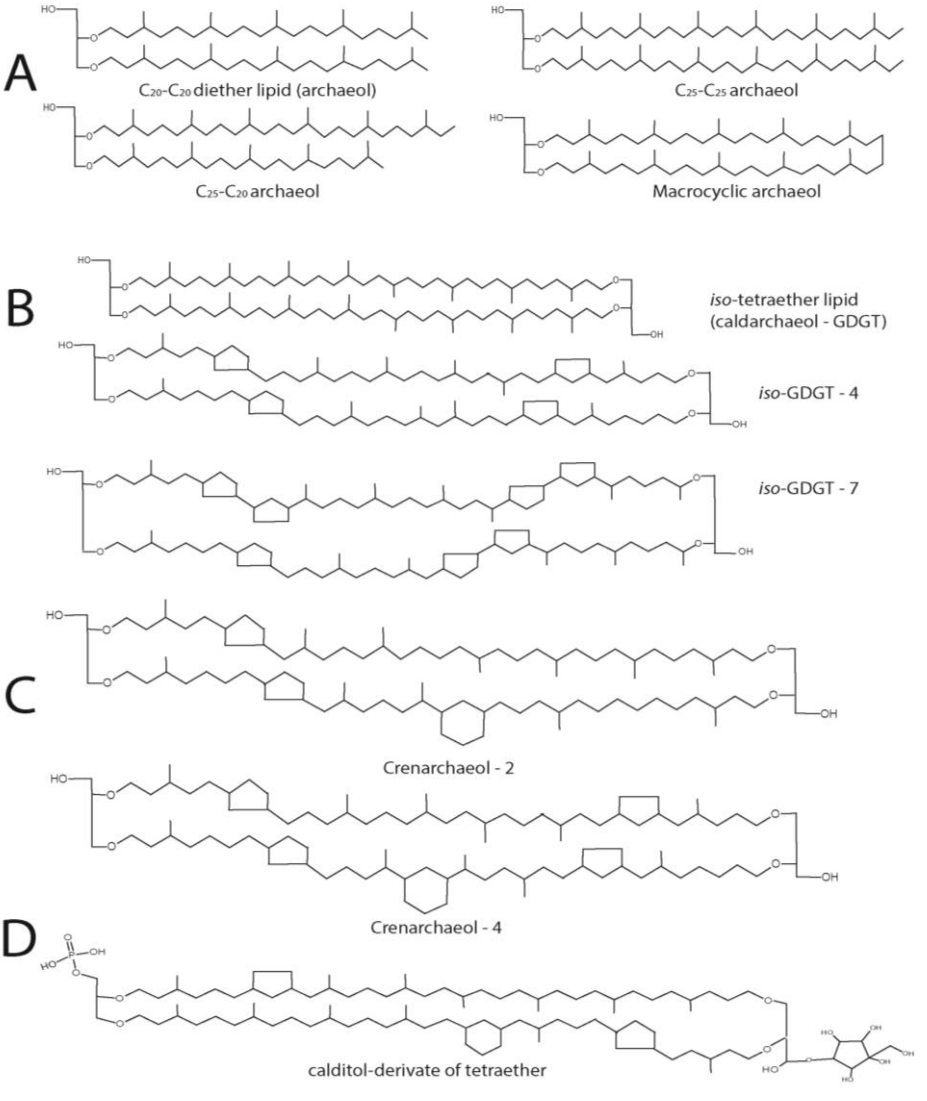


Figure 5| Structures of archaeal lipids. (**A**) Diether lipids (archaeol) with modifications in chain length and with macrocyclic ring structure are mostly found in Euryarchaeota (**B**) Tetraether lipids (caldarchaeol) can contain up to 8 cyclopentane ring moieties. (**C**) Crenarchaeol lipid structure with cyclopentane and cyclohexane rings found only in Thaumarchaeota. (**D**) Calditol derivative of tetraether lipid found in *Sulfolobus spp.*

In vitro based studies on the stability of archaeal lipids

Temperature impacts the membrane properties, influencing the ion permeability and phase behavior. It is believed that the unique membrane organization of archaeal tetraether lipid in a monolayer structure along with the presence of ether bonds, renders such membranes thermal resistant. Therefore, many studies have been conducted to understand the higher thermal stability of archaeal membrane lipids compared to the bacterial phospholipids. By comparing liposomes made of a polar lipid fraction from *S. acidocaldarius* and liposomes prepared from a bacterial lipid (POPC) or a synthetic lipid with a phytanyl chain (DPhPC), the importance of the methyl branched isoprenoid chain in membrane stability has been examined [80]. When incubated at 100°C for approximately half hour, archaeal liposomes showed a very low ion leakage compared to POPC and DPhPC liposomes that collapse after a few minutes incubation at that temperature. Likewise, only very slow release (8-10%) of the fluorophore carboxyfluorescin (CF) was observed with liposomes composed of *S. acidocaldarius* lipids compared to *E. coli* liposomes (50%) over a time interval of 62 days, while liposomes composed of a lipid extract from the thermophile bacterium *B. stearothermophilus* showed an intermediate stability [81].

The extremely high heat tolerance of archaeal liposomes open opportunities for biotechnological applications with the ability to be stable even after several autoclaving cycles, which might be exploited for biomedical uses. Autoclaving is a common and effective method for decontamination and the possibility to autoclave the archaeosomes vesicle for the drug delivery gives new prospects for the liposomes formulation. The effective archaeosomes stability against autoclaving was tested showing a remarkable strength against 2-3 cycles of autoclaving at pH 4.0-10.0 [82]. Besides the high heat tolerance, archaeal lipids are known to be resistant to conditions of extremely low pH and their low proton permeability contributes to maintaining a constant intracellular pH. Monolayer liposomes reconstituted from the lipid fraction of *S. acidocaldarius* indeed exhibited very low proton permeability even at higher temperatures [81,83]. However, at acidic conditions the archaeosomes appeared less resistant to autoclaving possibly because of a

higher protonation of the polar head groups, influencing hydrogen-bonds formation among the lipids. Liposomes containing long sugar chains linked to the phospholipids show much lower proton permeability than liposomes composed of lipids with only one sugar unit. The glycolipids amount in the polar lipid fraction of *Thermoplasma acidophilum* increases at lower pH and this seems a general mechanism for acidophiles against the chemically unstable conditions [84]. The low permeability of such liposomes can be exploited for drug-delivery with the added value of a high resistance against phospholipase A2, B and pancreatic lipase [85]. Due to their high pH tolerance they can easily pass the gastro-intestinal tract without damage. Overall, these studies confirm that the presence of tetraethers in archaeal membranes confers these membranes with a remarkable stability against high temperatures, low pH, and high salt. The degree of hydrocarbon chain saturation, the features of the polar head groups and the presence of cyclopentane ring [86] appear of secondary importance in providing stability.

The lipid divide

Prospect of LUCA with mixed membrane lipid composition

During the last decades many theories have been proposed about the origin of life and how the differentiation between the three domains of life occurred. All of these theories accept the existence of the last universal common ancestor (LUCA), also known as LUCAS or cenancestor from which organisms have diverged. Particular attention has been given on the membrane composition of LUCA. Isoprenoids are involved in a wide range of functions, and found both in archaea and bacteria. Whereas fatty acid metabolism is also widely distributed, fatty acid biosynthesis seems underdeveloped in archaea and in some organisms even absent. Therefore, one of the hypotheses is that early life forms were dependent on the presence of membrane lipids with isoprenoid hydrocarbon core. However, the most divergent feature that is at the base of the Lipid Divide is the glycerophosphate backbone. Archaea contains G1P as glycerophosphate moiety while bacteria depend on G3P. These two compounds are

synthetized by two different enzymes that are not evolutionary correlated [87].

Thus, different hypothesis were suggested to understand the process that brought the ancestor to archaeal, bacterial and eukaryotic organisms. According to Koga *et al.* [87] the evolution of the two phospholipid pathways occurred independently leading to the simultaneous appearance of bacterial and archaeal enzymes. This contrast the theory of Martin and Russel [88], according to whom the two different organisms evolved from an ancestor characterized by iron monosulphide compartment. Another hypothesis [89] proposed a three stage process from the pre-cell to the eukaryotic cell. It was suggested that there was a cenancestor with a chemically derived heterochiral membrane containing both the enantiomeric forms of the glycerophosphate backbone, which slowly diverged into a more stable homochiral membrane organism leading to the differentiation between Bacteria and Archaea. This idea implies that such heterochiral mixed membrane are intrinsically unstable leading to the emergence of chiral selective enzymes and a divide between bacteria and archaea. However, hybrid membranes were tested for their stability using a mixture of egg phosphatidylcholine and extracted lipid from *Solfolobus solfataricus* [90] and a higher stability compared to liposomes reconstituted of only archaeal lipid was observed. Moreover a heterochiral membrane consisting of bacterial G3P lipids and archaeal G1P lipids was analyzed [91] and the heterochiral membranes were found to be more stable at higher temperatures compared to liposomes prepared from only the bacterial lipids. Thus, based on these studies, the hypothesis of the existence of an ancestor with both the G1PDH and G3PDH would be possible but there must have been other factors that have driven the segregation into the two different domains. The theory of the coexistence of both the enzymes for the glycerophosphate enantiomers productions can be also extended to the other enzymes involved either in the synthesis of isoprenoid and fatty acid based phospholipids (**Figure 6**). In LUCA, four different membrane lipids may be obtained by the combination of the two glycerophosphate backbones with either isoprenoid or fatty acid hydrocarbon chains [43]. Environmental pressure and the need to adapt to extreme conditions may have induced archaea to evolve their membranes

[4,92]. In particular, the use of different hydrocarbon chains in response to the growth environment may have induced the segregation in Archaea and Bacteria and resulted in homochiral membrane formation [93].

Differentiation of membrane lipids in archaea and bacteria

The structural variability found in the membrane lipid composition of archaea and bacteria reflects differences and similarities in the respective biosynthetic pathways. When compared with the well-characterized bacterial ester-lipid biosynthetic pathway, several similarities with the

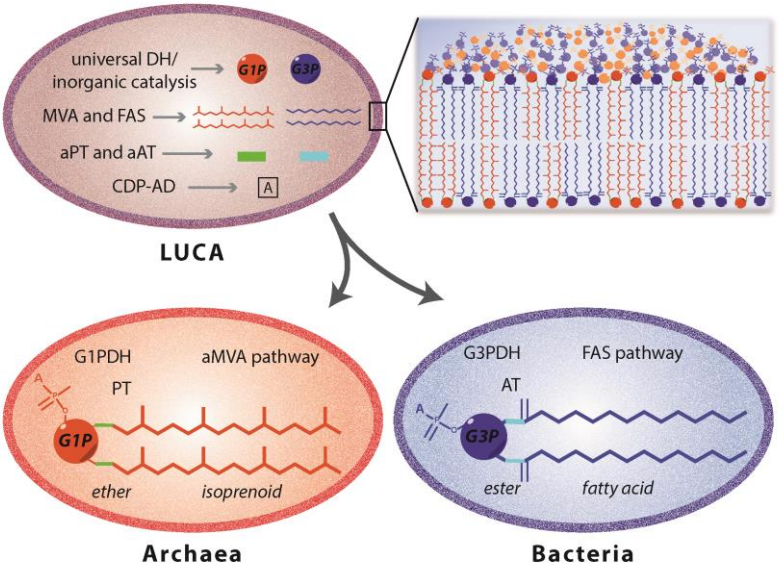


Figure 6| Proposed evolutionary model for Archaea and Bacteria differentiation. The existence of a cenancestor (LUCA) embordered by a mixed membrane is one of the main theories of organism's evolution. This primordial organism would be characterized by the presence of heterogeneous enzymes involved in the synthesis of archaeal and bacterial lipids. Enzyme separation occurred by environmental pressures leading to the differentiation of bacteria with a phospholipid biosynthetic pathway [94] and archaea with ether-lipid biosynthesis. The abbreviations are: DH: dehydrogenase; MVA: mevalonate pathway; aMVA: alternate mevalonate pathway; FAS: fatty acid synthesis pathway; PT: prenyltransferases; AT: acyltransferases; 'a' used as prefix for PT and AT: ancestral; CDP-AD: CDP-alcohol phosphatidyltransferase family of enzymes.

archaeal ether-lipid biosynthesis are evident. First, the sequence of reactions that yield the final membrane lipid from the building blocks is basically the same even though some enzymes involved in these reactions are equipped with specific features to the lipid of the two different domains. Second, the glycerophosphate backbone in both cases is synthetized by a reduction of DHAP at the 2-OH using NADH as cofactor while the two hydrocarbon chains are linked to the same position on the glycerol moiety. Third, the polar head attachment takes place via a CDP-activated intermediate [93]. In particular the peculiar features that distinguish archaeal lipids from bacterial ones occur in the first half of the biosynthetic pathway while the last stages, which involve the replacement of the CDP group with one of the polar groups, is essentially similar in these two domains of life. For the latter reactions, the enzymes belong to the same superfamily and share sequence similarity [7,43]. On the other hand, the isoprenoid building blocks synthesis differs in Bacteria and in Archaea since it takes place via two different pathway, the DOXP (1-deoxy-D-xylulose 5-phosphate) and the alternate MVA (mevalonate) pathway, respectively [2]. The other striking difference involved in the lipid divide is the enantiomeric conformation of the glycerophosphate backbone along with the saturation of the double bonds present on the isoprenoid chains and further modification of the hydrocarbon chains. The remarkable similarities of both biosynthetic pathways are an indication of the existence of a common ancestor with promiscuous enzymatic composition that sufficed for the synthesis of heterochiral membrane lipids.

Reversing evolution - synthesis of archaeal lipids in bacteria?

Did a complex heterochiral membrane ever exist? There are several speculations about the membrane lipid composition of LUCA [4] and one evolutionary way to approach this is to design a cell which has a heterochiral membrane. The prospect of engineering *E. coli* with membranes harboring archaeal lipids has been initiated in several studies. In the first study, five genes were overexpressed in *E. coli*, four from a hyperthermophilic archaeon *A. fulgidus* (G1P dehydrogenase, GGPP synthase, GGGP synthase and DGGGP synthase) and one from *E. coli* (IPP isomerase) [21]. The enzyme IPP isomerase catalyzes the isomerization of

IPP and DMAPP. Together with GGPP synthase, the carbon flux could be increased towards the synthesis of GGPP building blocks which had been demonstrated previously for carotenoid production in *E. coli* [95]. The activity of each enzyme was monitored by different methods. IPP isomerase and GGPP synthase were analyzed by lycopene based calorimetric assay, G1P dehydrogenase activity was measured spectrophotometrically to detect DHAP dependent NADH oxidation reaction and using radiolabeled substrates, GGGP synthase activity was measured by thin layer chromatography (TLC) and HPLC. To detect the formation of DGGGP in *E. coli*, lipids were extracted and dephosphorylated from a 24hr growing culture and analyzed by liquid chromatography mass spectrometry (LC-MS/MS) using electron spray ionisation (ESI-MS). This study showed that indeed the archaeal lipid biosynthesis pathway is functional in *E. coli*. Unlike G1P dehydrogenase and GGGP synthase, DGGGP synthase protein could not be detected and the amount of DGGGP formed was not quantitated.

In another study, the same four genes but now from the methanogen *M. acetivorans* were chosen as it grows at 37°C and are readily overexpressed in *E. coli* [40]. Activity of all the enzymes was monitored in a TLC based assay using radiolabeled substrates. The lipids were extracted from the recombinant *E. coli* cells and unlike in previous study, lipids were not dephosphorylated but analyzed directly using LC/ESI-MS. Interestingly, the result showed that DGGGP synthesized by the archaeal enzymes was further metabolized by *E. coli* endogenous enzymes to form a phosphatidylglycerol-type derivative of DGGGP which was named DGGGP-Gro. Two speculations were made regarding the endogenous enzymes that might have brought about the reaction. If it emerged from the phospholipid biosynthetic enzymes of *E. coli*, it would require three enzymes to recognize and accept the archaeal substrate, namely CDP-diacylglycerol synthase, *sn*-glycerol-3-phosphate transferase and the phosphatase. However, addition of CTP to their *in vitro* TLC based assay did not increase the amount of product formation and no other polar head group attachment was observed. The other possibility is that the *sn*-1-phosphoglycerol group from the osmoregulated periplasmic glucans was transferred to DGGGOH directly by the phosphoglyceroltransferase

system. The estimated amount of total archaeal membrane lipids extracted from *E. coli* cells in this study was only 60 μg/g wet cells and at these levels it is not possible to study the influence of archaeal lipids on the physical properties of the cytoplasmic membrane. In a follow up study, geranylgeranyl reductase and ferredoxin (see section saturation of double bonds) were introduced along with the other genes from *M. acetivorans* and expressed in *E. coli*. The formation of DGGGP-Gro was reduced yielding mostly saturated archaetidic acid in *E. coli* [53].

The archaeal lipids have also been synthesized *in vitro* using a set of five purified enzymes, two from bacteria and three from archaea. A mutant of FPP synthase of *E. coli* was used that was shown previously to synthesize GGPP, G1P dehydrogenase was from *B. subtilis*, GGGP synthase from methanogen *M. maripaludis*, DGGGP synthase and CarS from the hyperthermophilic *A. fulgidus*. All enzymes were purified, and by using substrates IPP, FPP, DHAP and NADH, the enzymes were shown to be able to synthesize CDP-archaeol in the presence of CTP, Mg^{2+} and detergent at 37°C [41]. The feasibility of synthesizing archaeal lipids in *E. coli* and *in vitro* are promising first steps towards deciphering the biosynthetic pathway further and eventually understanding the properties of a cell with a heterochiral membrane lipid composition.

Future Challenges

Although during the last decade, many of the intimate features of the archaeal lipid biosynthesis pathway have been resolved, there are still several important questions that need to be answered. Understanding the mechanism of tetraether formation and identifying the enzyme(s) involved in the reaction requires thorough investigations. *In vitro* analysis of such a reaction(s) would be great advancement in the field. Various other derivatives of diether and tetraether lipids like cyclopentane and macrocyclic ring formation, glycosylation and formation of crenarchaeol are also not well understood. The pathway refractory of archaeal lipid biosynthesis in *E. coli* is currently incomplete where the amount of archaeal lipids formed in comparison to the *E. coli* lipids is very low. Challenging aspect is to modulate and suppress the endogenous pathway and integrate the archaeal lipid biosynthesis pathway in the genome of the

host stable yielding the exclusive formation of these lipid species. Further structural and biochemical analysis of the enzymes of the pathway from different families of archaea would progress the field and bring it in par with the understanding of bacterial phospholipid biosynthesis. Also, regulation of the phospholipid metabolism is poorly understood and could be enhanced through the use of genetic studies, which now became feasible because of the rapid developments in genetic toolbox in archaea. However, to study essential genes using these techniques is still a challenge.

Recent studies have shown that in spite of the uniqueness of the archaeal membrane lipid structure, they are not as distinct as previously thought. The presence of fatty acids and isoprenoids in the three domains of life and the common mode of polar head group attachment in bacteria and archaea, and the presence of homologues of archaeal G1P dehydrogenase and GGGP synthase in bacteria are a few of the similarities. Other question that still remains unanswered is the exact reason why G1P is ether linked to the isoprenoid hydrocarbon chains and G3P is linked via ester bonds to the fatty acid chains; are there still organisms with such mixed membranes? In order to answer such questions, more biochemical and functional investigation are needed on the archaeal lipid biosynthetic pathway along with a deep phylogenetic analysis.

Acknowledgements

S.J. and A.C are supported by the research program of the biobased ecologically balanced sustainable industrial chemistry (BE-BASIC). We thank Ilja Kusters, John van der Oost, Melvin Siliakus and Servè Kengen for fruitful discussions.

Scope of the thesis

This thesis focusses on a further understanding of archaeal lipid biosynthetic pathway, and to use this knowledge to develop novel avenues for lipid engineering of the bacterial membrane.

Chapter 1 provides a general introduction on the biosynthesis of the archaeal membrane ether lipids. In particular, it discusses current knowledge on the enzymatic steps that lead to the biosynthesis of archaeal phospholipids from simple building blocks like glycerol 1-phosphate and isoprenoid chains. The diversity of ether lipid composition in different archaeal species is discussed with particular attention to the changes in membrane composition in response to the environmental stress which many of the extremophilic archaea encounter. The phylogenetic implications of the lipid membrane composition of the last universal common ancestor (LUCA) and the evolution of the phospholipid biosynthesis enzymes in the differentiation between bacteria and archaea are shortly presented.

Chapter 2 presents an investigation on a missing step in the archaeal lipid biosynthetic pathway. The gene encoding for CDP-archaeol synthase (CarS) was discovered by bioinformatics analysis involving remote sequence similarity with the analogous bacterial CDP-diacylglycerol synthases. The CarS protein was biochemically characterized for the formation of the product CDP-archaeol in presence of the chemically synthetized substrate DGGGP and CTP. Finally, the entire ether lipid biosynthetic pathway was reconstituted *in vitro* employing five purified bacterial and archaeal enzymes.

Chapter 3 further characterizes the ether lipid biosynthesis by a detailed study on the final steps in phospholipid biosynthesis: polar head group attachment. Various bacterial enzymes involved in the replacement of the CDP-moiety from CDP-archaeol with L-serine and glycerol were investigated for their substrate promiscuity towards this archaeal substrate. Using a combination of nine purified archaeal and bacterial enzymes, the two unsaturated archaeal lipids archaetidyglycerol (AG) and archaetidylethanolamine (AE) were synthetized *in vitro*. Moreover, the

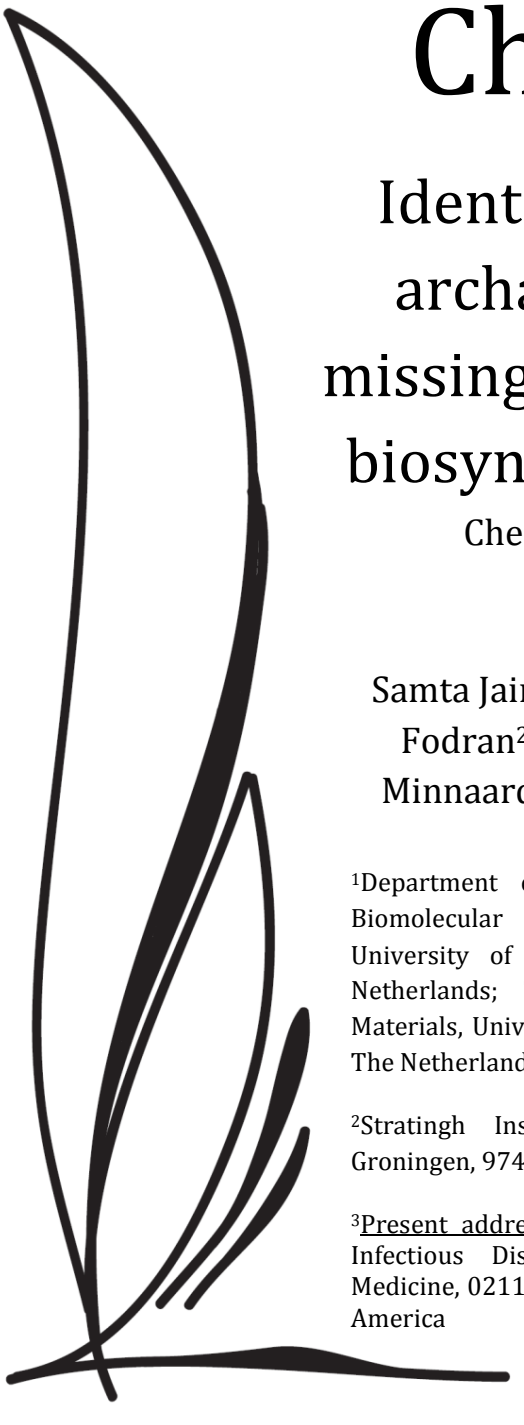
pathway was introduced in the bacterium *Escherichia coli* and biosynthesis of the archaeal AG and AE could be observed in a living cell.

Chapter 4 presents an extensive engineering approach to generate a hybrid heterochiral membrane in the bacterium *Escherichia coli*. By using a combination of metabolic engineering to boost isoprenoid biosynthesis and heterologous expression of the ether lipid biosynthetic pathway, a strain was obtained which produces a substantial (up to 32% of total phospholipid) amount of archaeal lipids among the bacterial lipidome. In addition, one of the key enzymes of the archaeal lipid biosynthetic pathway exhibits a remarkable promiscuity for glycerol-phosphate shed new light on the lipid divide process. Finally, the cells with a hybrid heterochiral membrane show altered cell growth, morphology and gained a certain degree of robustness towards environmental stress.

Chapter 5 presents the crystal structure of one of the key enzymes of the ether lipid biosynthetic pathway, CarS from the archaeon *Aeropyrum pernix* in a complex with CTP and Mg^{2+} at 2.38 Angstroms (Å). The data defined the molecular basis of CTP recognition by the CarS enzyme and sheds new light on the catalytic mechanisms of the CTP-transferase family enzymes. In particular, the CarS structure shows an unusual fold of five transmembrane segments that form a charged cavity at the cytoplasmic side of the membrane that accommodates the CTP. Binding of CTP and Mg^{2+} stabilize the enzymatic complex allowing the formation of two hydrophobic grooves for lipid substrate binding.

In **Chapter 6** a short description of the enzymatic mechanism of the saturation of double bonds in the isoprenoid chains is presented. Geranylgeranyl reductases (GGR) from two different archaeal organisms, *M. acetivorans* and *M. maripaludis*, were tested for their hydrogenation activity. Total lipid analysis of the engineered *E. coli* strains harboring a different combination of the enzymes of the ether lipid biosynthetic pathway was performed by LC-MS in order to detect the presence of the saturated archaeal AG and AE species.

Chapter 7 summarizes the work presented in the thesis and presents an outlook.



Chapter 2

Identification of CDP-archaeol synthase, a missing link of ether lipid biosynthesis in Archaea

Chemistry & Biology 2014, 21(10):1392-401

Samta Jain[1,3], Antonella Caforio[1], Peter Fodran[2], Juke S. Lolkema[1], Adriaan Minnaard[2] and Arnold J. M. Driessen[1]

[1]Department of Molecular Microbiology, Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, 9747 AG Groningen, The Netherlands; The Zernike Institute for Advanced Materials, University of Groningen, 9747 AG Groningen, The Netherlands

[2]Stratingh Institute for Chemistry, University of Groningen, 9747 AG Groningen, The Netherlands

[3]Present address: Department of Medicine, Section of Infectious Diseases, Boston University School of Medicine, 02118 Boston, Massachusetts, United States of America

Abstract

Archaeal membrane lipid composition is distinct from Bacteria and Eukarya, consisting of isoprenoid chains etherified to the glycerol carbons. Biosynthesis of these lipids is poorly understood. Here we identify and characterize the archaeal membrane protein CDP-archaeol synthase (CarS) that catalyzes the transfer of the nucleotide to its specific archaeal lipid substrate, leading to the formation a CDP-activated precursor (CDP-archaeol) to which polar head groups can be attached. The discovery of CarS enabled reconstitution of the entire archaeal lipid biosynthesis pathway *in vitro*, starting from simple isoprenoid building blocks and using a set of five purified enzymes. The cell free synthetic strategy for archaeal lipids we describe opens opportunity for studies of archaeal lipid biochemistry. Additionally, insight into archaeal lipid biosynthesis reported here allow addressing the evolutionary hypothesis of the lipid divide between Archaea and Bacteria.

Introduction

Glycerol linked hydrocarbon chains constitute the cellular membrane lipid composition of all living organisms. However, the versatility of this composition fundamentally distinguishes Archaea from Bacteria and Eukarya. The bacterial and eukaryotic membrane lipids are composed of an *sn*-glycerol-3-phosphate (*sn*G-3-P) backbone esterified to mostly linear fatty acids. On the other hand, the archaeal membrane lipids are characterized by an *sn*-glycerol-1-phosphate (*sn*G-1-P) backbone etherified to linear isoprenoids. This membrane lipid divide is considered evolutionarily very significant, and implicated in the differentiation of archaea and bacteria. It is not known why and how this differentiation occurred, and whether the two phospholipid biosynthesis pathway originated independently or from an ancestral cell with a heterochiral membrane lipid composition [4].

Understanding the biosynthetic pathways leading to the formation and regulation of membrane lipid composition would help decipher the unknown aspects of early evolution. Although ether phospholipids or fatty acid based phospholipids are also found in a few bacteria and eukarya [96], they differ from the isoprenoid derived archaeal ether lipids. The unique structure of archaeal membrane lipids is believed to be vital for the adaptation of these organisms to the extreme environmental conditions [2,97] but the basic lipid architecture is found in all archaea including the mesophilic Thaumarchaea. Among archaea, a great diversity of lipids exists derived from the basic diether structure *sn*-2,3-diphytanylglycerol diether called archaeol. A frequent modification of archaeol is the formation of the dimeric tetraether structure called caldarchaeol which is prevalent in hyperthermophilic archaea and mesophilic Thaumarchaeota [10]. Due to its unique membrane lipid composition, studies on the lipid biosynthetic pathway of archaea (**Figure 1**) are of particular interest. Crucial steps of the pathway from the isoprenoid biosynthesis leading to the formation of unsaturated archaetidic acid have been investigated in detail. However, certain key enzymes [6,8] have not yet been identified, precluding a complete reconstitution of the pathway, while *in vitro* studies are hampered by the difficulty to acquire specific substrates.

In archaea, the biosynthesis of the isoprenoid building blocks isopentenyl pyrophosphate (IPP) and dimethylallyl pyrophosphate (DMAPP) occurs via the mevalonate pathway [10]. The sequential condensation of IPP and DMAPP leads to the formation of an appropriate isoprenoid chain length of either geranylgeranyl (C20) and/or farnesylgeranyl (C25) diphosphate, involving enzymatic steps catalyzed by geranylgeranyl diphosphate (GGPP) synthase and farnesylgeranyl diphosphate (FGPP) synthase, respectively [19,44]. These enzymes belong to a wide-spread family of *E*-isoprenyl diphosphate synthases [16] that is also found in bacteria and plays a role in the biosynthesis of quinones and pigments [16,98]. The *sn*G-1-P backbone is synthesized by the enzyme glycerol-1-phosphate dehydrogenase (G1PDH) that converts dihydroxyacetone phosphate (DHAP) to *sn*G-1-P using nicotinamide adenine dinucleotide (NADH) or the reduce form of nicotinamide adenine dinucleotide phosphate as coenzyme [28,29]. The archaeal G1PDH enzymes share sequence similarity with alcohol and glycerol dehydrogenases, but belong to a different enzyme family than bacterial glycerol-3-phosphate dehydrogenases [2].

The two ethers that link the isoprenoid chains to carbon 3 and 2 of the *sn*G-1-P backbone are formed by two enzymes of the prenyl transferase family. The first enzyme is a cytoplasmic geranylgeranylglyceryl phosphate synthase (GGGP synthase) and the second enzyme is the membrane bound digeranylgeranylglyceryl phosphate synthase (DGGGP synthase). GGGP synthases are conserved in archaea, but also occur in some bacteria, they are phylogenetically divided into two groups [35,46]. They fulfill an evolutionarily central reaction by mediating the three characteristic features of the ether lipid structure, selectively joining the *sn*G-1-P enantiomer, rather than *sn*G-3-P, to the isoprenoid chain via an ether linkage. The crystal structure of the Group I GGGP synthase from *Archaeoglobus fulgidus* shows a dimeric structure with a TIM-barrel fold [36] bound to *sn*G-1-P, while the archaeal group II GGGP synthases show a similar structure but they form higher order oligomers with modified G1P binding pocket [35,45]. The mechanism forming the product 3-*O*-geranylgeranylglyceryl phosphate (GGGP) leads to the release of pyrophosphate and is Mg^{2+} dependent [36,99]. The second ether bond is formed by the membrane protein DGGGP synthase that belongs to the

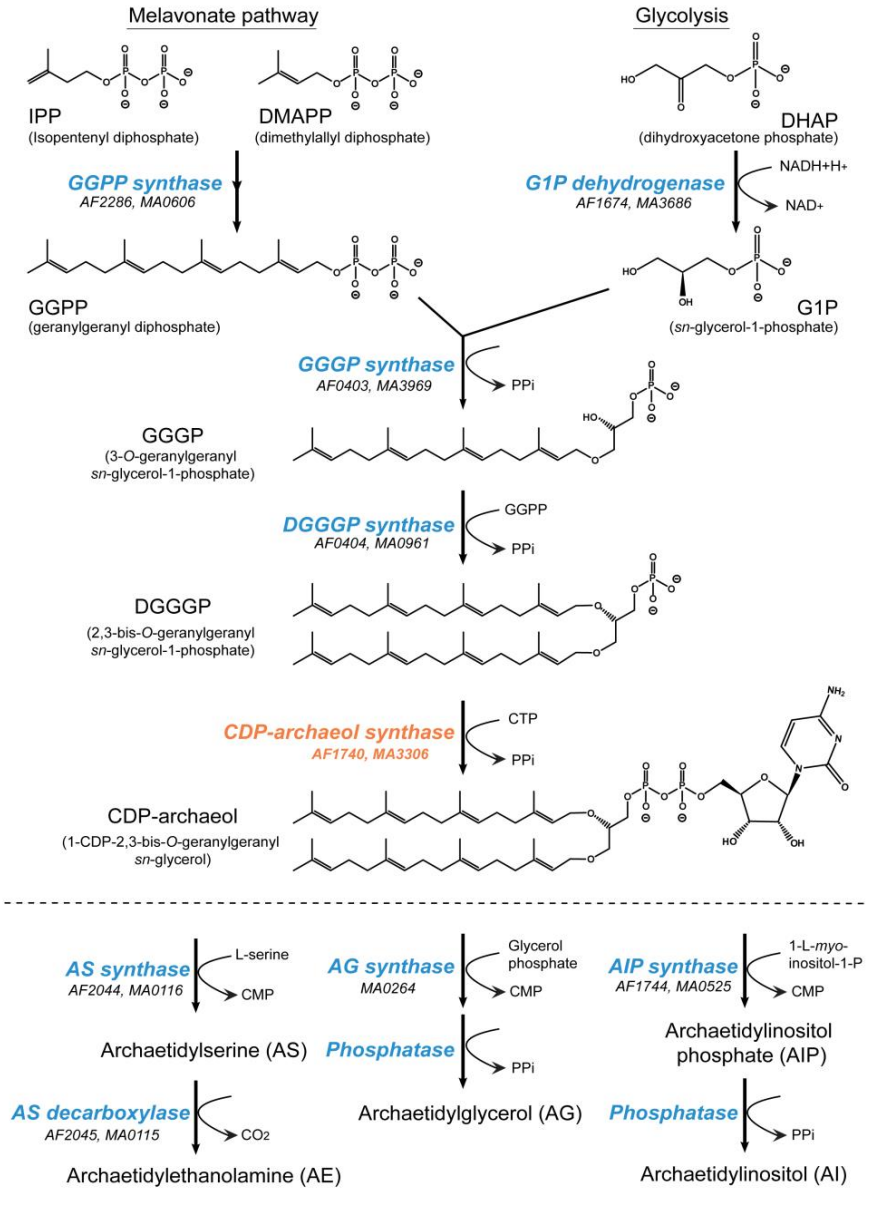


Figure 1| Archaeal lipid biosynthetic pathway. Enzymes of the pathway are colored in blue. CDP-archeaol synthase colored in orange is as identified in this study. Below the dashed line is the proposed pathway for polar head group attachment. The genes from Euryarchaeota *A. fulgidus* and *Methanosarcina acetivorans* are indicated. The enzyme for the hydrogenation of the double bonds (AF0464 and MA1484) is not included since it is unclear at which step of the pathway the hydrogenation occur.

UbiA prenyltransferase family. DGGGP synthase of *Sulfolobus solfataricus* was expressed in *E. coli*, and the enzymatic activity of the purified protein was found to be specific for the substrates GGPP and GGGP [38] synthesizing the product 2,3-di-*O*-geranylgeranylglyceryl phosphate (DGGGP). In a later study, DGGGP synthase was shown to accept both the S and R form of GGGP thus being rather enantio unselective [39].

For polar head group attachment, a CDP activated precursor is required. In Bacteria and Eukarya, the analogous reaction is catalyzed by the enzyme CDP-diacylglycerol synthase that binds the substrates CTP and phosphatidic acid and releases the products PP_i and CDP-diacylglycerol [100,101]. CDP-diacylglycerol is the central intermediate in the biosynthesis and regulation of phospholipids in Bacteria [102,103] and in Eukarya [104–106], in particular for the biosynthesis of phosphatidylinositol, phosphatidylglycerol, and in some organisms phosphatidylserine [107]. It is believed that in Archaea a similar reaction takes place in which CDP is transferred to the unsaturated archaetidic acid DGGGP, forming CDP-archaeol. Biochemical studies using membrane fractions of the archaeon *Methanothermobacter thermoautotrophicus* indicated an activity wherein [^{3}H]-CTP was incorporated into a lipid extractable fraction in the presence of DGGGP. The reaction was found to be Mg^{2+} dependent and was specific for the archaetidic acid substrates. Trace amounts of the compound CDP-archaeol were detected in cells of *M. thermoautotrophicus* [42]. Although this suggests the occurrence of CDP-archaeol synthase activity in the organism, the gene encoding the enzyme in *M. thermoautotrophicus* or other archaea has remained elusive.

Here we report the identification and characterization of the key enzyme CDP-archaeol synthase of the ether lipid biosynthesis pathway that is ubiquitously present in Archaea. In conjunction with the other enzymes of this pathway, the synthesis of CDP-archaeol could be reconstituted *in vitro* using purified enzymes and simple building blocks.

Results

Bioinformatics analysis of putative CDP-archaeol synthase

Except for the predicted CDP-archaeol synthase, all enzymes involved in the formation of archaeol have been identified, including enzymes for the hydrogenation of the double bonds [53] and the polar head group attachment (**Figure 1**). An National Center of Biotechnology Information-BLAST analysis was carried out using the bacterial CDP-diacylglycerol synthase (CdsA) as query against the archaeal kingdom to identify candidate proteins with possible CDP-archaeol synthase activity. Further analysis of the retrieved sequences resulted in an extended list of conserved hypothetical proteins. These proteins belong to a family of unknown function (DUF46), contain several predicted transmembrane segments and are found in all archaea except the phylum Nanoarchaeota and in three families of the phylum Thaumarchaeota. The domain of unknown function (DUF) region of the putative CDP-archaeol synthase spans more than two thirds of the entire protein sequence. Multiple sequence alignment indicates conserved residues mostly in the predicted cytoplasmic loop regions (C1 and C2) and the extreme C-terminus while the sequence conservation at the amino terminus is low (**Figure S1**). The DUF46 family is grouped in the same clade (CTP-transferase like superfamily of the Pfam database) as the CTP-transferase 1 family that encompasses the bacterial CdsA protein. This indicates sequence or structure conservation among the two families as also recognized in a recent bioinformatics study [8].

To this end, a secondary structure analysis was performed by aligning the family averaged hydropathy profile of the CdsA and DUF46 families. Hydropathy profiles are evolutionarily better conserved than the primary amino acid sequence, providing a measure of the structural similarity of a membrane protein family [108]. The alignment of the archaeal family profile showed a common pattern with five predicted transmembrane domains (TMDs) and an extracellular amino-terminus (**Figure 2A**). The bacterial CdsA sequences are longer than the archaeal ones, containing additional TMDs at the N-terminal end that are poorly conserved. However, six bacterial sequences were retrieved that are shorter and

display homology to the archaeal sequences (**Figure S1**). Interestingly, the alignment of the archaeal with the bacterial hydrophobicity profile indicates overlapping transmembrane segments and cytoplasmic loop regions at the carboxy-terminal half of the proteins (**Figure 2B**), in spite of low sequence similarity. Analysis of cytoplasmic loop C2 revealed a consensus sequence (G/S)Dxxx(S/A)xxKR conserved between the two families (**Figure 2C**). The hypothetical proteins displayed no conserved gene context and only a single copy per genome was retrieved for each sequence. The phylogenetic tree distribution (**Figure S2**) and a detailed list of the hypothetical proteins from all the sequenced genomes of archaea can be found in **Table S1**. The bioinformatics analysis suggests that the hypothetical protein may function as a CDP-archaeol synthase designated as CarS.

Identification and characterization of CDP-archaeol synthase

The amino acid sequence of CarS from *Archeoglobus fulgidus* was codon optimized for the overexpression in *E. coli*. An octa-histidine tag was introduced at the C-terminus of CarS and the protein was overexpressed in *E. coli* Lemo21-DE3 strain under the control of T7 promoter, and purified by Ni-nitrilotriacetic acid (NTA) affinity chromatography after solubilization of the membranes with the detergent n-dodecyl-β-D-maltopyranoside (DDM). On an SDS-PAGE gel, CarS showed a slightly anomalous running behavior than its theoretical molecular weight as expected for a polytopic membrane protein (**FigureS3 and 5A**). The protein was additionally identified by peptide mass fingerprinting using liquid chromatography-tandem mass spectrometry (LC-MS/MS) to confirm the amino acid sequence (**Figure S3**) and western blotting using α-his antibody (**Figure S3**). To examine the activity of the purified CarS protein, we chemically synthesized the substrate DGGGP and confirmed its structure using nuclear magnetic resonance (NMR) as described in **Supplemental experimental procedure**. In an *in vitro* reaction using purified CarS, the formation of the product CDP-archaeol from the substrate DGGGP was observed in the presence of CTP and Mg^{2+} by LC-MS (m/z= 1020.55 $[M-H]^-$), while the product was absent in the EDTA (+) control reaction (**Figure 3A and S3**). The nucleotide specificity of CarS for

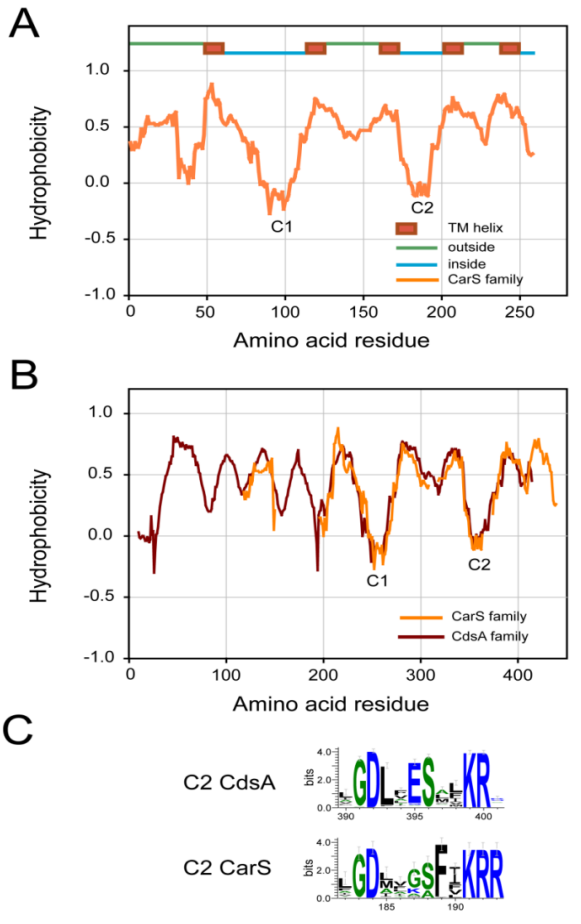


Figure 2| Hydrophobicity profile of the family of CarS and CdsA proteins. (**A**) The averaged hydropahty profile (orange line) of the CarS family (**Table S1**) based on a multiple sequence alignment of 45 members those share between 20% and 60% sequence identity (see **Figure S1** for multiple sequence alignment). The membrane topology model is depicted above the plot. Cytoplasmic loops (C1 and C2) are indicated. The structure divergence score (SDS) [108], which is a measure of the similarity of the individuals profiles, was 0.139. For phylogenetic tree distribution, see **Figure S2**. (**B**) Hydropathy profile alignment of the CdsA (maroon) and CarS (orange) families. The averaged hydropathy profile (maroon line) of the CdsA family was based on 234 sequences sharing identity between 20% and 69%. The profile alignment reveals a similar pattern at the C terminus half with an S score of 0.83 [108] (**Figure S1** shows sequence alignment of CarS with six bacterial homologs). (**C**) Conserved sequence motif of the C2 region of CdsA and CarS family is generated using WebLogo [109], where the overall eight of each amino acid indicates its relative frequency of occurrence at that position.

CTP was analyzed and no product formation was observed when ATP, guanosine-5'-triphosphate (GTP), or thymidine triphosphate (TTP) was used instead of CTP (**Figure 3B**). When the reaction was incubated in the presence of deoxycytidine triphosphate (dCTP), the spectral analysis revealed the presence of a new product identified as deoxycytidine diphosphate (dCDP)-archaeol (m/z= 1004.55 $[M-H]^-$), at almost the same retention time as CDP-archaeol (**Figure 3C and S3**). This result indicates that the absence of a hydroxyl group on the 2' position of the ribose moiety of CTP does not affect the enzymatic activity as reported earlier for the *E. coli* CdsA [101]. CarS used in this study was derived from a thermophile. Consequently, the rate of product formation was found to be higher at 65 ˚C when compared to lower temperatures (37, 45 and 55 ˚C) (**Figure 3D**). The activity of CarS was also determined using [^{3}H]-CTP where the chloroform extractable lipid fraction from the reaction was analyzed by thin layer chromatography [42]. A single spot of radiolabeled CDP-archaeol was observed only in the presence of DGGGP (**Figure 3E**). Next, the CarS activity was measured at different concentrations of substrate DGGGP using [^{3}H]-CTP. Normal saturation kinetic was obtained when the data were fitted to Michaelis-Menten equation using nonlinear regression and an apparent K_m = 0.12 ± 0.02 mM and k_{cat} = 0.55 ± 0.03 s^{-1} was observed (**Figure 3F**). Furthermore, no significant deviation from linearity was observed when the data was fitted using Lineweaver-Burk plot (**Figure S3**). Taken together, these data demonstrate that CarS functions as a CDP-archaeol synthase.

CarS has a distinct activity versus CdsA

CDP-diacylglycerol synthase (CDS) is a ubiquitous and essential enzyme as shown previously in bacterial and eukaryal systems [102,104,110]. A conditionally lethal strain of *E. coli* with a mutant *cdsA* gene (GN80) cannot grow at pH values above 8.0, showing greatly reduced CdsA activity and the accumulation of the substrate phosphatidic acid (PA) at levels around 30% of the total lipid composition compared to 0.2% in the wild type [111]. Previous studies have shown that the *Drosophila CDS* gene expression could rescue the *E. coli cdsA* mutant pH sensitive phenotype [102,112], although the eukaryotic CDS has a longer amino acid sequence

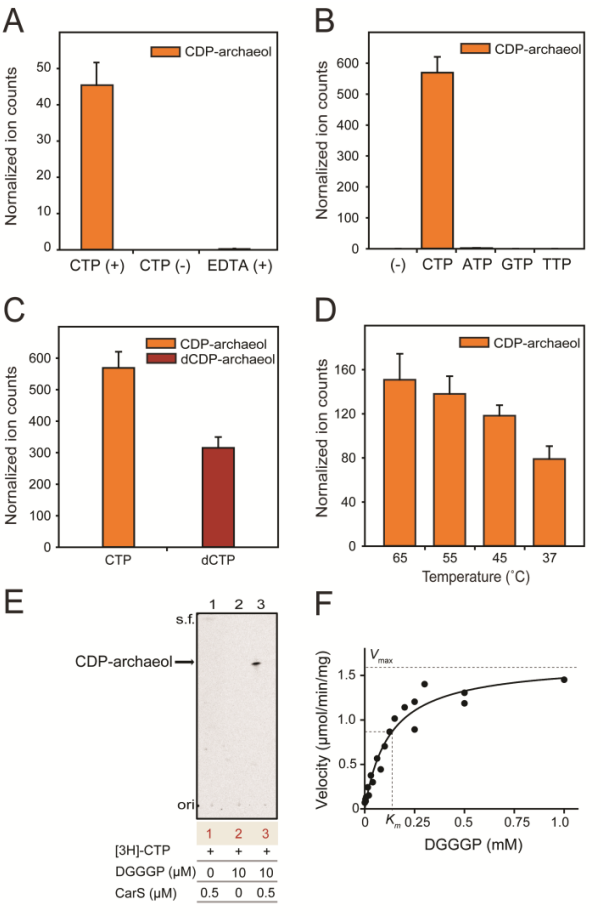


Figure 3| Characterization of CarS activity. (**A**) *In vitro* reactions performed using purified CarS (**Figure S3**) and chemically synthetized substrate DGGGP (10 μM), product CDP-archaeol formation (m/z = 1,020.55 $[M-H]^-$) was analyzed by LC-MS. (**B**) Nucleotide specificity of CarS examined for CTP, ATP, TTP and GTP in the reaction. (**C**) CarS activity measured using CTP or dCTP in a reaction to detect the formation of CDP-archaeol and dCDP-archaeol (m/z = 1,004.55 $[M-H]^-$). (**D**) Enzymatic activity of CarS measured at different temperatures; the reaction was performed for 5 min. (**E**) TLC autoradiogram of chloroform extractable lipid fraction from *in vitro* reactions, single spot of $[^3H]$-CDP-archaeol observed in lane3; solvent front (s.f.) and origin (ori). (**F**) Kinetic analysis of CarS using different substrate concentrations at 65°C for 5 min, the conversion of $[^3H]$-CTP to lipid (see **Figure S3** for Lineweaver-Burk plot). Unless specified, the reactions were performed for 1 hr at 37°C with 0.5 μM CarS, 100 μM DGGGP, and 2 mM nucleotides. Total ion counts from LC-MS data were normalized using DDM as internal standard. LC-MS data are the average of three experiments ± SE.

than the bacterial homolog. To examine this possibility also for the archaeal CarS, a complementation study was performed. *A. fulgidus carS* could not restore the growth defect of *E. coli* GN80 (*cds*$^-$) strain at pH 8.5 while the *E. coli cdsA* could complement the growth defect of the *E. coli* GN80 (*cds*$^-$) strain (**Figure 4A**). As a control, the isogenic wild type *E. coli* GN85 (*cds*$^+$) strain was also used and the expression level of CarS was found unaltered between the two strains (**Figure 4B**).

To measure the specificity of *A. fulgidus* CarS and *E. coli* CdsA for their respective lipid substrates DGGGP and PA *in vitro*, *E. coli* CdsA was overexpressed and purified (**Figure S3**), and the assay was performed using radiolabeled [^{3}H]-CTP. Since the enzymatic activity of CarS and CdsA is Mg^{2+}dependent, EDTA (+) was used as a control for each reaction and the chloroform extractable lipid fraction of the reaction was measured. When CarS was incubated either with PA 18:1 (oleic acid) or PA 16:0 (palmitic acid) (**Figure 4C**), only a very low background signale (<1%; likely spill over of radioactivity originating from incomplete phase separation) was detected in EDTA (+) control that did not differ from the reaction without EDTA. A very low activity was observed when CdsA was incubated with DGGGP in the reaction, whereas high levels of activity occurred with PA (18:1 and 16:0) (**Figure 4D**). These data demonstrate that CarS and CdsA have distinct substrate specificities.

Enzymatic conversion of the precursors Isopentenyl pyrophosphate and Farnesyl diphosphate to CDP-archaeol

To reconstitute the complete ether lipid biosynthesis pathway *in vitro*, enzymatic reactions were performed using a combination of five purified proteins of which three are encoded by archaea and two by bacteria as described in **Table S2**. The archaeal enzymes GGGP synthase, DGGGP synthase and CarS were codon optimized for the overexpression in the host *E. coli*. For the expression of DGGGP synthase, a synthetic ribosome-binding site was introduced that greatly enhanced the expression levels [113]. The various enzymes were purified from the DDM-solubilized membrane fraction (DGGGP synthase and CarS) or the cytosolic fraction (IspA(Y79H, S140T), G1P dehydrogenase and GGGP synthase) by Ni-NTA

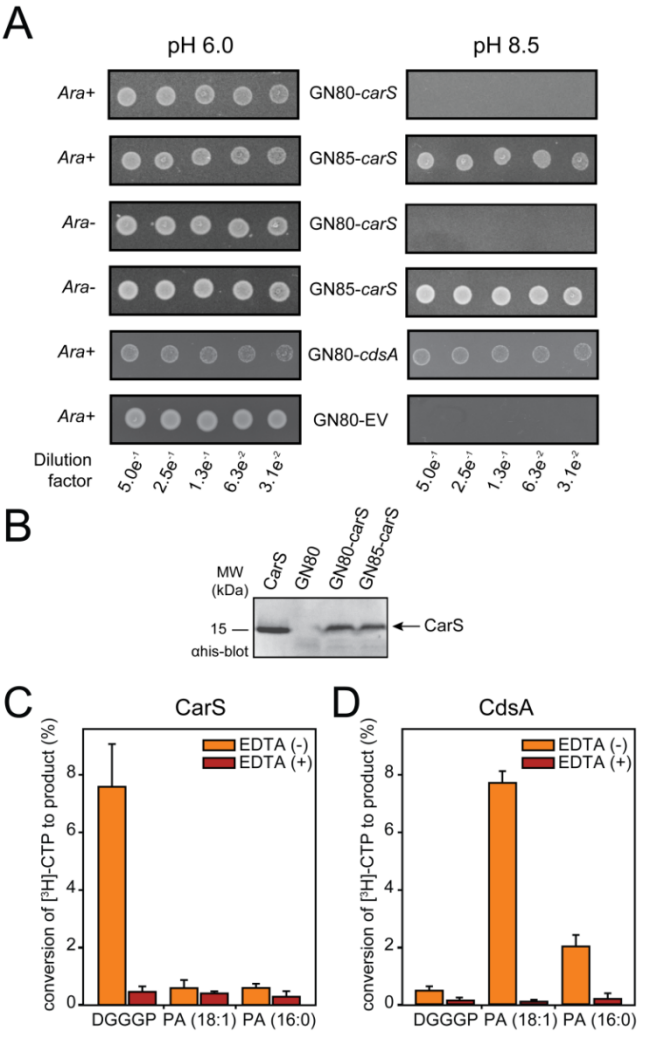


Figure 4| *carS* cannot complement *cdsA* lethal mutant phenotype. (**A**) Serial dilution (OD_{600} 0.3) of *E. coli* pH sensitive GN80 (*cdsA*$^-$) expressing *A. fulgidus carS*, *E. coli cdsA*, or EV were spotted on LB plates at pH 6.0 and pH 8.5. Arabinose (0.1%) was used to induce the expression. (**B**) The total membrane fraction of the specified strains were analyzed for expression of CarS (C-terminus his tag) by immunoblotting with α-his antibody, purified CarS was used as control. (**C** and **D**) *In vitro* reactions with purified CarS (**C**) or CdsA (**D**) and various lipid substrates were performed using radiolabeled [^{3}H]-CTP. EDTA (+) reactions were used as controls. Liquid scintillation counts from the chloroform extractable fraction were used as a measure of the product formation. Data are the average of three experiments ± SE.

affinity column (**Figure 5A**). Isoprenoid building blocks IPP and DMAPP undergo serial condensation to form oligomers of defined chain lengths (C10, C15, C20 etc.) [16]. In archaea, GGPP synthase catalyzes the formation of geranylgeranyl diphosphate (C20 and GGPP). Mutants of its *E. coli* homolog IspA (farnesyl diphosphate [FPP] synthase) were shown previously to synthesize GGPP instead of farnesyl diphosphate (C15 and FPP) [25]. There was one such mutant IspA(Y79H, S140T) that was generated in this study and shown to catalyze the conversion of IPP and FPP to GGPP (m/z= 449.19 $[M-H]^-$) as measured by LC-MS (**Figure 5B, lane 2**).

Recently, it was discovered that the enzyme G1P dehydrogenase, responsible for the stereo-specificity of archaeal lipids, is also encoded by certain bacteria [32]. Of these, AraM of *Bacillus subtilis* was functionally characterized and shown to perform the same reaction as the archaeal homolog, thereby synthesizing G1P from DHAP and NADH [31]. In our study, we thus expressed G1P dehydrogenase (AraM) of *B. subtilis* in *E. coli* and purified the enzyme accordingly (**Figure 5A, lane 2**). The G1P dehydrogenase was catalytically active in an NADH oxidation assay using DHAP and NADH as substrates as described previously [31] (**Figure S4**).

GGGP synthase catalyzes the first ether formation between the substrates GGPP and G1P, leading to the synthesis of the compound GGGP [36]. Like the other enzymes of the family [45], GGGP synthase of *M. maripaludis* also purifies as a higher order oligomer (data not shown). Its activity was established using an *in vitro* reaction where formation of the product GGGP (m/z= 443.26 $[M-H]^-$) was detected in LC-MS only after addition of the substrates GGPP and racemic G1P (rG1P) (**Figure S4**). When GGGP synthase was added to the reaction in combination with enzymes IspA(Y79H, S140T) and G1P dehydrogenase and in the presence of substrates IPP, FPP, DHAP and NADH, formation of the product GGGP and consumption of GGPP was observed (**Figure 5B, lane 3**).

The second ether is formed by the first integral membrane protein of the pathway, DGGGP synthase, using substrates GGGP and GGPP [38]. In this study, DGGGP synthase encoded by *A. fulgidus* was first monitored in a reaction added in different ratios to the *M. maripaludis* GGGP synthase, the common substrate GGPP and the substrate rG1P for the

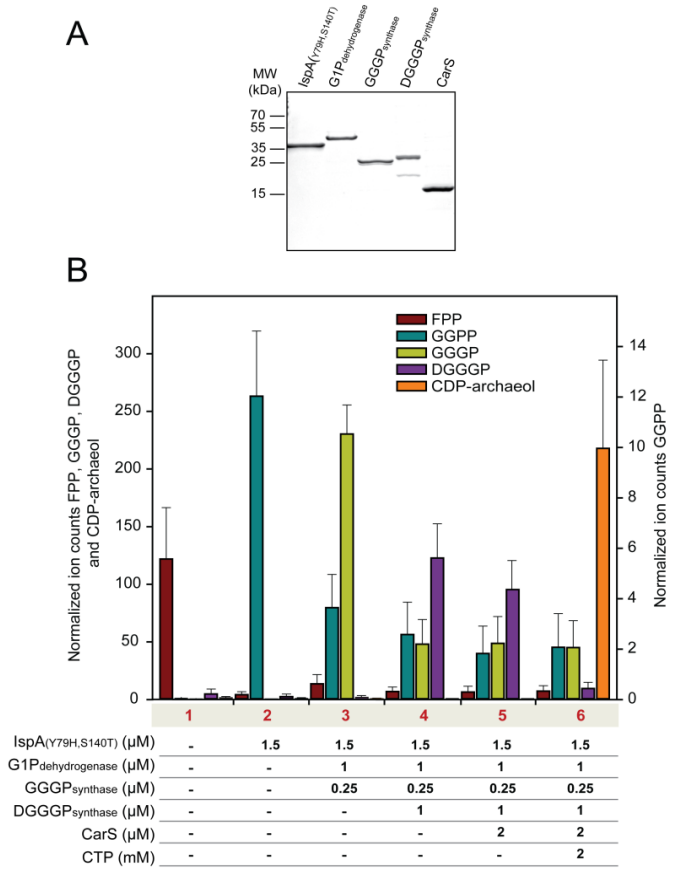


Figure 5| *In vitro* conversion of Isopentenyl pyrophosphate and Farnesyl diphosphate to CDP-archaeol. (**A**) Coomassie stained SDS-PAGE gels showing the Ni-NTA purified proteins from different sources overexpressed in *E. coli.* IspA(Y79H and S140T) mutant of *E. coli* (33kDa), G1P dehydrogenase of *B. subtilis* (44 kDa), GGGP synthase of *M. maripaludis* (29 kDa), DGGGP synthase of *A. fulgidus* (34 kDa), and CarS of *A. fulgidus* (20 kDa). For the description and activities of the individual proteins, see **Table S2** and **Figure S4**. (**B**) *In vitro* reactions were performed using purified proteins as specified and substrates IPP, FPP, DHAP, and NADH in the presence of Mg^{2+} and 0.1% DDM. The products were extracted and analyzed by LC-MS. The spectral data constituting total ion counts (measured as peak area using Thermo Scientific XCalibur processing software) of the products FPP (m/z = 381 $[M-H]^-$), GGPP (m/z = 449.19 $[M-H]^-$), GGGP (m/z = 443.26 $[M-H]^-$), DGGGP (m/z = 715.51 $[M-H]^-$), and CDP-archaeol (m/z = 1,020.55 $[M-H]^-$) were normalized using DDM (m/z = 509.3 $[M-H]^-$) as internal standard and plotted on the y axis. The graph represents average of three experiments ± SE, see **Figure S4** for the total ion counts and **Figure S5** for the TLC based quantitation of the metabolites.

formation of the product DGGGP (m/z= 715.51 $[M-H]^-$) (**Figure S4**). In a subsequent coupled reaction with the other enzymes of the pathway, DGGGP synthase dependent formation of DGGGP was observed (**Figure 5B, lane 4**).

The subsequent reactions were performed in the presence of CarS where a CTP dependent synthesis of CDP-archaeol was only detected in conjunction with other enzymes (**Figure 5B, lanes 5 and 6**). No other unique products were found while comparing the LC-MS spectral data of reaction in lane 5 and 6 indicating that CarS does not utilize GGGP as a substrate. The total ion counts acquired for each product during the LC-MS analysis of the assay are shown in **Figure S4**.

To estimate the amount of different metabolites formed in these reactions, a similar assay was performed using radiolabelled substrate [^{14}C]-IPP, and the products were analyzed by thin layer chromatography (TLC) (**Figure S5**). Quantitation of the spots revealed that picomole amounts of products were formed in a 100 µl reaction starting with 90 pmoles of [^{14}C]-IPP indicating an efficient conversion (**Figure S5**). The data showed a similar spectrum of metabolites being formed in the different reactions as observed for the LC-MS based analysis. TLC analysis further confirmed that CarS does not utilize GGGP as substrate.

The reaction described here thus constitutes the biosynthetic route of the archaeal lipid precursor CDP-archaeol formation (**Figure 1**), establishing the role of CarS in context with the other enzymes of the pathway.

Discussion

CDP-activated precursors are important intermediates in the biosynthesis of phospholipids in bacteria and eukarya and are subsequently modified in various pathways for polar head group attachment. It is synthetized by the conserved and essential enzyme CDP-diacylglycerol synthase [111]. In this study, we elucidated the formation of the CDP-activated precursor CDP-archaeol, the uncharacterized key step in the lipid biosynthetic pathway of archaea. The enzyme required for its synthesis in archaea, CarS, was identified and characterized *in vitro.*

Bioinformatics analysis of CarS indicated a conserved hypothetical protein in the kingdom archaea with remote homology to the bacterial CdsA as predicted previously [8]. Structure analysis by the hydropathy profile alignment revealed a partial fold similarity between the two families that share conserved aspartate, lysine and arginine residues in the cytoplasmic loop (C2) region where they may be involved with the CTP associated activity of the enzyme. Bacterial CdsA proteins are longer than the archaeal CarS, with their N-terminal region less conserved and of unknown function. Interestingly, the homologous eukaryotic CDS proteins are even longer, having an extended N-terminal hydrophilic region. However, six bacterial sequences with shorter length and high sequence similarity to archaeal CarS were retrieved suggesting an event of horizontal gene transfer.

The CarS sequence was retrieved from all sequenced genomes of archaea, with some exceptions. It is not present in the phyla Nanoarchaeota which is a symbiont that encodes no genes of the lipid biosynthetic pathway [33]. The lack of CarS sequence in the three families of the phylum Thaumarchaeota (where CarS is found only in the unclassified family *Candidatus Caldiarchaeum subterraneum*) is intriguing since they contain ether lipids. In our analysis, we retrieved GGGP synthase sequences from these families, but not the DGGGP synthase sequences which is recently reported to exhibit high sequence divergence in Thaumarchaeota where it might be related to the synthesis of cyclohexane moiety specific to Thaumarchaeota lipid structure [6]. We therefore hypothesize that the CarS is also poorly conserved in these organisms and thus difficult to identify. Another possibility could be that an additional enzyme with low sequence homology performs the same function, a situation similar to what was observed previously for folate biosynthesis in archaea [114].

In this study, we used the chemically synthesized unsaturated archaetidic acid DGGGP as substrate to examine the enzymology of CarS. Biochemical investigation performed with purified CarS and DGGGP clearly showed CTP dependent formation of the product CDP-archaeol. Similar to CdsA, CarS uses metal (Mg^{2+}) ions for co-ordination, accepts CTP and dCTP as substrate, and does not utilize ATP, GTP or TTP nucleotides in the

reaction. The apparent K_m for DGGGP (measured with constant detergent concentration) was calculated as 0.12 ± 0.02 mM, which only serves as an indication for the affinity towards the lipid substrate due to the presence of detergent in the reaction. The K_m for PA of CDS from various organisms has been studied before and in general ranges from 0.28 mM (*E. coli*) to 0.9 mM (*Plasmodium falciparum*) [107,115], and thus the K_m of CarS for DGGGP compares favorable with those studies. In some studies, the K_m was measured in constant lipid:detergent ratio, thereby taking surface dilution kinetics [116] into account. However, in those cases, the influence of detergent variation on the activity of the protein cannot be excluded. *A. fulgidus carS* could not complement the pH sensitive *E. coli cdsA*$^-$ phenotype *in vivo*, while the enzyme CarS also did not accept PA as a substrate in the *in vitro* reactions, thereby distinguishing the archaeal and bacterial enzymes for their lipid substrate specificity.

Having identified the fate of DGGGP in the pathway, we attempted to reconstitute the conversion of IPP to CDP-archaeol *in vitro*. This technically challenging aspect of the study was accomplished by selecting the appropriate combination of enzymes for the reaction. Several proteins were individually screened for expression, purification and activity. The *E. coli* IspA and archaeal GGPP synthase belong to the prenyl transferase family that have two highly conserved aspartate rich regions called the FARM and SARM domains. Mutations close to these domains lead to alteration in product length formation probably due to change in the hydrophobic pocket of the enzyme [16]. IspA(Y79H,S140T) mutant described previously [25] was therefore selected with the Y79 residue being five amino acids upstream of the FARM domain. The *B. subtilis* AraM is the only characterized bacterial G1P dehydrogenase, but homologs are found in several other bacteria as hypothetical proteins [31]. G1P in *B. subtilis* is further converted into an archaea type lipid (heptaprenylglyceryl phosphate), the physiological function of which is still unknown [32,35]. The *E. coli* IspA mutant and *B. subtilis* AraM were used in our study for the optimum overexpression in the host *E. coli* compared to their instable archaeal homologues that readily form inclusion bodies (unpublished data). Here we report on the activity of the purified GGGP synthase encoded by *M. maripaludis*, a methanogen that grows at 37 °C. Like the

other archaeal species [35,36,45,46], the GGGP synthase of *M. maripaludis* converts G1P and GGPP to GGGP and is purified as an oligomer. The archaeal membrane proteins DGGGP synthase and CarS encoded from *A. fulgidus*, a hyperthermophile, were selected for this study that display higher expression levels and stability than the methanogenic counterparts. The *in vitro* coupled reaction showing the conversion of archaeal lipid precursor IPP into CDP-archaeol demonstrates that all the enzymes are functional in detergent (DDM) and at 37 °C. This holistic approach further establishes the enzymatic role of CarS in the archaeal lipid biosynthetic pathway.

The subsequent step of the pathway is polar head group attachment. Since the polar head groups are shared among the three domains of life, their biosynthesis is thought to be mediated in a similar manner [2,8]. The enzymes for polar head group attachment that accept the CDP-activated precursors indeed belong to a large family called CDP-alcohol phosphatidyltransferase with homologs in bacteria and archaea, but their classification is unclear with mixed sequence distribution [8]. Interestingly, the archaeatidylserine synthase and the bacterial subclass II phosphatidylserine synthase were shown to even accept lipid substrates of each other, suggesting a broad substrate specificity and possibly common mechanism [47]. In archaea, it is unclear at what step of the pathway the saturation of the membrane lipids takes place [52].

The strategy described here to synthesize archaeal lipids in a cell free manner paves the way for future studies on archaeal lipid biochemistry like the mechanism of tetraether linkage, sugar group attachment, cyclopentane and macrocyclic ring formation, and various other derivatives of the archaeal lipids. A further challenge is to reprogram bacterial ester-based lipid biosynthesis into ether-based lipids to test the evolutionary hypothesis of the lipid divide between Archaea and Bacteria [4].

Experimental Procedures

Protein expression and purification

Proteins were overexpressed in *E. coli* BL21 strain and induced with 0.5 mM isopropyl β-D-1-thiogalactopyranoside (IPTG). For overexpression of CarS, the *E. coli* Lemo strain was used and induced using 0.4 mM of IPTG and 0.5 mM of L-rhamnose. After 2 hr of induction, the cells were harvested and washed with 50 mM Tris-HCl pH 7.5, and re-suspended in the same buffer supplemented with 0.5 mg/ml of RNAse and DNAse and complete EDTA free protease inhibitor tablet (Roche). The suspension was subjected to cell disruption at 13,000 psi and the cell lysate was centrifuged for 10 minutes at low spin (12,000 x*g*) to remove unbroken cells. The cytoplasmic and membrane fractions were separated with high-speed centrifugation at (235,000 x*g*) for 1 hour.

For the purification of soluble proteins (IspA(Y79H, S140T), G1P dehydrogenase and GGGP synthase), the cytoplasmic fraction was incubated with Ni-NTA beads (Sigma) in buffer A (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 10% glycerol) for 30 min at 4 °C. The beads were washed five times with 20 column volumes (CV) of buffer A supplemented with 20 mM Imidazole and eluted two times with 2 CV of buffer A supplemented with 250 mM imidazole.

For purification of membrane proteins (DGGGP synthase, CarS and CdsA), inner membrane vesicles (IMVs) of *E. coli* were isolated as described previously [117] and re-suspended in buffer A. The IMVs (1 mg/ml) were solubilized at 4 ˚C for 1 hr in 2% DDM detergent. Insolubilized materials were removed by centrifugation (173,400 x*g*) for 30 minutes and the supernatant was incubated with Ni-NTA beads for 30 min at 4 °C. The beads were washed five times with 40 CV of buffer B (0.2 % DDM, 50 mM Tris pH 7.5, 150 mM NaCl, 10% glycerol), supplemented with 10 mM imidazole and eluted three times with 0.5 CV of buffer B supplemented with 250 mM imidazole. The purity of the proteins was checked by 15% SDS-PAGE, stained with Coomassie Brilliant Blue. Absorbance was measured at 280 nm in a spectrophotometer to determine the concentration of purified protein.

In vitro assays for lipid synthesis

Reactions were performed using assay buffer with an end concentration of 50 mM Tris-HCl pH 7.5, 10 mM $MgCl_2$, 75 mM NaCl, 0.1 % DDM, 125 mM imidazole and 5% glycerol. Where specified, 2 mM CTP, 20 mM EDTA, 0.25 mM NADH, 0.1 mM DHAP, 0.1 mM IPP, 0.1 mM FPP, 4 mM rG1P and 0.1 mM GGPP were used in a 100 µl reaction volume. After incubation at specified temperature and time, the products were extracted two times with 0.3 ml *n*-butanol and evaporated under a stream of nitrogen gas. The samples were re-suspended in 0.05 or 0.1 ml methanol and analyzed by LC-MS.

Assays using radionucleotides

Purified CarS or CdsA (0.5 µM) was incubated with specified amounts of DGGGP or 200 µM PA in assay buffer with 5 µCi [^{3}H]-CTP (1.68 µM) and 2 mM cold CTP in a 100 µl reaction volume. The reaction for lipid biosynthesis had same composition as described above for the LC-MS assay with the addition of 0.005 µCi [^{14}C]-IPP (0.9 µM) in a 100 µl reaction volume. The reaction was incubated for 5 minutes at 65 °C (for kinetic measurements) or for 1 hr at 37 °C. Acidic Bligh and Dyer extraction [118] was performed and the chloroform extractable lipid fraction was evaporated. The radioactivity was counted after the addition of 10 ml Emulsion Scintillation liquid (PerkinElmer) using Packard scintillation counter. When analyzed by thin layer chromatography, the samples were re-suspended in 10 µl chloroform and spotted on Silica Gel 60 (Merck) plates. Solvents chloroform, methanol, 7 M ammonia in the ratio 60:35:8 [42] were used as mobile phase. Using phosphor screen or tritium sensitive phosphor screen, the autoradiograph was obtained by the phosphorimager (Roche). The spots were quantitated along with [^{14}C]-IPP dilution series using Image J software and the amounts were extrapolated using the [^{14}C]-IPP calibration curve.

Genetic complementation

E. coli strains GN80 and GN85 [111] were transformed with two plasmids, pSJ122 (CarS under the control of T7 promoter) and pTara [119] (under

the control of *araBAD* promoter). In this manner, the addition of arabinose to the culture led to the controlled expression of T7 polymerase that subsequently led to the expression of CarS. For the controls, pSJ122 was replaced by pSJ148 (*E. coli* CdsA under the control of T7 promoter) or empty vector (EV) and transformed in GN80. To perform the pH sensitivity assay, serial dilutions of the transformants with OD_{600} of 0.3 were plated on LB agar plates with pH 6.0 (pH adjusted using 50 mM 2-9N-morpholino ethanesulfonic buffer) or pH 8.5 (pH adjusted using NaOH) that were prepared as described previously [111]. Expression was induced by the addition of 0.1% arabinose.

Liquid chromatography mass spectrometry (LC-MS)

Samples were analyzed using an Accella1250 high-performance liquid chromatography system coupled with a bench top electrospray ionization mass spectrometry (ESI-MS) Orbitrap Exactive (Thermo Fisher Scientific). A sample of 5 or 10 µl was injected into a COSMOSIL 5SL-II packed C4 column with dimensions 4.6 mm I.D. x 150 mm (Nacalai tesque) operating at 40 °C with a flow rate of 500 µL/min. Eluents A (25 mM ammonium bicarbonate) and B (acetonitrile) were used as follows :- 97% A isocratic for 5 minutes, gradient 97% A to 100% B (Acetonitrile) for 30 minutes, 100% B isocratic for 25 min, gradient 100% B to 97% A for 10 minutes and 97% A isocratic for 5 minutes. The column effluent was injected directly into the Exactive ESI-MS Orbitrap operating in negative ion mode. Voltage parameters of 3 kV (spray), -52.5V (capillary) and -160 V (tube lens) was used. Capillary temperature of 300 °C, sheath gas flow 60, auxiliary gas flow of five was maintained during the analysis. For performing fragmentation, high-energy collision induced dissociation (HCD) gas was used at 35eV. Data analysis was performed using the Thermo XCalibur processing software for automated peak integration and quantification. The ICIS algorithm for component peak detection was applied in this analysis.

Chemical synthesis of unsaturated archaetidic acid, DGGGP

DGGGP was chemically synthesized according to the scheme in **Supplementary Experimental Procedures**. Compound **A** was prepared as described previously [86]. The (*S*)-1,2-isopropylidene glycerol was protected with a dimethoxybenzyl group, followed by hydrolysis of the acetonide and bis-alkylation with geranylgeranyl chloride. Oxidative deprotection with 2,3-dichloro-5,6-dicyanobenzoquinone afforded **A**. The phosphorylation of **A** proved to be surprisingly challenging, possibly due to the polyunsaturation of the geranylgeranyl side chains. Treatment of **A** with $POCl_3$ and subsequent hydrolysis led to decomposition. Phosphorylation using $(MeO)_2POCl$ as previously reported [42] gave the desired bismethoxy phosphoric ester **B**. However, the demethylation conditions described in that study using TMSBr led to decomposition of **B**. Other common P^V reagents could not be used because these require either hydrogenation or strongly acidic conditions for deprotection. An alternative viable approach towards the phosphorylated product **G** is phosphoramidite methodology [120]. From a plethora of reagents reported for various purposes, we chose those that require relatively mild basic conditions for cleavage. Application of phosphoramidite **C** gave the desired bisprotected product **D** in 70% yield. The cyanoethyl groups were reported to undergo facile deprotection with an excess of a secondary or tertiary amine. However, this did not occur, although partial deprotection was observed. This led to application of the more base-sensitive phosphoramidite **E** bearing two 9-fluorenylmethyl groups [121]. The desired bisprotected derivative **F** was obtained in an excellent 94% yield. The first 9-fluorenylmethyl group could be readily cleaved using an excess of Et_3N. After changing to more basic conditions, the remaining protecting group could be also removed. The desired phosphoric acid derivative **F** was obtained in 48% yield after column chromatography. Compared to the previously published synthesis with a yield of <6% [42], the procedure described here gave a higher yield (17%).

The detailed experimental protocol and NMR data are described in **Supplementary experimental procedures**.

Author contributions

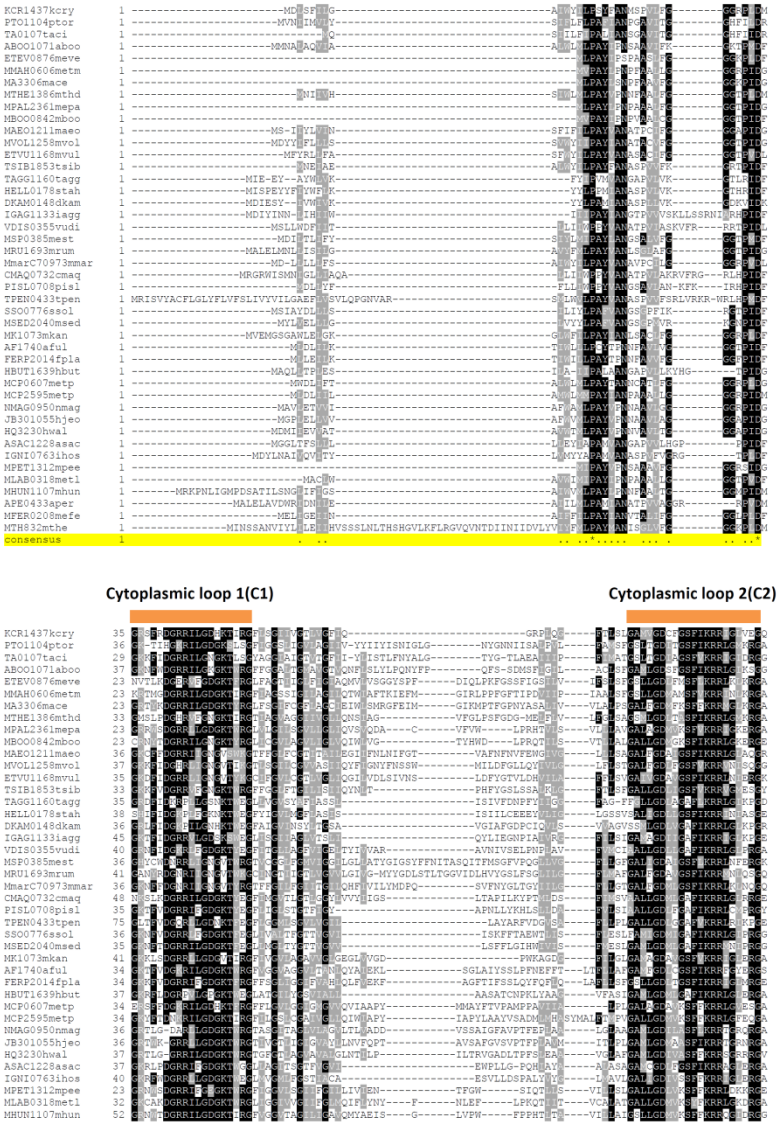
S.J. and A.D. conceived and designed the research; S.J. performed the *in vivo* complementation studies and *in vitro* reconstitution assays; A.C. purified and characterized CarS; J.L. performed the bioinformatics; A.M. designed the DGGGP synthesis, which was performed by P.F.; the manuscript was written by the contributions of all authors.

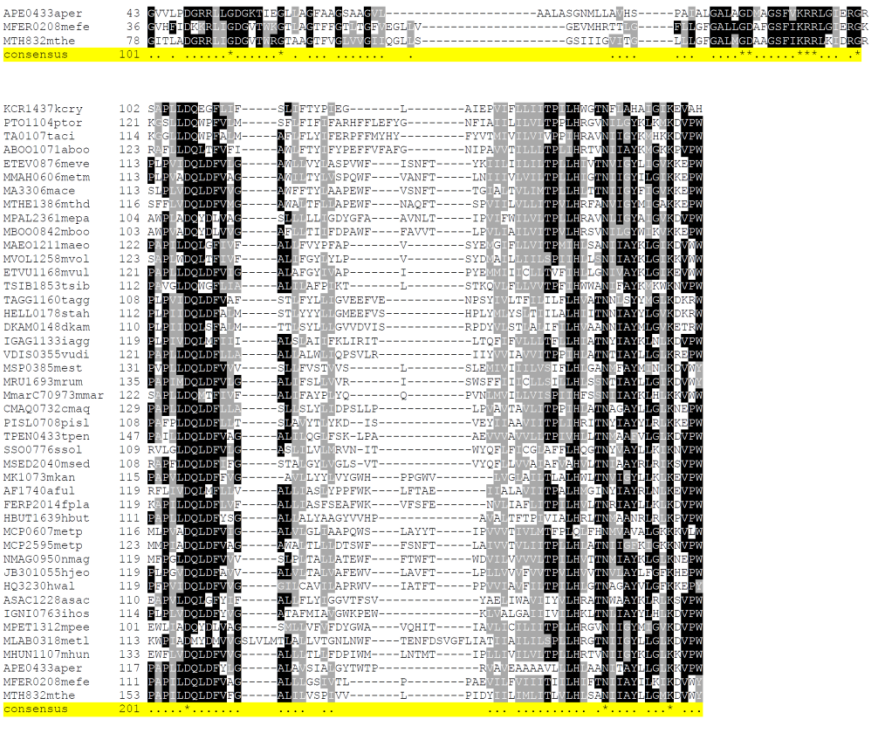
Acknowledgements

This project was carried out within the research program of the biobased ecologically balanced sustainable industrial chemistry (BE-BASIC). We are grateful to William Dowhan for providing us with the *E. coli* strains GN80 and GN85. We thank Oleksander Salo, Stephan Portheine and Ilja Kusters for technical assistance. We extend our gratitude to John van der Oost, Melvin Siliakus and Servè Kengen for fruitful discussions.

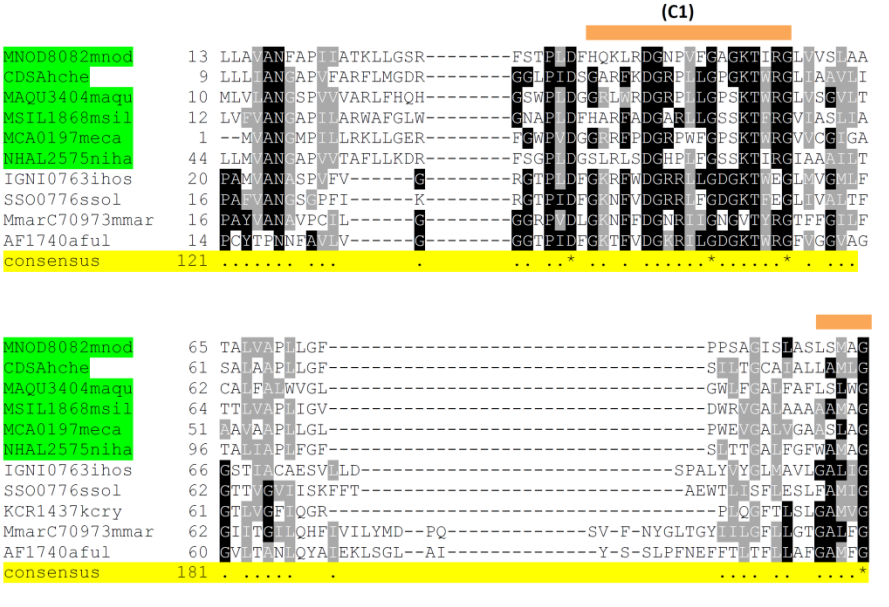
Supplementary information

A





B



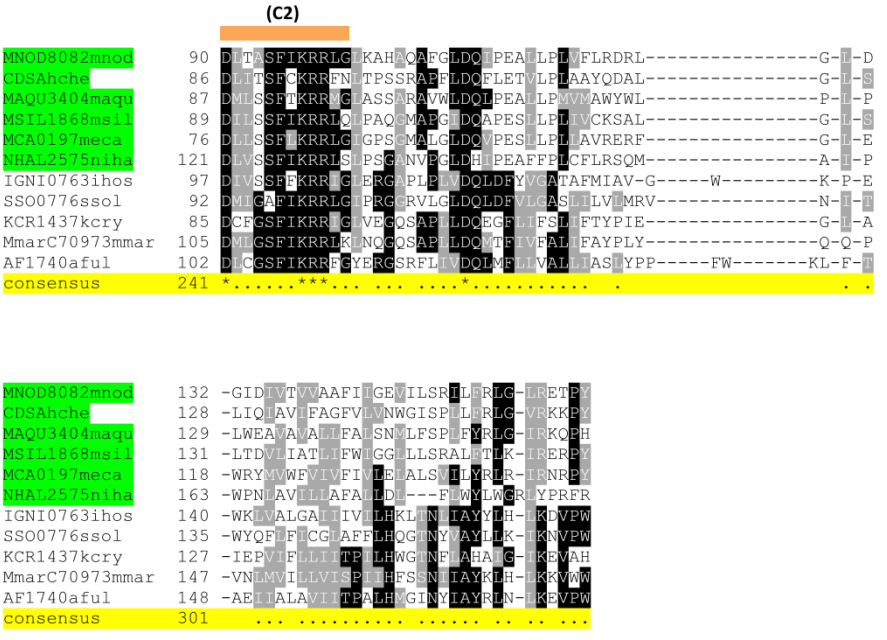


Figure S1 related to Figure 2| (**A**) Multiple sequence alignment of the archaeal CarS proteins**.** Parwise sequence alignment of 44 diverse sequence of CarS in archaea. Cytoplasmic loop C1 and C2 (orange bar) and consensus (yellow highlight) are indicated. Gene symbol/locus is denoted followed by the name of the organism in abbreviation (e.g. *Archaeoglobus fulgidus* as Aful). (**B**) Multiple sequence alignment of the archaeal CarS and bacterial CdsA. CdsA of bacteria *Methylobacterium nodulans, Hahella chejuensis, Mariinobacter aquaeolei, Methylocella silvestris, Methylococcus capsulatus and Nitrosococcus halophilus* (highlighted green) and CarS of archaea *Ignicoccus hospitalis, Sulfolobus solfataricus, Methanococcus maripaludis* and *Archaeoglobus fulgidus* are represented. Cytoplasmic loop C1 and C2 (orange bar) and consensus (yellow highlight) are indicated.

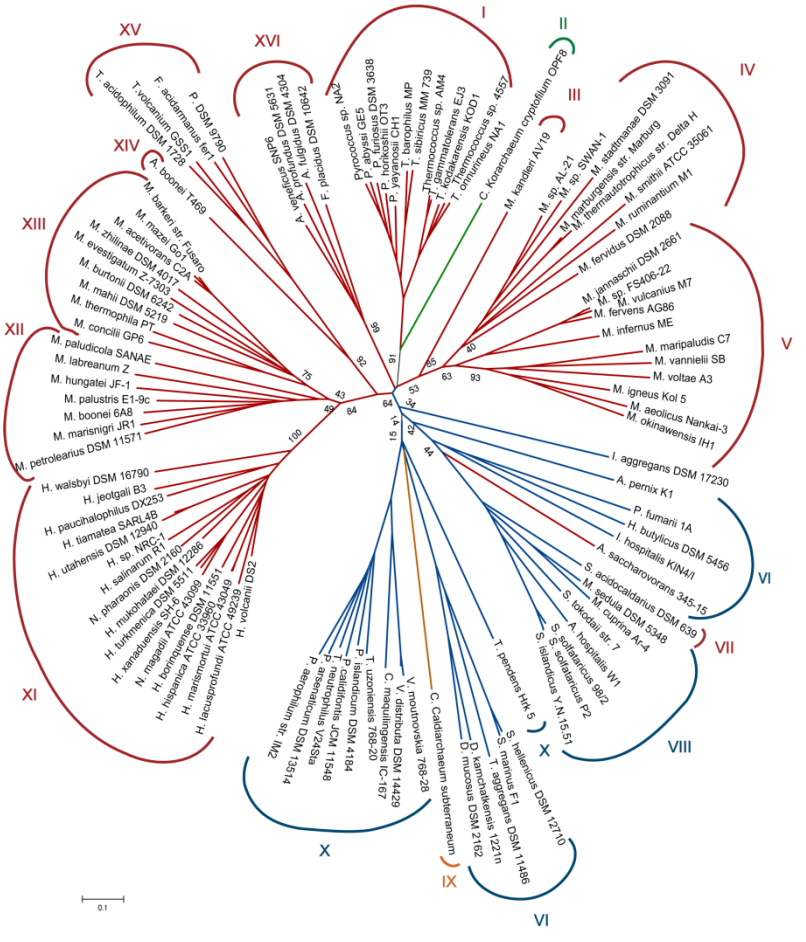
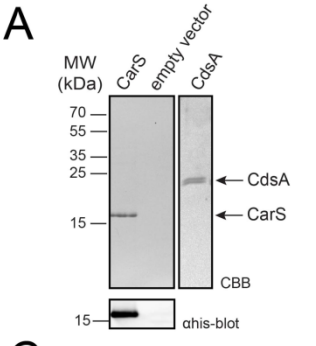
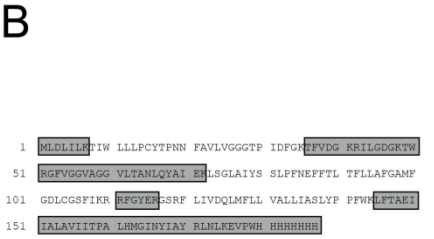


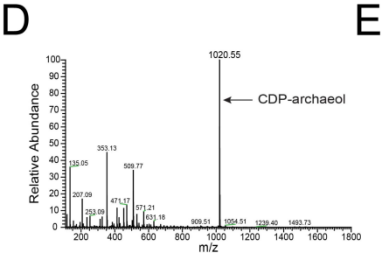
Figure S2 related to Figure 2| Phylogenetic tree distribution of CarS in Archaea. The tree was reconstructed using neighbourhood joining method (bootstrap value set to 1000) with pair wise alignment of 108 CarS representative sequences from archaea. The different colours highlight the phyla classification; Crenarchaeota (blue), Euryarchaeota (red), Korarchaeota (green) and Thaumarcheota (orange). The Roman numbers indicate the family distribution of the species; **I -** Thermococcaceae, **II -** Cenarchaeaceae, **III -** Methanopyraceae, **IV** - Mehtanobacteriaceae, **V -** Methanococcaceae, **VI -** Desulfurococcacaeae, **VII -** Acidolobaceae, **VIII -** Sulfolobaceae, **IX -** unclassified traumarchaeota, **X -** Thermoproteaceae, **XI -** Halobacteriaceae, **XII -** Methanomicrobiaceae, **XIII -** Methanosarcinaceae, **XIV -** unclassified euryarchaeota, **XV -** Thermoplasmaceae and **XVI -** Archaeoglobaceae.

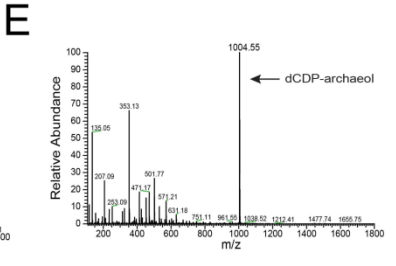




C

Peptides	Score (%)	-10lgP	Mass	ppm	m/z	z	RT	Scan	Start	End	PTM
R.GFVGGVAGGVLTANLQYAIEK.L	100	144	2063.105	0.8	1032.561	2	56.06	4180	52	72	
K.TWRGFVGGVAGGVLTANLQYAIEK.L	99.9	71.05	2506.333	0.5	836.4521	3	54.34	3885	49	72	
R.ILGDGKTWR.G	99.9	60.54	1044.572	-0.3	523.293	2	31.13	1765	43	51	
K.TFVDGKR.I	99.9	48.57	821.4395	-0.4	411.7269	2	25.43	1718	36	42	
MLDLILK.T	99.8	44.82	844.5092	-0.2	423.2618	2	46.96	3258	1	7	
M(+15.99)LDLILK.T	98	32.77	860.5041	-0.1	431.2593	2	42.19	2724	1	7	Oxidation (M)
K.RILGDGK.T	95.7	29.28	757.4446	0	379.7296	2	20.01	908	42	48	
R.LNLKEVPWHHHHHHHHH	95.2	28.42	2094.031	-0.7	419.8132	5	27.26	1421	172	187	
K.EVPWHHHHHHHH	94.7	27.56	1625.725	0	542.9156	3	16.93	645	176	187	
R.RFGYER.G	94.4	26.75	826.4086	-0.2	414.2115	2	27.4	1779	111	116	
K.LFTAEIIALAVIITPALHM(+15.99)GINYIAYR.L	88.8	24.1	3002.667	-0.1	751.6739	4	71.45	5078	145	171	Oxidation (M)





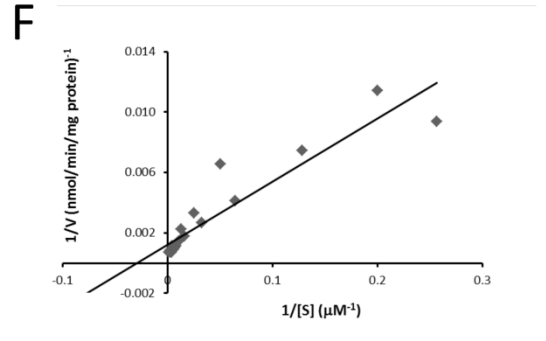


Figure S3 related to Figure 3 and 4| CarS and CdsA purification, mass spectra of the product formation and kinetic analysis. (**A**) SDS-PAGE showing the purified protein *A. fulgidus* CarS with C-terminus histidine tag and *E. coli* CdsA with N-terminus histidine tag. The proteins were overexpressed in *E. coli* and purified over Ni-NTA resin from DDM-solubilizd inner membrane vesicles. As a control for non-specific binding, empty vector IMVs was used for the purification. Expression of CarS was examined by immunoblotting with α-his antibody and coomassie brilliant blue staining (CBB). (**B**) Protein sequence coverage of CarS from the LC-MS/MS fragmentation analysis after tryptic digestion of the CarS protein band from A. (**C**) Description of the list of peptides that were fragmented and assigned to CarS sequence, RT, retention time; PTM, post translational modification CarS (Accession number: O28534|Y1740_ARCFU) was identified exclusively (except keratin) with a score of 99.1%. (**D**) Mass spectra from the LC-MS run of a product sample from **Figure 3A** depicting the CDP-archaeol peak (m\z = 1020.55 $[M-H]^-$). (**E**) Mass spectra from the LC-MS run of product sample from **Figure 3C** depicting the dCDP-archaeol peak (m\z = 1004.55 $[M-H]^-$). (**F**) A double reciprocal or the Lineweaver-Burk plot of CarS activity measured at different concentrations of substrate DGGGP using $[^3H]$-CTP. The unit of velocity is nmol/min/mg of CDP-archaeol formed and that of DGGGP concentration is μM. The $[^3H]$-CTP concentration was 1.68 μM (cold CTP concentration was 2mM), and CarS was added at a concentration of 0.01mg/ml. The Y intercept ($1/V_{max}$ =0.0012) and the slope (K_m/V_{max} =0.04).

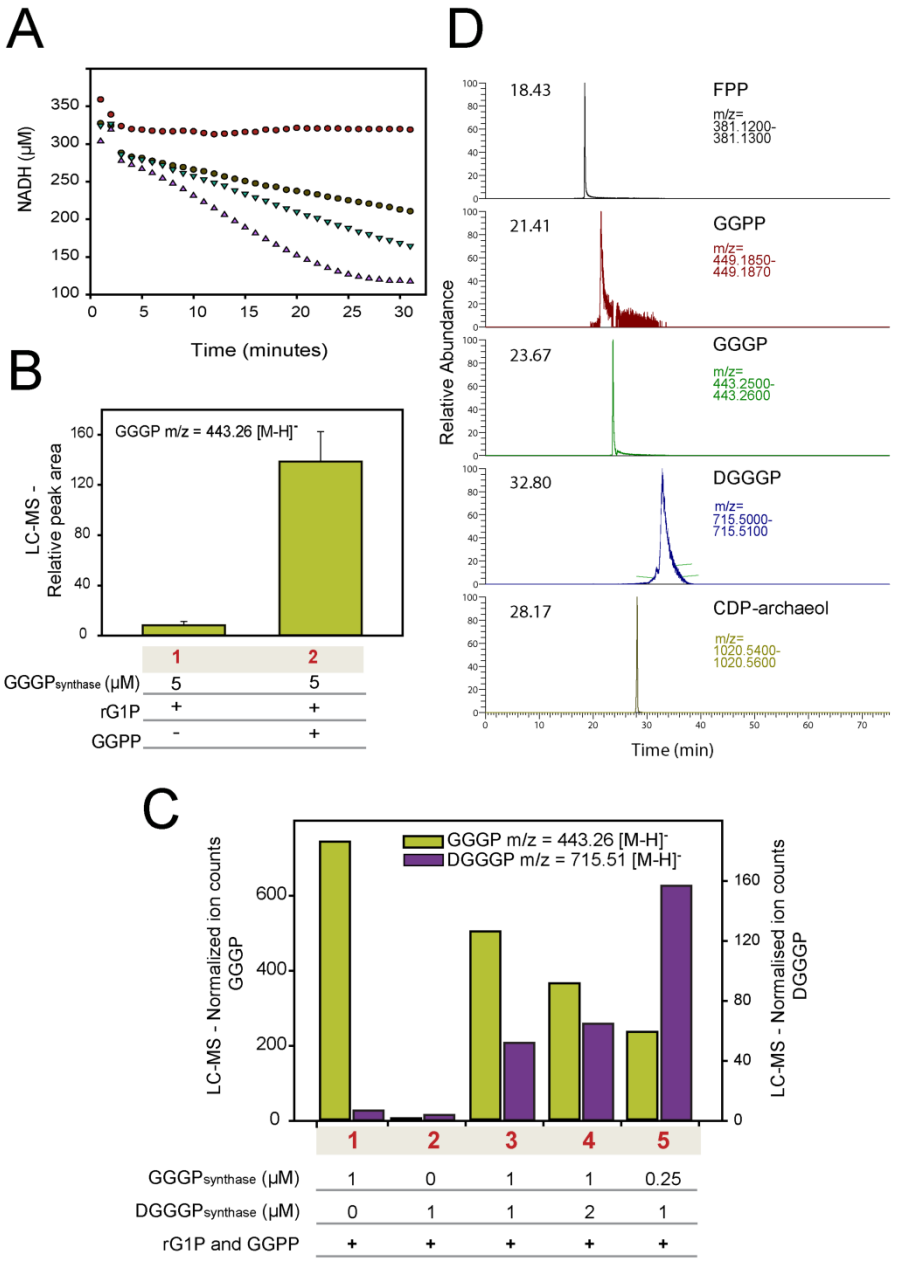

A
NADH (μM)
Time (minutes)
B
GGGP m/z = 443.26 [M-H]-
LC-MS - Relative peak area
GGGPsynthase (μM)
rG1P
GGPP
D
18.43
FPP
m/z= 381.1200-381.1300
21.41
GGPP
m/z= 449.1850-449.1870
23.67
GGGP
m/z= 443.2500-443.2600
32.80
DGGGP
m/z= 715.5000-715.5100
28.17
CDP-archaeol
m/z= 1020.5400-1020.5600
Relative Abundance
Time (min)
C
GGGP m/z = 443.26 [M-H]-
DGGGP m/z = 715.51 [M-H]-
LC-MS - Normalized ion counts GGGP
LC-MS - Normalised ion counts DGGGP
GGGPsynthase (μM)
DGGGPsynthase (μM)
rG1P and GGPP

Figure S4 related to Figure 5| Enzymatic activities of G1P dehydrogenase and total ion counts (**A**) G1P dehydrogenase activity measured by NADH oxidation assay. G1P dehydrogenase catalyzes the conversion of DHAP to G1P using cofactor NADH [31]. Using different concentrations of purified G1P dehydrogenase in a reaction (0 µM - red circle, 0.5 µM - brown circle, 1 µM - green triangle and 2 µM - purple triangle); the consumption of NADH over time was determined by measuring the decrease in absorbance at 340nm. The reaction was performed at 37°C in the presence of 0.1 M DHAP, 0.25 M NADH and 0.5 mM $NiCl_2$ using 50 mM Tris, pH 8 buffer in a 100 µl reaction volume. (**B**) GGGP synthase activity measured by LC-MS. *In vitro* reaction was performed using purified GGGP synthase and substrates 4 mM rG1P (racemic glycerol-1-phosphate, Santa Cruz) and 20 µM GGPP (Sigma) in the presence of Mg^{2+} in a 100 µl reaction volume. After incubation at 37 °C for 1hr, the product was extracted using *n*-butanol, separated over a C18 column (Shimadzu) and analysed by ESI-MS in a negative ion mode. As a control, reaction was performed in the absence of GGPP. A mass spectrum with a peak corresponding to singly charged ions of GGGP was obtained and the relative peak area was calculated (Thermo Scientific XCalibur™ processing and instrument control software). The graph represents average of duplicate experiments ± S.E. (**C**) DGGGP synthase activity measured by LC-MS**.** *In vitro* coupled reaction was performed using indicated amounts of purified GGGP synthase and DGGGP synthase, substrates 4 mM rG1P and 100 µM GGPP in the presence of Mg^{2+} in a 100 µl reaction volume. After incubation at 37 °C for 1 hr, the product was extracted using *n*-butanol and analysed by LC-MS as before. The product GGGP formed by the enzyme GGGP synthase is converted by DGGGP synthase to DGGGP. The spectral data constituting total ion counts (measured as peak area using Thermo Scientific XCalibur™ processing and instrument control software) of the products were normalized using DDM (m/z= 509 $[M-H]^-$) as internal standard and plotted on the Y axis. (**D**) LC-MS data depicting total ion counts. Using the mass range as denoted, the total ion counts acquired during the LC-MS analysis for compounds FPP, GGPP, GGGP, DGGGP and CDP-archaeol are shown with their respective retention time when they are separated over a C4 column (using samples from the *in vitro* reactions for lipid synthesis**, Figure 5**).

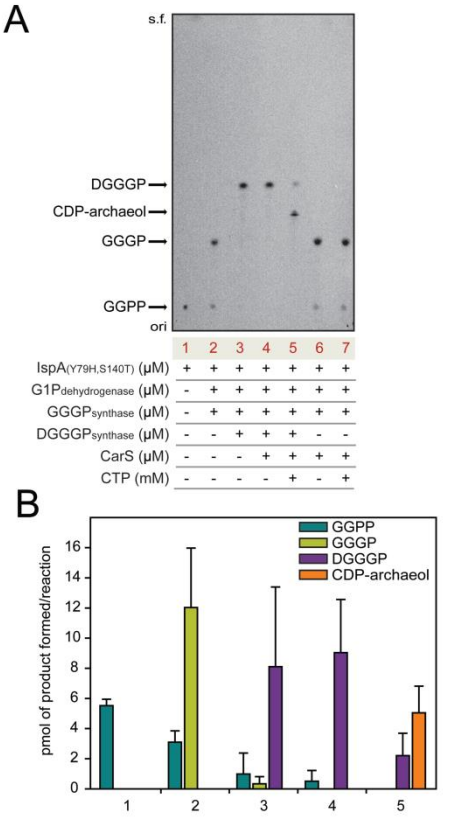


Figure S5 related to Figure 5| Thin Layer Chromatography (TLC) based quantitation of *in vitro* lipid synthesis. (**A**) TLC autoradiogram of chloroform extractable lipid fraction from *in vitro* reactions performed using purified proteins as specified and substrates [^{14}C]-IPP (0.9μM, 90 pmol/reaction), IPP (100 μM), FPP (100 μM), DHAP (2 mM) and NADH (2 mM) in a 100 μl reaction volume. Solvents chloroform, methanol and 7 M ammonia (60:35:8) were used as mobile phase. The retention factor (Rf) of GGPP and CDP-archaeol were compared with the standards [^{3}H]-GGPP (Perkin Elmer) and CDP-archaeol from lane3, **Figure 3E** spotted on the same TLC plate. Note that free [^{14}C]-IPP does not extract in the organic phase. s.f., solvent front, ori, origin. (**B**) Quantitation of the spots in (**A**) from lane 1 to lane 5. The spots were quantitated and analyzed using the [^{14}C]-IPP calibration curve and corrected for the number of IPP molecules in each metabolite. The graph represents average of two experiments ± S.E.

Table S1 related to Figure 2. CarS identified in the sequenced genomes of Archaea

Phylum	Family	Species	length (AA)	NCBI Reference
Crenarchaeota	Desulfuro-ccocaceae	*Aeropyrum pernix K1*	174	NP_147225.1
		Desulfurococcus kamchatkensis 1221n	170	YP_002427844.1
		Desulfurococcus mucosus DSM 2162	170	YP_004176488.1
		Hyperthermus butylicus DSM 5456	168	YP_001013798.1
		Ignicoccus hospitalis KIN4/I	171	YP_001435352.1
		Ignisphaera aggregans DSM 17230	177	YP_003859825.1
		Pyrolobus fumarii 1A	165	YP_004780313.1
		Staphylothermus hellenicus DSM 12710	172	YP_003668220.1
		Staphylothermus marinus F1	176	YP_001040632.1
		Thermosphaera aggregans DSM 11486	168	YP_003650381.1
	Sulfolobaceae	*Acidianus hospitalis W1*	164	YP_004457713.1
		Metallosphaera cuprina Ar-4	165	YP_004408834.1
		Metallosphaera sedula DSM 5348	165	YP_001192104.1
		Sulfolobus acidocaldarius DSM 639	163	YP_255553.1
		Sulfolobus islandicus Y.N.15.51	166	YP_002840495.1
		Sulfolobus islandicus L.D.8.5	166	YP_003419524.1
		Sulfolobus islandicus L.S.2.15	166	YP_002832112.1
		Sulfolobus islandicus M.14.25	166	YP_002829416.1
		Sulfolobus islandicus M.16.27	166	YP_002843342.1
		Sulfolobus islandicus M.16.4	166	YP_002914630.1
		Sulfolobus islandicus REY15A	166	YP_005648506.1
		Sulfolobus islandicus Y.G.57.14	166	YP_002837547.1
		Sulfolobus solfataricus 98/2	166	ZP_06388974.1
		Sulfolobus solfataricus P2	166	NP_342283.1
		Sulfolobus tokodaii str. 7	166	NP_376350.1
	Thermo-proteaceae	*Caldivirga maquilingensis IC-167*	187	YP_001540558.1
		Pyrobaculum aerophilum str. IM2	164	NP_558747.1

		Pyrobaculum arsenaticum DSM 13514	164	YP_001154512.1
		Pyrobaculum calidifontis JCM 11548	165	YP_001056967.1
		Pyrobaculum islandicum DSM 4184	164	YP_930229.1
		Pyrobaculum sp. 1860	164	YP_005085817.1
		Pyrobaculum oguniense TE7	133	YP_005261183.1
		Thermofilum pendens Hrk 5	204	YP_919845.1
		Thermoproteus neutrophilus V24Sta	164	YP_001795319.1
		Thermoproteus tenax Kra 1	165	YP_004893382.1
		Pyrobaculum uzoniensis 768-20	156	YP_004338600.1
		Vulcanisaeta distributa DSM 14429	179	YP_003900811.1
		Vulcanisaeta moutnovskia 768-28	179	YP_004244984.1
	Acidolobaceae	*Acidilobus saccharovorans 345-15*	167	YP_003816664.1
Euryarchaeota	unclassified	*Aciduliprofundum boonei T469*	186	YP_003483442.1
	Archaeo-globaceae	*Archaeoglobus fulgidus DSM 4304*	179	NP_070568.1
		Archaeoglobus profundus DSM 5631	179	YP_003400024.1
		Archaeoglobus veneficus SNP6	179	YP_004341422.1
		Ferroglobus placidus DSM 10642	179	YP_003436426.1
	Halobacteriaceae	*Halalkalicoccus jeotgali B3*	181	YP_003735396.1
		Haloarcula hispanica ATCC 33960	181	YP_004794952.1
		Haloarcula marismortui ATCC 43049	181	YP_137505.1
		Halobacterium salinarum	181	YP_001689902.1
		Haloferax volcanii DS2	180	YP_003534409.1
		Halogeometricum borinquense DSM 11551	180	YP_004037825.1
		Halomicrobium mukohataei DSM 12286	181	YP_003176940.1
		Halopiger xanaduensis SH-6	228	YP_004596780.1
		Haloquadratum walsbyi DSM 16790	181	YP_658923.1
		Halorhabdus utahensis DSM 12940	181	YP_003129842.1
		Halorubrum lacusprofundi ATCC 49239	180	YP_002565024.1
		Haloterrigena turkmenica DSM 5511	181	YP_003404486.1

		Natrialba magadii ATCC 43099	181	YP_003479096.1
		Natronomonas pharaonis DSM 2160	189	YP_326154.1
		Halobacterium sp. NRC-1	214	NP_280779.1
		Halophilic archaeon DL31	180	YP_004809073.1
	Methano-bacteriaceae	*Methanobacterium sp. SWAN-1*	172	YP_004519509.1
		Methanobrevibacter ruminantium M1	191	YP_003424435.1
		Methanobrevibacter smithii ATCC 35061	188	YP_001273423.1
		Methanosphaera stadtmanae DSM 3091	187	YP_447429.1
		Methanothermobacter marburgensis str. Marburg	171	YP_003850131.1
		Methanothermobacter thermautotrophicus str. Delta H	209	NP_275969.1
		Methanothermus fervidus DSM 2088	167	YP_004003773.1
	Methano-coccaceae	*Methanocaldococcus fervens AG86*	177	YP_003127650.1
		Methanocaldococcus infernus ME	170	YP_003615646.1
		Methanocaldococcus jannaschii DSM 2661	177	NP_248610.1
		Methanocaldococcus sp. FS406-22	177	YP_003457889.1
		Methanocaldococcus vulcanius M7	177	YP_003247508.1
		Methanococcus aeolicus Nankai-3	178	YP_001325400.1
		Methanococcus maripaludis C7	178	YP_001330191.1
		Methanococcus vannielii SB	178	YP_001323515.1
		Methanococcus voltae A3	179	YP_003707888.1
		Methanothermococcus okinawensis IH1	206	YP_004576579.1
		Methanotorris igneus Kol 5	179	YP_004484533.1
		[Methanococcus maripaludis C5]	178	YP_001098219.1
		[Methanococcus maripaludis C6]	178	YP_001549020.1
		[Methanococcus maripaludis X1]	178	YP_004743461.1
		[Methanococcus maripaludis S2]	178	NP_988818.1
	Methano-microbiaceae	*Methanocorpusculum labreanum Z*	185	YP_001029761.1
		Methanoculleus marisnigri JR1	166	YP_001046917.1
		Methanoplanus petrolearius DSM 11571	163	YP_003894510.1
		Candidatus Methanoregula boonei 6A8	165	YP_001404003.1

		Methanosphaerula palustris E1-9c	166	YP_002467366.1
		Methanospirillum hungatei JF-1	195	YP_502575.1
		Methanocella paludicola SANAE	185	YP_003357650.1
	Methano-pyraceae	*Methanopyrus kandleri AV19*	173	NP_614356.1
	Methano-sarcinaceae	*Methanococcoides burtonii DSM 6242*	175	YP_565734.1
		Methanohalobium evestigatum Z-7303	175	YP_003726570.1
		Methanohalophilus mahii DSM 5219	175	YP_003541776.1
		Methanosaeta concilii GP6	178	YP_004383635.1
		Methanosaeta thermophila PT	178	YP_843801.1
		Methanosalsum zhilinae DSM 4017	191	YP_004615711.1
		Methanosarcina acetivorans C2A	175	NP_618196.1
		Methanosarcina barkeri str. Fusaro	175	YP_306966.1
		Methanosarcina mazei Go1	175	NP_632165.1
	Methano-cellaceae	*Methanocella arvoryzae MRE50*	174	YP_685268.1
	Thermo-coccaceae	*Pyrococcus abyssi GE5*	171	NP_127296.1
		Pyrococcus furiosus DSM 3638	167	NP_578127.1
		Pyrococcus horikoshii OT3	170	NP_142321.1
		Pyrococcus sp. NA2	168	YP_004423875.1
		Pyrococcus yayanosii CH1	166	YP_004624735.1
		Thermococcus barophilus MP	168	YP_004072009.1
		Thermococcus gammatolerans EJ3	171	YP_002960495.1
		Thermococcus kodakarensis KOD1	171	YP_184532.1
		Thermococcus onnurineus NA1	170	YP_002308011.1
		Thermococcus sibiricus MM 739	168	YP_002995251.1
		Thermococcus sp. 4557	171	YP_004763035.1
	Thermo-plasmaceae	*Ferroplasma acidarmanus fer1*	185	ZP_05570787.1
		Picrophilus torridus DSM 9790	184	YP_023882.1
		Thermoplasma acidophilum DSM 1728	176	NP_393585.1
		Thermoplasma volcanium GSS1	177	NP_110705.1
Korarchaeota	unclassified	*Candidatus Korarchaeum cryptofilum OPF8*	158	YP_001737866.1

Thauma-rchaeota	unclassified	*Candidatus Caldiarchaeum subterraneum*	185	BAJ47591.1
	Cenarchaeaceae	*Cenarchaeum symbiosum*		Not found
	Nitroso-pumilaceae	*Nitrosopumilus maritimus SCM1*		Not found
	Nitroso-sphaeraceae	*Candidatus Nitrososphaera gargensis*		Not found
Nano-archaeota	unclassified	*Nanoarchaeum equitans Kin4-M*		Not found

Table S2 related to Figure 5. Enzymes used in this study for *in vitro* reconstitution of the ether lipid biosynthesis pathway

Locus (gene)	Source	Protein expressed	Function	Reference
ECK0415 (*ispA* with base mutation T235C and T418A)	*E. coli*	His_6-IspA(Y79H, S140T)	IPP+DMAPP→FPP, FPP+IPP→GGPP	[25]
BSU28760 (*araM*)	*B. subtilis*	G1P dehydrogenase-His_6	DHAP + NADH → G1P	[31]
MmarC7_1004	*M. maripaludis*	His_8-GGGP synthase (codon optimized)	G1P + GGPP → GGGP	[45]
AF0404 (*ubiA*)	*A. fulgidus*	His_8-DGGGP synthase (codon optimized)	GGPP + GGGP → DGGGP	[38]
AF1740	*A. fulgidus*	CarS-His_8 (codon optimized)	DGGGP + CTP → CDP-archaeol	This study

Table S3 related to Experimental procedures. Plasmids and primers used in this study

Plasmids	Description	References
pRSFDuet-1	Cloning and expression vector (Kan^R), T7 promoter	Novagen
pETDuet-1	Cloning and expression vector (Amp^R), T7 promoter	Novagen
pCDFDuet-1	Cloning and expression vector (Str^R), T7 promoter	Novagen

pTara	T7 RNA polymerase expression vector (Cm^R), *ara*BAD promoter	[119]
pSJ120	Synthetic gene encoding codon optimized CarS from *A. fulgidus* with C-terminus His_8 tag was cloned into pRSFDuet vector using restriction sites *NdeI* and *XhoI*	This study
pSJ103	Synthetic gene encoding codon optimized GGGP synthase from *M. maripaludis* with N-terminus His_8 tag was cloned into pRSFDuet vector using restriction sites *EcoRI* and *HindIII*	This study
pSJ122	Synthetic gene encoding codon optimized DGGGP synthase from *A. fulgidus* with N-terminus His_8 tag was and redesigned ribosome binding site AGGACGTTAACAT cloned into pRSFDuet vector using restriction sites *HindIII* and *XhoI*	This study
pSJ130	PCR product of *araM* gene with C-terminus His_6 tag was obtained from *B. subtilis* genomic DNA using primers 70 and 71 and cloned into *NdeI* and *XhoI* sites of pETDuet vector	This study
pSJ140	PCR product of *cdsA* gene with N-terminus His_6 tag was obtained from *E. coli* genomic DNA using primers 105 and 106 and cloned into *BamHI* and *NcoI* sites of pRSFDuet vector	This study
pSP002	PCR product of *ispA* gene with N-terminus His_6 tag was obtained from *E. coli* K12 genomic DNA using primers 56 and 57 and cloned into *NcoI* and *EcoRI* sites of pCDFDuet vector.	This study
pSP003	Using primers 58 and 59 and template pSP002, PCR product with *ispAS140T* mutation was obtained. The linear fragment was closed by blunt ligation.	This study
pSP005	Using primers 62 and 63 and template pSP003, PCR product with *ispAY79H, S140T* mutation was obtained. The linear fragment was closed by blunt ligation.	This study

Primer name	**Primer sequence 5'→3'**
56	GCCGCCATGGGCAGCCATCACCATCATCACCACAGCATGGACTTTCCGCAGCAACTC
57	GCGCGAATTCTTATTTATTACGCTGGATGATGTAG
58	ACGATGATTTCTGAACTGGCGAGC
59	AATTCTGTCGCGGTCCGACAC
62	CACTCATTAATTCATGATGATTTACCGGCAATGG
63	AGCGTGGATACACTCAACGGC
70	GCGCCATATGAATCGTATCGCAGCTGAC
71	GCGCCTCGAGTTAGTGATGATGGTGGTGATGTTCATATAGACCATGGTTGATCAGCG

105	GCGCGGATCCTTAAAGCGTCCTGAATACCAGTAAC
106	GCCGCCATGGGCAGCCATCACCATCATCACCACAGCCTGAAGTATCGCCTGATATCTGC

Supplemental Experimental procedures

Reagents

Isopentenyl pyrophosphate, IPP (CAS number: 116057-53-5), farnesylpyrophosphate, FPP (CAS number: 13058-04-3) and rac-glycerol-1-phosphate, rG1P (CAS number: 34363-28-5) were purchased from Santa Cruz. Dihydroxyacetone phosphate, DHAP (CAS number: 102783-56-2), β-nicotinamide adenine dinucleotide, NADH (CAS number: 606-68-8), cytidine 5′-triphosphate disodium salt, CTP (CAS number: 36051-68-0) were from Sigma and geranylgeranyl diphosphate, GGPP (CAS number: 104715-21-1) from Echelon Biosciences. Lipids 1,2-dipalmitoyl-sn-glycero-3-phosphate, 16:0 PA (CAS number: 169051-60-9) and 1,2-dioleoyl-sn-glycero-3-phosphate, 18:1 PA (CAS number: 108392-02-5) were purchased from Avanti polar lipids. Radionucleotides [5-^{3}H]-cytidine 5`-triphosphate and [4-^{14}C]-Isopentenyl pyrophosphate, triammonium salt, were obtained from Perkin Elmer with specific activity >20 Ci (740GBq)/mmol and 56.6 mCi (2.09GBq)/mmol, respectively.

Bacterial strains and plasmids

Escherichia coli strain DH5α (Invitrogen) was used for cloning and BL21 (DE3) or Lemo21 (DE3) [122] (New England Biolabs) were used for protein over-expression. *E. coli* strains GN80 and GN85 [111] were gifts from William Dowhan. A list of plasmids and primers used in this study are listed in **Table S3**. *E. coli* cultures were grown aerobically at 37 °C in Luria-Bertani (LB) medium supplemented with appropriate antibiotics; ampicillin (100 μg/ml), chloramphenicol (34 μg/ml), kanamycin (50 μg/ml) or streptomycin (50 μg/ml). Codon optimized synthetic genes of GGGP synthase (GeneScript), DGGGP synthase and CDP-archaeol synthase (Invitrogen) were cloned in the expression vectors as described in **Table S3**. Ribosome binding site for DGGGP synthase was redesigned using RBS

web calculator [113]. Site directed mutagenesis of *ispA* was performed using primers with required mutations as described in **Table S3**.

NADH oxidation assay

G1P dehydrogenase enzyme activity was measured as described previously [31]. Different concentrations of purified enzyme were incubated with 0.25 mM NADH (Sigma) and 0.1 mM DHAP (Sigma) in buffer containing 0.5 mM $NiCl_2$ and 50 mM Tris-HCl, pH 8.0 at 37°C in a 100 µl reaction volume. The absorbance at 340 nm was measured every one minute for 30 minutes in a microtitre plate reader (SynergyMX, BioTek).

Peptide mass fingerprinting of CarS

Purified CarS was run on an SDS-PAGE gel and stained with coomassie G250 **(Figure S3)**. The gel band was excised and in-gel digestion was performed with trypsin and the frgmented peptides were analyzed by LC-MS/MS. Instrument parameters were as follows- Ion sourse, ESI (nano spray); fragmentation mode, CID, CAD (y and b ions); MS scan mode, FT-ICR/Orbitrap; MS/MS scan mode, linear ion trap. The swissprot-fasta database was searched using search engine PEAKS 7.0 with parent mass error tolerance 10 ppm, fragment mass error tolerance 0.5Da, monoisotopic precursor mass, 1 missed cleavage, carbamidomethylation and oxidation as modifications.

Western blotting

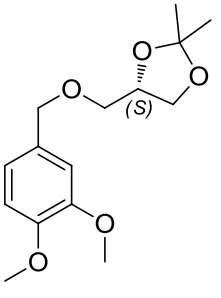
Tetra his antibody (5 PRIME) was used with 1:3000 dilution and Anti-mouse IgG alkaline phosphatase conjugated (Sigma) was used as the secondary antibody with 1:10000 dilution. The chemiluminescence signal was obtained using the CDP-star substrate (Roche) on a Luminescent image analyzer (Fujifilm).

Chemical synthesis of unsaturated archaetidic acid, DGGGP

Solvents for chemical reactions were dried according to the standard procedures. Solvents for flash chromatography were used without further purification. All reagents were obtained from Sigma-Aldrich and were used

without further purification unless noted otherwise. All the reactions were performed using Schlenk techniques. Glassware was dried by heating (150 °C) for at least 2 h and subsequent cooling under vacuum before use. Reactions were monitored by GC/MS (GC, HP6890: MS HP5973) equipped with an HP1 column (Agilent Technologies, Palo Alto, CA) or by TLC on silica coated aluminum foils (60 Å, 0.25 mm coating thickness). TLC was performed with the solvent chloroform, methanol, 7 M ammonia in the ratio 60:35:8 [42]. TLCs were visualized by the following stains: iodine stain, Seebach's stain, bromcresol green stain or Dittmer stain (for phospholipids). Flash column chromatography was performed on 230-430 mesh silica gel. ^{1}H, ^{13}C, ^{31}P NMR spectroscopy was recorded on Varian VXR300 or AMX400 spectrometers. Chemical shifts were determined relative to the residual solvent peaks ($CHCl_3$, δ = 7.26 ppm for ^{1}H NMR, δ = 77.16 ppm for ^{13}C NMR). The (–) sign stands for negative phase in APT (Attached Proton test). Optical rotations were measured on a Schmidt+Haensch polarimeter (Polartronic MH8) with a 10 cm cell. The mass spectra were recorded on a Thermoscientific LTQ OrbitrapXL spectrometer.

(*S*)-4-(((3,4-dimethoxybenzyl)oxy)methyl)-2,2-dimethyl-1,3-dioxolane



A dry three-necked round-bottom flask equipped with a reflux condenser was charged with (*S*)-1,2-isopropylidene glycerol (744 mg, 4 mmol) and tetra-n-butylammonium iodide (10 mol%, 147 mg). Solids were degassed in three cycles before dry THF (12 ml) was added. To the obtained solution, KH (50% in paraffin, 370 mg, 1.15 equiv) was added in small portions. The mixture was stirred for 10 min before 4-(chloromethyl)-1,2-dimethoxybenzene[123] (860 mg, 1.15 equiv) was added in one portion. The so-obtained reaction mixture was immersed into a preheated oil bath

(87 °C) and refluxed for 16 h. After removal from the oil bath and cooling down to rt, solid NH_4Cl (1 g) was added. The mixture was stirred for 15 min, filtered and the collected filtrate was evaporated to dryness. The yellow liquid residue was further purified by flash chromatography using 50% Et_2O in pentane. Fractions with an R_f = 0.37 (50% Et_2O in pentane) were collected and concentrated to afford 1.04 g of the desired compound as colorless thick liquid (92%).

^{1}H NMR (400 MHz, $CDCl_3$) δ 6.93 – 6.70 (m, 3H), 4.54 – 4.43 (m, 2H), 4.27 (m, 1H), 4.03 (dd, *J* = 8.2, 6.5 Hz, 1H), 3.71 (dd, *J* = 8.2, 6.4 Hz, 1H), 3.51 (dd, *J* = 9.8, 5.8 Hz, 1H), 3.44 (dt, *J* = 12.3, 4.7 Hz, 1H), 1.40 (s, 3H), 1.34 (s, 3H).

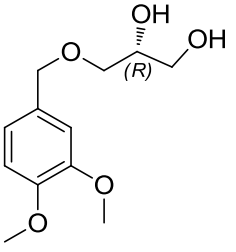
^{13}C NMR (101 MHz, $CDCl_3$) δ 149.20, 148.83, 130.66, 120.51, 111.21, 111.02, 109.54, 74.93, 73.56, 70.96, 67.00, 56.06, 55.99, 26.95, 25.54.

α^D =+15.9 (c =0.067, $CHCl_3$).

Elemental analysis calculated for $C_{15}H_{22}O_5$ (C: 63.81%, H: 7.85%) found (C: 63.51%, H:7.88%).

The spectroscopic data correspond to previously published [124]

(*R*)-3-((3,4-dimethoxybenzyl)oxy)propane-1,2-diol



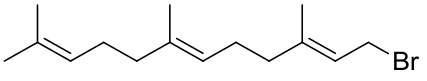
The corresponding acetonide (950 mg, 3.36 mmol) was dissolved in CH_2Cl_2/MeOH (10ml/10ml). Amberlite IR120 (acid form, 100 mg) was added and the mixture was stirred at RT until full conversion (36 h). The catalyst was filtered off and the filtrate was concentrated *in vacuo*.

The reaction afforded 815 mg of desired product (>99%) as colorless very thick liquid.

^{1}H NMR (400 MHz, $CDCl_3$) δ 6.88 – 6.81 (m, 3H), 4.47 (s, 2H), 3.88 (s, 3H), 3.87 (s, 3H), 3.69 (dd, *J* = 11.4, 3.9 Hz, 1H), 3.61 (dd, *J* = 11.4, 5.5 Hz, 1H), 3.57 – 3.48 (m, 2H), 2.31 (s, 2H).

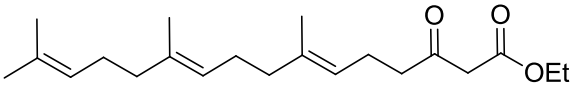
^{13}C NMR (101 MHz, $CDCl_3$) δ 149.02, 148.76, 130.18, 120.47, 111.16(–), 110.94(–), 73.45, 71.44, 70.71(–), 64.02, 55.89(–), 55.86 (–).

Farnesyl bromide



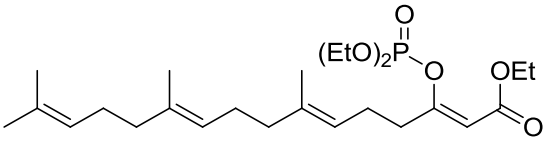
A dry flask was charged with farnesol (2.0 g, 8.9 mmol). Dry THF (30 ml) was added and resulting solution was immersed into a -47 °C bath (ethanol, cryostat). After stirring for 10 min, freshly distilled MsCl (900 µl, 1.3 equiv) was added via syringe over 5 min. Subsequently, Et_3N (2.5 ml, 2.0 equiv) was added over another 5 min. After complete addition, the mixture was stirred for 45 min at -47 °C. To the resulting suspension a solution of LiBr (3.0 g, 4.0 equiv) in dry THF (10 ml) was added dropwise over 5 min. After complete addition, the reaction vessel was transferred to a 0 °C bath (ice/water) and stirred for 1 h. The reaction mixture was poured into chilled saturated $NaHCO_3$ solution. The organic layer was separated, the aqueous layer was extracted with cold Et_2O (a mixture of Et_2O with pieces of ice, 3x25 ml), the combined organic layers were washed with cold water, brine, dried over $MgSO_4$ and evaporated. The crude farnesyl bromide was obtained as a yellow liquid and used without further manipulation.
The reaction afforded 1.91 g of the desired product as a light yellow oil (75% yield) which was stored at -80 °C, and used within 1 day.

(6E,10E)-ethyl 7,11,15-trimethyl-3-oxohexadeca-6,10,14-trienoate



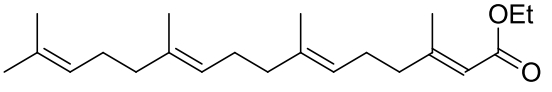
A dry Schlenk flask was charged with NaH (60% dispersion, 880 mg, 3.3 equiv). The mineral oil was removed by three washings with pentane (3x10 ml). The resulting white solid was dried in vacuum. Dry THF (16 ml) was added and the resulting white suspension was cooled to 0 °C (ice/water bath). To this suspension, freshly distilled ethyl acetoacetate (2.6 ml, 3 equiv) was added dropwise over 5 min. After complete addition

the suspension turned into a light yellow solution. To this solution was added a solution of ^{n}BuLi in hexanes (2.5 M, 8.3 ml, 3.1 equiv) over 15 min. The resulting orange solution was stirred for additional 15 min at 0 °C before a solution of farnesyl bromide (from the previous experiment 1.91 g, 6.7 mmol) in dry THF (3.5 ml) was added dropwise over 5 min. The resulting solution was stirred at 0 °C for 15 min, during which formation of a precipitate was observed. The reaction was quenched by careful addition of chilled aqueous HCl (1 M, 10 ml, **EXOTHERMIC**). The mixture was transferred into a separatory funnel, where the organic layer was separated, the aqueous layer was extracted with Et_2O (3x10 ml), the combined organic layers were washed with brine, dried and evaporated.
The title compound was obtained after column chromatography using 10% Et_2O in pentane as 1.67 g of a light yellow liquid (55% starting from farnesol, lit 70%).
^{1}H NMR (400 MHz, CDCl3) δ 12.09 (s, 0.2H), 5.08 (d, *J* = 6.5 Hz, 3H), 4.19 (dt, *J* = 7.2, 5.3 Hz, 2H), 3.42 (s, 2H), 2.56 (t, *J* = 7.4 Hz, 2H), 2.38 – 2.17 (m, 3H), 2.05 -1.98 (m, 8H), 1.68 (s, 3H), 1.64 – 1.58 (m, 9H), 1.28 (t, *J* = 7.1 Hz, 3H). This corresponds to the data reported in the literature[125].

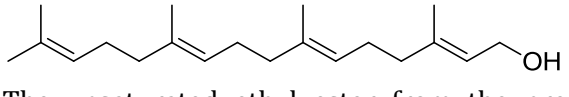


A dry Schlenk flask was charged with NaH (60% dispersion, 228 mg, 1.15 equiv). The mineral oil was removed by three successive washings with hexane (3 x 5 ml) and the resulting white solid was suspended in dry Et_2O (21 ml). The suspension was immersed in an ice/water bath of 0 °C and a solution of (6*E*,10*E*)-ethyl 7,11,15-trimethyl-3-oxohexadeca-6,10,14-trienoate (1.65 g, 4.95 mmol) was added as a solution in dry Et_2O (7 ml) over 15 min. After the addition was complete, the resulting light yellow solution was stirred for 15 min at 0 °C and for 15 min at 21 °C (RT). Then the solution was again cooled in the ice bath and neat $(EtO)_2P(O)Cl$ (1.1 ml, 1.5 equiv) was added dropwise. The resulting reaction mixture was stirred for 15 min at 0 °C and subsequently quenched by addition of saturated aqueous NH_4Cl solution (15 ml). The organic layer was separated, and the

aqueous layer was extracted with Et_2O (3 x 15 ml). The combined organic layers were washed with saturated aqueous $NaHCO_3$ (3 x 15 ml), brine (2 x 15 ml), dried over $MgSO_4$ and the solvent was removed *in vacuo*.
The resulting crude (2.08 g) was obtained as a yellow liquid and used without further purification in the following step.



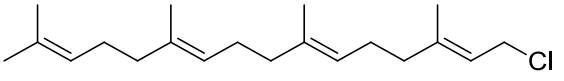
A dry Schlenk flask was charged with CuI (1.7 g, 1.8 equiv) which was suspended in dry Et_2O (5.5 ml) and cooled to 0 °C. The resulting suspension was treated with MeLi (1.6 M in Et_2O, 3.6 equiv). The suspension turned yellow after initial addition of MeLi, after complete addition the CuI fully dissolved affording a nearly colorless solution.
The reaction vessel with Me_2CuLi was immersed in a cryostat at -78°C and a solution of the phosphate (2.08 g, previous step) in Et_2O (dry, 7 ml) was added dropwise via the cold wall of the Schlenk flask (in order to chill the solution of the phosphate). After complete addition, the color changed to orange/red and the resulting solution was stirred at -78 °C. After 1 h the bath was allowed to warm to -47 °C and the reaction mixture was stirred at -47 °C for 2 h. After this time TLC showed full conversion of the phosphate and a new spot had appeared on TLC. MeI (630 µl) was added to quench the unreacted cuprate. After stirring for 10 min, the reaction mixture was carefully poured into a solution of NH_4Cl (24 ml) and NH_4OH (6 ml) (**can be exothermic with gas evolution**). The mixture was stirred until all solids dissolved. Layers were separated, the aqueous layer was extracted with Et_2O (3 x 20 ml) and the combined organic layers were washed with NH_4OH (10%, 2 x 40 ml), brine (2 x 40 ml), dried and evaporated.
The reaction afforded 1.15 g of a yellow liquid which was used without further purification.



The unsaturated ethyl ester from the previous step was dissolved in toluene (p.a. grade, 17 ml). This solution was cooled to -78 °C (N_2/acetone

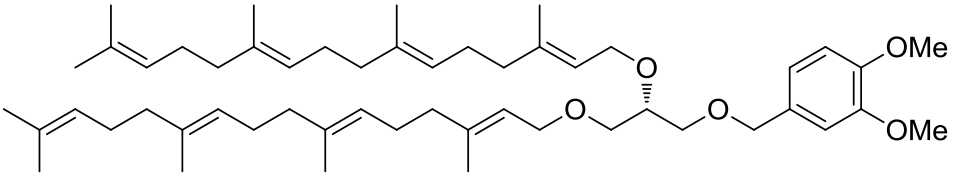
bath) and a solution of DIBAL (2.6 ml, 3 equiv) in hexane (15 ml) was added dropwise. The mixture was stirred until complete consumption of the starting material (TLC). The reaction was quenched by careful addition of MeOH (3 ml, added over 10 min, **EXOTHERMIC, GAS EVOLUTION**). When gas evolution ceased, the mixture was removed from the bath and stirred for 10 min at rt. The reaction mixture was poured into saturated NH_4Cl (50 ml)/HCl (50 ml) solution and stirred until clear separation of the layers took place (ca 30 min). The aqueous layer was extracted with Et_2O (3 x 50 ml). The combined organic layers were washed with water (2 x 50 ml) and brine (2 x 50 ml), dried over $MgSO_4$ and evaporated. The residual thick liquid was further purified by flash chromatography on silica using 30% Et_2O in pentane as the eluent.

The reaction afforded 857 mg of geranylgeraniol (60% over three steps) with >99% double bond isomer purity according to GC analysis. The spectral data correspond to those previously reported[125].



N-chlorosuccinimide (1.2 g, 1.3 equiv) was suspended in dry CH_2Cl_2 (15 ml). The suspension was cooled to -30 °C (acetone/liquid N_2 bath) and after stirring for 5 min, dimethyl sulfide (750 µl, 1.5 equiv) was added. The reaction was stirred for 10 min at -30 °C and 10 min at 0 °C. Then the solution was cooled to -40 °C. A solution of geranylgeraniol (2 g, 6.9 mmol) in CH_2Cl_2 (dry, 5 ml) was added dropwise. The resulting suspension was allowed to warm during 150 min to 0 °C, turning into a cloudy solution at -15 °C. The reaction mixture was poured into pentane (150 ml), The organic layer was washed with water (2x50 ml), brine (50 ml), dried over $MgSO_4$ and evaporated.

The reaction afforded 1.94 g of geranylgeranyl chloride as a colorless liquid which was used without further purification.



A dry Schlenk flask was charged with NaH (60% dispersion in mineral oil, 2.1 equiv). The mineral oil was removed by washing with pentane (3 x 5 ml) and the white solid was dried in high vacuum before suspending in DMSO (5 ml). The obtained suspension was immersed in a preheated oil bath (70 °C) and stirred for 40 min during which the suspension turned into a pale yellow solution. The flask was removed from the bath and allowed to cool to rt (21 °C). To this solution, a solution of (R)-3-((3,4-dimethoxybenzyl)oxy)propane-1,2-diol in dry DMSO (5 ml) was added carefully. After the complete addition, the reaction mixture was stirred for 1 h at rt (21 °C). Then a mixture of geranylgeranyl chloride (1.94 g from the previous experiment) in a small amount of DMSO (2 ml) was added. The resulting solution was stirred for 16 h before pouring into saturated aqueous NH_4Cl solution (20 ml). The aqueous layer was extracted with Et_2O (3 x 50 ml). The combined organic layers were washed with brine, dried and evaporated. The crude residue was further chromatographed using 30% Et_2O in pentane to afford the desired product and the product of the mono-alkylation.

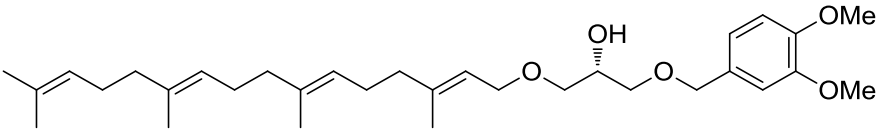
The dialkylated product 1.52 g (61%) was obtained as a colorless liquid

^1H NMR (400 MHz, $CDCl_3$) δ 6.97 – 6.75 (m, 3H), 5.35 (dt, *J* = 13.5, 6.6 Hz, 2H), 5.10 (d, *J* = 5.8 Hz, 6H), 4.49 (s, 2H), 4.16 (d, *J* = 6.7 Hz, 2H), 4.01 (d, *J* = 6.7 Hz, 2H), 3.87 (d, *J* = 2.6 Hz, 6H), 3.68 (dt, *J* = 10.0, 5.1 Hz, 1H), 3.62 – 3.46 (m, 5H), 2.16 – 1.90 (m, 23H), 1.68 (s, 6H), 1.65 (s, 26), 1.59 (s, 16H)

^{13}C NMR (101 MHz, $CDCl_3$) δ 149.06, 148.60, 140.20, 139.95, 135.43, 135.40, 135.06, 131.39, 131.10, 124.51 (–), 124.32 (–), 124.31 (–), 124.03 (–), 123.99 (–), 121.33 (–), 120.99 (–), 120.31 (–), 111.07 (–), 110.91 (–), 77.03(–), 73.40, 70.35, 70.19, 68.02, 66.95, 56.03 (–), 55.92 (–), 39.86, 39.84, 39.78, 26.89, 26.78, 26.53, 26.50, 25.85(–), 17.83 (–), 16.69(–) , 16.66 (–), 16.15 (–) .

α^D =+5.2 (c =1.0, $CHCl_3$)

NMR data correspond to those previously published[86].



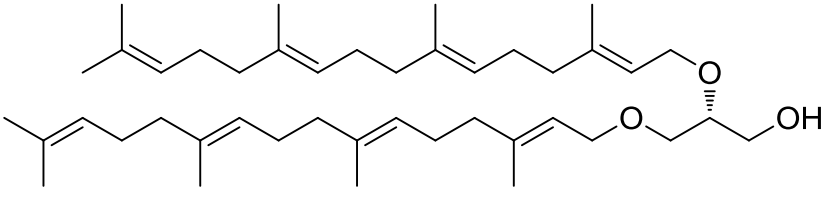
The monoalkylated product 256 mg (16%) was obtained as a colorless liquid.

^{1}H NMR (400 MHz, $CDCl_3$) δ 6.84 (m, 3H), 5.34 (d, *J* = 7.1 Hz, 2H), 5.10 (s, 3H), 4.48 (d, *J* = 5.0 Hz, 2H), 4.01 (m, 3H), 3.87 (2x s, 6H), 3.76 – 3.40 (m, 5H), 2.05 (m, 12H), 1.66 (d, *J* = 8.3 Hz, 6H), 1.59 (s, 9H).

^{13}C NMR (101 MHz, $CDCl_3$) δ 149.11, 148.76, 148.73, 140.72, 140.67, 135.49, 135.45, 135.05, 131.36, 130.62, 124.47 (–), 124.26 (–), 123.88 (–), 123.85 (–), 120.79 (–), 120.59 (–), 120.48 (–), 120.38 (–), 111.16 (–), 111.02 (–), 110.95 (–), 110.93 (–), 77.61 (–), 73.53, 73.48, 71.34, 71.22, 69.93, 69.70 (–), 67.94, 66.67, 63.01, 56.01 (–), 55.93 (–), 55.92 (–), 39.83, 39.80, 39.72, 26.86, 26.72, 26.43, 25.82 (–), 17.80 (–), 16.66 (–), 16.63 (–), 16.13 (–) 16.12 (–).

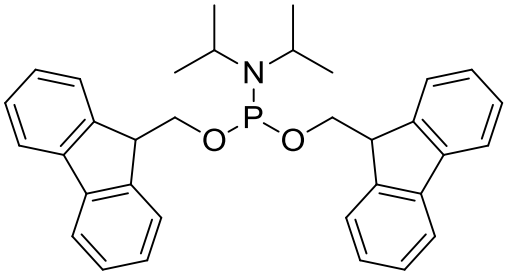
α^{D} = +9.2 (c =1.0, $CHCl_3$).

HRMS calculated for $C_{32}H_{50}O_5Na$ [M+Na]: 538.354 found 538.355



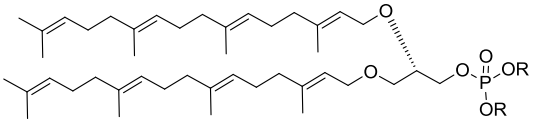
Protected (*R*)-2,3-bisgeranylgeranyl glycerol (1.44 g, 1.83 mmol) was dissolved in CH_2Cl_2 (30 ml). To this solution, water (0.75 ml) was added. The obtained biphasic solution was cooled in an ice bath (0 °C) and DDQ (830 mg, 2 equiv) was added. The reaction was stirred for 4 h at 0 °C until TLC showed full conversion of the starting material. The crude reaction mixture was filtered over a small silica pad and washed with CH_2Cl_2 (200 ml). The washings were combined and evaporated. The obtained yellow liquid was further purified by column chromatography (30% Et_2O in pentane).

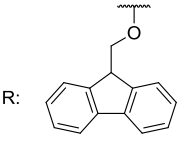
The reaction afforded 705 mg of the desired product as yellow thick liquid containing traces of unknown co-eluting impurities.
^{1}H NMR (400 MHz, $CDCl_3$) δ 5.35 (dt, *J* = 13.5, 6.7 Hz, 2H), 5.11 (t, *J* = 6.5 Hz, 6H), 4.26 – 4.06 (m, 2H), 4.02 (d, *J* = 6.7 Hz, 2H), 3.77 – 3.41 (m, 5H), 2.16 – 1.93 (m, 24H), 1.68 (s, 12H), 1.60 (s, 21H)
^{13}C NMR (101 MHz, $CDCl_3$) δ 140.68, 135.49, 135.47, 135.07, 131.37, 124.52 (–), 124.32 (–), 123.9 (–)3, 120.91 (–), 120.69 (–), 77.61 (–), 70.16, 68.08, 66.66, 63.21, 39.88, 39.86, 39.83, 39.75, 26.90, 26.77, 26.50, 26.46, 25.83 (–), 17.82 (–), 16.68 (–), 16.65 (–), 16.15 (–), 16.14 (–).
HRMS calculated for $C_{43}H_{72}O_3Na$ [M+Na]: 659.536, found: 659.537.



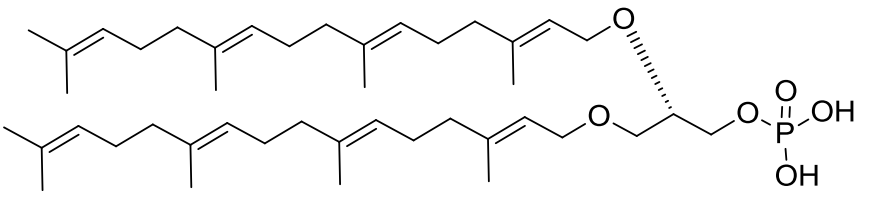
PCl_3 (7 ml, 80 mmol) was dissolved in pentane (700 ml). Via an addition funnel, a solution of diisopropylamine (distilled from CaH_2, 23 ml, 2 equiv) in pentane (100 ml) was added dropwise over 15 min. After this time, a significant amount of white precipitate had formed. The suspension was stirred for 2 h. The mixture was transferred to a separatory funnel. The pentane layer was washed with acetonitrile (the pentane layer stays on top, 5 x 100 ml of acetonitrile, after the washing the pentane layer was fully transparent). Pentane was subsequently evaporated. The reaction afforded 1,1-dichloro-N,N-diisopropylphosphinamine (6 g, 38%) as colorless liquid (^{1}H NMR (400 MHz, $CDCl_3$) δ 4.02 – 3.83 (m, 1H), 1.28 (d, *J* = 6.8 Hz, 6H), ^{31}P NMR (162 MHz, $CDCl_3$) δ 169.59).
From the obtained liquid, an aliquot was taken (2.0 g, 10 mml). This was dissolved in dry THF (40 ml) and DIPEA (3.5 ml, 2 equiv) was added. The solution was cooled in an ice bath and (9H-fluoren-9-yl)methanol (3.92 g, 2 equiv) in dry THF (10 ml) was added. The solution was stirred for 10 h during which a white precipitate formed. The reaction was poured into aqueous phosphate buffer (1 M, pH = 7) and extracted with ethyl acetate (4

x 50 ml). The combined organic extracts were washed with the same phosphate buffer, brine, dried and evaporated. The crude product was used without further purification due to its sensitivity.
^{31}P NMR (162 MHz,CDCl3) δ 148.01.





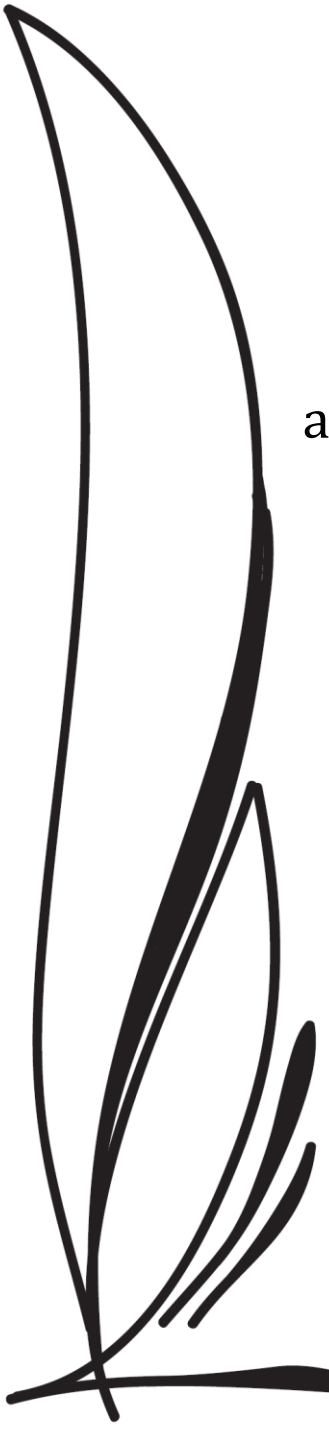
(*R*)-2,3-bisgeranylgeranyl glycerol was dissolved in CH_2Cl_2/CH_3CN (0.5 ml/0.5 ml) and bisfluorenylmethyl phosphoramidite (154 mg, 3 equiv) was added. The mixture was cooled to 0 °C and tetrazole (21 mg, 3 equiv) was added. The mixture was allowed to gradually warm to rt (21 °C) and stirred overnight. When full conversion of starting material was observed (TLC), the mixture was cooled and a solution of ^{t}BuOOH (5 M in decane, 3.2 equiv) was added in one portion followed by stirring for 45 min. The mixture was subsequently poured into aqueous phosphate buffer (1 M, pH = 7) and extracted with Et_2O (4 x 20 ml). The combined extracts were washed with brine, dried and concentrated. The crude residue was purified by flash chromatography using (50% Et_2O in pentane). Fractions with an Rf = 0.4 (50% Et_2O in pentane) were collected to afford 102 mg (94%) of the desired compound.
^{1}H NMR (400 MHz, $CDCl_3$) δ 7.71 (dd, *J* = 12.2, 4.3 Hz, 4H), 7.64 – 7.47 (m, 4H), 7.46 – 7.29 (m, 4H), 7.29 – 7.14 (m, 4H), 5.27 (q, *J* = 6.8 Hz, 2H), 5.20 – 4.98 (m, 6H), 4.34 – 4.21 (m, 4H), 4.20 – 4.07 (m, 3H), 4.08 – 3.90 (m, 5H), 3.59 (dd, *J* = 9.7, 4.9 Hz, 1H), 3.54 – 3.53 (m, 2H), 2.18 – 1.88 (m, 24H), 1.68 (s, 6H), 1.59 (dd, *J* = 11.2, 6.0 Hz, 24H)
^{13}C NMR (101 MHz, $CDCl_3$) δ 147.46, 147.40, 145.60, 144.69, 144.56, 139.59, 139.57, 139.21, 135.52 (-), 132.10 (-), 131.36 (-), 129.48 (-), 129.43 (-), 128.66 (-), 128.47 (-), 128.09 (-), 124.89 (-), 124.86 (-), 124.25 (-), 124.21 (-), 80.44 (-), 80.36 (-), 73.59, 73.53, 73.07, 72.22, 71.45, 71.40, 71.07, 52.24 (-), 52.16 (-), 44.00, 43.98, 43.87, 31.04, 30.93, 30.70, 30.65, 29.98 (-), 21.97 (-), 20.78 (-), 20.29 (-).
HRMS calculated for $C_{71}H_{93}O_6PNa$ [M+Na]: 1095.660, found 1095.660.



Bisprotected phosphoric ester (102 mg, 100 µmol) was dissolved in acetonitrile (5 ml). To this solution, Et_3N (20 equiv) was added and the resulting mixture was stirred overnight. All volatiles were evaporated and the crude mono deprotected phosphoric ester (as assumed from the TLC) was suspended in aqueous NaOH (1 M, 5 ml) until full conversion of the monoprotected ester was observed (TLC, 3 h). The mixture was acidified with HCl (1 M) to pH = 1 and extracted with Et_2O (3 x 20 ml). The combined extracts were dried over $MgSO_4$ and concentrated. The crude residue was purified on a silica column using a carefully established gradient of 2% MeOH/$CHCl_3$ → 33% MeOH/$CHCl_3$.
The reaction afforded 34.3 mg (48%) of the desired compound as a colorless liquid.
NOTE: The final product was stored in the freezer (- 20 °C), due to its relative instability.
^{1}H NMR (400 MHz, $CDCl_3$) δ 5.32 (m 2H), 5.09 (m, 6H), 4.15 (m, 2H), 4.00 (m, 4H), 3.57 (m, 3H), 2.01 (m, 24H), 1.77 – 1.49 (m, 30H).
^{13}C NMR (101 MHz, $CDCl_3$, the spectrum shows a considerable number of overlapping signals) δ 135.23, 134.85, 131.16, 124.38, 124.22, 123.88, 120.71, 39.75, 39.72, 26.76, 26.61, 25.67, 17.66, 16.53, 16.00, 15.97.
^{31}P NMR (162 MHz, $CDCl_3$) δ 1.37.
HRMS: calculated for $C_{43}H_{72}O_6P$ [M-1]: 715.507 found: 715.506

NMR Spectra

The NMR spectras of the chemical synthesis can be found with article online at http://dx.doi.org/10.1016/j.chembiol.2014.07.022.



Chapter 3

Formation of the ether lipids archaetidylglycerol and archaetidylethanolamine in *Escherichia coli*

Biochemical Journal 2015, 470, 343-355

Antonella Caforio[1] Samta Jain[1,4], Peter Fodran[2], Melvin Siliakus[3], Adriaan Minnaard[2], John van der Oost[3] and Arnold J. M. Driessen[1]

[1]Department of Molecular Microbiology, Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, 9747 AG Groningen, The Netherlands; The Zernike Institute for Advanced Materials, University of Groningen, 9747 AG Groningen, The Netherlands

[2]Stratingh Institute for Chemistry, University of Groningen, 9747 AG Groningen, The Netherlands

[3]Department of Microbiology, Wageningen University, Dreijenplein 10, 6703 HB Wageningen, The Netherlands

[4]Present address: Department of Medicine, Section of Infectious Diseases, Boston University School of Medicine, 02118 Boston, Massachusetts, United States of America

Abstract

In archaea, the membrane phospholipids consist of isoprenoid hydrocarbon chains that are ether-linked to a *sn*-glycerol-1-phosphate backbone. This unique structure is believed to be vital for the adaptation of these microorganisms to extreme environments, but it also reflects an evolutionary marker that distinguishes archaea from bacteria and eukaryotes. CDP-archaeol is the central precursor for polar head group attachment. We examined various bacterial enzymes involved in the attachment of L-serine and glycerol as polar head groups for their promiscuity in recognizing CDP-archaeol as a substrate. Using a combination of mutated bacterial and archaeal enzymes, archaetidylethanolamine (AE) and archaetidylglycerol (AG) could be produced *in vitro* using nine purified enzymes while starting from simple building blocks. The ether lipid pathway constituted by a set of archaeal and bacterial enzymes was introduced into *E. coli*, which resulted in the biosynthesis of AE and AG. This is a further step in the reprogramming of *E. coli* for ether lipid biosynthesis.

Introduction

The cytoplasmic membrane is an essential constituent of cells. It forms a barrier that separates the cytosol from the external milieu. In conjunction with energy transducing complexes and transporter proteins, the phospholipid bilayer allows homeostasis of the intracellular concentration of nutrients and other metabolites within the cell [126]. The lipid composition of the cytoplasmic membrane differs between organisms and is one of the elements that distinguishes archaea from bacteria and eukarya. This marked diversity of the lipid composition between archaea and all other organisms is also termed "the lipid divide" that may finds its origin in the early stages of the evolution of life [88,89,91,127].

Archaeal lipids are composed of highly branched isoprenoid chains ether-linked to a glycerol-1-phosphate (G1P) backbone, compared to fatty acid chains ester-linked to the glycerol-3-phosphate (G3P) backbone as typically found in Bacteria and Eukarya. Besides these main characteristics, archaeal membranes display a further diversity in their lipid composition consisting of different modifications of the two major structures: *sn*-2,3-diphytanylglycerol diether, called archaeol and *sn*-2,3-diphytanylglycerol tetraether lipid, known as caldarchaeol[57,75,127]. The biosynthetic pathway leading to the formation of archaeal lipids has been studied in some detail [2,5,6,10,21,37] and most of the enzymes involved in the biosynthesis have been identified and characterized. However, the entire pathway is not completely understood, nor is it clear how caldarchaeol is formed. The isoprenoid building blocks isopentenyl pyrophosphate (IPP) and dimethylallyl pyrophosphate (DMAPP) are synthesized via the mevalonate pathway in archaea [10] and are combined through sequential condensation reactions catalyzed by geranylgeranyl diphosphate (GGPP) synthase and farnesylgeranyl diphosphate (FGPP) synthase, depending on the length of the isoprenoid chain product [19,44]. G1P in archaea and G3P in bacteria and eukarya are formed by similar reactions although the enzymes involved, i.e., glycerol-1-phosphate dehydrogenase (G1PDH) [28,29] and glycerol-3-phosphate dehydrogenase (G3PDH), are not evolutionarily related and belong to different protein families [2]. The elongated isoprenoid chain and G1P are subsequently

linked together through ether linkages by two prenyltransferases. A cytosolic protein geranylgeranylglyceryl phosphate synthase (GGGP synthase) selectively attaches the isoprenoid chain to the G1P [35,36] leading to the first ether bond formation. Next, the di-*O*-geranylgeranylglycerylphosphate synthase (DGGGP synthase) [38,39,41] catalyzes the second ether bond formation linking another isoprenoid chain to the lipid precursor yielding DGGGP. The following step concerns the activation of DGGGP via a CTP-transferring reaction by CDP-archaeol synthase (CarS) that was recently discovered [41]. CDP-archaeol is an important intermediate for the successive steps in lipid biosynthesis where the CDP group is replaced by a polar head group. Serine, ethanolamine, glycerol and *myo*-inositol are common polar head groups found throughout the three domains of life. The enzymes involved in the initial replacement of CMP from a CDP-alcohol with different types of polar head groups share mostly a common mode of action among eukarya, bacteria and archaea (**Figure 1**). A different mechanism to produce phosphatidylethanolamine (PE) has been described for Eukarya that involves a direct replacement of the CDP group with ethanolamine [128,129].

The *E. coli* membrane is composed of 70-80% of PE, 20-25% of phosphatidylglycerol (PG) and 5% or less of cardiolipin (diphosphatidylglycerol) as the major phospholipid components [126,130,131]. PG represents the main anionic phospholipid and is important for various cellular processes such as the initiation of DNA synthesis [126,132–134] and protein translocation [135–137], whereas PE is the main zwitterionic lipid. In order to re-program *E. coli* for the synthesis of archaeal ether lipids, it will be essential to produce the two archaeal lipids archaetidylethanolamine (AE) and archaetidylglycerol (AG) to accommodate at least the polar head group composition. For polar head group modification, the branch point in bacteria lies with CDP-diacylglycerol (CDP-DAG), produced by CDP-diacylglycerol synthase encoded by the *cdsA* gene [102]. CDP-DAG acts as a substrate for two sets of enzymes. For the formation of PE, the CDP group of CDP-DAG is replaced with L-serine by phosphatidylserine synthase (Pss) leading to the production of phosphatidylserine (PS). Next, PS is decarboxylated by

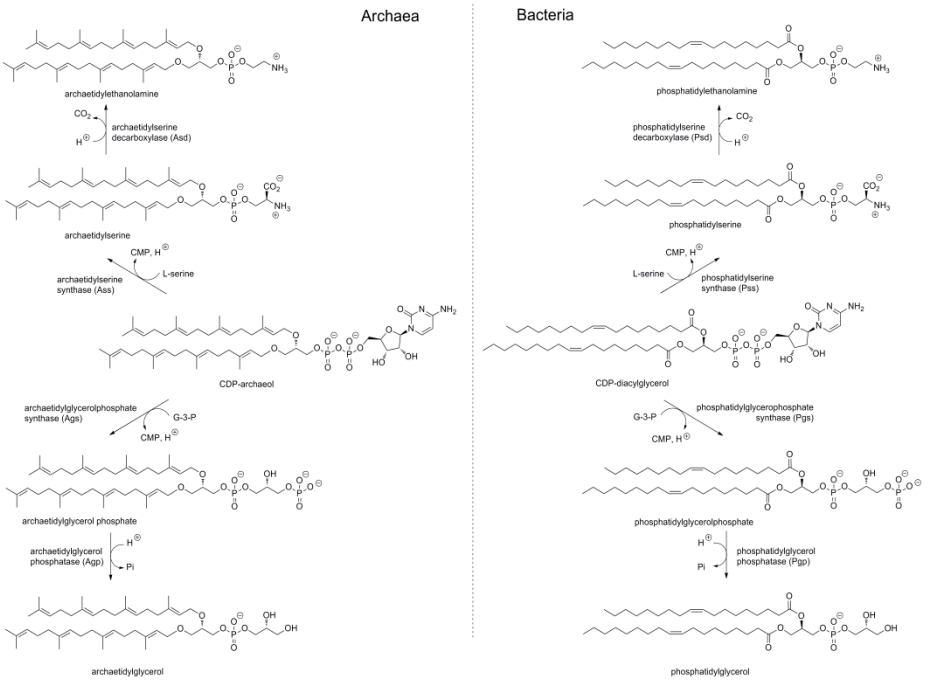


Figure 1 | Polar head group attachment in bacteria and archaea. The scheme represents the enzymatic steps of the lipid biosynthetic pathway for polar head group attachment in Archaea and Bacteria. The archaeal and the bacterial enzymes involved in the replacement of CDP-group with L-serine or glycerol from CDP-archaeol or CDP-diacylglycerol are indicated.

phosphatidylserine decarboxylase (Psd) converting it into PE. In the other biosynthetic pathway the CDP group of CDP-DAG is initially replaced by a glycerol-3-phosphate (G3P) moiety by PG synthase (Pgs) leading to the formation of phosphatidylglycerol phosphate (PGP). The enzyme phosphatidylglycerol phosphatase (Pgp) removes the phosphate resulting in the formation of PG [70,126,130,131]. In archaea, the formation of the corresponding archaeal lipids, AE and AG, seem to take place via very similar mechanisms (**Figure 1**). The two biosynthetic branches diverge from the CDP-archaeol towards the formation of AE or AG. The former is produced by the sequential action of two enzymes, archaetidylserine synthase (Ass) [47] and archaetidylserine decarboxylase (Asd) which replace the CDP group with L-serine whereupon a decarboxylation

reaction yields the AE. Likewise, AG synthesis involves the enzyme AG phosphate (AGP) synthase (Ags) which attaches a G3P to CDP-archaeol forming AGP which is subsequently dephosphorylated into AG by the action of archaetidylglycerol phosphatase (Agp)[7,93].

The bacterial and archaeal synthase enzymes involved in polar head group attachment all contain a well conserved domain, identified as D-G-x(3)-D-x(3)-D in the PROSITE database, that classifies these proteins as members of the CDP-alcohol phosphatidyltransferase family (**Supplementary Figure S1**) [7,43]. Previous bioinformatics analysis [7,8,93] revealed a wide distribution of these enzymes in bacteria and archaea suggesting the existence of an ancestral enzyme in the last universal common ancestor (LUCA) [43] able to produce both archaeal- and bacterial-like lipids. Therefore, this could indicate substrate promiscuity among these enzymes and possibly that bacterial enzymes are able to catalyze polar head group attachment to the CDP-archaeol. In this study, we have tested various bacterial enzymes for the formation of unsaturated AE and AG from CDP-archaeol and *in vitro* reconstituted the entire pathways employing simple building blocks and up to nine purified enzymes of archaeal and bacterial origin. In addition, we have introduced the archaeal ether lipid biosynthetic pathway for AE and AG synthesis into *E. coli* making use of the substrate promiscuity of bacterial enzymes catalyzing the final polar head group conversion steps.

Results

Archaetidylglycerol formation by bacterial PgsA and PgpA enzymes

PG synthesis in *E. coli* proceeds via two enzymatic steps. The first reaction is catalyzed by phosphatidylglycerophosphate synthase. *E. coli* contains two individual genes that encode for this enzyme activity [138,139]. The *pgsA* gene exclusively functions in phosphatidylglycerol formation [140], whereas the *pgsB* gene is also involved in the synthesis of the Lipid A core of Lipopolysaccharide (LPS) [103]. The resulting PGP is dephosphorylated to PG. In *E. coli*, three main PGPs are found, encoded by the *pgpA*, *pgpB* and *pgpC* genes [141]. PgpA exhibits a narrow substrate

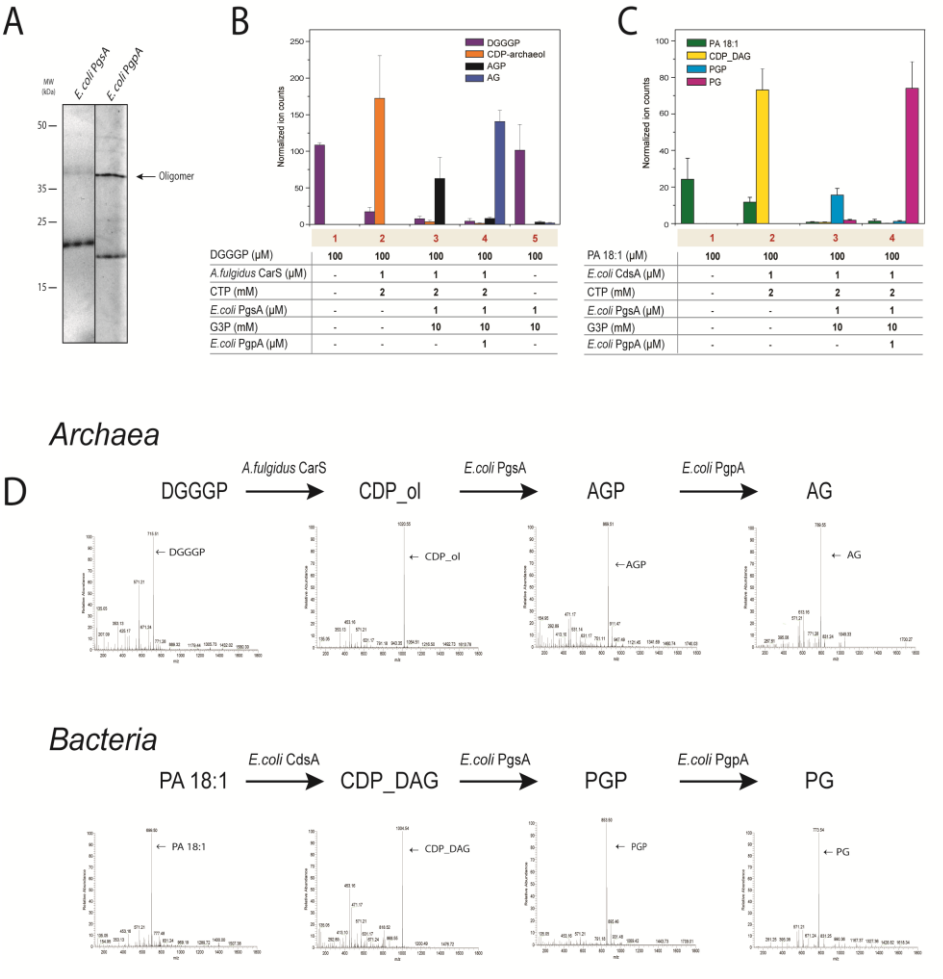


Figure 2| *In vitro* demonstration of AG biosynthesis involving the *E. coli* PgsA and PgpA. (**A**) Coomassie Blue-stained SDS-PAGE gels showing the Ni-NTA purified proteins PgsA (21 kDa) and PgpA (20.5 kDa) from *E. coli*. *In vitro* reactions using (**B**) DGGGP or (**C**) PA ($C_{18:1}$) and the purified proteins as specified, to test the substrate specificity of the *E. coli* PgsA and PgpA. Total ion counts from LC-MS data were normalized using DDM as internal standard. The data are the average of three experiments ± SE. (**D**) Schematic representation of the *in vitro* reactions. The purified enzymes used in the experiments are indicated above the arrows and the mass spectra from the LC-MS runs of the corresponding products are shown: DGGGP (m/z= 715.51 [M-H]-), CDP_ol (m/z= 1020.54 [M-H]-), AGP (m/z= 869.51 [M-H]-) and AG (m/z= 789.55 [M-H]-) in (**B**) and PA 18:1 (m/z= 699.50 [M-H]-), CDP_DAG (m/z= 1004.54 [M-H]-), PGP (m/z= 853.50 [M-H]-) and PG (m/z= 773.54 [M-H]-) in (**C**).

specificity dephosphorylating only PGP [142] while PgpB also shows phosphatase activity towards DAG pyrophosphate [147], phosphatidic acid and lysophosphatidic acid [144,145]. Bioinformatics analysis [7,8] revealed the occurrence of Pgs homologs in archaea that belong to the CDP-alcohol phosphatidyltransferase family (**Supplementary Figure S1**), suggesting functional and structural conservation of these enzymes among bacteria and archaea. The same analysis revealed a weak sequence conservation of Pgp in archaea, limited to PgpA and PgpB only (**Supplementary Figure S2**). To investigate the ability of bacterial enzymes to accept CDP-archaeol as a substrate for AG formation, the *pgsA* and *pgpA* genes of *E. coli*, were cloned and overexpressed in *E. coli* BL21. Both overexpressed proteins localized to the membrane fraction after cell fractionation [141]. Upon membrane solubilization with the detergent n-dodecyl-β-D-maltopyranoside (DDM), PgsA and PgpA could be purified by Ni-NTA affinity chromatography (**Figure 2A**). The activity of the enzymes was analyzed in *in vitro* coupled reactions in detergent solution using LC-MS for detection (**Figure 2D**). The activity of the *E. coli* PgsA towards the archaeal substrates CDP-archaeol (CDP-ol) was tested using chemically synthetised DGGGP (unsaturated archaetidic acid) as a substrate that was converted into CDP-ol in the presence of the *A. fulgidus* CarS and CTP as described previously [41] (**Figure 2B, lane 2**). Upon the addition of the *E. coli* PgsA and G3P, the formation of AGP could be demonstrated (m/z= 869.51 $[M-H]^-$) (**Figure 2B, lane 3**). In the subsequent reaction the conversion of AGP into AG (m/z= 789.55 $[M-H]^-$) was observed when also the *E. coli* PgpA was included in the reaction (**Figure 2B, lane 4**). The *E. coli* PgsA was unable to use archaeal DGGGP as substrate (**Figure 2B, lane 5**). As a control, the activities of PgsA and PgpA were also tested towards the bacterial substrate CDP-DAG. The latter was produced by incubating the purified *E. coli* CdsA protein with its substrates PA ($C_{18:1}$) and CTP (**Figure 2C, lane 2**). The formation of PGP (m/z= 853.50 $[M-H]^-$) was detected only in the presence of the *E. coli* PgsA and G3P (**Figure 2C, lane3**), which was further converted into the final product PG (m/z= 773.54 $[M-H]^-$) upon the addition of the *E. coli* PgpA (**Figure 2C, lane 4**). These results demonstrate that the *E. coli* PgsA and PgpA recognizes and

converts the archaeal substrate CDP-archaeol and AGP respectively, forming the archaeal polar lipid AG.

Archaetidylethanolamine formation by bacterial PssA and Psd enzymes

Like PG synthesis, PE synthesis also requires the action of two enzymes: Pss for the synthesis of phosphatidylserine (PS) which is further decarboxylated to PE by Psd [146]. Two different subclasses of Pss exist: Pss-I, a cytoplasmic protein present mainly in Gram-negative bacteria such as *E. coli*, and Pss-II, a membrane protein that is found in Gram-positive bacteria, yeast [147] and archaea [148]. The Pss-II enzymes contain a highly conserved domain present in CDP-alcohol phosphatidyltranferases [148]. Previously, Ass activity was identified in the membrane fraction of *Methanothermobacter thermautotrophicus* [47]. Secondary structure analysis of the archaeal members of the CDP-alcohol phosphatidyltransferase family indicates the presence of eight conserved transmembrane domains (TMDs) comparable to the corresponding bacterial protein family (**Figure 3A**). Interestingly, the PssA sequence of *B. subtilis* is substantially smaller than the other members of this family (only 177 amino acids instead of 451 amino acids of the *E. coli* Pss-I), but it shows the core of five TMDs (**Figure 3B**) but lacking two C-terminal TMDs. A membrane fraction of *B. subtilis* incubated with the different archaea-like substrates showed AS formation whereas such activity could not be demonstrated with *E. coli* membranes [47] suggesting a more narrow substrate specificity of the *E. coli* Pss-I than Pss-II enzymes. Therefore, the *pssA* gene of *B. subtilis* was cloned and overexpressed in *E. coli* BL21 strain under the control of T7 promoter. The protein was solubilized from the membrane with DDM and purified by Ni-NTA affinity chromatography (**Figure 3C**). Coupled *in vitro* reactions were performed using DGGGP as initial substrate along with the *A. fulgidus* CarS and CTP. Products were extracted with *n*-butanol and analyzed by LC-MS (**Figure 3F**). In the presence of the *B. subtilis* PssA, AS (m/z= 802.53 [M-H]-) formation was observed in the presence of L-serine and Mg^{2+} (**Figure 3D, lane 3**). PssA showed no activity towards DGGGP indicating the strict requirement of a

CDP-activated intermediate for the reaction (**Figure 3D, lane 4**). Moreover, no AS was formed in presence of EDTA (**Figure 3D, lane 5**) which chelates divalent cations, consistent with a requirement of Mg^{2+} for enzymatic activity [149]. As a control, formation of phosphatidylserine by *B. subtilis* PssA was demonstrated in the presence of PA ($C_{18:1}$), CTP and the *E. coli* CdsA (**Figure 3D, lane 6**). To examine the activity of the *E. coli* PssA towards CDP-archaeol, the enzyme was also overexpressed and purified (**Figure 3C**). In the presence of CDP-archaeol as substrate, no AS formation was observed demonstrating that the *E. coli* PssA (the Pss-I enzyme) indeed does not recognize the archaeal precursor (**Figure 3E, lane 1**). The enzyme, however, converted CDP-DAG into PS (m/z= 786.53 [M-H]-) in a coupled enzyme assay using PA as substrate (**Figure 3E, lane 4**).

In archaea and bacteria, AE and PE are produced by a decarboxylation reaction of L-serine. In *E. coli*, the Psd is encoded by the *psd* gene that specifies a membrane-associated pro-enzyme which undergoes an autocatalytic internal cleavage [150] leading to two subunits, the α subunit containing a pyruvoyl prosthetic group and a β subunit [151,152]. Previous bioinformatics analysis [7] identified an archaeal hypothetical protein as a potential Asd showing sequence similarity to the bacterial Psd (**Supplementary Figure S3**). Moreover, a similar operon conservation of the two genes *pss/psd*, typical of several bacterial species, was found in some archaea [7]. However, the archaeal Asd has not been biochemically characterized. Because of the general mechanism of the decarboxylation reaction [128], the possibility exists that the endogenous Psd of *E. coli* is able to recognize AS. Therefore, the *E. coli psd* gene was overexpressed in *E. coli* BL21 strain under the control of T7 promoter. Membranes bearing overexpressed levels of Psd were solubilized with DDM and the enzyme was purified by Ni-NTA affinity chromatography (**Figure 4A**). The α-subunit showed a slower migration on SDS-PAGE as expected on the basis of its predicted molecular mass. This is probably due to the presence of a covalently attached pyruvoyl prosthetic group that may affect the folding state in SDS-PAGE [151]. The enzymatic activity of the *E. coli* Psd was tested by *in vitro* coupled reactions as described above and product formation was detected by LC-MS (**Figure 4C**). Conversion of AS to AE (m/z= 758.55 $[M-H]^-$) by the *E. coli* Psd was observed in the presence of

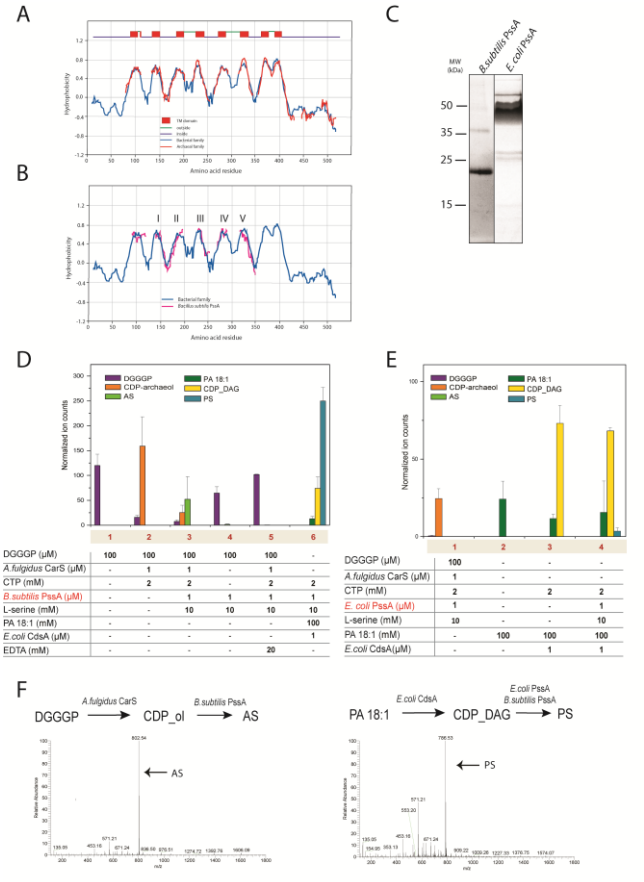


Figure 3| *In vitro* AS synthesis by the *B. subtilis* PssA. (**A**) Averaged hydropathy profile of the bacterial Pss proteins (blue line) and archaeal ones (red line) is aligned highlighting their conservations in the two kingdoms. The bacterial hydrophobicity profile is based on a multiple sequence alignment of 260 sequences sharing between 40 % and 80 % sequence identity. Likewise the archaeal hydrophobicity plot is based on a multiple sequence alignment of 38 archaeal sequences having a sequence identity between 30% and 70%. The membrane topology prediction is depicted above the plot. (**B**) Hydropathy profile alignment of *B. subtilis* PssA (purple line) and the averaged hydropathy profile of the bacterial Pss protein family (blue line). The conserved transmembrane domains (TMDs) are indicated by Roman numbers. (**C**) Coomassie Blue-stained SDS-PAGE gels showing the Ni-NTA purified proteins *B. subtilis* PssA (21 kDa) and *E. coli* PssA (53.6 kDa). Specificity of the bacterial *B. subtilis* PssA (**D**) and *E. coli* PssA (**E**) towards DGGGP and PA ($C_{18:1}$) as assessed by means of an *in vitro* assay using the purified enzymes. Total ion counts from LC-MS data were normalized using DDM as internal standard. The data are the average of three experiments ± SE. (**F**) Schematic representation of the performed *in vitro* reactions. The mass spectra from the LC-MS runs of the two products AS (m/z= 802.54 [M-H]-) and PS (m/z= 786.53 [M-H]-).

Mg^{2+}, L-serine, *B. subtilis* PssA, *A. fulgidus* CarS, CTP and DGGGP (**Figure 4B, lane 4**).

In vitro reconstitution of archaetidylethanolamine and archaetidylglycerol formation

Previously, we have described the *in vitro* reconstitution of CDP-archaeol formation starting from the precursors IPP and DMAPP/FPP [41] using two bacterial enzymes and three archaeal enzymes that were overexpressed and purified from *E. coli* upon codon optimization. Due to the wider substrate specificity described above, the extension of this pathway for the reconstitution of the formation of AE and AG requires four additional bacterial enzymes described in the previous section.

The ether lipid biosynthetic pathway reconstitution [41] starts from the two isoprenoid building blocks IPP and FPP which undergo several cycles of condensation leading to isoprenoid chains with the required C_{20} length [16]. GGPP (m/z= 449.19 $[M-H]^-$) formation was detected by LC-MS with

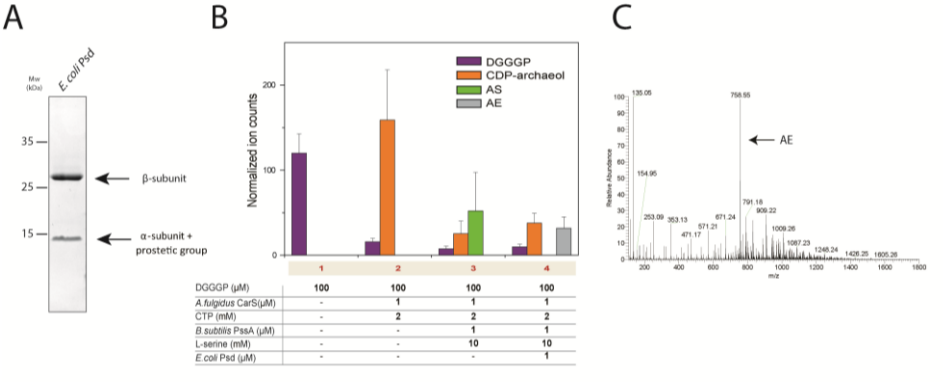
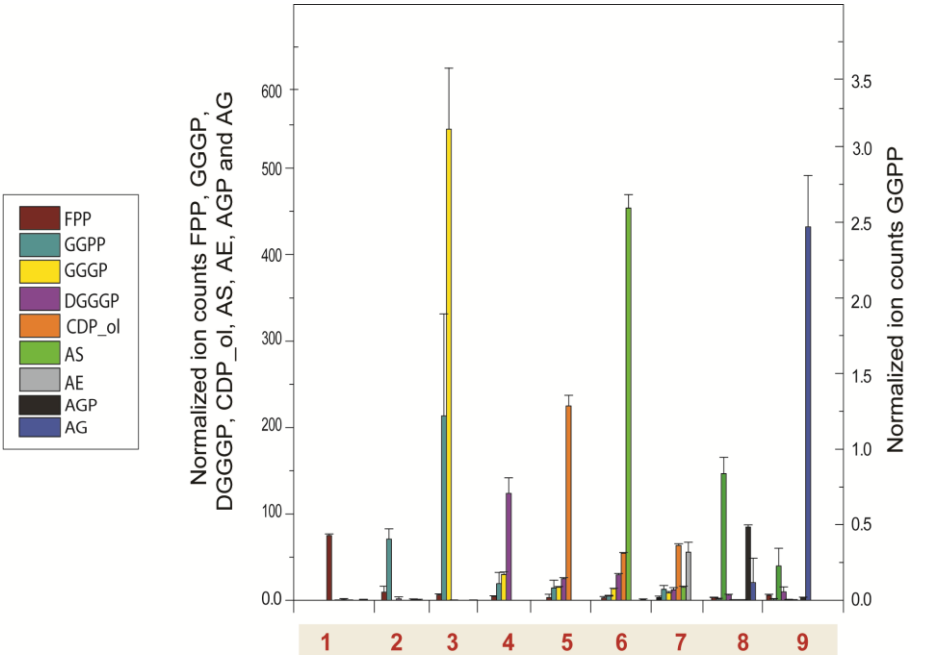


Figure 4| *In vitro* AE formation involving the *E. coli* Psd. (**A**) Coomassie Blue-stained SDS-PAGE gel showing the Ni-NTA purified protein *E. coli* Psd. The α-subunit with a molecular weight of 14.4 kDa and the β-subunit of 28.6 kDa. (**B**) *In vitro* reactions using the chemically synthesized substrate DGGGP, and the purified proteins as specified. Total ion counts from LC-MS data were normalized using DDM as internal standard. The data are the average of three experiments ± SE. (**C**) The mass spectra from the LC-MS runs of the products AE (m/z= 758,55 $[M-H]^-$).

an *E. coli* IspA [25] mutant as described previously (**Figure 5A, lane2**) [41]. Despite the unique feature of archaeal G1P dehydrogenases to synthesize the G1P, some bacterial enzymes are also able to perform this reaction [32]. The enzyme AraM (G1PDH) from *B. subtilis* [31] was used that produces the required glycerophosphate backbone which can be attached to the GGPP via the first ether bond leading to the synthesis of GGGP (m/z= 443.26 $[M-H]^-$). *M. maripaludis* GGGP synthase (MmarC7_1004) [36] was used to catalyze the latter reaction which in combination with IspA and G1PDH and the substrates IPP,FPP, DHAP and NADH, leads to the conversion of GGPP into GGGP (**Figure 5A, lane 3**). DGGGP production (m/z= 715.51 $[M-H]^-$) was observed in a subsequent reaction with the *A. fulgidus* DGGGP synthase (AF0404) [38] (**Figure 5A, lane 4**). Next, DGGGP was converted into CDP-archaeol (m/z= 1020.54 $[M-H]^-$) by the *A. fulgidus* CarS (AF1740) in the presence of CTP [41] (**Figure 5A, lane 5**). CDP-archaeol is the precursor for the formation of AS (m/z= 802.51 $[M-H]^-$) in the presence of the *B. subtilis* PssA and L-serine (**Figure 5A, lane 6**), which was further converted into AE (m/z= 758.54 $[M-H]^-$) (**Figure 5A, lane 7**) by the *E. coli* Psd. In another reaction, the CDP-archaeol was converted into AGP (m/z= 869.51 $[M-H]^-$) by the *E. coli* PgsA in the presence of G3P (**Figure 5A, lane 8**). AGP was then converted into AG (m/z= 789.54 $[M-H]^-$) by the addition of the *E. coli* PgpA to the *in vitro* reactions (**Figure 5A, lane 9**). Taken together, the reactions described here employing purified enzymes represent the *in vitro* reconstitution of the entire archaeal lipid pathway using a set of archaeal and bacterial enzymes.

Archaetidylglycerol and archaetidylethanolamine formation in *E. coli*

To reconstruct the entire archaeal ether lipid biosynthetic pathway into *E. coli,* a system of four compatible expression vectors was used to co-express seven ether lipid genes into *E. coli*. In this system, the vectors containing one or two genes each, as listed in the **Table 1**, allowed the simultaneous expression of three archaeal enzymes (*M. maripaludis* GGGPS, *A. fulgidus* DGGGPS and *A. fulgidus* CarS) and four bacterial



	1	2	3	4	5	6	7	8	9
E.coli IspA mutant (μM)	-	1.5	1.5	1.5	1.5	1.5	1.5	1.5	1.5
B.subtilis AraM (μM)	-	1	1	1	1	1	1	1	1
M.maripaludis GGGPS(μM)	-	-	0.26	0.26	0.26	0.26	0.26	0.26	0.26
A.fulgidus DGGGPS(μM)	-	-	-	1	1	1	1	1	1
A.fulgidus CarS(μM)	-	-	-	-	1	1	1	1	1
CTP (mM)	-	-	-	-	2	2	2	2	2
B.subtilis PssA (μM)	-	-	-	-	-	0.5	0.5	-	-
L-serine (mM)	-	-	-	-	-	10	10	-	-
E.coli Psd (μM)	-	-	-	-	-	-	0.5	-	-
E.coli PgsA (μM)	-	-	-	-	-	-	-	0.5	0.5
G3P (mM)	-	-	-	-	-	-	-	10	10
E.coli PgpA (μM)	-	-	-	-	-	-	-	-	0.5

Figure 5| *In vitro* reconstitution of the archaeal ether lipid pathway. *In vitro* reactions were performed using a combination of purified enzymes as specified and the substrates IPP, FPP, DHAP and NADH in presence of Mg^{2+} and 0.2% of DDM. Each product was detected by LC-MS: FPP (m/z= 381 $[M-H]^-$), GGPP (m/z= 449.19 $[M-H]^-$), GGGP (m/z= 443.26 $[M-H]^-$), DGGGP (m/z= 715.51 $[M-H]^-$), CDP-archaeol (m/z= 1020.54 $[M-H]^-$), AS (m/z= 802.51 $[M-H]^-$), AE (m/z= 758.54 $[M-H]^-$), AGP (m/z= 869.51 $[M-H]^-$) and AG (m/z= 789.54 $[M-H]^-$). The total ion counts were normalized using DDM as internal standard. The graph represents average of two experiments ± SE.

enzymes (*E. coli* Idi, *E. coli* mutant IspA, *B. subtilis* AraM and *B. subtilis* PssA). All genes were expressed under the control of the T7 promoter and induced with 0.25 mM of ITPG for 3.5 h. Total lipids were extracted from the *E. coli* membrane fraction using the Bligh and Dyer method [153] and analyzed by LC-MS. Four different engineered *E. coli* strains were compared, containing a different combination of archaeal lipid enzymes: (1) control harboring only the empty vectors; (2) **–** five ether lipid enzymes (*E. coli* Idi, *E. coli* mutant IspA, *B. subtilis* AraM, *M. maripaludis* GGGPS and *A. fulgidus* DGGGPS); (3)**–** six ether lipid enzymes (*E. coli* Idi, *E. coli* mutant IspA, *B. subtilis* AraM, *M. maripaludis* GGGPS, *A. fulgidus* DGGGPS and *A. fulgidus* CarS) and (4) **–** seven ether lipid enzymes (*E. coli* Idi, *E. coli* mutant IspA, *B. subtilis* AraM, *M. maripaludis* GGGPS, *A. fulgidus* DGGGPS, *A. fulgidus* CarS and *B. subtilis* PssA). When the total phospholipid composition of *E. coli* was compared for the different strains, no major differences were observed for the four differently engineered *E. coli* strains (**Figure 6A**). However, a detailed analysis showed the presence of a peak corresponding to AE (unsaturated) in *E. coli* strain 4 (**Figure 6B, lane 4**) which contains all the ether lipid genes required for synthesis compared to the other strains that lack the *B. subtilis* PssA (**Figure 6B, lane 3**) or both the *A. fulgidus* CarS and *B. subtilis* PssA (**Figure 6B, lane 2**) or the control (**Figure 6B, lane 1**). Another archaeal specific product digeranylgeranylglyceryl phopshoglycerol (DGGGP-Gro) (m/z= 789.5 $[M-H]^-$) was detected in the engineered *E. coli* strains, which is absent in the control strain (**Figure 6B, lane 1**). This compound was also detected in a previous study [40] and measured as the unsaturated form of AG. This lipid was observed in the *E. coli* strains that contain the basic set of five ether lipid genes up to the genes required for the formation of DGGGP (**Figure 6B, lane 2**). However, an enhanced production of unsaturated AG was observed when the strain also contains the *A. fulgidus* CarS (**Figure 6B, lane 3**). Upon the introduction of the AS-producing *B. subtilis* PssA enzyme strain 4, the amount of AG decreased and instead, AE was produced (**Figure 6B, lane 4**). Formation of unsaturated AG has been attributed to an endogenous reaction in *E. coli* which directly attaches the glycerol moiety to DGGGP [40], as it was assumed that the *E. coli* PgsA and PgpA enzymes are not able to accept the archaeal intermediate as a

substrate. As shown in the previous section, these two enzymes do recognize CDP-archaeol and convert it into AG. Therefore, the presence of unsaturated AG in the *E. coli* strain expressing the ether lipid genes is due to the low, but significant activity of the endogenous CDP-DAG synthase towards DGGGP. To confirm this hypothesis, an *in vitro* reaction was performed where DGGGP was incubated in presence of the purified CDP-DAG synthase (*E. coli* CdsA) in the presence of CTP. Under those conditions, a significant fraction of the DGGGP was converted into CDP-archaeol (**Figure 6C, lane 2**) demonstrating that CdsA is active with the archaeal substrate. When the *E. coli* PgsA and PgpA and the substrate G3P were added to the reaction, the CDP-archaeol was converted into AG (**Figure 6C, lane 3**). Taken together, these data demonstrate that the low activity of the endogenous *E. coli* CdsA (*in vitro* less than 1% compared to CarS in our previous work [41]) towards DGGGP results in substantial AG production *in vivo*. It is concluded that with a limited set of genes the archaeal lipids AG and AE can be produced in *E. coli*.

Discussion

For several decades, *E. coli* has been used for metabolic engineering such as the improvement of isoprenoid or carotenoid production [95,154–156]. Several attempts have been made to reconstruct the archaeal ether lipid biosynthetic pathway into *E. coli*. Gunsalus *et al.* [21] demonstrated production of DGGGP upon the overexpression of the endogenous *E. coli* Idi and expression of four enzymes (G1PDH, GGPPS, GGGPS and DGGGPS) from the hyperthermophilic archaeon *Archaeoglobus fulgidus*. Likewise, Yokoi *et al.* [40] employed ether lipid genes from the mesophilic archaeon *Methanosarcina acetivorans* to produce DGGGP and another compound that was identified as the PG-type derivative of DGGGP, named DGGGP-Gro. Isobe *et al*. [53] in addition expressed the *M. acetivorans* geranylgeranyl reductase for double bond reduction in conjunction with the four aforementioned archaeal enzymes [40]. They observed the *in vivo* production of archaeal lipids with a fully saturated isoprenoid chain only when the cofactor *M. acetivorans* ferredoxin was co-expressed.

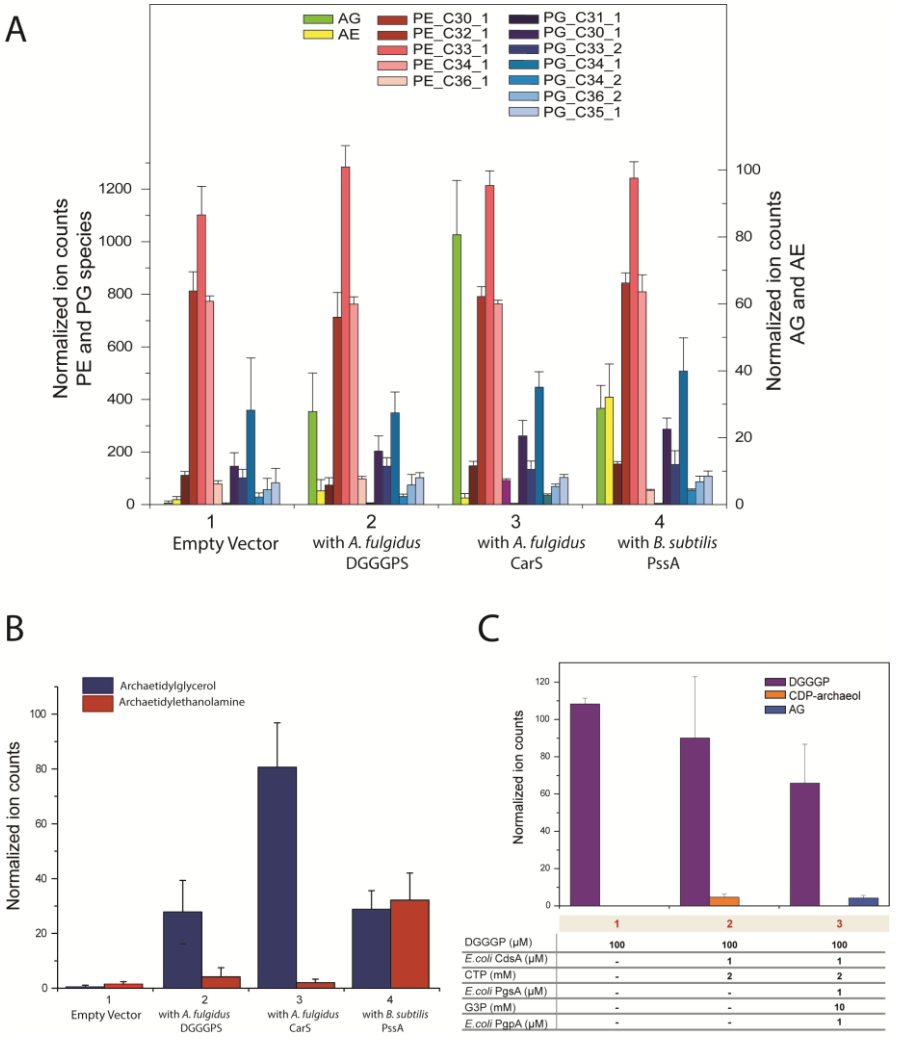


Figure 6| Archaeal lipid production in *E. coli*. (**A**) Total lipid analysis of four different engineered *E. coli* strains containing a different combination of ether lipid enzymes. The lipids were extracted from the membrane fraction and analyzed by LC-MS. The total ion counts of the several PE and PG species and of the archaeal lipids were normalized using Eicosane (m/z= 281.55 $[M-H]^-$) as internal standard. The species as classified according to the number of carbon atoms and number of unsaturated bonds in the acyl chains at the sn-1 and sn-2 positions. (**B**) Archaeal lipids production by different strains of *E. coli* upon the expression of the ether lipid biosynthetic genes. Results are the average of three biological replicates ± SE. (**C**) *In vitro* reactions using DGGGP and the purified *E. coli* CdsA, PgsA and PgpA proteins to determine the ability of CdsA to convert DGGGP into CDP-archaeol. Total ion counts from LC-MS data were normalized using DDM as internal standard. The data are the average of three experiments ± SE.

In this study, we aimed to produce the two major archaeal-like lipids AE and AG in *E. coli* as these lipids have the same polar head group as the major phospholipids PE and PG present in the *E. coli* membrane. However, this required a further understanding of the enzymatic steps required for polar head group attachment. Importantly, we have previously shown the production of CDP-archaeol *in vitro* by also including the CarS into the pathway that yielded DGGGP. Expression of CarS provides a means to substantially increase the production of the endogenously produced DGGGP-Gro that corresponds to AG (**Figure 6B**). During polar head group attachment, the CDP group from the CDP-activated precursor is replaced by a different polar head group. In bacteria and in archaea, the reactions involved in this process are very similar and mediated by the enzymes belonging to the CDP-alcohol phosphatidyltransferase superfamily. Given the high sequence homology among archaeal and bacterial phosphatidyltransferase [7,8], we investigated the substrate promiscuity of the enzymes PgsA and PgpA that are involved in PG formation, and their ability to recognize CDP-archaeol. Using *in vitro* reactions, purified *E. coli* PgsA was able to produce AGP from CDP-archaeol which in turn was produced by the *A. fulgidus* CarS from DGGGP, CTP and G3P. By means of the *E. coli* phosphatase PgpA, the AGP was readily converted into AG. This demonstrates a high substrate promiscuity of these bacterial enzymes and alleviates the need to introduce archaeal enzymes into *E. coli* to perform these reactions. Indeed, when the ether lipid biosynthesis pathway up to the formation of CDP-archaeol is introduced into *E. coli*, AG formation is observed. However, CarS is not essential, as, even in its absence, some AG can be formed. The origin of this AG was previously unknown, but we now show that this is due to a low activity of the endogenous *E. coli* CdsA for DGGGP resulting in the formation of CDP-archaeol that is further converted by *E. coli* PgsA and PgpA into AG.

In contrast to AG formation, AE formation has not been observed previously in *E. coli* which must imply that the *E. coli* Psses are unable to recognize CDP-archaeol in order to produce the intermediate AS that by decarboxylation should be further converted into AE. Indeed, the purified *E. coli* PssA was inactive with this substrate. Structural analysis of the bacterial and archaeal members of CDP-alcohol phosphatidyltransferase

family indicated a high level of secondary structure conservation of these enzymes with eight transmembrane segments. The *B. subtilis* PssA is a truncated version of these enzymes, being significantly shorter but still containing the highly conserved protein core. Importantly, this enzyme was previously shown to be active with the archaeal-like substrates [47]. Indeed, the purified *B. subtilis* PssA catalyzed the production of AS from CDP-archaeol and L-serine. The next step, the decarboxylation of the serine moiety, seems less specific, as purified *E. coli* Psd mediated the decarboxylation of AS with the concomitant formation of AE. Thus, for the production of AE in *E. coli* cells, only the *B. subtilis* PssA needs to be introduced (**Figure 6C**).

Having established the exact sequence of reactions needed for the archaeal ether lipid biosynthetic pathway [41] including the polar head group attachment, the reconstitution of AE and AG formation *in vitro* could be realized. Using a combination of archaeal and bacterial enzymes, and a breakdown in individual reactions, the synthesis of the unsaturated archaeal lipids AE and AG from the initial building blocks IPP, FPP and DHAP was achieved with nine purified enzymes. This defined the conditions needed for the reconstitution of AE and AG formation in *in vivo*, using *E. coli* as a host. Since three endogenous *E. coli* enzymes (Psd, PgsA and PgpA) recognize the archaeal precursors, the *in vivo* reconstitution depends on the expression of seven ether lipid genes. This includes overexpression of the *E. coli* Idi to boost IPP formation, and expression of a mutant IspA to generate GGPP; the *B. subtilis* AraM for G1P formation, and *B. subtilis* PssA for AS formation, the key ether lipid biosynthetic proteins, the *M. maripaludis* GGGPS, the *A. fulgidus* DGGGPS and the *A. fulgidus* CarS. Although the conversion reactions appear efficient *in vitro*, the amounts of AE and AG produced *in vivo* are still low and compared to the total *E. coli* lipid content they are likely less than 1%. However, the expression of CarS elevated the levels three-fold as compared to a previous report on AG formation [40].

The work described here represents a unique strategy to synthesize archaeal ether lipids in bacteria. Although the levels are still low, it is important to realize that the production of AE and AG as reported in this study, is performed in the presence of a fully functional phospholipid

biosynthetic pathway. High-level AE and AG production likely also requires the up-regulation of the entire pathway leading to IPP production. In addition, the pathway resulting in the reduction of the isoprenoid chains needs to be introduced to produce the saturated archaeatidyl compounds. Future studies should be directed towards a gradual down-regulation of the endogenous ester-bonded phospholipid biosynthetic pathway so that the endogenous lipids can be replaced by ether lipids. Such a bacterial strain could be used as an experimental model to examine the impact of the 'lipid divide' on the physiology and robustness of bacteria.

Experimental procedures

Bacterial strain and cloning procedures

Escherichia coli and *Bacillus subtilis* genomic DNA was used as template for the amplification of genes encoding the bacterial enzymes. *E. coli* DH5α (Invitrogen) was used for cloning. The primers and the plasmids used in this study are listed in **Table 1** and **2**. *E. coli* BL21 (DE3) or Lemo21 (DE3) [122] was used as protein overexpression host strain and grown in aerobic condition at 37°C in LB medium supplemented with the required antibiotics, Kanamycin (50 µg/ml), chloramphenicol (34 µg/ml), streptomycin (50 µg/ml) and ampicillin (50 µg/ml) in conjunction with 0.2% glucose added when necessary.

Expressions and purification of ether lipid enzymes

The bacterial proteins *B. subtilis* PssA and *E. coli* Psd were expressed in *E. coli* BL21 strain and induced with 1 mM of IPTG. *E. coli* PssA was induced with 0.5 mM of IPTG in the same over-expression strain *E. coli* BL21, whereas *E. coli* PgsA and PgpA were expressed in *E. coli* Lemo strain and induced with 0.4 mM of IPTG and 0.5 mM of L-rhamnose. After 2,5 hours of induction the cytoplasmic and membrane fractions were separated as described in a previous study [41]. The inner membrane vesicles (IMVs) of *E. coli* expressing the membrane proteins (*B. subtilis* PssA and *E. coli* Psd, PgsA and PgpA) were isolated as previously described [117]. The *E. coli* IMVs harboring the *B. subtilis* PssA and *E. coli* Psd were resuspended in buffer A (50 mM Tris-HCl pH 7.5, 300 mM NaCl and 10% glycerol) and 0.5

mg/ml of IMVs were solubilized in 2% of DDM (n-dodecyl-β-D-maltopyranoside) detergent for 1 hour at 4 °C. A centrifugation (173,400 x*g*) step of 30 min at 4 °C removed the insolubilized materials and the supernatant was incubated with Ni-NTA (Ni^{2+}-nitrilotriacetic acid) beads (Sigma) for 1 hour at 4 °C. The Ni-NTA beads were washed ten times with 40 column volumes (CV) of buffer B (50 mM Tris-HCl pH 7.5, 300 mM NaCl, 10% glycerol and 0.2% DDM) supplemented with 20 mM imidazole, and the proteins were eluted three times with 0.5 CV of buffer B supplemented with 250 mM imidazole. The *E. coli* IMVs (1 mg/ml) containing the *E. coli* PgsA and PgpA were resuspended in buffer C (50 mM Tris-HCl pH 7.5, 150 mM NaCl and 10% glycerol). The solubilization steps were performed as above. The *E. coli* PgsA-bound beads were washed five times with 40 CV of buffer D (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 10% glycerol and 0.2% DDM) supplemented with 10 mM imidazole and eluted three times with 0.5 CV of buffer D supplemented with 250 mM imidazole. The Ni-NTA beads containing *E. coli* PgpA were washed ten times with40 CV of buffer D supplemented with 10 mM imidazole and the protein was eluted with 0.5 CV of buffer D supplemented with 300 mM imidazole.

For the purification of the soluble protein *E. coli* PssA, the cytoplasmic fraction was incubated with Ni-NTA beads in buffer C overnight at 4 °C. The beads were washed three times with 40 CV of buffer C supplemented with 10 mM imidazole, once with 40 CV of buffer C supplemented with 60 mM Imidazole end eluted with 2 CV of buffer C supplemented with 300 mM imidazole. The purity of the eluted proteins were assessed on 12% SDS-PAGE gel stained with Coomassie Brilliant Blue and the protein concentration was determined by measuring the Absorbance at 280 nm. The other cytosolic proteins (*E. coli* isopentenyl diphosphate isomerase (Idi), *E. coli* mutant farnesyl diphosphate synthase (IspA), *B. subtilis* glycerol-1-phosphate dehydrogenase (AraM) and *Methanococcus maripaludis* GGGPS) and the membrane proteins (*Archaeoglobus fulgidus* DGGGPS, *A. fulgidus* CarS and *E. coli* CDP-diacylglycerol synthase (CdsA)) used in the present study were expressed and purified as described previously [41].

In vitro assays for archaeal lipids production

In vitro reactions were performed in 100 µl of assay buffer containing a final concentration of 50 mM Tris-HCl pH7.5, 10 mM $MgCl_2$, 52.5 mM NaCl, 87.5 mM Imidazole, 0.07% DDM and 3.5% glycerol. Where specified, 100 µM synthetic DGGGP, 100 µM PA (phosphatidic acid) ($C_{18:1}$), 2 mM CTP, 10 mM G3P, 10 mM L-serine, 20 mM EDTA and the indicated amount of purified enzymes were added to the reaction mixture. For the *in vitro* reconstitution of AE and AG, 100 µl of reaction volume was used containing the following assay buffer: 50 mM Tris-HCl pH 7.5, 10 mM $MgCl_2$, 15 mM NaCl, 25 mM Imidazole, 0.02% DDM, 1% glycerol, 0.5 mM DTT, 0.1 mM farnesyl pyrophosphate (FPP), 0.1 mM IPP, 2 mM nicotinamide adenine dinclueotide (NADH) and 2 mM dihydroxyacetone phosphate (DHAP). Reactions were incubated at 37 °C for 1 hour as described previously [41] and the products were extracted two times with 0.3 ml of *n*-butanol. Extracted lipids were evaporated under a stream of nitrogen gas and resuspended in 50 µl of methanol for the LC-MS analysis.

In vivo archaeal lipids synthesis

Engineered *E. coli* strains were aerobically grown at 37 °C in 200 ml of LB medium supplemented with required antibiotics [kanamycin (25 µg/ml), chloramphenicol (17 µg/ml), streptomycin (25 µg/ml) and ampicillin (25 µg/ml), 0.2% of glucose and 1 mM $NiCl_2$]. The cells were induced with 0.25 mM IPTG and after 3.5 hours, the total membrane fractions were isolated as previously described [41]. The internal standard Eicosane (20 µM) was added to the total membrane fractions (8 mg/ml) and lipids were extracted by means of the Bligh and Dyer method [153]. The chloroform extractable lipid fraction was washed with the aqueous phase of a blank Bligh and Dyer extraction solution and evaporated under a stream of nitrogen gas. The evaporated samples were resuspended with 0.3 ml 1:2 chloroform:MeOH, evaporated under a stream of nitrogen and finally resuspended in 100 µl methanol for LC-MS analysis.

Table 1. Expression vectors used in the present study

Plasmid	Description	Reference
pRSF-Duet-1	Cloning and expression vector (Kan^R), T7 promoter	Novagen
pET-Duet-1	Cloning and expression vector (Amp^R), T7 promoter	Novagen
pCDF-Duet-1	Cloning and expression vector (Str^R), T7 promoter	Novagen
pACYC-Duet-1	Cloning and expression vector (Cm^R), T7 promoter	Novagen
pSJ122	Synthetic gene encoding codon optimized DGGGP synthase from *A. fulgidus* with N-terminal His-tag and redesigned ribosome binding site AGGACGTTAACAT cloned into pRSF-Duet vector using the primers 41 and 42	Jain *et al.* [41]
pSJ135	PCR product of *ispA* gene with N-terminal His-tag from *E. coli K12* genomic DNA containing a double mutation Y79H and S140T. PCR product of *idi* gene with His-tag at the N-terminal from *E. coli K12* genomic DNA. Both genes were cloned into pCDF-Duet vector using the primers 62, 63, 24 and 57	This study
pSJ138	Synthetic gene encoding codon optimized GGGP synthase from *M. maripaludis* with N-terminal His-tag. PCR product of *araM* with C-terminus His-tag from *B. subtilis* genomic DNA. Both genes were cloned into pET-Duet vector using the primers 70, 71, 11 and 12	This study
pSJ140	Synthetic gene encoding codon optimized DGGGP synthase from *A. fulgidus* with N-terminal His-tag and redesigned ribosome binding site AGGACGTTAACAT. Synthetic gene encoding codon optimized CDP-archaeol synthase from *A. fulgidus* with C-	This study

	terminus His-tag. Both synthetic genes are cloned into pRSF-Duet vector using the primers 32, 20, 84 and 86	
pSJ148	PCR product of *cdsA* with N-terminus His-tag from *E. coli* genomic DNA cloned into pACYC-Duet vector using the primers 103 and 106	This study
pAC004	PCR product of *pss* with C-terminus His-tag from *B. subtilis* genomic DNA cloned into pACYC-Duet vector using the primers 89 and 90	This study
pAC008	PCR product of *psd* with C-terminus His-tag from *E. coli* genomic DNA cloned into pACYC-Duet vector using the primers 533 and 534	This study
pAC011	PCR product of *pss* with C-terminus His-tag from *E. coli* genomic DNA cloned into pET-Duet vector using the primers 542 and 543	This study
pAC015	PCR product of *pgsA* with C-terminus His-tag from *E. coli* genomic DNA cloned into pRSF-Duet vector using the primers 551 and 552	This study
pAC017	PCR product of *pgpA* with C-terminus His-tag from *E. coli* genomic DNA cloned into pET-Duet vector using the primers 562and 563	This study

Table 2. Oligonucleotide primers used in the study.

Primers name	Primer sequence 5' ⟶ 3'	Restriction site
11	GCGCGAATTCATGCATCACCACCACC	EcoRI
12	GCGCAAGCTTTCATTTTTTGGACAGC	HindIII
20	GCGCCTCGAGGACAGGTTTCCCGACTGGAAAG	XhoI
24	GATATACCATGGGCAGCCATCACCATC	NcoI
32	GGCGCCATATGCTGGATCTGATTCTGAA	NdeI
41	GACCAAGCTTGCGGCCGCATAATGC	HindIII
42	GATGCTCGAGTTAGAATGCACCGGCG	XhoI
57	GCGCGAATTCTTATTTATTACGCTGGATGATGTAG	EcoRI
62	CACTCATTAATTCATGATGATTTACCGGCAATGG	blunt

63	AGCGTGGATACACTCAACGGC	blunt
70	GCGCCATATGAATCGTATCGCAGCTGAC	NdeI
71	GCGCCTCGAGTTAGTGATGATGGTGGTGATGTTCATATAGACCAT GATCAGCG	XhoI
84	GCCGCCATGGGTAGTCATCATCACCACCATC	NcoI
86	GCGCGAATTCTTAGAATGCACCGGCGA	EcoRI
89	GCGCCATATGAATTACATCCCCTGTATGATTACG	NdeI
90	GCGCCTCGAGTTAGTGATGGTGATGGTGGTGATGATGATTCCATC GACTCCAG	XhoI
103	GCGCCTCGAGTTAGTGATGGTGATGGTGGTGATGATGAAGCGTCC TACCAGTAAC	XhoI
106	GCCGCCATGGGCAGCCATCACCATCATCACCACAGCCTGAAGTATC GATATCTGC	NcoI
533	GCGCCTCGAGAAAACAATGGCCTGGAGGCTACCTTGTTAAATTCA ACTTTCGCTAC	XhoI
534	GCGCTTAATTAATTAGTGATGGTGATGGTGGTGATGATGGACCTG TTTTTGTCGTCAAC	PacI
542	CGGCCATATGCTGTCAAAATTTAAGCGTAATAAAC	NdeI
543	GCGCCTCGAGTTAGTGATGATGGTGGTGATGCAGGATGCGGCTAA TC	XhoI
551	GCGCCATATGCAATTTAATATCCCTACGTTGC	NdeI
552	CGCGCTCGAGTCAGTGATGGTGATGGTGGTGATGATGCTGATCAA ATCTGCACGC	XhoI
562	CGCGGAATTCATGACCATTTTGCCACGCCATAAAG	EcoRI
563	CGGCGCGGCCGCCTAGTGATGGTGATGGTGGTGATGATGCGACAG CCCAGCG	NotI

LC-MS analysis

The samples from *in vitro* reactions were analyzed using an Accela1250 high-performance liquid chromatography system coupled with an ESI-MS Orbitrap Exactive (Thermo Fisher Scientific). A volume of 5 µl of each sample was used for the analysis. The LC-MS method parameters used in this study were the same as described previously [41]. The capillary and the tube lens voltage were set to -75 V and -190 V, respectively. For the samples from *in vivo* experiments, a sample volume of 5 µl was injected into a Shim-pack XR-ODS/C8/Phenyl column with dimension 3.0 mm I.D. x 75 mm (Shimadzu) operating at 55 °C with a flow rate of 400 µl/min. Mobile phase A [10 mM ammonium formate with 0.1% formic acid in water:acetonitrile (40:60 v/v)] and mobile phase B [(10 mM ammonium formate with 0.1% formic acid in acetonitrile:isopropanol (10:90 v/v)] were used as follows: initial condition started with 40% B, a linear gradient was started in 2 min from 40% to 43% B, gradient from 43% to

50% B in 0.1 min, 54% B isocratic for the next 9.9 min, linear gradient from 54% to 70% B in 0.1 min, 99% B isocratic for the following 5.9 min, gradient from 99% to 40% B in 0.1 min and 40% B isocratic for the last 1.9 min [157]. The MS settings used for this analysis were the same as described above. The Thermo XCalibur processing software was used for the data analysis and Genesis algorithm for automated peak detection and integration was applied to this analysis.

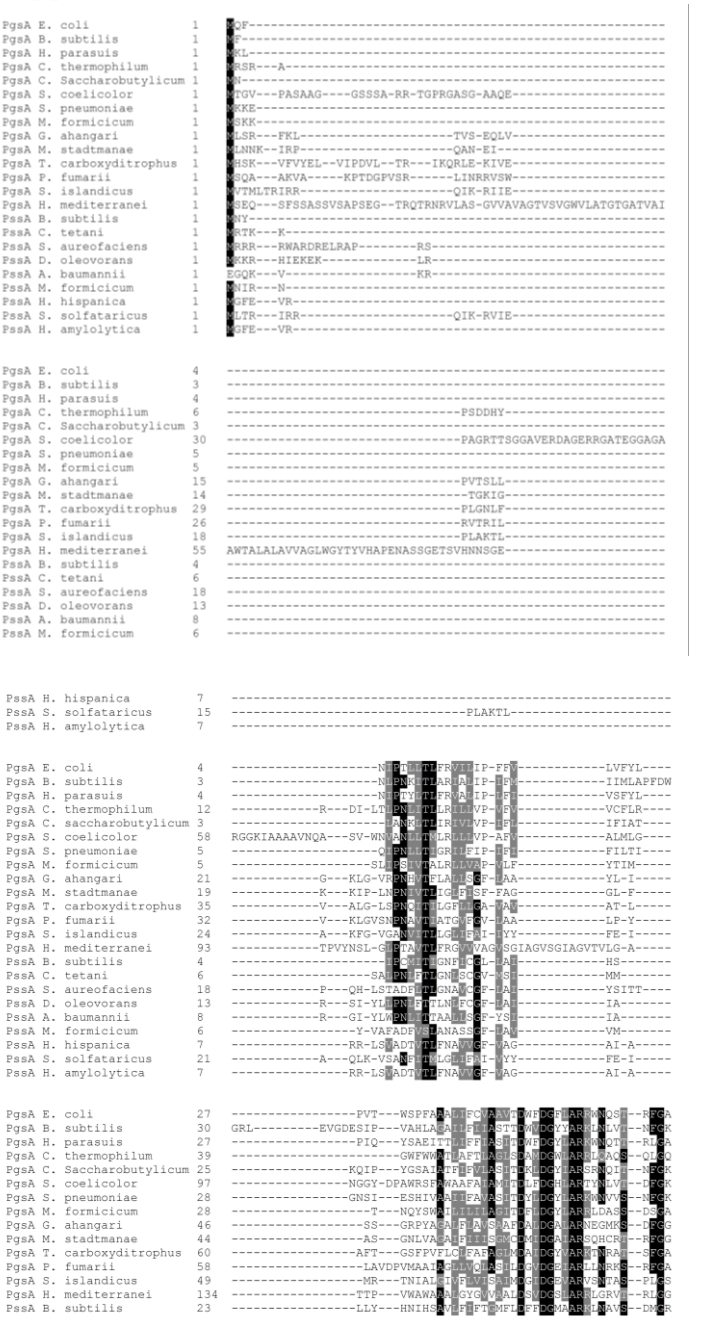
Author Contribution

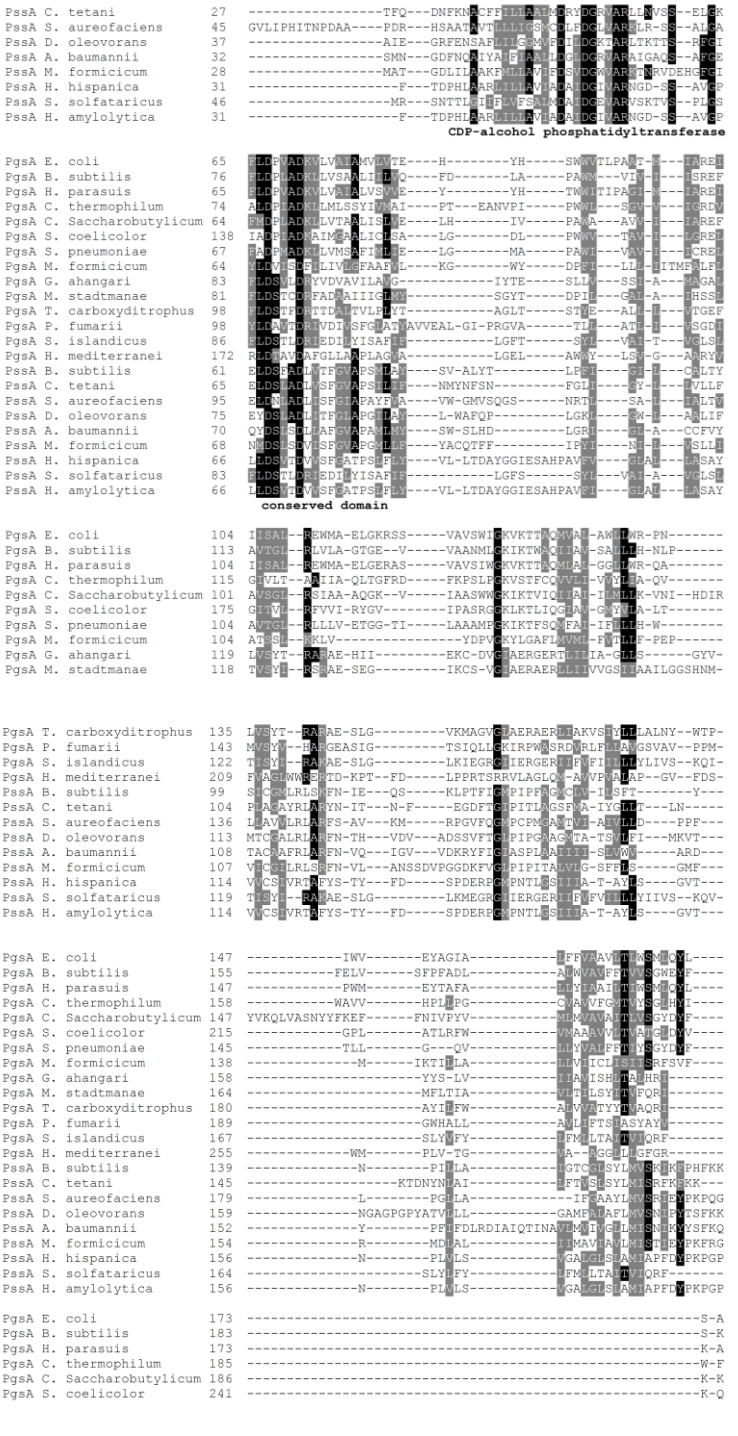
Antonella Caforio, Samta Jain and Arnold Driessen conceived and designed the research. Antonella Caforio cloned the genes, purified the enzymes and performed the experiments. Melvin Siliakus assisted in the cloning of genes and John van der Oost coordinated the pathway design. Adriaan Minnard design the DGGGP synthesis, which was performed by Peter Fodran. The manuscript was written by the contribution of all the authors.

Acknowledgments

This project was carried out within the research program of the biobased ecologically balanced sustainable industrial chemistry (BE-BASIC). We thanks Oleksander Salo for technical assistance and Juke Lolkema for the assistance with the hydrophobicity analyisis.

Supplementary Information

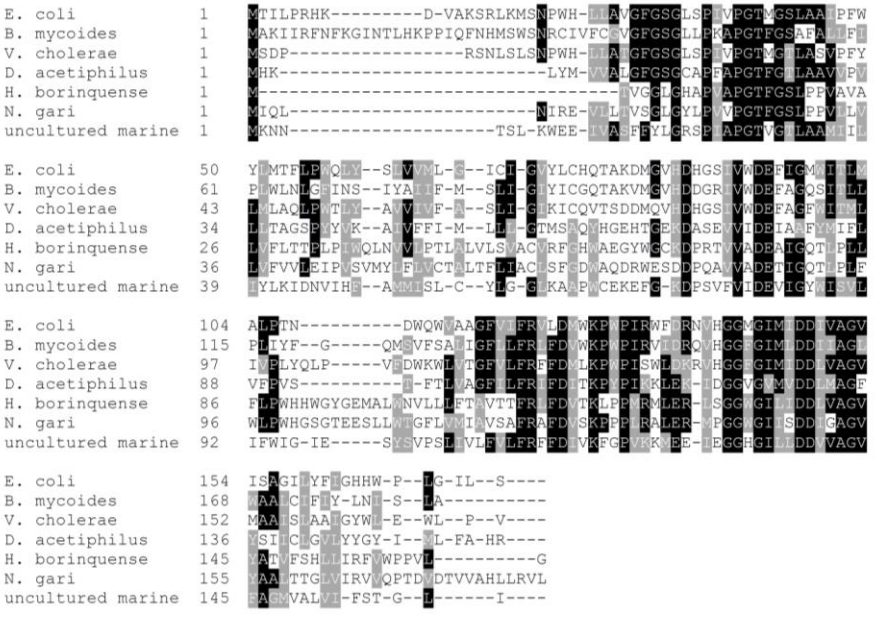



CDP-alcohol phosphatidyltransferase
conserved domain

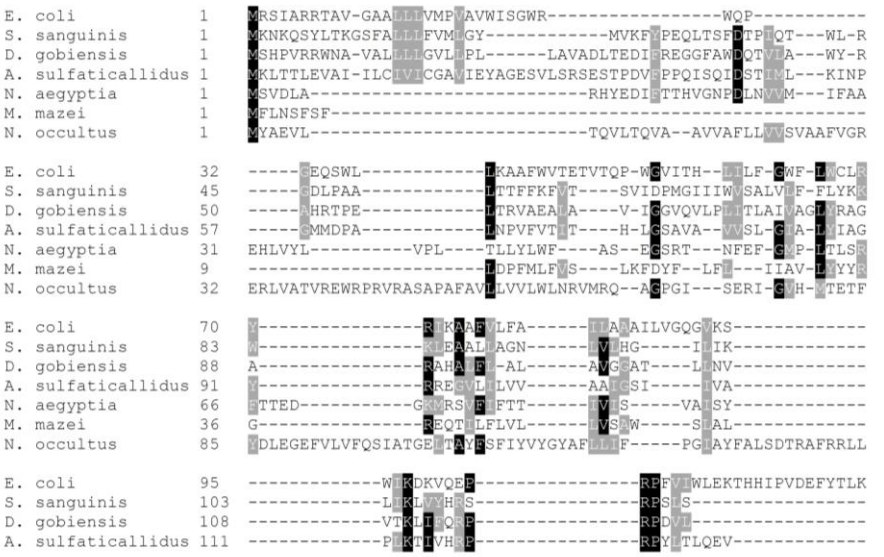
```
PgsA S. pneumoniae        168 ------------------------------------------------------------K-G
PgsA M. formicicum        163 ------------------------------------------------------------L-N
PgsA G. ahangari          177 ------------------------------------------------------------Y-H
PgsA M. stadtmanae        184 ------------------------------------------------------------H-H
PgsA T. carboxyditrophus  200 ------------------------------------------------------------T-H
PgsA P. fumarii           209 ---------------------------------------------------VAKTV-E
PgsA S. islandicus        187 ------------------------------------------------------------H-A
PgsA H. mediterranei      274 ------------------------------------------------------------D-W
PssA B. subtilis          166 H-----------------------------------------------------------
PssA C. tetani            172 ------------------------------------------------------------
PssA S. aureofaciens      204 LLA-T-ATLCWVVVAIGCLAAWAT---GLP--GGDTLM--HVGAFAQITLAA-MAPLL-V
PssA D. oleovorans        194 APPVLVRAN-----------------------NFNILVIAVLML----VFIVQEPVV-A
PssA A. baumannii         190 MDR-K-RVPF-VVMLPVVLIFAAI---TYN--IPMGIL--TV-SI----IYA-LSGFVTT
PssA M. formicicum        181 MML-M-IGG--SVLIIGTILPQNILSSIGY--LPAKLL--FIFAI----VYVIIVPLM-E
PssA H. hispanica         183 KHA-I-P--MGVVQAVAVVAPAAF---FRA--GPRAML--AIS----LLFFA-LGPWV-Y
PssA S. solfataricus      184 ------------------------------------------------------------Y-A
PssA H. amylolytica       183 KHA-I-P--MGVVQAVAVVAPTAL---FRA--GPRAML--AIA----LLFFA-LGPWV-Y


PgsA E. coli              175 AR---ADLLDQ---------------------
PgsA B. subtilis          185 NW---EALKTSN--------------------
PgsA H. parasuis          175 SK---GSLLKS---------------------
PgsA C. thermophilum      187 VTREVREFRRRTGD------SL---------P
PgsA C. saccharobutylicum 188 NK---DAVAIDK--------------------
PgsA S. coelicolor        243 AI---VLRRRGIAE---RRAALKE----TE-V
PgsA S. pneumoniae        170 SA---YVFKGTFGS-----------------K
PgsA M. formicicum        165 RN---N-----G---------------------
PgsA G. ahangari          179 TY---SQS-----------------------R
PgsA M. stadtmanae        186 VW---IRL-----------------------N
PgsA T. carboxyditrophus  202 AY---KTLT-R-A----PTS----------A
PgsA P. fumarii           215 LW---LLWR-R-G----EI-----------G
PgsA S. islandicus        189 VY---SLLR----------------------K
PgsA H. mediterranei      276 LY-VSGRLR-PVD----AGSQVTS----V-EN
PssA B. subtilis          167 -----------------AAENLES----GRWN
PssA C. tetani            172 -------------------------------F
PssA S. aureofaciens      253 I------RR-KVGQ---KVGDVRARRAESRLG
PssA D. oleovorans        225 L-----------------FT----------L
PssA A. baumannii         234 LF-----AR-K------SNETLK--------T
PssA M. formicicum        228 LY---GRLH-R------SGPHV---------R
PssA H. hispanica         226 ----------S-G-----------------E
PssA S. solfataricus      186 VY---SLLR----------------------K
PssA H. amylolytica       226 ----------S-G-----------------E
```

Figure S1| Multiple sequence alignment of archaeal and bacterial PgsA and PssA. The multiple sequence alignment indicates the presence of a conserved domain typical of the CDP-alcohol phosphatydiltransferase superfamily. Species list: Bacteria: *Escherichia coli, Bacillus subtilis, Haemophilus parasuis, Chloracidobacterium thermophilum, Clostridium saccharobutylicum, Streptomuyces coelicolor, Streptococcus pneumonia, Clostridium tetani, Streptomuyces aureofaciens, Desulfococcus oleovorans, and Acinetobacter bbaumannii;* Archaea: *Methanobacterium formicicum, Geoglobus ahangari, Methanosphaera stadtmanae, Thermofilum carboxyditrophus, Pyrolobus fumarii, Sulfolobus islandicus, Haloarcula mediterranei, Methanobacterium formicicum, Haloarcula hispanica, Sulfolobus solfataricus, and Haloarcula amylolytica*.

A


E. coli 1
B. mycoides 1
V. cholerae 1
D. acetiphilus 1
H. borinquense 1
N. gari 1
uncultured marine 1
E. coli 50
B. mycoides 61
V. cholerae 43
D. acetiphilus 34
H. borinquense 26
N. gari 36
uncultured marine 39
E. coli 104
B. mycoides 115
V. cholerae 97
D. acetiphilus 88
H. borinquense 86
N. gari 96
uncultured marine 92
E. coli 154
B. mycoides 168
V. cholerae 152
D. acetiphilus 136
H. borinquense 145
N. gari 155
uncultured marine 145

B


E. coli 1
S. sanguinis 1
D. gobiensis 1
A. sulfaticallidus 1
N. aegyptia 1
M. mazei 1
N. occultus 1
E. coli 32
S. sanguinis 45
D. gobiensis 50
A. sulfaticallidus 57
N. aegyptia 31
M. mazei 9
N. occultus 32
E. coli 70
S. sanguinis 83
D. gobiensis 88
A. sulfaticallidus 91
N. aegyptia 66
M. mazei 36
N. occultus 85
E. coli 95
S. sanguinis 103
D. gobiensis 108
A. sulfaticallidus 111

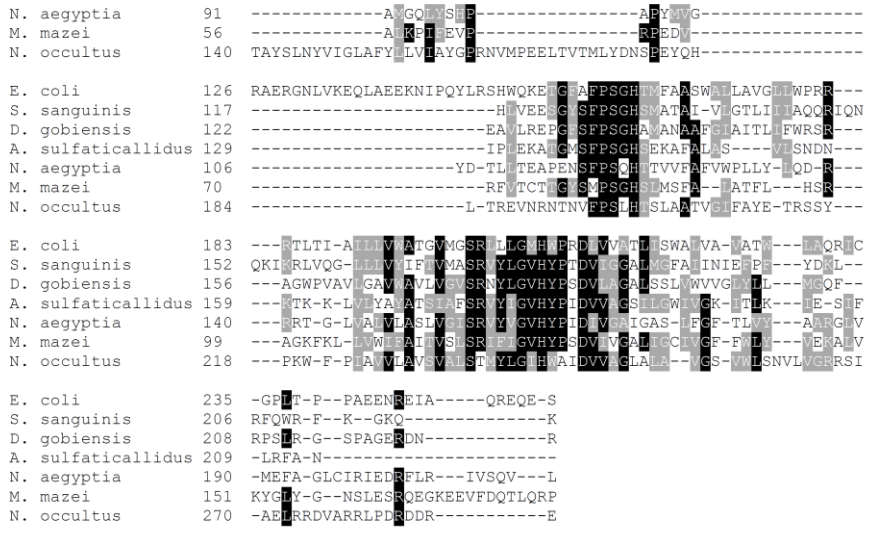
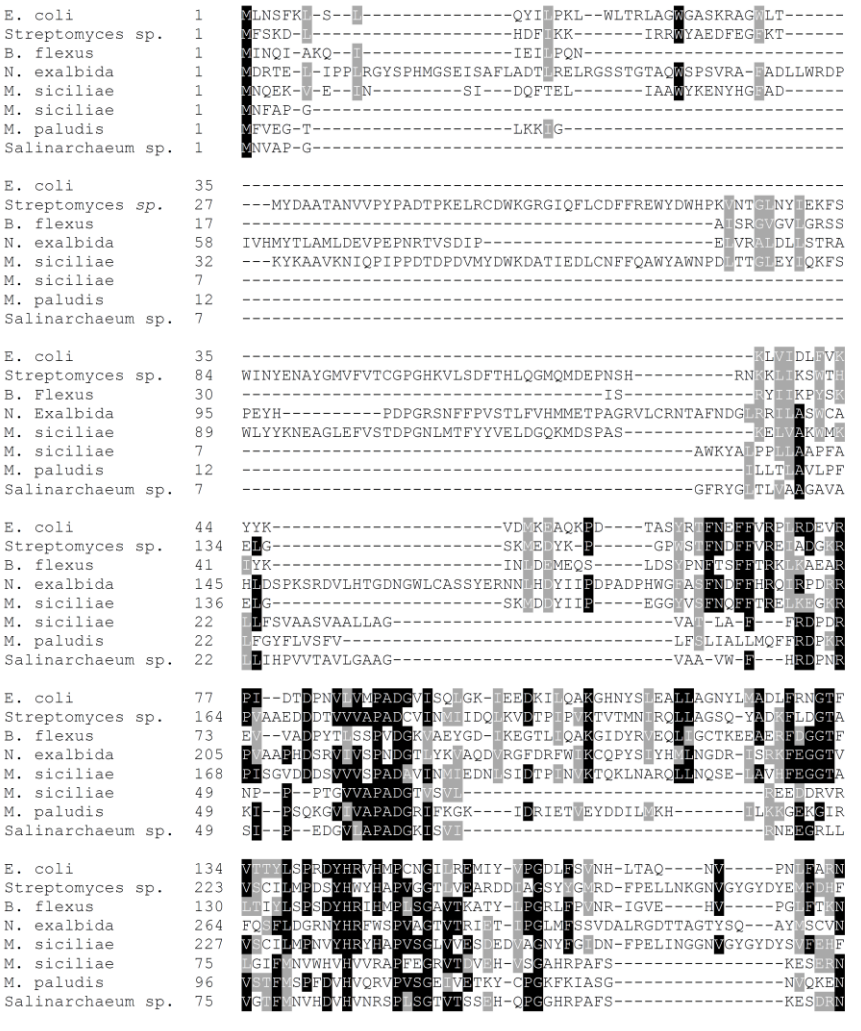


Figure S2| Multiple sequence alignment of a diverse group of bacterial and archaeal PgpA and PgpB proteins. (**A**) Alignment of bacterial and archaeal PgpA protein sequences showing a high degree con conservation. Species list: Bacteria: *Escherichia coli*, *Bacillus mycoides*, *Vibrio cholera*, and *Denitrovibrio acetiphilus*; Archaea: *Halogeometricum borinquense*, *Natrinema gari*, and an uncultured marine archaeon (Euryarchaeota). (**B**) Alignment of bacterial and archaeal PgpB protein sequences. Species list: Bacteria: *Escherichia coli*, *Streptococcus sanguinis,* and *Deinococcus gobiensis*; Archaea: *Archaeoglobus sulfaticallidus*, *Natrialba aegyptia, Methanosarcina mazei*, and *Natronococcus occultus.*



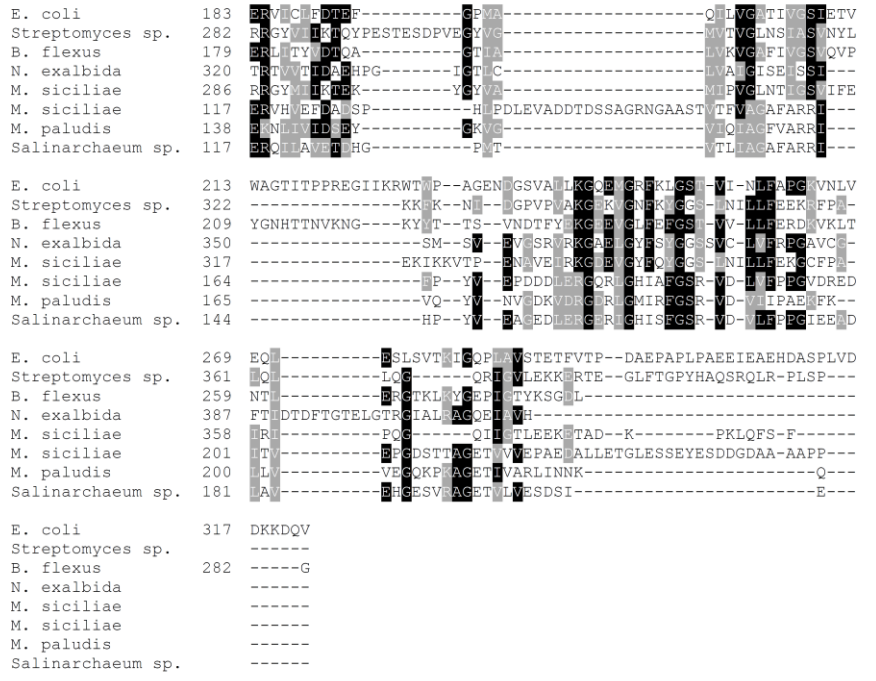
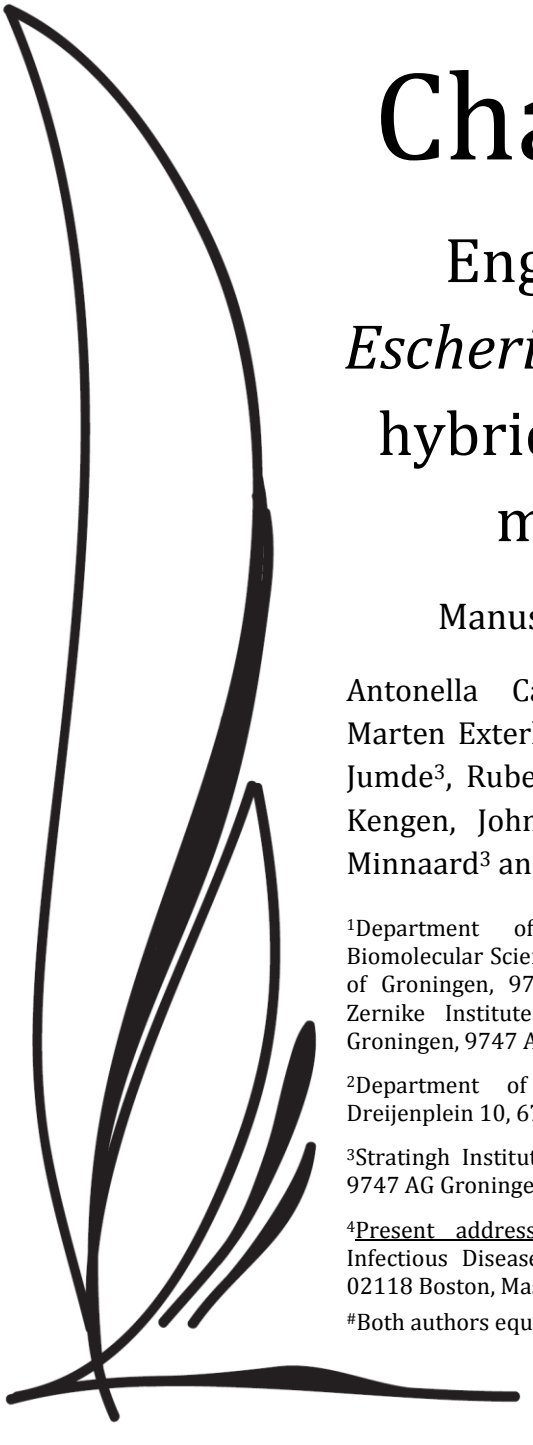


Figure S3. Multiple sequence alignment of a selection of diverse archaeal and bacterial Psd proteins. Species list: Bacteria: *Escherichia coli*, Streptomyces sp.*, Bacillus flexus,* and *Nocardia exalbida;* Archaea: two sequences from *Methanosarcina siciliae, Methanobacterium paludis* and a *Salinarchaeum sp.*



Chapter 4

Engineering of *Escherichia coli* with a hybrid heterochiral membrane

Manuscript in preparation

Antonella Caforio[1,#], Melvin Siliakus[2,#], Marten Exterkate[1], Samta Jain[1,4], Varsha R. Jumde[3], Ruben L.H. Adringa[3], Servé W. M. Kengen, John van der Oost[2], Adriaan J. Minnaard[3] and Arnold J.M. Driessen[1*]

[1]Department of Molecular Microbiology, Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, 9747 AG Groningen, The Netherlands; The Zernike Institute for Advanced Materials, University of Groningen, 9747 AG Groningen, The Netherlands

[2]Department of Microbiology, Wageningen University, Dreijenplein 10, 6703 HB Wageningen, The Netherlands

[3]Stratingh Institute for Chemistry, University of Groningen, 9747 AG Groningen, The Netherlands

[4]Present address: Department of Medicine, Section of Infectious Diseases, Boston University School of Medicine, 02118 Boston, Massachusetts, United States of America

[#]Both authors equally contributed to the work

Abstract

The last universal common ancestor (LUCA) is the most recent organism from which all organisms now living on Earth have a common descent. The membrane composition of LUCA represents an unresolved aspect in the differentiation of Bacteria (and Eukarya) and Archaea. The driving force behind this segregation has often been attributed to the chemical instability of a mixed membrane composed of a racemic mixture of glycerol-1P ether and glycerol-3P ester based lipids. However, such mixed membranes have never been reproduced in living cells. Here, we present for the first time a stable hybrid heterochiral membrane through lipid engineering of the bacterium *Escherichia coli*. By using a combination of metabolic engineering to boost isoprenoid biosynthesis and heterologous expression of the archaeal ether lipid biosynthetic pathway genes, an *E. coli* strain was obtained with up to 32% of archaeal lipids in the lipidome with the expected chirality. This resulted in viable cells but with altered cell growth, morphology and robustness towards environmental stress. The hybrid heterochiral membrane bacterial strain sheds new light on the lipid divide and opens novel possibilities for bio-industrial applications.

Introduction

A widely accepted hypothesis on the separation of the three domains of life Archaea, Bacteria and Eukarya is the existence of a common living ancestor, known as LUCA (last universal common ancestor) or cenancestor, from which the archaea and bacteria have diverged. Based on genomic analysis, predictions have been made about the organization of the transcriptional and translational machinery present in LUCA during the early stages of evolution [1]. The cell membrane of LUCA has attracted particular attention. Although the acellular theory states that the ancestor cell was characterized by the absence of a membrane [87] or surrounded by mineral membranes [88], most theories claim the presence of a cenancestor with a defined cellular membrane as a consequence of the need of compartmentation and self replication [158]. The existence of a phospholipid based membrane in the ancestor cell is further supported by phylogenetic studies that revealed a high conservation of the mevalonate pathway for the synthesis of the isoprenoid building blocks in archaea, eukarya and some bacterial species [5,9]. Also the presence of conserved membrane proteins such as the ATPase [159], redox proteins for respiration [160] and proteins involved in secretion like Sec and YidC [161] suggests that LUCA was embroidered by a phospholipid based cellular membrane. However, the chemical identity of the membrane lipids of LUCA remains an unresolved question considering the remarkable differences between archaeal and bacterial lipids and the hereinto "lipid divide" that differentiates these two domains of life. Archaeal lipids are composed of isoprenoid chains, connected via ether linkeages to a glycerol-1-phosphate (G1P) backbone while in bacteria and eukarya, phospholipids are based on straight-chain fatty acid esters linked to the enantiomeric glycerol-3-phosphate (G3P) backbone. Therefore, the stereochemical configuration of the lipids represents a crucial aspect to take into consideration during the divide between archaea and bacteria.

Assuming that the pre-cell was characterized by non-stereoselective enzymes [89] and abiotic catalysis, a racemic mixture of both G1P and G3P must have been present. Thus, the downstream enzymes of the lipid biosynthetic pathway should have had the capability of recognizing both

substrates leading to the formation of a heterochiral membrane. However, the existence of ancestral pre-cells with heterochiral membranes is assumed to be an unstable situation and expected to evolve towards a more stable homochiral membrane upon the occurrence of stereoselective enzymes and the differentiation between archaea and bacteria [87,89]. In contrast, *in vitro* experiments using liposomes composed of a mixture of archaeal and bacterial lipids showed a higher stability of mixed liposomes than those composed of only archaeal [90] or bacterial lipids in term of the temperature dependent permeability [91]. Although some bacteria are known to produce small quantities of ether lipids as well [162], so far no consistent evidence of the coexistence of substantial amounts of two phospholipids with opposite chirality has been observed in the membrane of any living cell. Few studies attempted to reproduce an *in vivo* heterochiral mixed membrane by introducing the partial [21,53] or almost entire [163] ether lipid biosynthetic pathway into the bacterium *Escherichia coli*, but the levels of ether lipids produced were minor and less than 1% compared to the endogenous *E. coli* lipid content.

Here we report the engineering of a hybrid heterochiral membrane in a viable bacterial cell. Via the upregulation of the synthesis of the isoprenoid building blocks and the co-expression of the archaeal lipids biosynthetic pathway genes, archaeal lipids with the G1P configuration were produced in *E. coli*, replacing nearly the complete endogenous pool of phosphatidylglycerol for archaetidylglycerol. In addition, we uncover substrate promiscuity of key enzymes of the archaeal lipid biosynthetic pathway, which supports the existence of a common ancestor with a heterochiral membrane from which archaea and bacteria diverged.

Results

Lipid biosynthesis engineering in *E. coli*

Almost all enzymes involved in the biosynthesis of diether lipids in Archaea have been identified and characterized [2,6,21,41,127,163] (**Figure S1**). The first metabolic step is the isoprenoid building blocks biosynthesis. The two isoprenoid units, isopentenyl-diphosphate (IPP) and

dimethylallyl-diphosphate (DMAPP), are widespread in nature and used in many other biosynthetic processes such as carotenoids, steroids or quinones. Their condensation by the enzyme geranylgeranyl diphosphate (GGPP) synthase leads to an isoprenoid chain of twenty carbon atoms [19,44]. The glycerophosphate backbone (G1P) is synthetized in Archaea by the glycerol-1-phosphate dehydrogenase (G1PDH) [28,29]. Even though bacteria have a different glycerophosphate configuration (G3P) conferred by the evolutionary unrelated enzyme glycerol-3-phosphate dehydrogenase (G3PDH) [2], G1PDH is also found in some bacteria [31]. GGPP and G1P are linked via two ether bonds catalyzed by two different archaeal enzymes: geranylgeranylglycerol phosphate (GGGP) synthase [35,46] and digeranylgeranylglycerol phosphate (DGGGP) synthase [38,41]. CDP-archaeol formation from DGGGP involves the recently discovered CarS enzyme [41]. Next, the replacement of the CDP moiety present in CDP-archaeol with a polar head group such as glycerol-3-phosphate or L-serine [163] leads to the formation of archaetidylglycerol (AG) and archaetidylethanolamine (AE), respectively. In Archaea, the isoprenoid chains are further saturated, but the exact mechanism of this process has not yet been fully resolved.

Thus, in order to reproduce a hybrid heterochiral membrane in the bacterial host *E. coli* strain JM109DE3, a composite pathway was developed that consists of both bacterial and archaeal enzymes (**Table S1**) in order to yield the unsaturated archaetidylglycerol (AG) and archaetidylethanolamine (AE) which are counterparts of the bacterial phosphatidylglycerol (PG) and phosphatidylethanolamine (PE), respectively. In previous work, this resulted in less than 1% of unsaturated AG and AE of the total lipidome [163]. To achieve much higher amounts of ether lipids, the endogenous MEP-DOXP pathway, responsible for IPP and DMAPP synthesis in *E. coli* was upregulated. Two synthetic operons composed of only the native *IDI* gene and *IDI*, *IspD*, *IspX* and *DSX* genes were integrated at the 'ori' macrodomain of the *E. coli* chromosome [164,165] (**Figure S2 A**). The two obtained strains (listed in **Table 1**) containing the single *IDI* gene (*E. coli* IDI$^+$) and the entire operon *IDI*, *IspD*, *IspX* and *DSX* (*E. coli* MEP/DOXP$^+$), respectively. These were tested for inducible MEP-DOXP upregulation using a reporter that converts IPP and

DMAPP to the red carotenoid lycopene [166]. The upregulation with 10 µM of Isopropil-β-D-1-tiogalattopiranoside (IPTG), resulted in a 2.5 and 5.3-fold increase in lycopene production by the engineered strains, respectively (**Figure S2 B**). The successful increase of isoprenoid building block production represented the starting point for archaeal ether lipids production. Using a system of two compatible vectors (**Figure S2 A**), up to six ether lipid genes were introduced into the *E. coli* IDI$^+$ and *E. coli* MEP/DOXP$^+$ strains leading to *E. coli* IDI$^+$EL$^+$ and *E. coli* MEP/DOXP$^+$EL$^+$ strains, respectively. This concerned the *crtE* gene from *Pantoea ananatis* [166,167] encoding a GGPP synthetase, *araM* from *Bacillus subtilis* [31] specifying the G1P dehydrogenase, *MmarC7_1004* [35,46] and *MmarC7_RS04845* [38,39] from *Methanococcus maripaludis* encoding the GGGP and DGGGP synthases, respectively, *AF1749* (CarS) from *Archaeoglobus fulgidus* [41] encoding the CDP-archaeol synthase and *pssA* from *Bacillus subtilis* encoding a phosphatidylserine synthase (PssA) [47,163] (**Table S1**). For the attachment of glycerol as polar head group and the conversion of L-serine into ethanolamine, the endogenous enzymes Psd, PgsA and PgpA of *E. coli* were exploited due their ability to recognize the archaeol derivatives [163] (**Figure S1**).

Increased IPP and DMAPP production dramatically stimulated the synthesis of unsaturated AG (**Figure S2 C**). A higher amount of this lipid was observed in the *E. coli* MEP/DOXP$^+$EL$^+$ strain, harboring the entire MEP-DOXP operon compared to the *E. coli* IDI$^+$EL$^+$ strain containing only the *IDI* gene and the control strain (**Figure S2 C**). A second operon, containing the *crtE*, *araM* and *MA3969* genes, was further integrated into the bacterial chromosome, but this did not improve ether lipids production compared to plasmid based expression (data not shown). Despite the presence of the *B. subtilis pssA* gene for the synthesis of unsaturated AE, only a very low amount of this lipid was detected. The use of alternative Ribosome Binding Sites (RBS) or the addition of L-serine to the growth medium [168] did not improve AE synthesis (data not shown). As previously shown [169], overexpression of the *B. subtilis* PssA impairs cell growth likely because of elevated levels of the non-bilayer lipid PE which is lethal to the cells [148]. Thus, the study further focuses on the increased AG levels.

Ether lipid production optimization

The obtained engineered *E. coli* strain (*E. coli* MEP/DOXP^{+}EL^{+}) was further optimized in terms of growth and induction to achieve the highest amounts of ether lipids possible. Herein, ether lipid synthesis was analyzed in a defined minimal medium (OPT1), optimized for increased isoprenoid production [170]. The total lipid analysis revealed a 2.1-fold increased AG production when the strain was grown in OPT1 medium compared to growth in rich LB medium (**Figure S3 A**). Further optimization was performed by inducing the engineered *E. coli* strain with 100 µM of IPTG at different growth phases. Induction at the beginning of growth (OD_{600} = 0.0) yielded the highest amounts of AG compared to induction at the early (OD_{600}= 0.3) and mid-exponential (OD_{600}= 0.6) growth phase (**Figure S3 B**). Next, different IPTG concentrations were used to induce AG production. The distribution of the main bacterial phospholipids PE, PG, and cardiolipin (CL) and the archaeal lipid AG was compared among the different strains using LC-MS [163] (**Figure S3 C**), TLC (**Figure 1A**) and lipid quantitation via phosphorous determination. Lower amounts of inducer resulted in a higher AG lipid production. From the quantitative TLC analysis, the PG content decreased from 52% in the wild type to less than 5% in the strain induced with 10 µM IPTG (**Figure 1B**), while the AG content increased up to 32%. Higher amounts of IPTG (50-100 µM) resulted in less AG production and higher amounts of PG. With the decrease in PG content, we noted triacylglyceride accumulated in the cells consistent with a their decreased demand because of the increased levels of AG production (**Figure 1A**). These data demonstrate that essentially the entire PG pool can be replaced by AG, resulting in a hybrid membrane.

Ether lipid chirality and archaeal enzyme substrate promiscuity

The G1P configuration is the most striking features that distinguishes archaeal and bacterial phospholipids. However, to ascertain that the correct archaeal lipid chirality was present, a strain was constructed that lacks the G1PDH due to an *araM* gene deletion (*E. coli* MEP/DOXP^{+}EL^{+} AraM^{-}). Surprisingly, in the absence of G1PDH, still high levels of AG were detected, although lower as compared to the strain bearing a functional *araM* (**Fig. S3 C**). This finding raised the question if *E. coli* can produce G1P

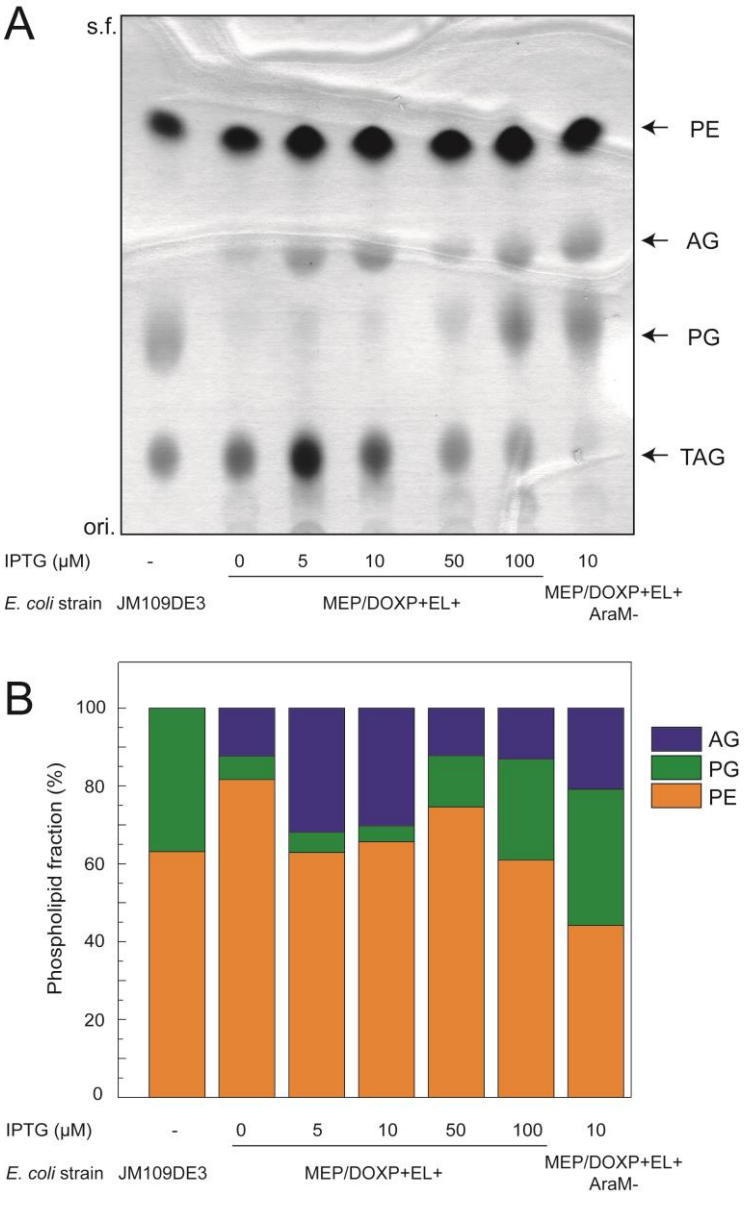


Figure 1| Thin Layer Chromatography (TLC) based quantitation of *in vivo* archaeal lipids synthesis. (**A**) TLC of lipids extracts from wild type *E. coli*, heterochiral mixed membrane *E. coli* induced with different IPTG concentrations and the *E. coli* strain harboring the entire ether lipid pathway but lacking the *araM* gene. (**B**) Relative quantitation of the spots detected in the TLC. Each lipid species was calculated as percentage of the total amount of lipid phosphorous detected in each lane.

or whether the ether lipid pathway contains non stereoselective enzymes. The first ether lipid enzyme involved in recognition of G1P is GGGPS, which links the long isoprenoid chain GGPP to the glycerophosphate backbone of G1P. This enzyme is considered to be highly stereoselective, as GGGPS of *Thermoplasma acidophilum* exhibited only a low enzymatic activity with G3P [46] and as confirmed by the resolved G1P and GGPP binding sites in the protein structures of GGGPS from *A. fulgidus* [36] and *Methanothermobacter thermoautotrophicus* [35]. However, analysis of the activity of GGGPS from *M. maripaludis in vitro* using GGPP as substrate revealed a remarkable non selectivity towards G1P and G3P (**Figure 2A, lane 2 and 3**). Also, purified G3P acyltransferase from *E. coli* (PlsB), involved in the attachment of the glycerophosphate backbone to the fatty acid chain [94,171], was tested for its specificity towards G3P and G1P. To this end an *in vitro* system was used to synthetize acyl-CoA by condensation of oleic acid and CoA by the *E. coli* FadD enzyme [172,173] (**Figure 2B, lane 1**) which is subsequently converted by PlsB into lyso-phosphatidic acid (LPA). LPA production was observed only in presence of G3P (**Fig. 2B, lane 2**) and no product was detected with G1P (**Figure 2B, lane 3**) demonstrating a very high stereoselectivity of PlsB. Kinetic analysis of the GGGPS enzyme with G1P and G3P showed 9 times higher preference of the enzyme towards G1P (K_m = 5.8 ± 1.6 μM) as compared to G3P (K_m = 46.7 ± 6 μM) (**Figure 2C and D**). The weaker chiral specificity of GGGPS could potentially account for AG formation in the absence of G1PDH. Therefore, to conclusively establish the configuration of the diether lipids in the engineered *E. coli* strains, both enantiomers of AG were prepared chemically and compared with the AG produced in *E. coli*. In short, saponification of the total lipid extract allowed the subsequent purification of the ether lipids by chromatography on silica. Chemical synthesis of AG with G1P (**Figure 2E, panel II**), and AG with G3P (**Figure 2E, panel I**) configuration was carried out according to our previous work [41]. All three samples were converted into their corresponding Mosher's ester and analyzed by ^{1}H- and ^{19}H-NMR [174]. Readily distinguished diastereotopic shifts in the ^{1}H-NMR (**Figure 2E**) showed unambiguously that the AG produced by the engineered strains both with (**Figure 2E, panel IV**) and without (**Figure 2E, panel V**) G1PDH have the archeal G1P

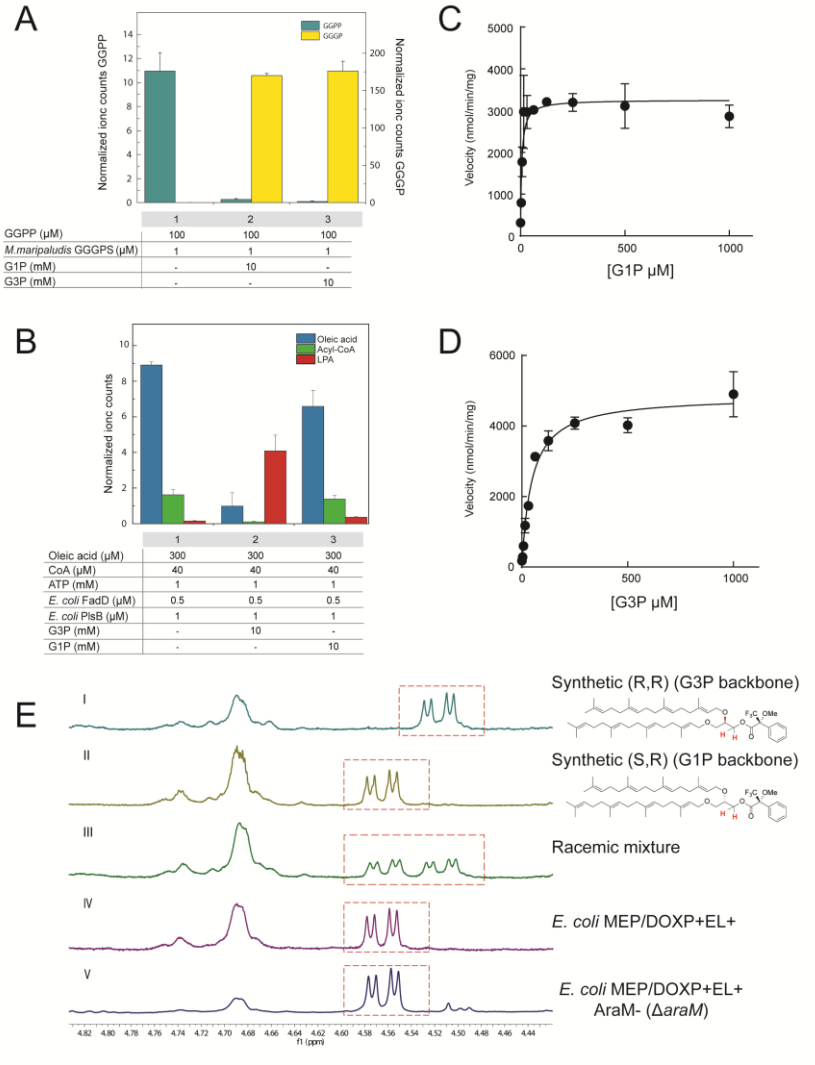


Figure 2| Chirality of the ether lipid biosynthesis in *E. coli* and stereoselectivity of the archaeal GGGPS. Specificity of archaeal *M. maripaludis* GGGPS (**A**) and the bacterial *E. coli* PlsB (**B**) enzymes towards G1P and G3P. Kinetic analysis of *M. maripaludis* GGGPS using different concentration of G1P (**C**) or G3P (**D**). Total ion counts are normalized using DDM as internal standard. Results are the averages of two experiments ± S.E.M. (**E**) NMR spectra of Mosher's ester derivatized AG. Synthetich AG with G3P configuration (**I**), synthetic AG with G1P configuration (**II**), a mixture of both (**III**), AG from the *E. coli* strain expressing the whole ether lipid biosynthetic pathway (**IV**) and from the *E. coli* strain harboring the *AraM* gene deletion (**V**). The red boxes highlight the diagnostic signals.

configuration. This confirms a high selectivity of the ether lipid enzymes for G1P *in vivo*, but also indicates that *E. coli* harbors an endogenous mechanism of G1P production.

Growth and cell morphology

The replacement of the endogenous PG pool for AG impacted bacterial growth (**Figure 3A**). The *E. coli* MEP/DOXP^{+}EL^{+} strain showed a long lag phase of ~16 hours before growth commenced with a growth rate similar to the parental strain. Both the non-induced and induced (10 μM) cells showed a similar growth behavior likely because of leakage of the promoter used to express the archaeal lipid enzymes as evidenced by the presence of AG in non-induced cells (**Figure S3 C**). With increased IPTG concentration (up to 100 μM), the lag phase shortened to ~8 hours but growth proceeded with a slower rate. Ether lipid biosynthesis caused an elongation of the cell length (**Figure 3B, panel II**) as evidenced by Scanning Electron Microscopy (SEM). The observed phenotype affects the majority of the cells with cell lengths ranging between 2 and 12 μm, compared to 500 nm of control *E. coli* cells. In particular, at higher IPTG levels, the engineered cells exhibited lobular appendages which extrude from the cell surface, ranging between 100 and 500 nm in diameter. These bulges occur at the cell poles or on the side. It appears that cell division takes place at these appendages sites, leaving at time a scar on the mother cell and the formation of small daughter cells (**Figure 3B, panel III**). At higher inducer levels, the phenomenon is more frequent and filamentous extrusions are formed connecting cells and cellular aggregates. Released extrusions could readily be isolated from the supernatant after high speed centrifugation of a cell culture. Total lipid analysis on the isolated appendages revealed a mixture of archaeal and bacterial lipids similar to the lipid content of the mother cells, (**Figure S3 D**) excluding the hypothesis that these structures are the result of lipid segregation. Also, the SDS-PAGE protein profile of cells and the isolated appendages was similar (data not shown). We hypothesize that these extrusions are formed as a result of high level lipid production.

To further examine the aberrant division mechanism, *E. coli* cells were stained with the dyes FM4-64 and DAPI that stain lipids and DNA,

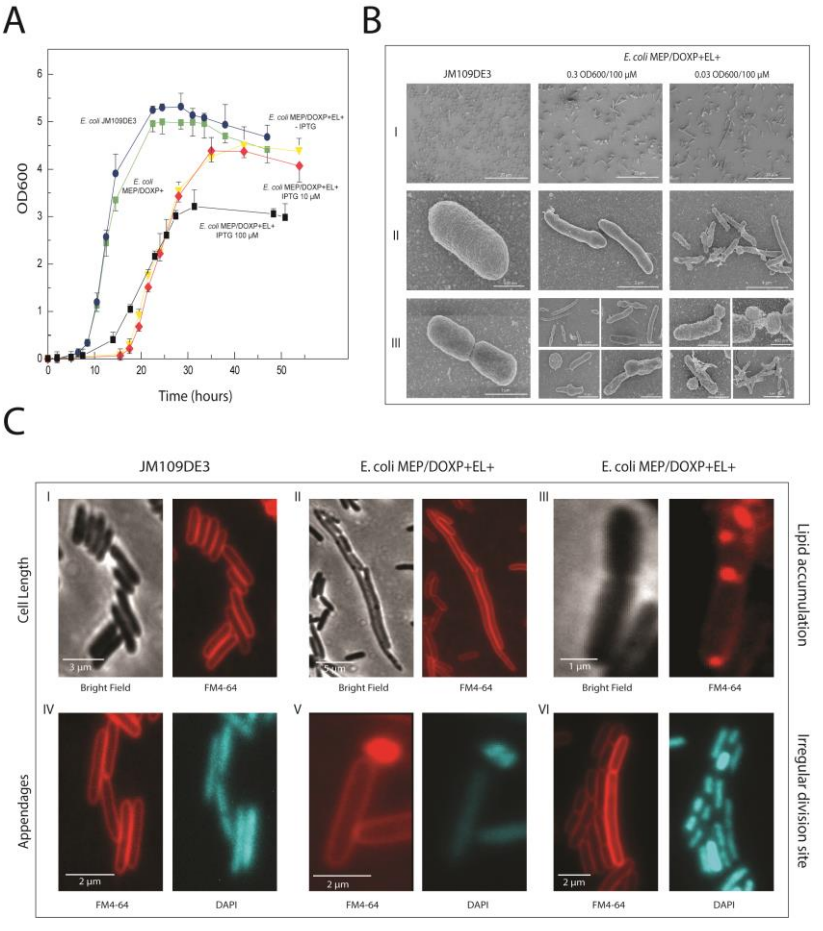


Figure 3| Growth and cell morphology analysis of the heterochiral mixed membrane strains. (**A**) Growth of the *E. coli* MEP/DOXP^{+}EL^{+} strain with all the ether lipids enzymes (not induced (—▼—), induced with 10 µM (—◆—) and induced with 100 µM (—■—) of IPTG) compared with two negative control strains (*E. coli* JM109DE3 wild type (—●—) and *E. coli* MEP/DOXP^{+} strain with the integrated MEP-DOXP operon (—■—)). The data are the averages of three biological replicates ± S.E.M. (**B**) Scanning Electron Microscopy (SEM) of wild type *E. coli*, the heterochiral mixed membrane strain induced at a later (0.3 OD_{600}) and earlier (0.03 OD_{600}) growth phase using 100 µM of IPTG. (**I**). Field of cells. (**II**) Altered cell shape and length. (**III**) Aberrant cell division and formation of bulges and shreds. (**C**) Effect of mixed heterochiral membranes on *E. coli* cells detected by double staining with FM4-64 and DAPI. (**I-II**) Lipid staining showing elongated and thinner cells in the engineered strain compared to the control. (**III**) Presence of membrane associated spots in the engineered strain. (**V**) Double staining with FM4-64 and DAPI showing the presence of appendages surrounded by a lipid layer and the presence of DNA. (**IV-VI**) Presence of irregular division sites in engineered cells compared to the symmetrical division septum present in the wild type cells.

respectively. The FM4-64 staining confirmed the presence of elongated and thinner cells in the engineered strain compared to wild type cells (**Figure 3C, I-II**). Furthermore, the lipid staining also signified the presence of intense membrane associated spots in the induced strain (**Figure 3C, III**) that possibly correspond to accumulation of anionic lipids in highly induced cells. Interestingly, the appendages contain genetic material as evidenced by the DAPI staining (**Figure 3C, V**). Finally, the double staining revealed the presence of irregular division sites in the elongated cells (**Figure 3C, VI**) compared to the typical mid-cell septum present in growing wild type cells (**Figure 3C, IV**). These data suggest that a high level of induction of archaeal lipid biosynthesis result in aberrant cell division.

Robustness of cells harboring a heterochiral mixed membrane

Archaeal ether lipids have been associated with extremophilicity and robustness, even though not all Archaea are extremophiles. Therefore, the survival of the strains with a heterochiral mixed membrane upon a heat and cold shock was tested. Three different engineered *E. coli* strains (JM109DE3, MEP/DOXP$^+$ and MEP/DOXP$^+$EL$^+$) were exposed to elevated temperatures for two minutes and recovered for one hour at 37 °C. The non-induced and induced (10 µM IPTG) *E. coli* MEP/DOXP$^+$EL$^+$ strain, showed an overall higher survival and ability to survive exposure to 55 °C and 58 °C compared to the two control strains JM109DE3 and *E. coli* MEP/DOXP$^+$ that do not survive when exposed to temperatures above 50 °C (**Figure 4A**). Cells were also exposed to freezing at -80 °C. The cells containing an induced heterochiral mixed membrane were remarkable more tolerant to this treatment than the control strains (**Figure 4B**) as evidenced by the higher CFU count but only when cells were induced with 10 µM IPTG. Finally, the tolerance to the organic solvent butanol was tested by exposing the strains for two minutes to different concentrations. A higher resistance of the non-induced and induced (10 µM IPTG) *E. coli* MEP/DOXP$^+$EL$^+$ strain harboring the entire ether lipid biosynthetic pathway was observed compared to the controls (**Figure 4C**). This was most notable when the cells were treated with 2% of butanol. Taken together these data demonstrate that the presence of archaeal lipids in the

bacterial membrane renders the engineered cells more resistant to different types of environmental stress.

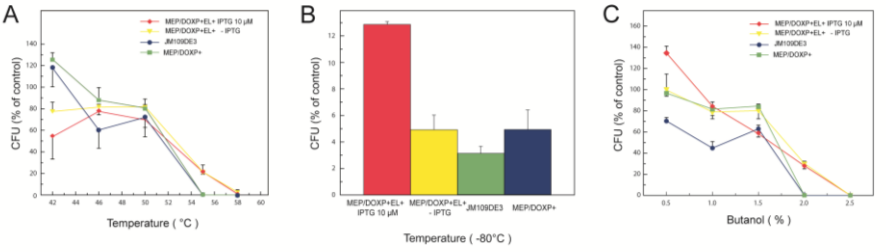


Figure 4| Robustness of *E. coli* with a heterochiral mixed membrane. The *E. coli* strain with all ether lipids enzymes MEP/DOXP^{+}EL^{+} (not induced – yellow; and induced with 10 µM IPTG - red) was compared with the wild type strain JM109DE3 (blue) and the strain harboring the integrated MEP-DOXP operon MEP/DOXP^{+} (green) for survival against exposure to different environmental stresses. (**A**) Heat shock, (**B**) Freezing at -80 °C, and (**C**) Butanol tolerance. The data were normalized against the CFU of untreated samples. The results are the averages of four biological replicates ± S.E.M.

Discussion

The "Lipid Divide" represents a critical event during the differentiation of the two domain of life Bacteria and Archaea, both originating from the last universal common ancestor (LUCA). According to the discordant hypothesis, the instability of a heterochiral mixed membrane in the common ancestor triggered the segregation of archaea and bacteria towards a more stable homochiral membrane [4,91]. While it is inherently difficult to test such a hypothesis *in vivo*, as the conditions of early evolution would need to be replicated, *in vitro* data using pure lipid liposomes failed to demonstrate the assumed instability. Also, so far no biological evidence has been reported for instable mixed heterochiral membranes in a living cell. Here we reproduced a viable bacterial cell with a heterochiral mixed membrane composed of bacterial and archaeal lipids through the introduction of the archaeal ether lipid biosynthetic pathway into *E. coli*. Such a heterochiral mixed membrane may be a biological

model for the coexistence of these two lipid species which might have characterized the membrane of the common cenancestor.

We have previously reported the introduction of a fully functional ether lipid pathway into the bacterium *E. coli* and the synthesis of the two archaeal lipids AG and AE [163]. However, the level of the ether lipids was very low compared to the bacterial lipidome (less than 1%) as also encountered in other studies [40,53]. In the present work, a higher level of isoprenoid units (IPP and DMAPP) was accomplished by a combination of the chromosomal integration of an inducible MEP-DOXP pathway [175] and the use of a statistically optimized medium [170]. Further strain optimization yielded an engineered bacterial strain in which the nearly complete PG pool is replaced by the archaeal AG. Importantly, the remarkable decrease of the PG content in favor of a high amount of newly synthetized AG demonstrates the functional integration of the ether lipid biosynthetic pathway in these cells.

A critical element of the introduced pathway and the generation of a mixed heterochiral membrane is the validation of the proper stereochemical configuration of the introduced ether lipids. The configuration of the glycerophosphate backbone represents one of the most distinctive differences between bacterial and archaeal lipids. With no exception, bacterial membranes are characterized by G3P-based lipids while archaea have G1P-based lipids [91]. The enzymes involved in the synthesis of G3P and G1P, G3PDH and G1PDH respectively, do not share any sequence and functional homology being members of evolutionary different protein families [2]. Moreover, since there is no mechanism known in *E. coli* for the production of G1P, the engineered *E. coli* strain lacking the introduced G1PDH, should not produce archaeal lipids. However, the *araM* gene was found to be redundant which raises questions on the stereoselectivity of the archaeal enzymes. Biochemical analysis using purified GGGPS from *M. maripaludis* suggest a preference for G1P over G3P, exhibiting a nine times higher affinity for G1P than G3P, and in the presence of saturating amounts of G3P, high levels of GGGP could be detected. In contrast, the analogous enzyme PlsB from *E. coli* produces LPA only in presence of G3P exhibiting a high stereoselectivity. Despite this lower stereoselectivity, even in the absence of the G1PDH AraM, the *in vivo*

synthesized archaeal lipids were derived from the G1P configuration. Thus, our data indicate that there must be a mechanism of G1P formation in *E. coli.* A possible mechanism of G1P formation is the phosphorylation of glycerol by glycerolkinase and the reductive phosphorylation of dihydroxyacetone by glycerol phosphate dehydrogenase. Although these enzymes are known to generate G3P, for none the chiral specificity has been examined in detail. Taken together, the data demonstrate that the lipid biosynthesis engineering resulted in the formation of a heterochiral mixed membrane in *E. coli*. The high stereoselectivity of the bacterial enzyme PlsB compared to the weak stereoselectivity of the archaeal enzyme GGGPS that carries out an analogous reaction, raises the possibility that the primordial insurgence of archaeal organisms was followed by the differentiation into bacteria. In this way the appearance of higher stereoselective enzymes as PlsB could have triggered the differentiation of bacterial organisms from the ancient cells, which further evolved in archaeal cell, keeping the primordial ability to survive in extreme environments and acquiring a specific membrane lipid composition. We never detected the formation of a possible archaeal counterpart of cardiolipin, di-archaetidylglycerol which suggests that the bacterial cardiolipin synthetase does not recognize the different chirality of the archaeal lipid.

A major question is if such heterochiral mixed membrane affects the cell characteristics. The engineered bacterial strains show a long lag phase of approximately 16 hours before growth commenced at growth rates comparable to the wild type. Genome sequencing of the adapted strain did not reveal any apparent mutation (data not shown) as expected for such short adjustment period. The restoration of growth could result from a metabolic adaptation of the bacterial strain and/or a tailoring of the expression of heterologous enzymes for the viable production of the archaeal ether lipids. On the other hand, strong induction of the archaeal lipid pathway causes severe cell stress as growth slows down and the cell morphology changes. Whereas the majority of engineered cells show elongated and thinner cells compared to the wild type strain, high induction also causes the formation of lobular appendages that are eventually released from the cells. Lipid analysis on these isolated

extrusions revealed the presence of a mixture of archaeal and bacterial lipids much akin the mother cell excluding the hypothesis of immiscibility and segregation of the archaeal lipids with the endogenous lipids as possible cause of this phenomenon. The extrusions also contain genetic material and likely originate from non-symmetrical cell division caused by the high level of archaeal lipid biosynthesis. The suggestion that archaeal lipids interfere with cell division is consistent with an important role of lipids in this process [176,177]. As the introduced ether lipid biosynthetic pathway is not fully integrated in the cellular and phospholipid homeostasis, we speculate that the shredding as seen under conditions of high induction is the result of high level overproduction of lipids that does not keep pace with other processes of cellular growth resulting in the formation of irregular division sites thereby clearing the cells from excess lipids.

Importantly, under conditions of moderate induction that lead to the nearly complete replacement of PG with AG, the archaeal lipids did not confer any toxicity to the bacterial cell. Archaea are well known to be able to survive under extreme conditions such as high temperatures [3], thus one may expect that the presence of archaeal lipids into a bacterial cell membrane could partially confer this ability. Indeed, a higher tolerance to heat treatment compared to control strains was observed. It should be stressed that the archaeal lipids are unsaturated and possibly, saturation will further enhance the survival to heat stress. Cells were also found to be more tolerant to freezing at -80 °C, a feature that can be attributed to the presence of the high concentration of unsaturated archaeal lipids which confers increased membrane fluidity needed to survive extreme cold temperatures [178,179]. Finally, the cells with the engineered membranes exhibited a higher tolerance against the organic solvent butanol. Although the acquired features of robustness are subtle, they are significant and demonstrate that bacteria gain properties by the presence of archaeal lipids rather than being detrimental to the cell's physiology.

The work described in the present study represents a unique approach to address a possible coexistence of archaeal and bacterial phospholipids as a heterochiral mixed membrane in a living bacterial cell. Despite the fact that the bacterial integral membrane proteins have evolved to function in

an ester-bond based phospholipid membrane, the near to complete replacement of one of the key lipid species of *E. coli*, phosphatidylglycerol for its archaeal counterpart, resulted in viable cells showing growth rates indistinguishable from that the parental strain. Our findings contrast the hypothesis of the instability of such membranes. The strategy described here may be applied to microorganism of industrial relevance to render them more robust with a higher tolerance to toxic products, organic solvents or byproducts without loss of productivity in bio-industrial processes. Further, it will be of interest to exploit the *E. coli* strains with archaeal phospholipids for the functional overproduction of archaeal membrane proteins.

Materials and Methods

Operon integration and cloning procedures

E. coli MG1655 genomic DNA was used as template for the amplification of the *IDI*, *IspDF* and *DXS* genes encoding for the MEP-DOXP operon. The primers and the plasmids used for the integration of the operon into *E. coli* are listed in **Table 2 and 3**. The three genes were cloned into the same plasmid vector, which was used as a template for the integration of the *lox71-kanR-lox66* selection marker cassette. The selection marker cassette along with the MEP-DOXP operon or the single *IDI* gene was amplified by PCR in order to get a DNA fragment for the integration into *E. coli* JM109DE3 competent cells via electroporation. *E. coli* cells containing the integrated operon were transformed with a plasmid expressing the Cre recombinase to remove the selection marker. The obtained *E. coli* strains (**Table 1**) containing the integrated *IDI* gene and the MEP-DOXP operon were used as basic strains for the following strain engineering. The primers and plasmids used for expressing the ether lipids genes in the engineered *E. coli* strains are listed in **Table 2 and 3**.

Table 1. *E. coli* strains used in this study.

Strain name	Genome integration	Plasmids
JM109DE3	none	pETduet and pRSF-duet

E. coli IDI$^+$	*IDI*	pETduet and pRSF-duet
E. coli MEP/DOXP$^+$	*IDI-IspDF-DSX*	pETduet and pRSF-duet
E. coli MEP/DOXP$^+$EL$^+$ (Ether Lipid)	*IDI-IspDF-DSX*	pMS148 and pAC029
E. coli MEP/DOXP$^+$EL$^+$AraM$^-$	*IDI-IspDF-DSX*	pMS148Δ and pAC029
E. coli IDI$^+$EL$^+$	*IDI*	pMS148 and pAC029

Table 2. Cloning and expression vectors used in this study.

Plasmid	Description	Reference
pGFPuv	Cloning vector expressing *Aequorea Victoria* GFP	Clontech
pCR2.1 TOPO	Cloning vector with *lox71-kanR-lox66* gene cassette	ThermoFisher
pKD46	Cre-recombinase expressing vector	[164]
pRSF-Duet-1	Cloning and expression vector (Kan^R), T7 promoter	Novagen
pPET-Duet-1	Cloning and expression vector (Amp^R), T7 promoter	Novagen
pACYC-Duet-1	Cloning and expression vector (CM^R), T7 promoter	Novagen
pMS003	*IDI* gene from *E. coli* MG1655 cloned into pGFPuv vector using primers BG3606 and BG3599	This study
pMS008	*IspDF* genes from *E. coli* MG1655 cloned into pMS003 vector using primers BG3600 and BG3601	This study
pMS011	*DXS* gene from *E. coli* MG1655 cloned into pMS008 vector using primers BG3602 and BG3603	This study
pMS051	*lox71-kanR-lox66* gene cassette from pCR2.1 TOPO vector cloned into pMS003 vector using primers BG4429 and BG4430	This study
pMS053	*lox71-kanR-lox66* gene cassette from pCR2.1 TOPO vector cloned into pMS011 vector using primers BG4429 and BG4430	This study
pMS016	*crtE* gene from *Pantotea ananatis* cloned into the pACYC-duet vector using primers BG3899 and BG3900	This study
pMS017	*crtB* and *crtI* genes from *Pantotea ananatis* cloned into the pMS016 vector using primers BG3901 and BG3902	This study
pSJ130	*araM* gene from *Bacillus subtilis* cloned into pET-duet vector using primers 70 and 71	[41]
pMS148	c*rtE* gene from *Pantotea ananatis* digested with EcoRI and cloned into the pSJ130 vector	This study
pMS148Δ	pMS148 vector containing a deleted version of the *araM* gene using the EcoRV and BmgBI	This study
pSP001	Codon-optimized GGGPS and DGGGPS genes from *M. maripaludis* cloned into pRSF-duet vector using the primers 11, 12, 39 and 40	[41]
pAC027	Codon optimized *carS* gene from *A. fulgidus* cloned into pSP001 vector using primers 583 and 584	This study
pAC029	*pssA* gene from *B. subtilis* cloned into the pAC027 vector using the vector using primers 585 and 586	This study

pSJ103	Codon optimized GGGPS and DGGGPS genes from *M. maripaludis* cloned into pRSF-duet vector using the primers 11 and 12	[41]
pME001	*fadD* gene from *E. coli* MG1655 cloned into pRSF-Duet-1 vector using the primers PrME001 and PrME002	This study
pME002	*plsB* gene from *E. coli* MG1655 cloned into pet28b vector using the primers PrME003 and PrME004	This study

Table 3. Oligonucleotide primers used in the present study.

Primers name	Primer sequence 5' → 3'	Restriction enzyme site
BG3606	GGCATGCCATGCAAACGGAACACG	SphI
BG3599	GCTGCAGTTATTTAAGCTGGGTAAATGCA	PstI
BG3600	GCTGCAGAGGAGATATACATATGGCAACCACTCATTTGG	PstI
BG3601	GTCTAGATCATTTTGTTGCCTTAATGAGTAGCG	XbaI
BG3602	GTCTAGAGGAGATATACTGATGAGTTTTGATATTGCCAAATAC	XbaI
BG3603	GCGGTACCTTATGCCAGCCAGGCC	KpnI
BG4429	GACGCGTACGGTGTCTTTTTTTACCTGTTTGACC	BsiWI
BG4430	GACGCTTAAGCTACCTCTGGTGAAGGAGTTGG	AflII
BG3899	GAACGAATTCAGCCCGAATGACGGTCTGC	EcoRI
BG3900	GAATCTTAAGGCGCGACCAGTTCCTGAG	AflII
BG3901	GCTGAGATCTGATGAAACCAACTACGGTAATTGG	BglII
BG3902	CTTACTCGAGAAAGACATGGCGCTAGAG	XhoI
70	GCGCCATATGAATCGTATCGCAGCTGAC	NdeI
71	GCGCCTCGAGTTAGTGATGATGGTGGTGATGTTCATATAGACCATGGTTGATC AGCG	XhoI
11	GCGCGAATTCATGCATCACCACCACC	EcoRI
12	GCGCAAGCTTTCATTTTTTGGACAGC	HindIII
39	TCTTTACCTCTCTTATACTTAACTAATATACTAAGATGGG	blunt
40	CATATGGGCAGCCATCACCATCATCACCACAGCG	blunt
583	CGCGCCATGGATGCTGGATCTGATTCTGAAAACCATTTG	NcoI
584	CGCGGGATCCTTAGTGATGGTGATGGTGGTGATG	BamHI
585	GCGCGCGGCCGCATGAATTACATCCCCTGTATG	NotI
586	CGCGCTTAAGTTAGTGATGGTGATGGTGGTG	AflII
PrME001	TACTAGGAATTCATGAAGAAGGTTTGGCTTAACCG	EcoRI
PrME002	AGTCATGCGGCCGCTCAGTGGTGGTGGTGGTGGGCTTTATTGTCCACTTT GC	NotI
PrME003	CATCCTGCCCATGGCCGGCTGGCCACGAATTTACTAC	NcoI
PrME004	CTTCATGATTCTCGAGCCCTTCGCCCTGCGTCGCAC	XhoI

Bacterial strains and growth conditions

Engineered *E. coli* strains were grown under aerobic conditions at 37 °C in 200 ml of LB medium supplemented with the antibiotics ampicillin (100 µg/ml) and kanamycin (50 µg/ml). OPT1 medium [170] was prepared by autoclaving a solution based on glycerol 1% (v/v), KH_2PO_4 2.4% (w/v), $(NH_4)_2HPO_4$ 0.4% (w/v), citric acid 0.17% (w/v) and by adding sterile 1

mM $NiCl_2$, 0.12 mM $MgSO_4$ and 1x MEM vitamin solution (Sigma-Aldrich). When not specified the cells were induced with 0.1 mM IPTG for 3 hours or overnight.

Lycopene quantification

Lycopene quantification was done as described by Yoon *et al.* [166] An aliquot of a growing culture (10 ml) was centrifuged at 4,700 x*g* for 10 minutes. The obtained pellet was washed with 1 ml of milli-Q water and further centrifuged at 12,000 x*g* for 5 minutes. The residual cell pellet was suspended in 500 µl of acetone and incubate 30 minutes at 55 °C to promote the lycopene extraction and then centrifuged for 20 minutes at 16,000 x*g* at 4 °C as previously described [166]. The extraction was repeated twice and the obtained lycopene extracts was additionally centrifuged at 16,000 x*g* for 2 minutes to remove possible impurities. Samples of 250 µl were then diluted with 750 µl of acetone and the amount of extracted lycopene was determined by measuring the absorbance at 472 nm. The lycopene concentration was calculated by means of a calibration curve and normalized to the dry cell weight (DCW).

Thin Layer Chromatography

An aliquot (3 µl) of lipid extracts from the different *E. coli* strains was spotted on Silica Gel 60 (Merck) plates. A solvent mixture of chloroform, methanol and water (50:10:1) was used as mobile phase for the separation of the different lipid species which were detected by molybdenum blue [180]. A solvent system chart from Avanti (http://avantilipids.com/tech-support/analytical-procedures/tlc-solvent-systems/) was used as reference for the lipid identification. The spots were relatively quantified using ImageJ software.

Expression and purification of GGGPS, FadD and PlsB enzymes

The archaeal protein GGGPS from *M. maripaludis* was expressed and purified as previously described [41]. The bacterial FadD and PlsB proteins from *E. coli* were overexpressed in *E. coli* BL21 induced with 1 mM IPTG as will be detailed elsewhere. After 2 hrs of induction, the cells were harvested (8,754 x*g*) and washed with buffer A containing 50 mM Tris-HCl

pH 8.0, 100 mM KCl and 20% glycerol. After re-suspension, the cells were supplemented with 0.5 mg/ml of DNAse and a complete EDTA free protease inhibitor tablet (Roche). The suspension was subjected to cell disruption at 13,000 psi and the cell lysate was centrifuged for 15 minutes at low spin (12,000 x*g*) to remove unbroken cells.

The purification of the cytoplasmic protein FadD was performed by separation of the cytosolic fraction from membranes by a centrifugation step at 43,667 x*g* for 15 min. The supernatant was incubated with Ni-NTA beads (SigmaAldrich) in buffer A for 60 min at 4 °C. The beads were washed 5 times with 20 column volumes (CV) of buffer A supplemented with 10 mM imidazole and eluted 2 times with 2 CV of buffer A supplemented with 300 mM imidazole.

The membrane protein PlsB was purified by a high-speed centrifugation at (235,000 xg) for 1 hour to isolate the membrane fraction. Total membrane (pellet) were suspended in buffer A and solubilized at 4 ˚C for 1 hour in 2% of n-dodecyl-β-D-maltopyranoside (DDM) detergent. Insolubilized materials were removed by centrifugation (15,800 xg) for 10 minutes and the supernatant was incubated with Ni-NTA beads for 90 min at 4 °C. The beads were washed 5 times with 40 CV of buffer B (0.05 % DDM, 50 mM Tris pH 8.0, 100 mM KCl, 20% glycerol) supplemented with 10 mM imidazole and eluted 3 times with 0.5 CV of buffer B supplemented with 300 mM imidazole. The purity of the proteins was checked by 15% SDS-PAGE, stained with Coomassie Brilliant Blue. Absorbance was measured at 280 nm to determine the concentration of purified protein.

In vitro enzyme reactions

In vitro reactions were performed in 100 µl of buffer containing a final concentration of 50 mM Tris-HCl pH 7.5, 10 mM $MgCl_2$, 60 mM NaCl, 100 mM Imidazole, 0.08% DDM and 4% glycerol. Where specified, 100 µM GGPP, 10 mM G3P, 10 mM G1P and the indicated amount of purified enzymes were added to the reaction mixture. The reactions were incubated at 37 °C for 1 hour. Kinetic assays were performed using the same reaction mixture but the reactions were incubated at 37 °C for 2 hours. The coupled FadD-PlsB *in vitro* assay was performed in a 100 µl reaction volume containing: 50 mM Tris-HCl pH 8, 10 mM $MgCl_2$, 100 mM

KCl, 20% glycerol, 0.05% DDM, 2mM DTT, 2.67 mM lipids (DOPC: DOPG:DOPE, 1:1:1). Where specified, 300 µM oleic acid, 40 µM CoA, 1 mM ATP, 0.5 µM purified FadD, 1 µM purified PlsB, 10 mM G3P and 10 mM G1P were added to the reaction mixture. Reactions were incubated at 37 °C for 4 hours. The products were extracted two times with 0.3 ml of *n*-butanol. Extracted lipids were evaporated under a stream of nitrogen gas and resuspended in 50 µl of methanol for the LC-MS analysis.

Lipid analysis

E. coli strains induced for the archaeal lipids synthesis were grown as described above. The total membrane fractions were isolated and the total lipid content was extracted according to the Bligh and Dyer method [163]. Samples were then resuspended in 100 µl of methanol for LC-MS analysis, or total lipid quantitation by a colorimetric assay, based on the formation of a complex between phospholipids and ammonium ferrothiocyanate [181]. Samples (20 µl) were evaporated under a nitrogen stream and resuspended in 500 µl of chloroform; 250 µl of ferrothiocyanate reagent was then added to the chloroform layer, mixed for 1 min and allowed to phase separate for 5 minutes. The lower red phase was collected and the absorbance at 490 nm was measured and calibrated against standards. The obtained values were also used to normalized the LC-MS ion counts for amounts of individual lipids.

LC-MS analysis

The lipid extracts and the samples from *in vitro* reactions were analyzed using an Accela1250 high-performance liquid chromatography system coupled with an electrospray ionization mass spectrometry (ESI-MS) Orbitrap Exactive (Thermo Fisher Scientific). A volume of 5 µl of each sample was used for the analysis. The LC-MS method parameters used in this study to analyzed both type of samples were the same as described previously [163].

Analysis of the configuration of the ether lipids

All chemical reactions were carried out under a nitrogen atmosphere using oven-dried glassware and using standard Schlenk techniques. Reaction

temperature refers to the temperature of the oil bath. All reagents and catalysts were purchased from Sigma-Aldrich, Acros, J&K Scientific and TCI Europe and used without further purification unless otherwise mentioned, any purification of reagents was performed following the methods described by Armarego et al. [182]. TLC analysis was performed on Merck silica gel 60/Kieselguhr F254, 0.25 mm. Compounds were visualized using either Seebach's reagent (a mixture of phosphomolybdic acid (25 g), cerium (IV) sulfate (7.5 g), H_2O (500 mL) and H_2SO_4 (25 mL)), 2,4-DNP stain (2,4-dinitrophenylhydrazine (12 g), conc. sulfuric acid (60 ml), water (80 ml), ethanol (200 ml)) or elemental iodine. Flash chromatography was performed using SiliCycle silica gel type SiliaFlash P60 (230 – 400 mesh) as obtained from Screening Devices. GC-MS measurements were performed with an HP 6890 series gas chromatography system equipped with an HP1 or HP5 column (Agilent Technologies, Palo Alto, CA), and equipped with an HP 5973 mass sensitive detector. High resolution mass spectra (HRMS) were recorded on a Thermo Scientific LTQ Orbitrap XL. (ESI+, ESI- and APCI). ^{1}H-, ^{13}C- and ^{19}F-NMR spectra were recorded on a Varian AMX400 (400, 101 and 376 MHz, respectively) using $CDCl_3$ as solvent unless stated otherwise. Chemical shift values are reported in ppm with the solvent resonance as the internal standard ($CDCl_3$: δ 7.26 for ^{1}H, δ 77.16 for ^{13}C). Data are reported as follows: chemical shifts (δ), multiplicity (s = singlet, d = doublet, dd = double doublet, ddd = double double doublet, td = triple doublet, t = triplet, q = quartet, b = broad, m = multiplet), coupling constants *J* (Hz), and integration. Enantiomeric excesses were determined by chiral HPLC analysis using a Shimadzu LC- 10ADVP HPLC instrument equipped with a Shimadzu SPD-M10AVP diode-array detector. Optical rotations were measured on a Schmidt+Haensch polarimeter (Polartronic MH8) with a 10 cm cell (c given in g/mL) at ambient temperature (±20 °C).

Scanning Electron Microscopy and Bright Field Microscopy

For the scanning electron microscopy analysis 150 µl of cell suspension was immobilized on poly-L-lysine coated cover slips (Corning art. 354085) for 1 hour. 2.5% glutaraldehyde in 0.05 M sodium cacodylate buffer pH 7.2 was added to the glass at room temperature for 1 hour. The sample was

rinsed three times in the same buffer and fixed for 1 hour in 1% OsO4 (w/v) in the same buffer. Two washes with water were performed, followed by a dehydration in a graded ethanol series (10, 30, 50, 70, 90, 100%) and dried with carbon dioxide (Leica EM CPD 300). The glasses were attached on a sample holder by carbon adhesive tabs (EMS Washington USA), sputter coated with tungsten (Leica EM SCD 500) and analyzed and digitally imaged with a field emission scanning electron microscope (FEI Magellan 400). Sample preparation, imaging and measurements were performed by the Wageningen Electron Microscopy Centre (WEMC) facility.

The bright field microscopy was performed on cells grown until exponential phase. Aliquots of 1 ml were centrifuged at max speed for 30 s on the top bench centrifuge. The obtained pellet was resuspended in 100 µl of Phosphate Buffered Saline (PBS, 58 mM Na_2HPO_4, 17 mM NaH_2PO_4 and 68 mM NaCl pH 7.3). The FM4-64 and DAPI dyes were added to the solution at the final concentration of 0.8 µM and 36 nM respectively [183]. The solution was incubated at room temperature for 10 minutes and centrifuged at max speed for 30 s. The stained cell pellet was suspended in 40 µl of PBS and spotted on agarose pad (1% w/v in PBS). Cells were imaged using a Nikon Ti-E-microscope (Nikon Instruments) equipped with a Hamamatsu Orca Flash 4.0 camera. The image analysis was performed by the software ImageJ [184].

Robustness tests

The engineered *E. coli* strains and the controls strains were grown and induced as described above. A dilution into fresh medium was performed to reach the OD_{600} = 1.0 and the obtained culture was diluted again for a dilution factor of 1000x in order to have approximately 10^5- 10^6 cells/ml. For the freezing survival, the cells aliquots of 20 µl were frozen in liquid nitrogen and kept at -80 °C for 4 days. An untreated cell sample was plated in a 100x dilution to be used as reference. Heat shock treatment was performed by exposing the strains to different temperatures (37 °C, 42 °C, 46 °C, 50 °C, 55 °C and 58 °C) for 2 minutes. The cells were then recovered by adding 980 µl of LB medium and incubated at 37 °C for 1 hour and platted for CFU counts. Butanol tolerance was tested by incubating cells in

LB supplemented with different butanol concentration (0%, 0.5%, 1%, 1.5%, 2% and 2.5%) for 2 minutes and recovering the treated cells at 37°C for 1 hour. After all the treatments 100 µl of cells were plated on LB agar plate supplemented with the proper antibiotics (Ampicillin 100 µg/ml and kanamycin 50 µg/ml) and incubated at 37°C overnight. The colony counting was performed using a developed plugin for the software ImageJ.

Supplemental Information

Supplementary information includes one table, three figures and supplemental experimental procedures.

Acknowledgments

This work was carried out within the research program of the biobased ecologically balanced sustainable industrial chemistry (BE-Basic). We thank Tiny Franssen Verheijen from the Wageningen Electron Microscopy Centre (WEMC) for providing the electron micrographs and Teunke van Rossum for providing the *Lox-KanR-lox* integration cassette. We also thank Anne-Bart Seinen for the graphic assistance and the development of a colony counting software and thank Anabela de Sousa Borges for the bright field microscopy assistance.

Authors contributions

A.C., M.S., A.D., S.K. and J.vd.O. conceived and designed the research. M.S. performed the operon integration, lycopene quantification, robustness tests and strain optimization. A.C. performed the total lipid analysis, the *in vitro* biochemical analysis of GGGPS, the robustness tests and the bright field microscope analysis. M.E. cloned the genes, purified the enzymes and performed the *in vitro* experiments for the PlsB activity assay. A.M. designed the synthesis of the synthetic standards for the total lipid stereochemical configuration analysis, which was performed by V. J. and R. H. The manuscript was written by the contributions of all the authors.

Competing financial interests

The authors declare that the research was conducted in absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Supplemental Information

Table S1 Combination of bacterial and enzymes used in this study for the *in vivo* ether lipids production.

Locus (gene)	Source	Protein expressed	Function	Ref
CrtE	*P. ananatis*	GGPP synthesis	IPP+DMAPP →GGPP	[167]
BSU28760 (araM)	*B. subtilis*	G1P dehydrogenase – His_6	DHAP+NADH→G1P	[31]
MmarC7_1004	*M. maripaludis*	His8-GGGP synthase (codon optimized)	G1P+GGPP→ GGGP	[35]
MmarC7_RS04845	*M. maripaludis*	His8-DGGGP synthase (codon optimized)	GGPP+GGGP → DGGGP	[21]
AF1740	*A. fulgidus*	CarS-His8 (codon optimized)	DGGGP+CTP → CDP---archaeol	[41]
pssA	*B. subtilis*	PssA-His8	CDP---archaeol→ AS	[163]
psd	*E. coli*	Psd-His8	AS→AE	[163]
pgsA	*E. coli*	PgsA-His8	CDP---archaeol→ AGP	[163]
pgpA	*E. coli*	PgpA---His8	AGP→AG	[163]

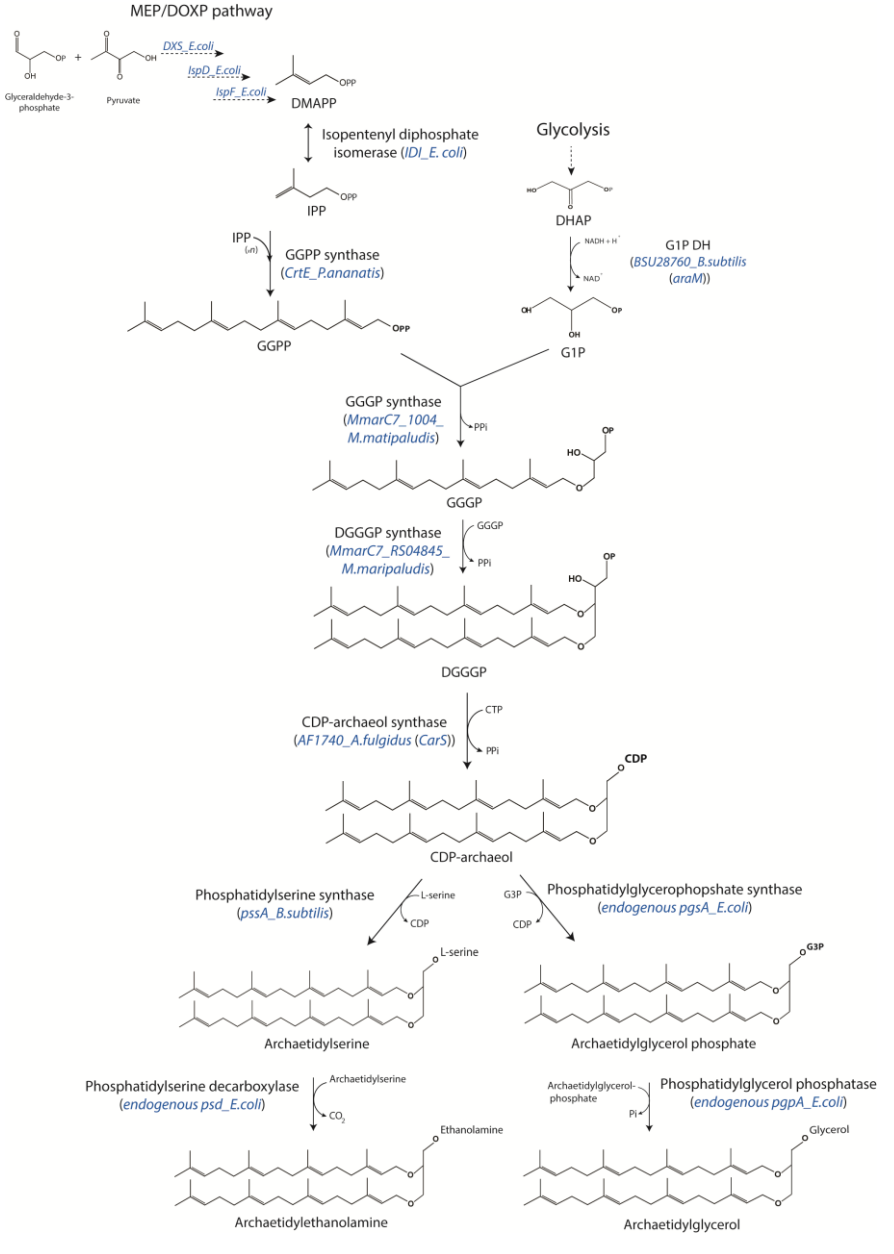


Figure S1 related to Figure 1| Schematic representation of the biosynthetic pathway introduced into the bacterium *E. coli* for archaeal lipids synthesis. The bacterial MEP/DOXP pathway enzymes used to overproduce the isoprenoid building blocks IPP and DMAPP and the genes encoding the enzymes for the ether lipids synthesis are highlighted in blue. The scheme indicates all the biosynthetic steps introduced in the bacterium *E. coli* for the production of a heterochiral mixed membrane.

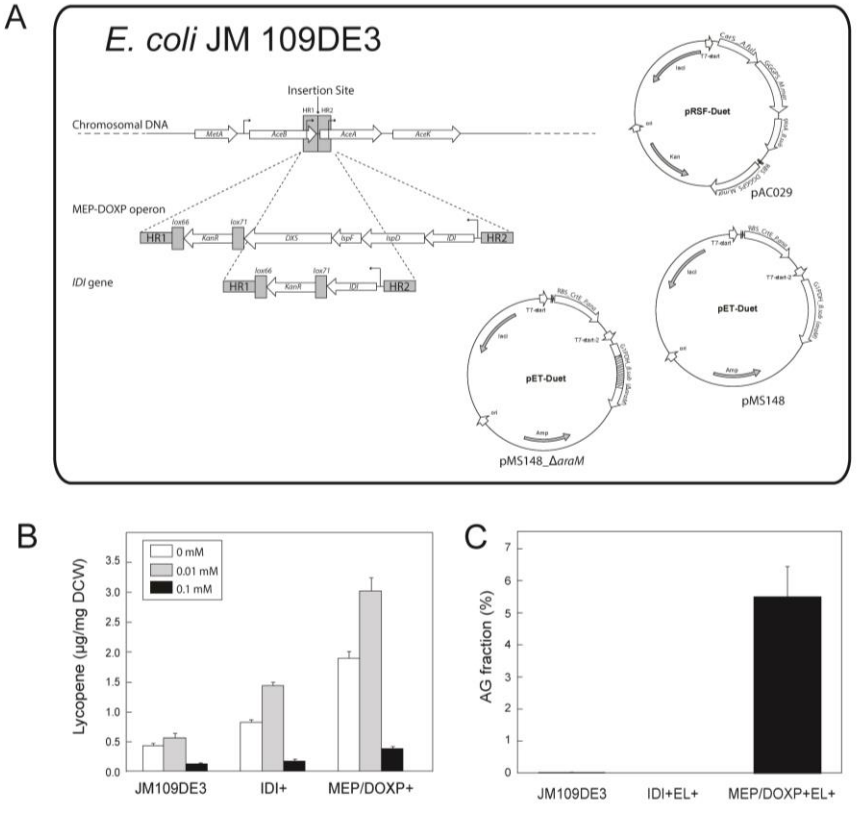


Figure S2 related to Figure 1| *E. coli* metabolic engineering. (**A**) Schematic representation of the engineering of *E. coli* JM109DE3 showing the integration of the MEP-DOXP operon or the *IDI* gene into the chromosome and the three vectors harboring the ether lipids enzymes. (**B**) Effect of the chromosomal integration of the *IDI* gene and the *IDI-IspDF-DXS* operon on the synthesis of isoprenoid building blocks as monitored through the production of lycopene using 0 (white bars), 10 (grey bars) and 100 (black bars) µM IPTG for induction. (**C**) *In vivo* production of AG by engineered *E. coli* strains (JM109DE3, IDI^{+}EL^{+} and MEP/DOXP^{+}EL^{+}) with improved IPP and DMAPP synthesis. Total ion count from LC-MS were normalized for the total amount of lipids present in each sample. The data are the averages of three biological replicates ± S.E.M. (AG: archaetidylglycerol)

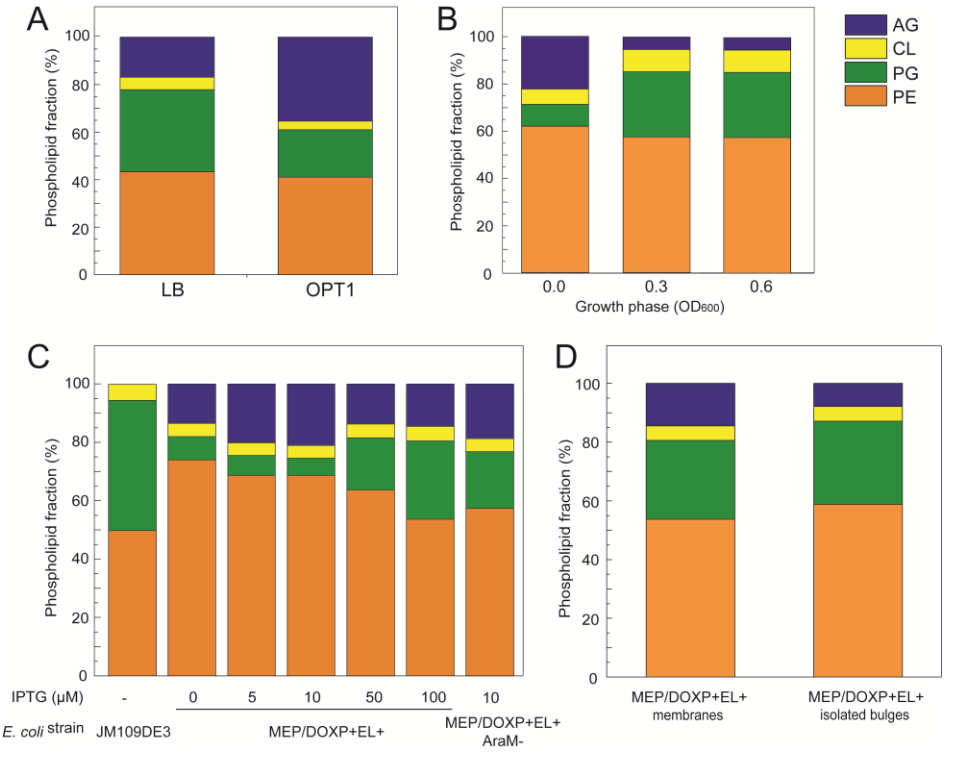
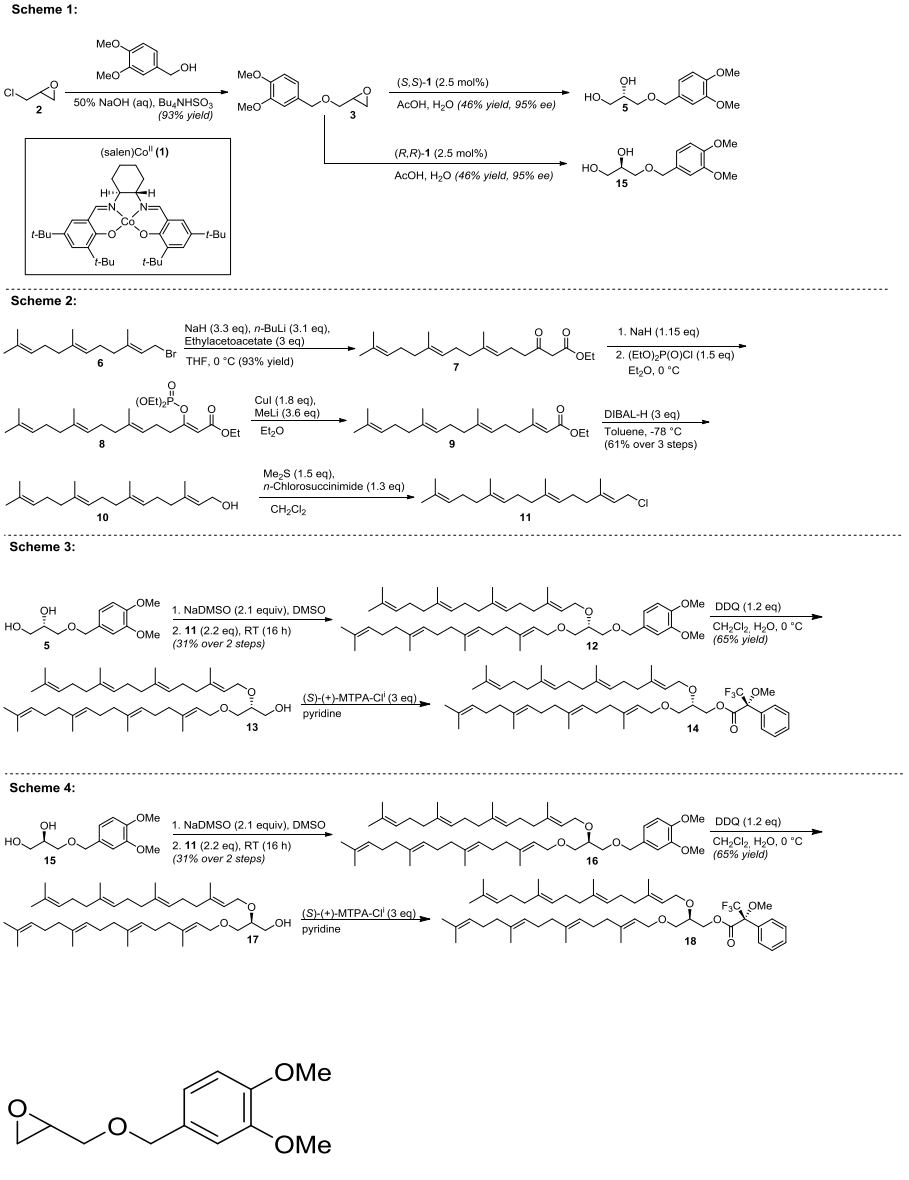


Figure S3 related to Figure 1| Optimization of archaeal lipid production in *E. coli.* Total lipid analysis of the *E. coli* MEP/DOXP^{+}EL^{+} strain harboring the entire ether lipid biosynthetic pathway grown and induced under different conditions. (**A**) Comparison between rich LB medium and a defined minimal medium (OPT) optimized for the isoprenoid production. (**B**) Different growth phases and (**C**) IPTG concentrations including the JM109DE3 wild type strain, the engineered *E. coli* MEP/DOXP^{+}EL^{+} strain and *E. coli* MEP/DOXP^{+}EL^{+}AraM^{-} strain lacking the *araM* gene. Total ion counts are normalized for the total amount of lipids present in each sample. The data are the averages of three biological replicates ± S.E.M. (AG: archaetidylglycerol, CL: cardiolipin, PG: phosphatidylglycerol and PE: phosphatidylethanolamine). (**D**) Total lipid analysis of the heterochiral mixed membrane *E. coli* strain and the isolated bulges. The lobular appendages were separated from the bacterial cells by a centrifugation step at low speed (5403 *xg*) and high speed (235,000*xg*). The obtained pellet was resuspend and the lipid analysis was performed and compared with the total lipidome of the isolated membranes from the same *E. coli* strain. The different total amount of lipids reflects the low starting material of the isolate bulges compared the total membranes. The total ion counts from LC-MS were normalized using eicosane as internal standard.

Chemical synthesis of the standard used for the phospholipid chirality analysis

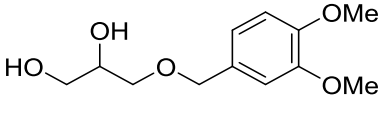


2-(((3,4-dimethoxybenzyl)oxy)methyl)oxirane[185] **(3).** To a 100 mL 3-necked flask equipped with magnetic stirrer bar was added 25 mL of a 50% NaOH solution, epichlorohydrin (18.5 g, 15.6 mL, 0.2 mol) and

Bu_4NHSO_4 (1.5 mmol, 525 mg, 4 mol%). The resulting solution was cooled to 0 ˚C (ice/water-bath) after which neat 3,4-dimethoxybenzyl alcohol (37.5 mmol, 5.5 mL, 6.3 g) was added dropwise over 30 min while the solution was stirred vigorously. The resulting turbid mixture was allowed to warm up over a 5 h period, after which complete conversion was observed by TLC. The entire content of the flask was poured into 100 mL of ice water which was subsequently extracted with diethyl ether (3 x 50 mL). The combined organic layers were washed with brine (2 x 50 mL) dried over $MgSO_4$ and concentrated *in vacuo*. The resulting crude was further purified by column chromatography (1:3 EtOAc/pentane) to give 2-(((3,4-dimethoxybenzyl)oxy)methyl)oxirane as a pale yellow oil (94% yield, 7.9 g).

^{1}H NMR (400 MHz, Chloroform-*d*) δ 6.94 – 6.78 (m, 3H), 4.52 (q, *J* = 11.6 Hz, 2H), 3.89 (s, 3H), 3,87 (s, 3H), 3.75 (dd, *J* = 11.5, 2.9 Hz, 1H), 3.41 (dd, *J* = 11.4, 5.9 Hz, 1H), 3.19 (td, *J* = 6.3, 3.2 Hz, 1H), 2.80 (t, *J* = 4.6 Hz, 1H), 2.61 (dd, *J* = 4.9, 2.7 Hz, 1H).

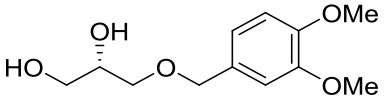
^{13}C NMR (101 MHz, Chloroform-*d*) δ 149.0, 148.7, 130.4, 120.4, 111.1, 110.9, 73.2, 70.6, 55.9, 55.8, 50.8, 44.3.



3-((3,4-dimethoxybenzyl)oxy)propane-1,2-diol (4). Epoxide **3** (200 mg, 0.9 mmol) with 6 mL water was added to a 10 mL round-bottomed flask equipped with magnetic stirrer bar. To this mixture was added 0.2 mL of 10% aqueous sulfuric acid followed by stirring for 5 h at rt. The resulting acidic solution was neutralized with 1 M NaOH and extracted with ethyl acetate (3 x 5 mL). The combined organic layers were washed with brine (2 x 5 mL), dried over $MgSO_4$ and concentrated *in vacuo* which yielded the desired product as a colorless thick oil (98% yield, 210 mg).

^{1}H NMR (400 MHz, Chloroform-*d*) δ 6.87 – 6.77 (m, 3H), 4.44 (s, 2H), 3.85 (s, 3H), 3.83 (s, 3H), 3.67 – 3.60 (m, 1H), 3.55 (dd, *J* = 11.5, 5.9 Hz), 3.52 – 3.42 (m, 2H), 3.03 (br s, 2H).

^{13}C NMR (101 MHz, Chloroform-*d*) δ 149.1, 148.8 , 130.3 , 120.5 , 111.2 , 111.0 , 73.5 , 71.5 , 70.8 , 64.1 , 55.6 , 55.9.



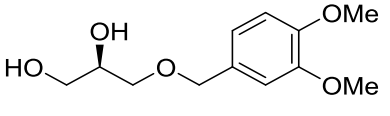
(*S*)-3-((3,4-dimethoxybenzyl)oxy)propane-1,2-diol[185] **(5).** A 25 mL flask equipped with a magnetic stirrer bar was charged with (*S,S*)-**1** (70 mg, 0.005 equiv). The catalyst was exposed to 2-(((3,4-dimethoxybenzyl)oxy)methyl)oxirane (5 g, 22.3 mmol) and AcOH (25 µL, 0.2 equiv). The resulting red mixture was allowed to stir for 30 min in order to oxidize the catalyst. To the resulting brown mixture was added H_2O (220 µL, 0.55 equiv) and was stirred rt for 48 h. The final product was isolated as a brown oil by flash column chromatography (100% EtOAc) (45% yield, 2.2 g).

Chiral HPLC analysis on a Lux® 5 µm Cellulose-3 column, *n*-heptane : *i*-PrOH = 90 : 10, 40 ˚C, flow = 1 mL/min, UV detection at 274 nm, t_R(major): 25.29 min, t_R(minor): 29.06 min, 97% ee

^{1}H NMR (400 MHz, Chloroform-*d*) δ 6.87 – 6.77 (m, 3H), 4.44 (s, 2H), 3.85 (s, 3H), 3.83 (s, 3H), 3.67 – 3.60 (m, 1H), 3.55 (dd, *J* = 11.5, 5.9 Hz, 1H) 3.52 – 3.42 (m, 2H), 3.03 (br s, 2H).

^{13}C NMR (101 MHz, Chloroform-*d*) δ 149.1, 148.8, 130.3, 120.5, 111.2, 111.0, 73.5, 71.5, 70.8, 64.1, 55.6, 55.9.

$[\alpha]_D^{20}$= -2.4 (c = 0.1 g/mL, $CHCl_3$).



(*R*)-3-((3,4-dimethoxybenzyl)oxy)propane-1,2-diolError! Bookmark not defined. (15). This compound was prepared with the same synthetic procedure that was used for **(*S*)-3-((3,4-**

dimethoxybenzyl)oxy)propane-1,2-diol (5), using (*R,R*)-**1** as catalyst (45% yield).**Error! Bookmark not defined.**

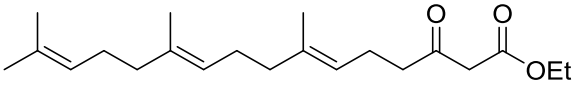
Chiral HPLC analysis on a Lux® 5 µm Cellulose-3 column, *n*-heptane : *i*-PrOH = 90 : 10, 40 ˚C, flow = 1 mL/min, UV detection at 274 nm, t_R(minor): 26.01 min, t_R(major): 29.21 min, 95% ee.

^{1}H NMR (400 MHz, Chloroform-*d*): Same as reported for compound **5**

^{13}C NMR (101 MHz, Chloroform-*d*): Same as reported for compound **5**

$[\alpha]_D^{20}$= +2.4 (c = 0.1 g/mL, $CHCl_3$).

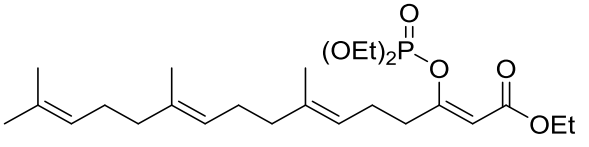
The spectral data correspond to those previously reported[124]



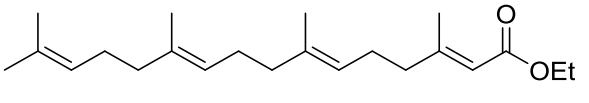
ethyl (6*E*,10*E*)-7,11,15-trimethyl-3-oxohexadeca-6,10,14-trienoateError! Bookmark not defined. **(7).** An oven dried Schlenk flask equipped with magnetic stirrer bar was charged with NaH (60% dispersion, 136 mg, 3.3 equiv). The mineral oil was removed by 3 successive washings with pentane. The remaining white solid was dried in vacuum, suspended in dry THF (2.5 mL) and cooled to 0 °C (ice/water-bath). To the resulting suspension, freshly distilled ethyl acetoacetate (400 mg, 3 equiv) was added dropwise over 15 min after which the solution turned light yellow. After stirring for an additional 15 min at 0 °C, a solution of *n*-BuLi in hexanes (1.6 M, 1.95 mL, 3 equiv) was added over 15 min. The resulting dark yellow solution was allowed to stir further for 15 min at 0 °C. Farnesyl bromide (**6**) (286 mg, 1 mmol) in 0.55 mL of dry THF was added dropwise over 10 min. The resulting orange suspension was quenched by the addition of HCl (1 M, 1.5 mL). The aqueous layer was separated and extracted with Et_2O (3 x 2 mL), the organic layers were combined, washed with brine, dried over $MgSO_4$ and concentrated *in vacuo*. The obtained crude oil was further purified by flash column chromatography after which the pure aceto-ester was obtained as a pale yellow oil (10% Et_2O in pentane) (yield 93%).

^{1}H NMR (400 MHz, Chloroform-*d*) δ 12.06 (s, 0.2H), 5.10 – 4.99 (m, 3H), 4.22 – 4.08 (m, 2H), 3.38 (d, 2H), 2.52 (t, *J* = 7.4 Hz, 2H), 2.24 (q, *J* = 7.4 Hz, 2H), 2.08 – 1.88 (m, 8H), 1.63 (s, 3H), 1.56 (m, 9H), 1.23 (t, *J* = 7.1 Hz, 3H).

The spectral data correspond to those previously reported[125]

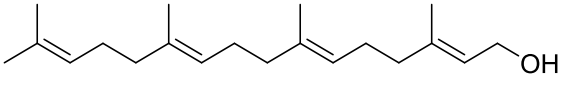


ethyl (2*Z*,6*E*,10*E*)-3-((diethoxyphosphoryl)oxy)-7,11,15-trimethylhexadeca-2,6,10,14-tetraenoateError! Bookmark not defined. (8). An oven dried Schlenk flask equipped with magnetic stirrer bar was charged with NaH (60% dispersion, 46 mg, 1.15 mmol, 1.15 equiv). The mineral oil was removed by 3 washings with pentane. The remaining white solid was dried in vacuum and suspended in dry Et_2O (4.5 mL). The suspension was cooled to 0 °C (ice/water-bath) and a solution of ethyl (6*E*,10*E*)-7,11,15-trimethyl-3-oxohexadeca-6,10,14-trienoate (**7**) in dry Et_2O (1.5 mL) was added over 15 min. The resulting yellow homogeneous mixture was stirred for 15 min at 0 °C and for 15 min at rt. The solution was again cooled to 0 °C and neat diethylchlorophosphate was added over 5 min. The resulting mixture was stirred for 15 min at 0 °C after which the reaction was quenched by addition of saturated aqueous NH_4Cl solution (3 mL). The organic layer was separated and the aqueous layer was extracted with Et_2O (3 x 3 mL). The combined organic layers were washed with saturated aqueous $NaHCO_3$ (3 x 3 mL), brine (3 x 3 mL), dried over $MgSO_4$ and the solvent was removed *in vacuo*. The resulting yellow oil (400 mg) was used without further purification in the successive step.



ethyl (2*E*,6*E*,10*E*)-3,7,11,15-tetramethylhexadeca-2,6,10,14-tetraenoateError! Bookmark not defined. **(9).** An oven dried Schlenk flask

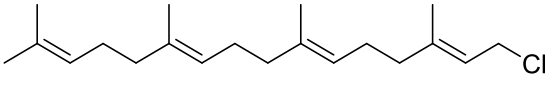
equipped with magnetic stirrer bar was charged with CuI (340 mg, 1.8 mmol, 1.8 equiv) and 1.1 mL Et_2O. The resulting suspension was treated with MeLi (1.6 M in Et_2O, 2.25 mL 3.6 equiv) upon which the mixture turned bright yellow. After complete addition of the MeLi the mixture was homogeneous and colorless. The mixture was cooled to -78 °C (cryostat, acetone-bath) and a solution of the phosphate **8** (400 mg) in dry Et_2O (1.4 mL) was added dropwise such that the phosphate solution was cooled by the cold wall of the Schlenk flask, during which the color of the reaction mixture turned to yellow. After complete addition the resulting bright orange solution was stirred at -78 °C for 1 h, after which the mixture was allowed to warm up to -45 °C and was stirred for an additional 2 h. The resulting dark red mixture was quenched by adding 130 µL of MeI and after stirring for 10 minutes the entire content of the Schlenk flask was carefully poured into a saturated aqueous solution of NH_4Cl (5 mL) and NH_4OH (25%, 6 mL) during which gas evolution was observed. The resulting mixture was stirred until it became homogeneous. The layers were separated, the aqueous layer was extracted with Et_2O (3 x 5 mL), the organic layers were combined and washed with an aqueous solution of NH_4OH (25%, 2 x 8 mL), brine (2 x 8 mL), dried and concentrated *in vacuo*. 290 mg of a yellow oil was obtained and used in the successive reaction without further purification.



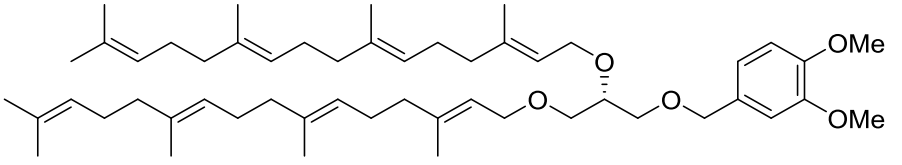
(2*E*,6*E*,10*E*)-3,7,11,15-tetramethylhexadeca-2,6,10,14-tetraen-1-ol (10). To an oven dried Schlenk flask under N_2 atmosphere was added ethyl (2*E*,6*E*,10*E*)-3,7,11,15-tetramethylhexadeca-2,6,10,14-tetraenoate (**9**), from the previous step in dry toluene (3 mL). The resulting solution was cooled to -78 °C (cryostat, acetone-bath) and a solution of DIBAL (1 M in CH_2Cl_2, 2.6 mL, 3 equiv) was added dropwise over 20 min. The resulting mixture was stirred at -78 °C for 1h after which full consumption was observed by TLC (30% Et_2O in pentane, I_2-stain, $R_f \sim 0.3$). The reaction was quenched by dropwise addition of MeOH over 10 min (gas evolution observed) after which the mixture stirred until gas evolution ceased. The

solution was allowed to warm up to ambient temperature and was allowed to stir for 10 min. The resulting mixture was poured into a 1:1 saturated aqueous solution of NH_4Cl/1N HCl solution (20 mL) and was stirred until a clear separation of layers was observed. The aqueous layer was extracted with Et_2O (3 x 10 mL) and the combined organic layers were washed with water (2 x 10 mL), brine (3 x 10 mL), dried over $MgSO_4$ and concentrated *in vacuo*. The residual yellow oil was purified by flash column chromatography (30% Et_2O in pentane) which yielded 181 mg of a yellow oil (63% yield over 3 steps)

The spectral data correspond to those previously reported**Error! Bookmark not defined.**



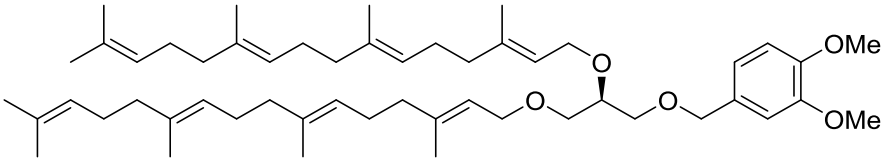
(2*E*,6*E*,10*E*)-1-chloro-3,7,11,15-tetramethylhexadeca-2,6,10,14-tetraeneError! Bookmark not defined. **(11).** *N*-chlorosuccinimide (145 mg, 1.1 mmol, 1.3 equiv.) was suspended in dry CH_2Cl_2 (2.5 mL) under N_2 atmosphere in a pre-dried Schlenk flask. The turbid mixture was cooled to -30 ˚C (acetone/liquid N_2 bath) after which dimethyl sulfide (90 µL, 1.25 mmol, 1.5 equiv.) was added dropwise. The reaction mixture was stirred for 10 min at -30 ˚C after which it was allowed to warm up to 0 ˚C for 10 min. The resulting solution was cooled to -40 ˚C and geranylgeraniol (**10**, 240 mg, 0.83 mmol) in dry CH_2Cl_2 (1 mL) was added dropwise over 15 min. The resulting suspension was allowed to warm up over 1 h to -20 ˚C after which the acetone/liquid N_2 bath was replaced by an ice/water bath and the resulting suspension was stirred for another h at 0 ˚C after which it was poured into pentane (20 mL). The pentane mixture was decanted from the white precipitate, the white crystals were washed with pentane (20 mL) and decanted again, this process was repeated three times. The pentane extracts were combined and evaporated obtaining geranylgeranyl chloride as a yellow oil quantitatively which was used immediately for the dialkylation of (*S*)-3-((3,4-dimethoxybenzyl)oxy)propane-1,2-diol.



1-(3,4-dimethoxybenzyl)-2,3-bisgeranylgeranyl-*sn*-glycerolError! Bookmark not defined. **(12).** An oven dried Schlenk flask equipped with magnetic stirrer bar was charged with NaH (60% dispersion in mineral oil, 33 mg, 825 μmol, 2 equiv). The mineral oil was removed by three successive washings with pentane and then dried under vacuum. The resulting white solid was suspended in DMSO (0.6 mL) and stirred at 60 ˚C (oil bath) until a clear solution was observed (45 min). The pale yellow solution was allowed to cool down to rt after which (*S*)-3-((3,4-dimethoxybenzyl)oxy)propane-1,2-diol **5**, (83 mg, 340 μmol) in dry DMSO (freshly distilled from CaH, 0.6 mL) was added in a dropwise manner. After the solution was stirred for 1.5 h at rt, the crude mixture of geranylgeranyl chloride **11**, (250 mg, 2.2 equiv) in dry DMSO (0.2 mL) was added dropwise over 20 min. The resulting reaction mixture was stirred for 16 h after which it was poured into a saturated aqueous solution of NH_4Cl (5 mL). The aqueous layer was extracted with diethyl ether (3 x 5 mL). The combined organic layers were washed with brine, dried and concentrated *in vacuo*. The resulting crude was further purified by flash column chromatography (20 % diethyl ether in pentane) to afford the dialkylated product (31% yield over 2 steps, 41 mg).

HRMS-ESI+ (*m/z*): $[M + Na]^+$ calculated for $C_{52}H_{82}O_5Na$, 809.6055; found, 809.6041.

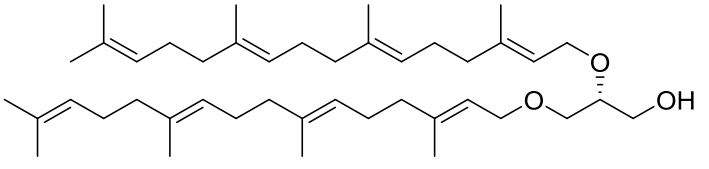
The spectral data correspond to those previously reported.[186]



1-(3,4-dimethoxybenzyl)-2,3-bisgeranylgeranyl-*sn*3-glycerol (16). This compound was prepared with the same synthetic procedure that was used for **(1-(3,4-dimethoxybenzyl)-2,3-bisgeranylgeranyl-*sn*-glycerol**, **(12).**

^{1}H NMR (400 MHz, Chloroform-*d*): Same as reported for **12**

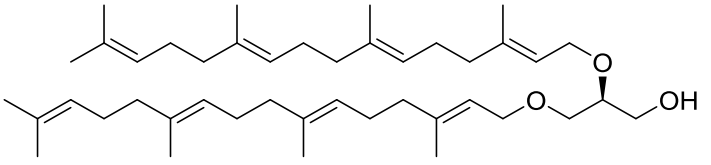
^{13}C NMR (101 MHz Chloroform-*d*): Same as reported for **12**



2,3-bisgeranylgeranyl-*sn*-glycerol (13). 1-(3,4-dimethoxybenzyl)-2,3-bisgeranylgeranyl-*sn*-glycerol (**12**) (35 mg, 45 μmol) was dissolved in CH_2Cl_2 (0.9 mL) to which water (0.1 mL) was added in order to form a biphasic system. The solution was cooled to 0 ˚C (ice/water bath) and DDQ (12.5 mg, 55 μmol, 1.2 equiv) was added in portions. The resulting green mixture was stirred for 2 h at 0 ˚C during which it turned light brown. The entire mixture was filtered over a small silica pad with CH_2Cl_2 and purified by flash column chromatography (20% diethyl ether in pentane, I_2 and 2,4-DNP stain) which afforded the desired product as a viscous yellow oil (70% yield, 20 mg).

HRMS-ESI+ (*m/z*): $[M + Na]^+$ calculated for $C_{43}H_{72}O_3Na$, 659.5374; found, 659.5369.

The spectral data correspond to those previously reported.**Error! Bookmark not defined.**



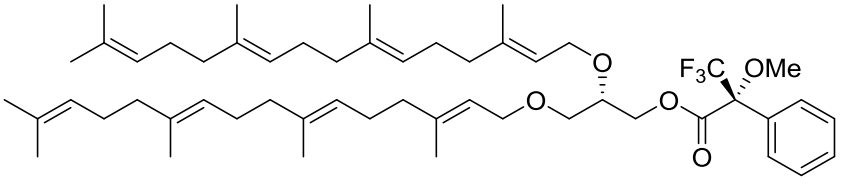
2,3-bisgeranylgeranyl-*sn*3-glycerol (17). This compound was prepared with the same synthetic procedure that was used for **2,3-bisgeranylgeranyl-*sn*-glycerol 13.**

^{1}H NMR (400 MHz, Chloroform-*d*): Same as reported for **13**

HRMS-ESI+ (*m/z*): $[M + Na]^+$ calculated for $C_{43}H_{72}O_3Na$, 659.5374; found, 659.5371.

Saponification of ester-lipids from natural lipid extract mixture. 480 mg NaOH was added to a mixture of MeOH (0.6 mL) and CH_2Cl_2 (4.4 mL) and stirred until all NaOH was dissolved. To this resulting stirred solution was added the natural lipid extract (40 mg) from the modified *E. coli* cultures in CH_2Cl_2 (1 mL) such that the resulting mixture had a 1:9 MeOH/CH_2Cl_2 ratio and a 2 N NaOH concentration. The solution was allowed to stir for 80 h at rt during which it became turbid. The mixture was filtered over celite and the residue was washed with CH_2Cl_2. The filtrate was neutralized with 1 M aqueous HCl, transferred to a separatory funnel and separated. The aqueous layer was extracted with CH_2Cl_2 (3 x 5 mL) and the combined organic layers were washed with brine (2 x 5mL), dried over $MgSO_4$ and concentrated *in vacuo*. The obtained residue was further purified by column chromatography (20% Et_2O in pentane) in order to yield 7 mg of the natural di-ether glycerol lipid (**20**) as a yellow oil with some co-eluted impurities.

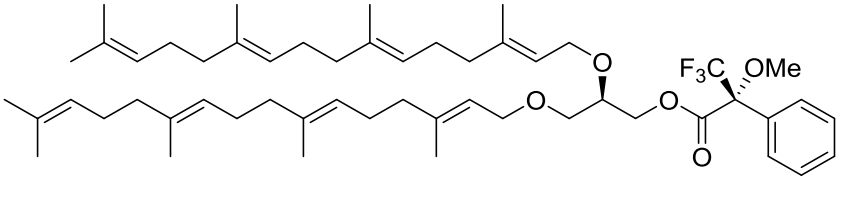
HRMS-ESI+ (*m/z*): $[M + Na]^+$ calculated for $C_{43}H_{72}O_3Na$, 659.5374; found, 659.5371.



2,3-bisgeranylgeranyl-*sn*-glycerol (+)-Mosher's ester derivative (14). Alcohol **13** (5 mg, 7.8 µmol) was dissolved in pyridine (0.2 mL) under N_2 atmosphere. (*S*)-(+)-MTPA-Cl (6 mg, 5 µL, 23.4 µmol, 3 equiv, Mosher's chloride) was added dropwise and the reaction mixture was stirred for 2 h at rt before it was quenched with 0.2 mL saturated $NaHCO_3$ solution. A small quantity of water (1 mL) was added and the aqueous mixture was extracted with ethyl acetate (3 x 1 mL), the organic layers were combined, washed with brine (2 x 1 mL) dried over $MgSO_4$ and concentrated *in vacuo*. The residue was filtered over silica with pentane and the residue was washed with a large amount of pentane. The filtrate was concentrated *in vacuo* in order to yield a yellow crude mixture containing the desired Mosher's ester (6 mg).

^{1}H NMR (599 MHz, Chloroform-*d*) δ 7.57 – 7.50 (m, 2H), 7.42 – 7.37 (m, 3H), 5.42 – 5.23 (m, 1H), 5.16 – 5.05 (m, 6H), 4.69 (s, 1H), 4.51 (dd, J = 11.5, 3.5 Hz, 1H), 4.37 – 4.32 (m, 1H), 4.09 – 4.05 (m, 2H), 4.00 – 3.96 (m, 2H), 3.74 (m, 1H), 3.55 (s, 3H), 3.49 (m, J = 10 Hz, 1H), 3.42 (m, J = 10.0 Hz, 1H), 2.11 – 1.94 (m, 24H), 1.68 (s, 6H), 1.60 (s, 18H), 1.55 (s, 6H).

HRMS-ESI+ (*m/z*): $[M + Na]^+$ calculated for $C_{53}H_{79}F_3O_5Na$, 875.5772; found, 875.5772.



2,3-bisgeranylgeranyl-*sn*3-glycerol (+)-Mosher's ester derivative (18). This compound was prepared with the synthetic route, identical to that used for **2,3-bisgeranylgeranyl-*sn*-glycerol Mosher's ester derivative** (**14**)

^{1}H NMR (599 MHz, Chloroform-*d*) δ 7.57 – 7.51 (m, 2H), 7.41 – 7.36 (m, 3H), 5.29 (m, 1H), 5.18 – 5.07 (m, 6H), 4.71 – 4.65 (m, 1H), 4.56 (dd, J = 11.5, 3.5 Hz, 1H), 4.37 – 4.32 (m, 1H), 4.07 (d, J = 6.5 Hz, 2H), 3.96 (d, J = 6.5 Hz, 2H), 3.75 – 3.71 (m, 1H), 3.56 (s, 3H), 3.49 (m, J = 10 Hz, 1H), 3.42 (m, J = 10 Hz, 1H), 2.12 – 1.94 (m, 24H), 1.64 (s, 6H), 1.59 (s, 18H), 1.56 (s, 6H).

HRMS-ESI+ (*m/z*): $[M + Na]^+$ calculated for $C_{53}H_{79}F_3O_5Na$, 875.5772; found, 875.5773.

Naturally derived di-ether lipid (+)-Mosher's ester derivative (19). This compound was prepared with the synthetic route, identical to that used for **2,3-bisgeranylgeranyl-*sn*-glycerol Mosher's ester derivative** (**14**)

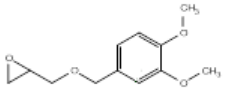
^{1}H NMR (599 MHz, Chloroform-*d*): Same as reported for **14**

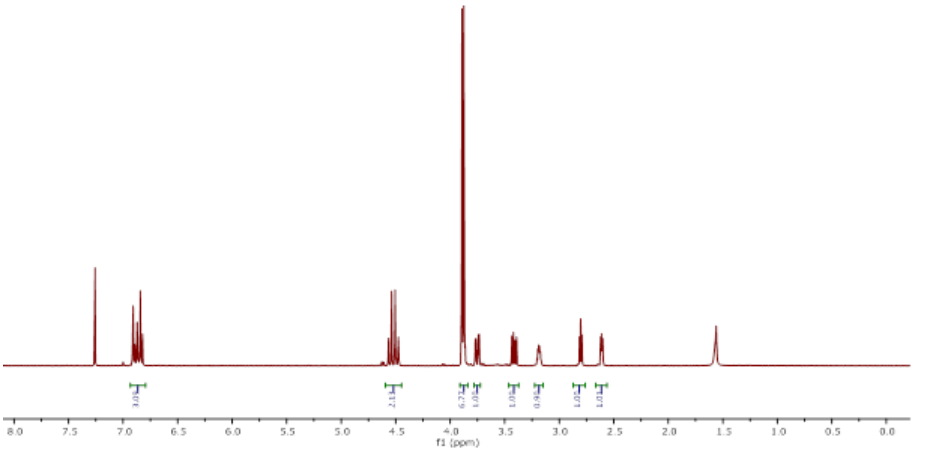
HRMS-ESI+ (*m/z*): $[M + Na]^+$ calculated for $C_{53}H_{79}F_3O_5Na$, 875.5772; found, 875.5773.

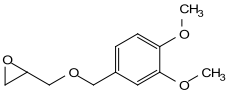
Naturally derived di-ether lipid (+)-Mosher's ester derivative from G1P deficient strain (20). This compound was prepared with the synthetic route, identical to that used for **2,3-bisgeranylgeranyl-*sn*-glycerol Mosher's ester derivative** (**14**)

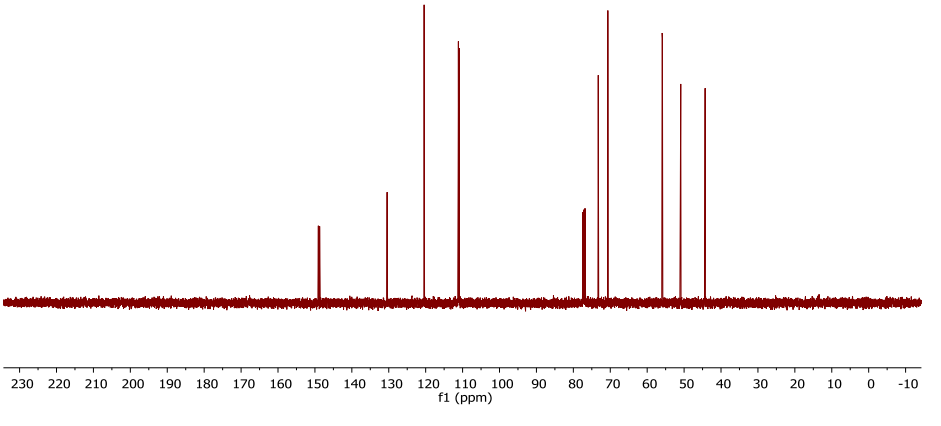
^{1}H NMR (599 MHz, Chloroform-*d*): Same as reported for **14**

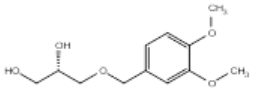
HRMS-ESI + (*m/z*): $[M + Na]^+$ calculated for $C_{53}H_{79}F_3O_5Na$, 875.5772; found, 875.5773.

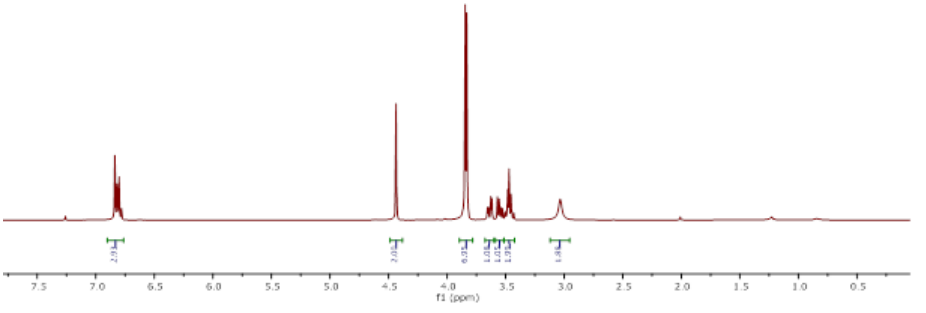


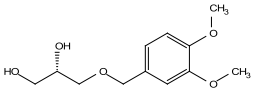

8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
f1 (ppm)

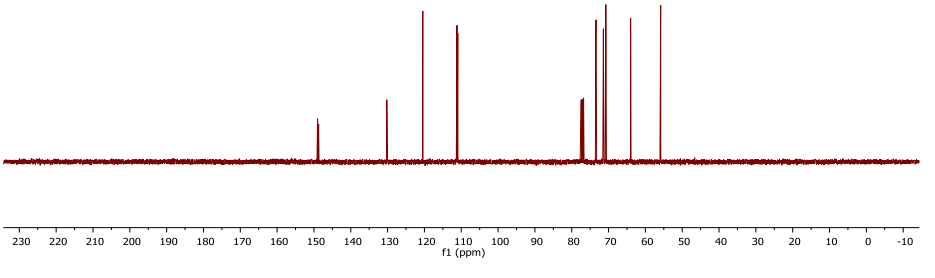



230
220
210
200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
-10
f1 (ppm)





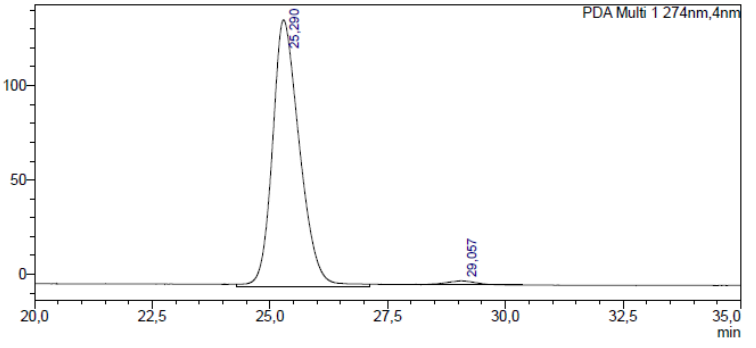




(*S*)-enantiomer:

<Chromatogram>

mAU



<Peak Table>

PDA Ch1 274nm

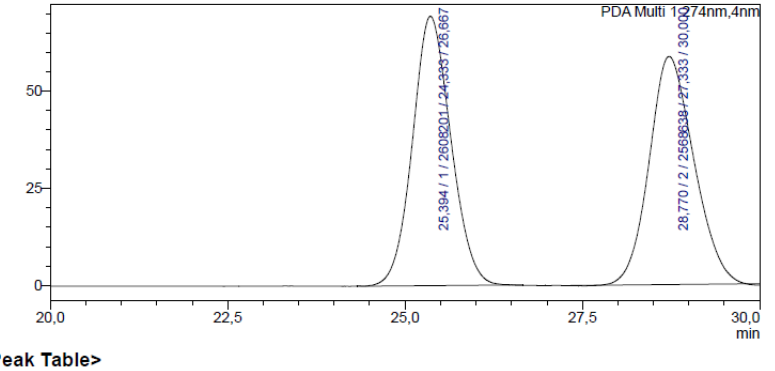
Peak#	Ret. Time	Area	Height	Conc.	Unit	Mark	Name
1	25,290	5674807	141288	98,654		M	
2	29,057	77410	1956	1,346		M	
Total		5752217	143244				

Peak Start	Peak End	Area%
24,280	27,112	98,654
28,264	30,368	1,346
		100,000

racemic mixture:

<Chromatogram>

mAU

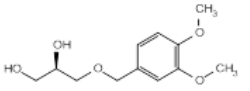


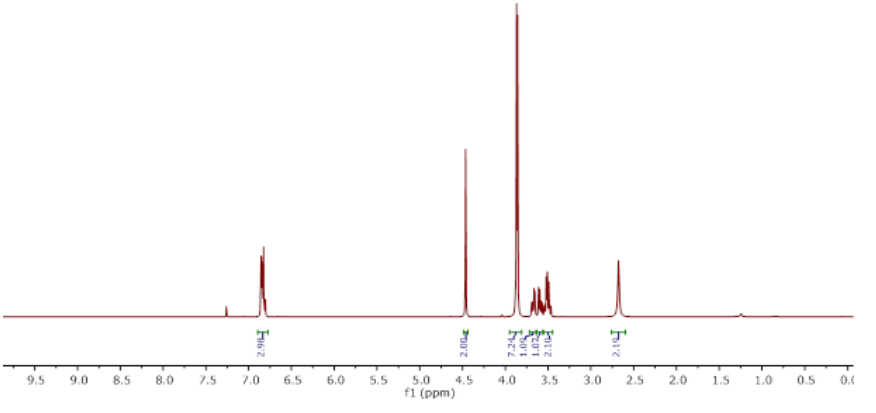
<Peak Table>

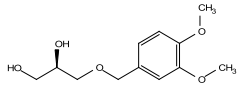
PDA Ch1 274nm

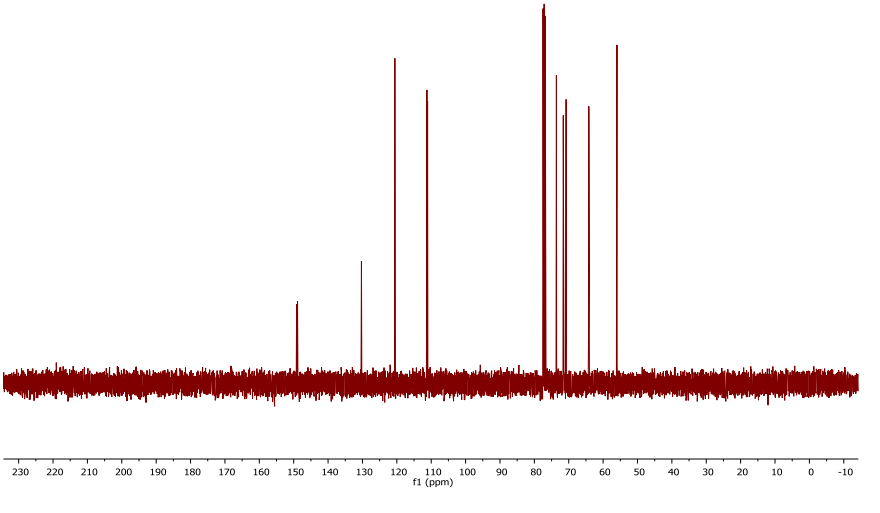
Peak#	Ret. Time	Area	Height	Conc.	Unit	Mark	Name
1	25,394	2608201	56409	50,382		M	
2	28,770	2568638	52337	49,618		M	
Total		5176839	108746				

Area%	Peak Start	Peak End
50,382	24,333	26,667
49,618	27,333	30,000
100,000		



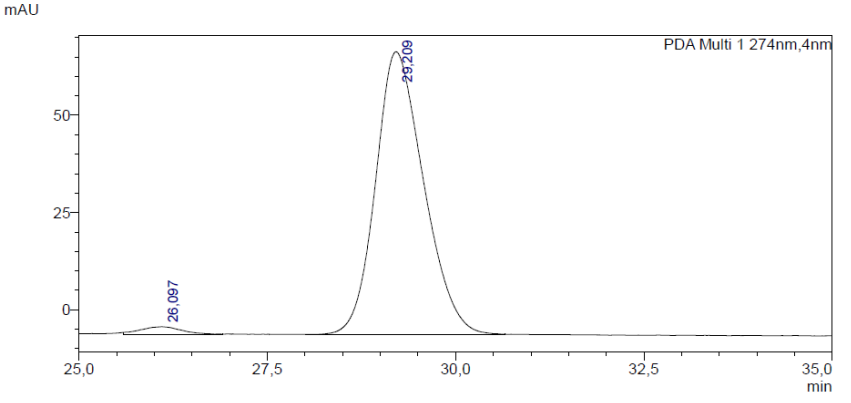






(*R*)-enantiomer:

<Chromatogram>



<Peak Table>

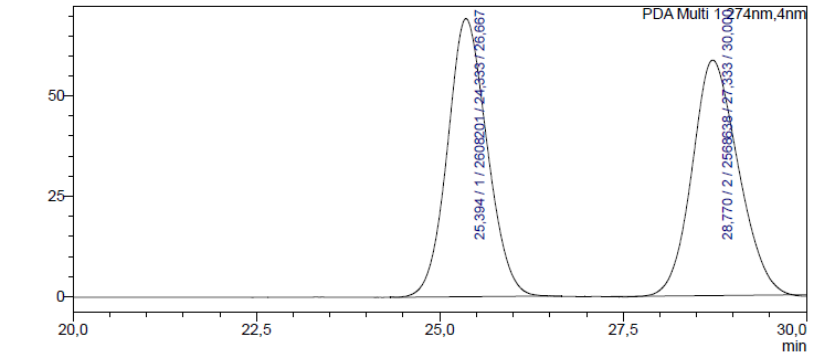
PDA Ch1 274nm

Peak#	Ret. Time	Area	Height	Conc.	Unit	Mark	Name
1	26,097	81222	2052	2,445		M	
2	29,209	3240443	72698	97,555		M	
Total		3321665	74750				

Area%	Peak Start	Peak End
2,445	25,592	26,904
97,555	28,008	30,656
100,000		

Racemic mixture:

<Chromatogram>

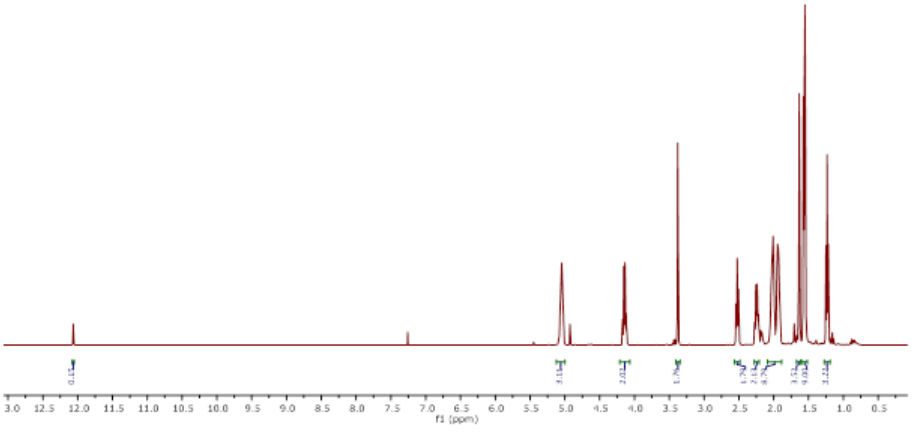


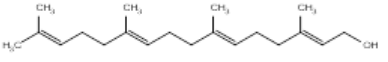
<Peak Table>

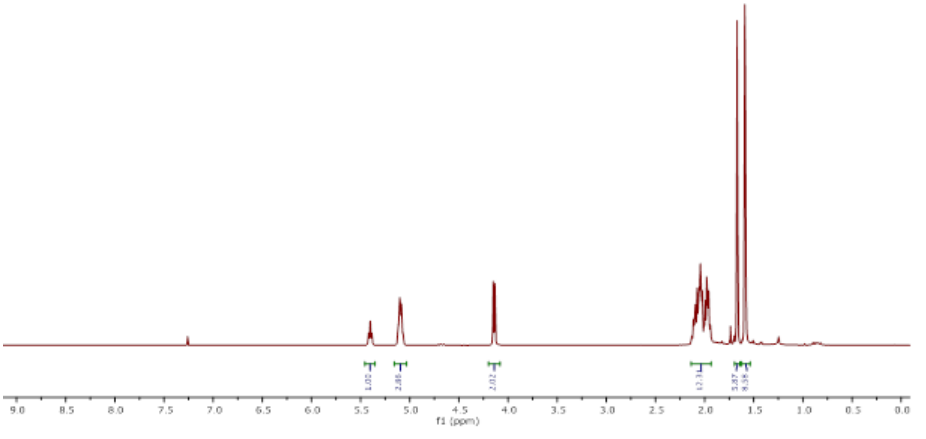
PDA Ch1 274nm

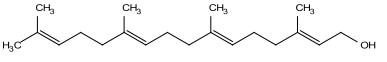
Peak#	Ret. Time	Area	Height	Conc.	Unit	Mark	Name
1	25,394	2608201	56409	50,382		M	
2	28,770	2568638	52337	49,618		M	
Total		5176839	108746				

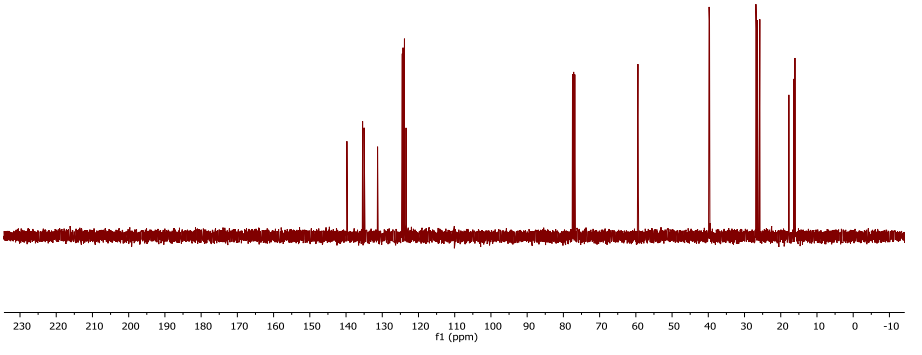
Area%	Peak Start	Peak End
50,382	24,333	26,667
49,618	27,333	30,000
100,000		

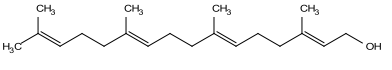


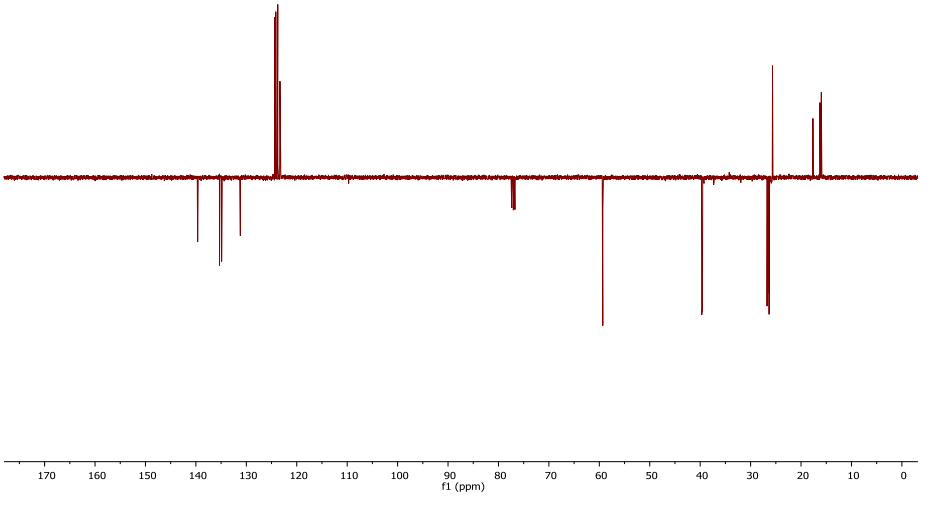


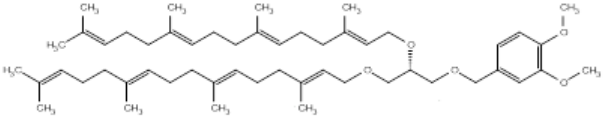


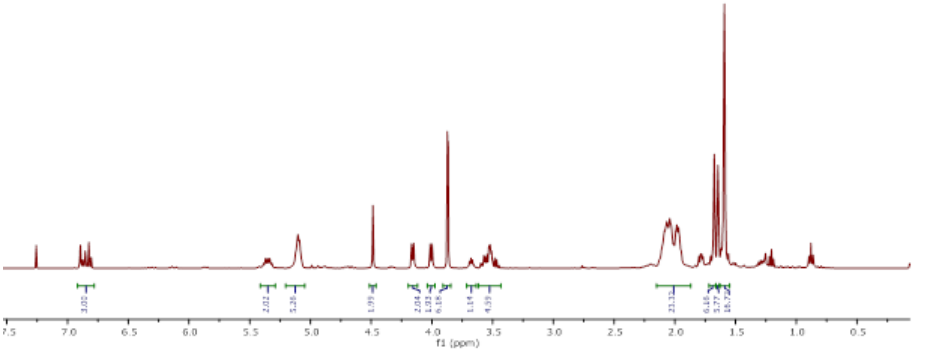

CH3
CH3
CH3
CH3
H3C
OH

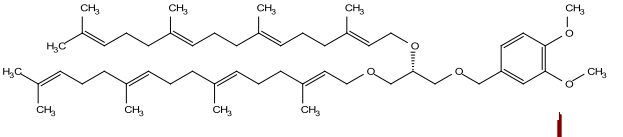

230 220 210 200 190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0 -10
f1 (ppm)

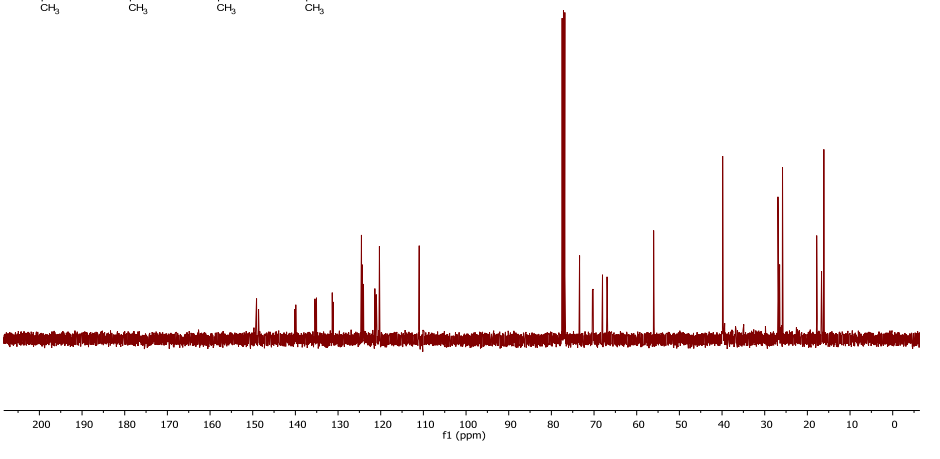

CH3
CH3
CH3
CH3
H3C
OH

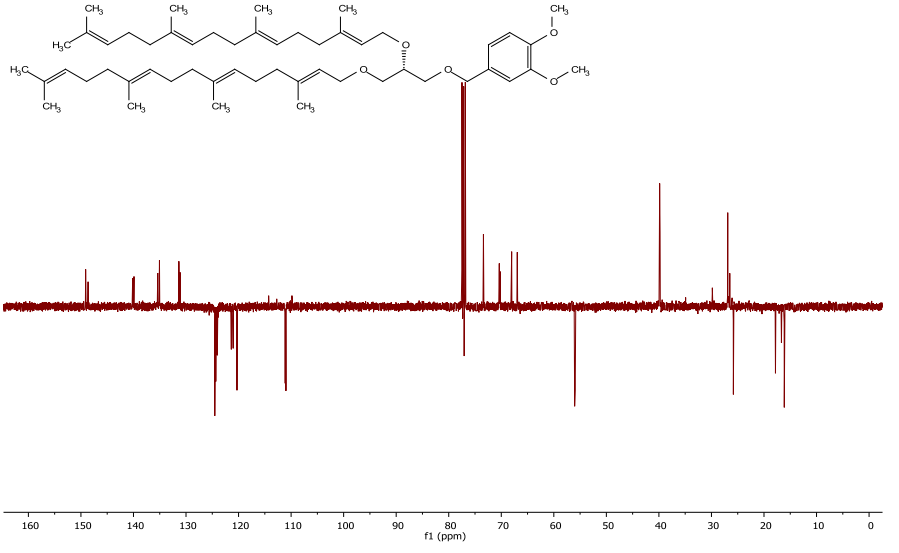

170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0
f1 (ppm)

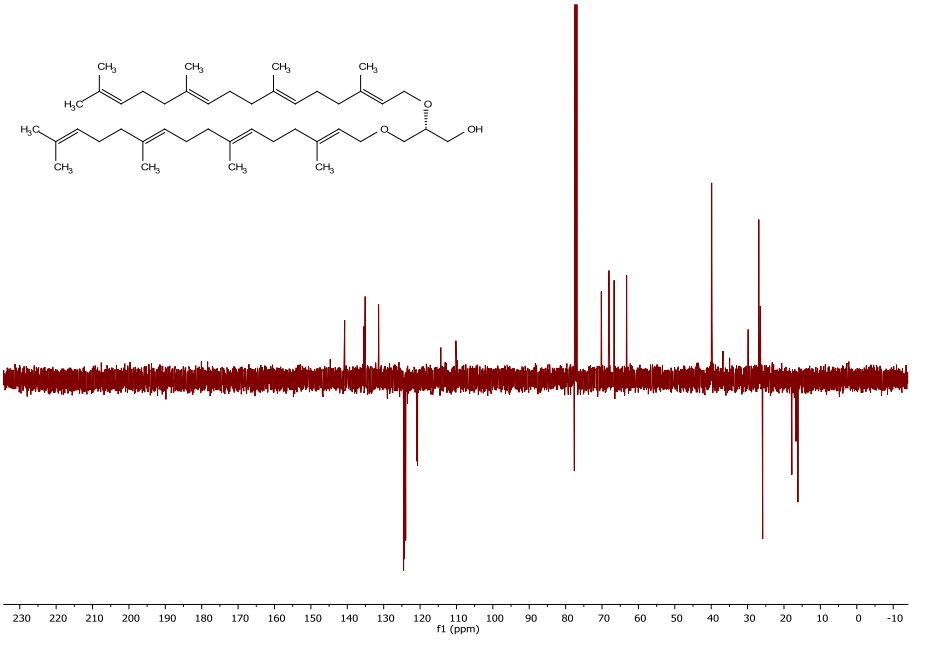


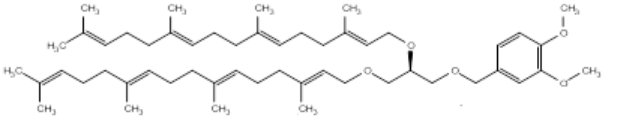

7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
f1 (ppm)

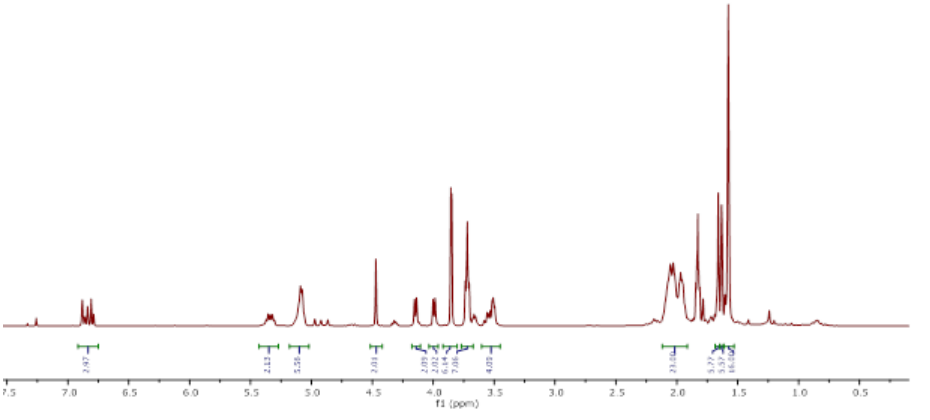


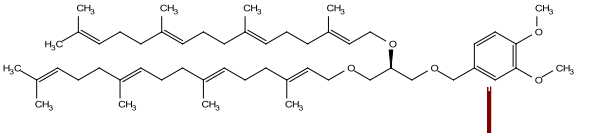

200
190
180
170
160
150
140
130
120
110
100
90
80
70
60
50
40
30
20
10
0
f1 (ppm)

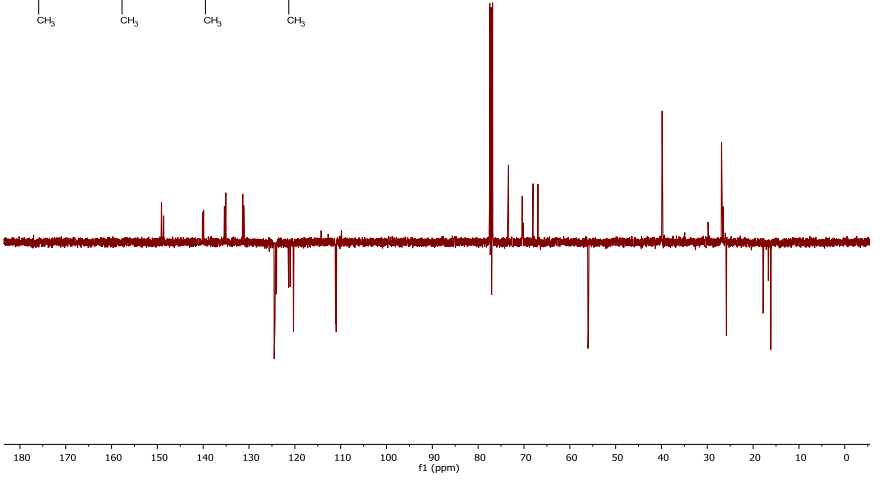

160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0
f1 (ppm)

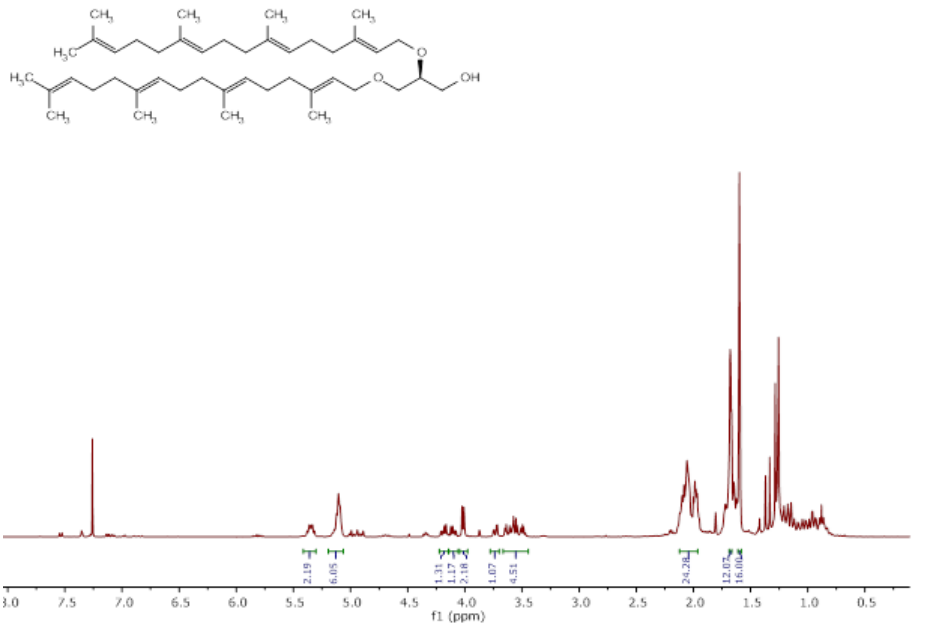

230 220 210 200 190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0 -10
f1 (ppm)



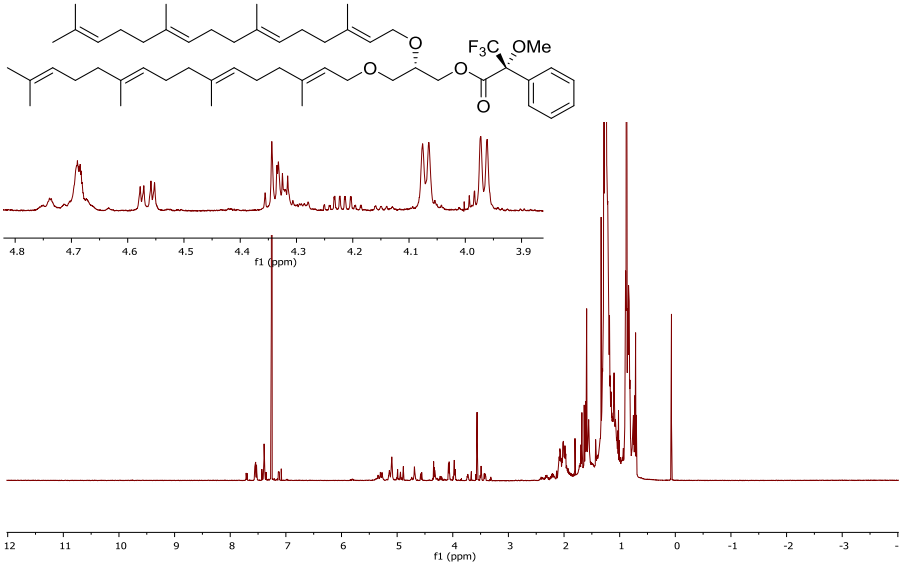

f1 (ppm)



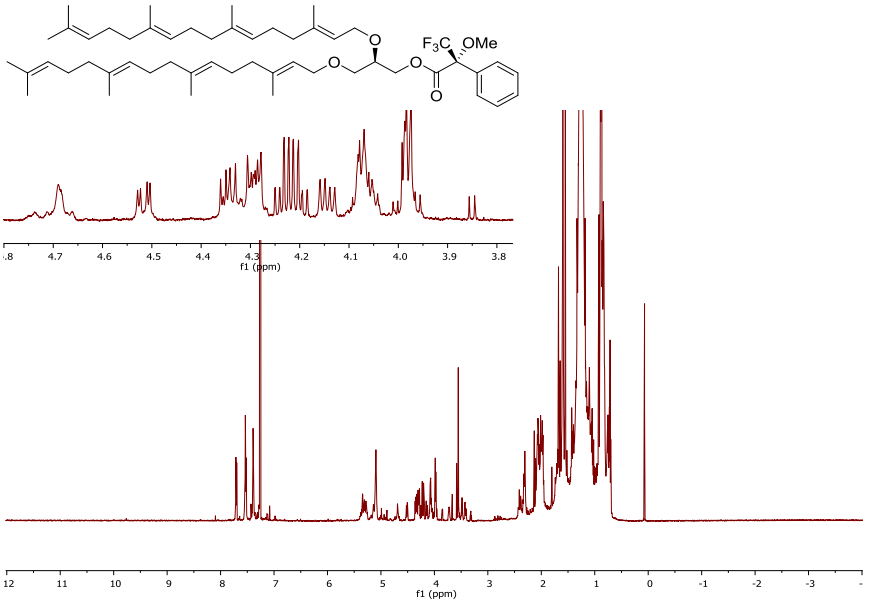

f1 (ppm)



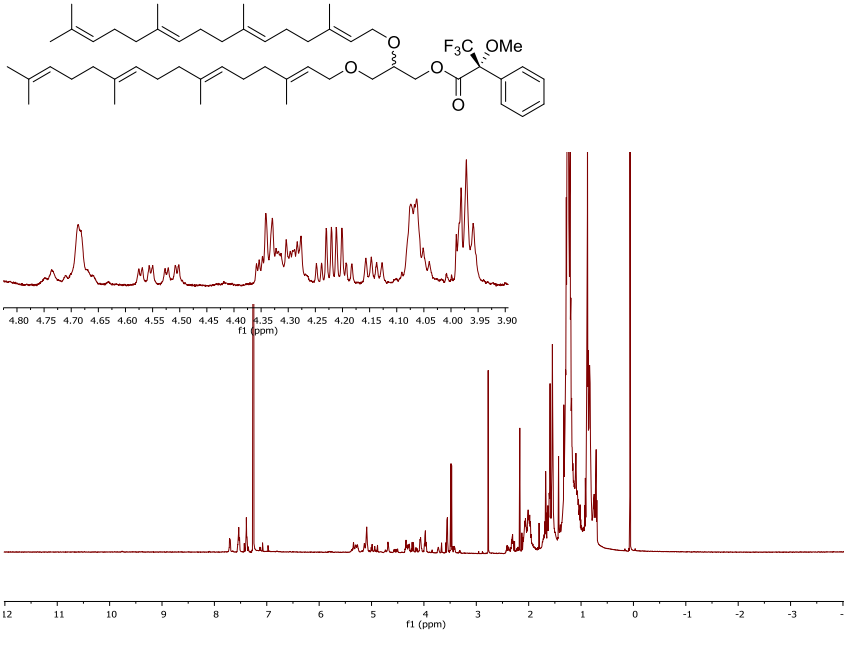
Synthetic *sn*-1 Mosher's ester lipid (**14**)



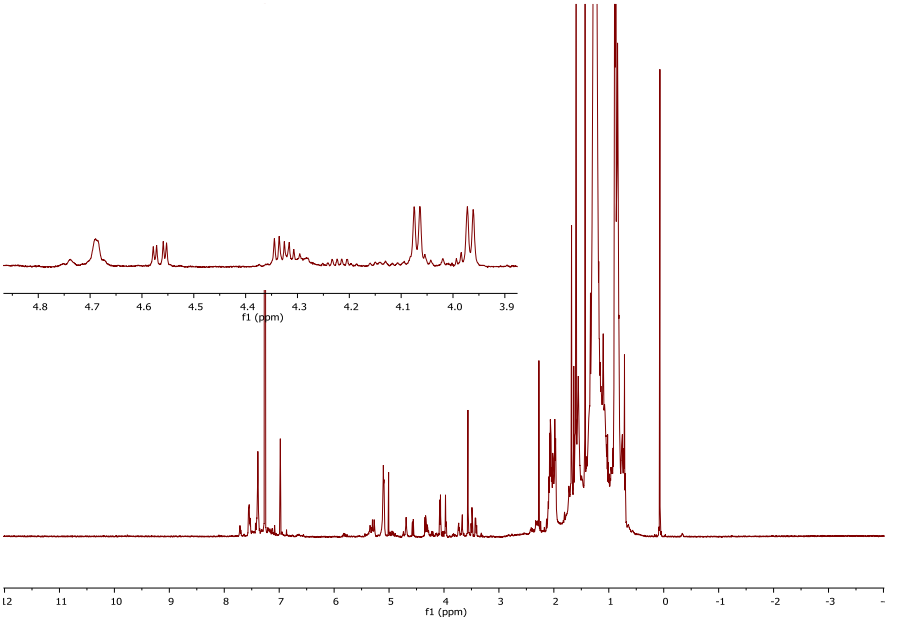
Synthetic *sn*-3 Mosher's ester lipid (**18**)



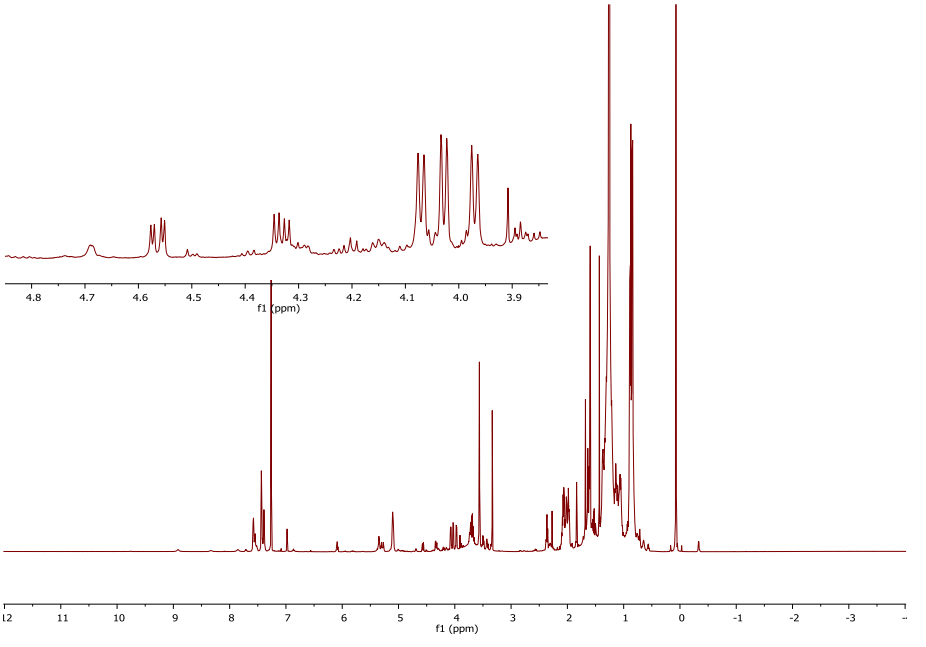
Mixture of *sn*-1 and *sn*-3 Mosher's ester lipids



Naturally derived Mosher's ester lipid (**19**)



Naturally derived Mosher's ester lipid (G1P-deficient strain) (**20**)



Chapter 5

Structural insights into lipid biosynthesis in archaeal membranes

Manuscript in preparation

Sixue Ren[1#], Bo Sun[2#], Antonella Caforio[3,4#], Qin Yang[4#], Feng Yu[2], Xiaofeng Zhu[1] , Yao Guo[1] , Jian Jiang[1], Jinjing Wang[1], Chao Dou[1] , Jianhua He[2], Arnold. J.M. Driessen[3,4] and Wei Cheng[1]

[1]Division of Respiratory and Critical Care Medicine, State Key Laboratory of Biotherapy, 7 West China Hospital of Sichuan University and Collaborative Innovation Center of 8 Biotherapy, Chengdu, 610041, China

[2]Department of Life Science, Shanghai Institute of Applied Physics, Chinese Academy of 10 Sciences, Shanghai, China

[3]Department of Molecular Microbiology, Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, 9747 AG Groningen, The Netherlands;

[4]The Zernike Institute for Advanced Materials, University of Groningen, 9747 AG Groningen, The Netherlands

[#]These authors contributed equally to this work

Abstract

Lipids play vital roles in cell membranes are involved in human metabolic diseases and associated with many fundamental cellular processes. The divergence of archaea, bacteria and eukaryotes is a fundamental step in evolution and reflected by major differences in membrane lipids chemistry. Archaeal lipids are characterized by ether linked isoprenoid chains while bacterial and eukaryotic lipids have ester linked fatty acid chains. The mechanisms underlying the biosynthesis of the archaeal lipids still remain elusive. Here, we report an essential enzyme structure of CDP-archaeol synthase CarS of *Aeropyrum pernix* (ApCarS) in a CTP- and Mg^{2+} -bound state, at a resolution of 2.38 angstroms (Å). The enzyme comprises a distinctive fold of five transmembrane helices and extensive cytoplasmic loops that form a large charged cavity. Specific recognition of CTP binding site reveals the enzymatic catalytic mechanismof CarS. The hydroxyl group on the phosphorous atom of the substrate DGGGP activates the α-phosphate of CTP upon Mg^{2+} interaction, triggering the realease of pyrophosphate group, stabilized then by a network of positively charged residues. Archaeol binds in unique hydrophobic membrane-embedded grooves that are formed by the structurally flexible transmembrane helix 5 (TM5) together with TM1 and TM4. Overall, the reported structural and fuctional studies not only reveal the structural basis for the general catalytic mechanism of CTP-transferase superfamily, but also provide insights into the origin of evolution of life.

Introduction

Phospholipids are key components of the cell membranes of all living organisms. They play vital roles in the formation and stabilization of the lipid bilayer, maintain the permeability and fluidity of the barrier [1,2,127], and provide an essential compartment for biological activity, particularly for lipid and membrane protein biogenesis, transport, and energy transduction [187–189]. In addition, many types of phospholipids play important regulatory roles in cell signaling, membrane trafficking, apoptosis, and immunity [190–192].

A key step in membrane phospholipid production is catalyzed exclusively by transmembrane enzymes of the CTP-transferase superfamily, which catalyze the conjugation of CDP to a glycerol-phosphate backbone, resulting in the formation of CDP-diacylglycerol (in bacteria) or CDP-archaeol (in archaea) [2,64,127]. Archaea have evolved to produce a variety of unique lipids comprised of a glycerol-1-phosphate (G1P, not G3P) backbone linked to isoprenoid hydrocarbon side chains via an ether bond [3]. This striking feature distinguishes archaea from bacteria, and most early evolution hypotheses propose that archaea and bacteria diverged directly from a common cenancestor that may have had a mixed heterochiral membrane [4]. Since the associated "lipid divide" that occurred during the divergence of archaea and bacteria from the cenancestor is considered evolutionarily significant, an intriguing question is how the ether- and ester-based phospholipid biosynthesis pathways evolved in these respective organisms. Among archaea, members of the membrane-embedded CTP-transferase superfamily share considerable sequence similarity (**Supplementary Fig. 1**); however, these proteins are not well conserved in bacteria and eukaryotes [8] (**Supplementary Fig. 2**). Previous studies characterized representative Cds (CDP diacylglycerol synthase) proteins, known as integral membrane enzymes, of *Escherichia coli*, *Saccharomyces cerevisiae*, mice, and humans by their preference for activated CTP, deoxy-CTP (dCTP), or other nucleotides as polar head groups for phospholipid biosynthesis [100,101,112,193,194]. Meanwhile, the first archaeal CDP-archaeol synthase (CarS) was characterized in 2014. While CarS was shown to catalyze an essential step in CDP-archaeol

formation, transferring CTP to its specific archaeal lipid substrate [41] (**Fig. 1A**), the enzymatic mechanism of CTP-transferases remains poorly characterized. Structural and biochemical studies of CarS are therefore necessary to reveal how this intramembrane CTP transfer step is catalyzed in membrane bilayers.

To address this question, we generated a crystal structure of the CarS of *Aeropyrum pernix* (ApCarS) in the CTP- and Mg^{2+}-bound state, at a resolution of 2.38 angstroms (Å). The structure revealed an unusual fold within the intramembrane enzyme containing a cytoplasmic domain, a transmembrane domain, and a periplasmic region. Moreover, our data provide a structural basis for the mechanism of CTP-recognition, as well as insights into the catalytic mechanism of this CTP-transferase family member. Lastly, structural comparisons of ApCarS and related dual-substrate enzymes suggest that the catalytic activities of CarS family members in archaea are determined by the binding of specific lipophilic substrates.

Results

ApCarS is a CTP-transferase

To facilitate the structural analysis of CDP-archaeol synthase (CarS), we purified and attempted to crystalize CarS proteins from various archaeal species. In doing so, we obtained x-ray-diffracting crystals from the *Aeropyrum pernix k1* CarS protein (ApCarS) after detergent trials. ApCarS shares 37% sequence identity with the functionally characterized CarS of *Archaeoglobus fulgidus* (AfCarS) [41]. We therefore compared the CTP-binding activity of ApCarS and AfCarS in the presence or absence of magnesium, which was previously shown to be essential for optimal enzymatic activity of AfCarS [110], by isothermal titration calorimetry (ITC). Purified ApCarS protein interacted strongly with CTP, with dissociation constants of 0.2 μM and 1.67 μM in the presence (**Figure 1B**) and absence (**Figure 1C**) of magnesium, respectively. To further test whether ApCarS exhibits CTP-transferase activity, we performed *in vitro* catalytic assays. Herein, purified ApCarS was incubated with CTP and 2,3-

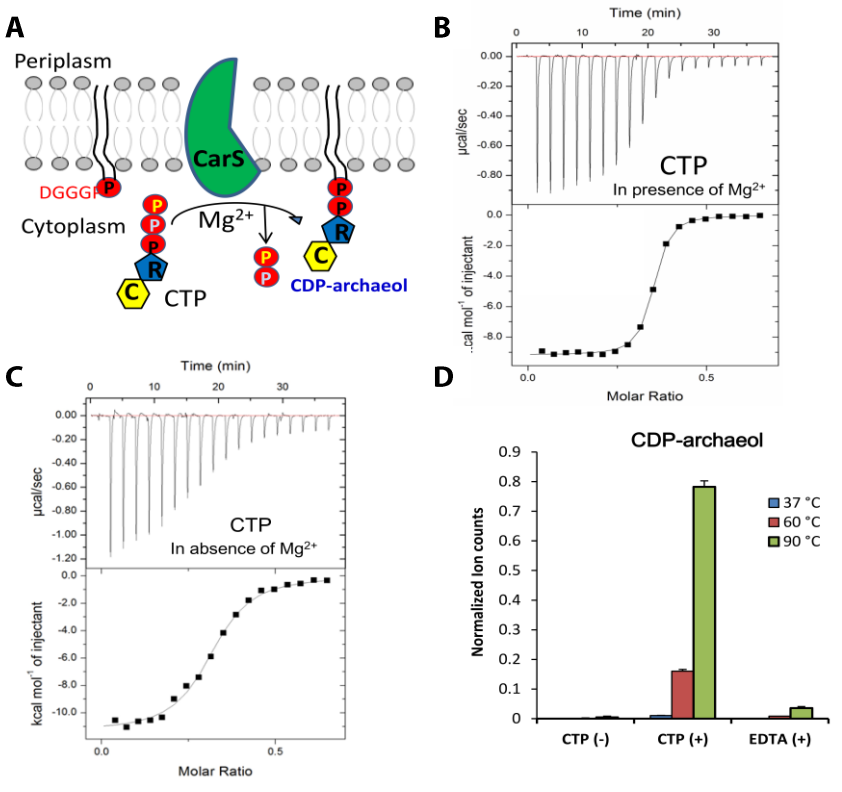


Figure 1| ApCarS is a CTP-transferase. (**A**) Cartoon representation of the reaction catalyzed by ApCarS, generating CDP-archaeol from CTP and DGGGP (2,3-bis-O-geranylgeranyl sn-glycerol-phosphate). The C-labeled hexagon and R-labeled pentagon represent the cytosine and ribose portions of CTP, respectively. Meanwhile, the phosphate groups are depicted as red circles, and ApCarS is colored in green. (**B**) Isothermal titration calorimetry (ITC) raw data and binding trace for CTP titrated into a solution containing ApCarS and Mg^{2+}: Kd = 0.20 µM; ΔH = -9.180 kcal mol^{-1}; ΔS = -2.18 cal mol^{-1}. The data shown are representative of three independent experiments. See details in **Supplementary Figure 7** and **Supplementary Table 2**. (**C**) Representative ITC raw data and binding trace for CTP titrated into a solution containing ApCarS inthe absence of Mg^{2+}: Kd = 1.67 µM; ΔH = -1.134 kcal mol^{-1}; ΔS = -11.6 cal mol^{-1}. The data shown are representative of three independent experiments. See details in **Supplementary Figure 7** and **Supplementary Table 2**. (**D**) Graphic depiction of ApCarS activity. The activity of purified ApCarS was assessed with chemically synthesized DGGGP as a substrate at different temperatures and in the absence or presence of 2 mM CTP and in the presence of Mg^{2+} and EDTA. The production of CDP-archaeol (m/z = 1,020.55 $[M-H]^-$) was verified by high performance liquid chromatography-mass spectrometry (HPLC-MS) analysis (**Supplementary Figure 3A,B**).

bis-O-geranylgeranyl sn-glycerol-phosphate (DGGGP) at 37 °C for 1 hr, and the resulting product, CDP-archaeol, was monitored by liquid chromatography-mass spectrometry (LC-MS; **Figure 1B and Supplementary Figure 3A,B**). As anticipated, CDP-archaeol was detected in the presence, but not in the absence, of CTP. Moreover, similar to AfCarS, ApCarS activity was Mg^{2+}-dependent, as only low levels of CDP-archaeol were observed in mixtures containing EDTA (**Figure 1D**). Interestingly, ApCarS-mediated production of CDP-archaeol was markedly enhanced at increased temperatures, with the most robust catalytic activity being observed at 90 °C. These results are consistent with the fact that *Aeropyrum pernix k1* is a thermophilic archaeon that grows at extreme temperatures. Together, our data indicate that ApCarS is a functional homolog of AfCarS.

Characterization of the structure of CTP-bound ApCarS

To understand the structural mechanism underlying the production of CDP-archaeol by ApCarS, we crystallized the CTP-ApCarS complex using the lipid cubic phase method (LCP), and determined its structure at a resolution of 2.38 Å (**Supplementary Table 1**). ApCarS is primarily composed of five transmembrane helices (TMs), with a large charged cavity at the cytoplasmic face (**Figure 2A,B**). One surface of the cavity is formed by TM1, TM2, TM3, and TM4, and is loosely occluded by TM5. The remaining portion of the cavity is formed by two cytoplasmic loops (CLs): CL1 and CL2, which connect TM1 to TM2 and TM3 to TM4, respectively (**Figure 2B**). Collectively, CL1 and CL2 comprise the cytoplasmic domain (CPD), which caps the transmembrane domain (TMD) composed of the five TMs. Moreover, CL1, which is longer than CL2, packs against one side of TM3, enabling extensive interactions within the loop (**Supplementary Figure 4A**).

As defined by electron density, CTP appears to bind to one side of the central cavity of ApCarS (**Figure 2B and Supplementary Figure 5A**), making tight contacts with the CPD, and therefore likely plays a role in the stabilization of this domain. Consistent with this conclusion, limited proteolysis assays indicated that purified ApCarS protein exhibited increased trypsin resistance in the presence of CTP (**Figure 2C**).

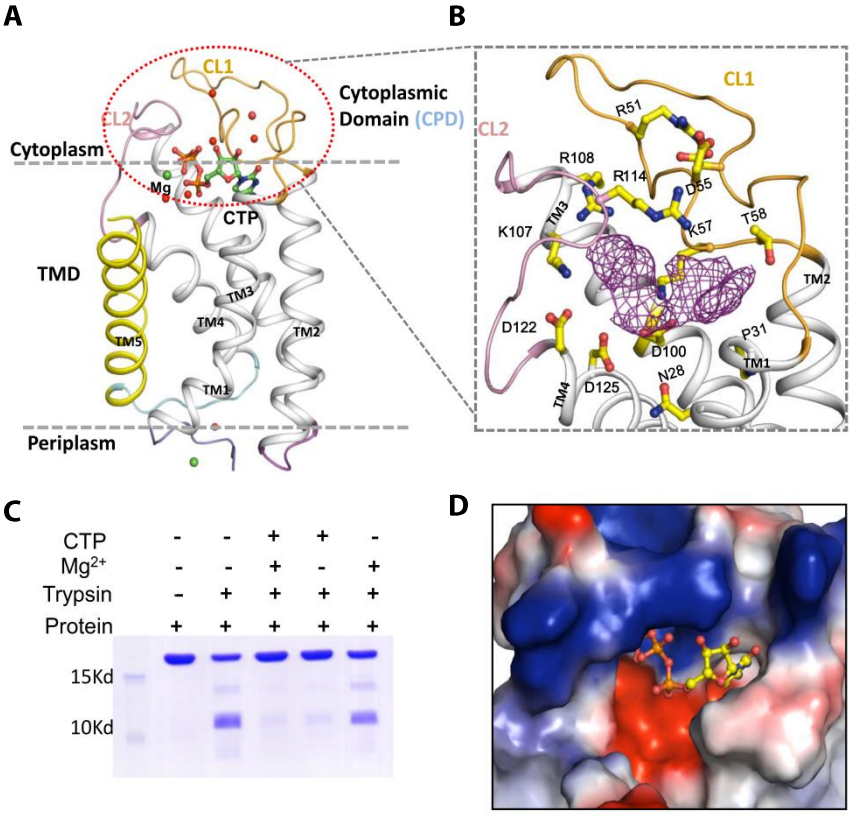


Figure 2| Structure of ApCarS-CTP complex. (**A**) Cartoon of the ApCarS-CTP complex along the membrane plane. The cytoplasmic domain (CPD) and the transmembrane domain (TMD) are indicated with different colors; TM5 is highlighted in yellow. (**B**) Stick diagram depicting the amino acid residues within the active site; cytoplasmic loops 1 (CL1) and 2 (CL2) are highlighted in orange and magenta, respectively. 2Fo – Fc electron density of CTP shown in pink mesh is contoured at 2.2σ. (**C**) Trypsin digestion of ApCarS in the presence of CTP and/or Mg^{2+}. Proteolytic fragments were visualized by SDS-PAGE and Coomassie staining. (**D**) Electrostatic surface representation of ApCarS bound to CTP. CTP is shown as a ball-and-stick representation within the active site. The large central cavity contains a polar pocket (red and blue), a hydrophobic pocket (gray) for binding to the phosphate groups, and the cytosine ring of CTP, respectively.

Meanwhile, ApCarS-bound CTP adopts a curved conformation, with the phosphate portion being nearly perpendicular to that of the nucleobase. Interestingly, half of the phosphate-binding pocket exhibits strong positive charges from the CPD, while the other half carries negative charges from the TMD. The β-and γ-phosphate groups of CTP are buried, whereas the remaining α-phosphate group is partially solvent-exposed. Compared to the phosphate groups, the nucleobase of CTP fits within a hydrophobic pocket located between the CPD and the TMD (**Figure 2D, Figure 6B, and Supplementary Figure 7B).**

Mechanism of CTP recognition

Recognition of CTP by ApCarS occurs through a combination of extensive polar and hydrophobic interactions. Specifically, the γ-phosphate group of CTP establishes salt bridges with Lys107 and Arg108 of ApCarS, as well as hydrogen bonds with the side chain of Ser104 and the amide nitrogen of Gly115. Additionally, this phosphate group forms water-mediated hydrogen bonds with the side chains of Asp100, Asp122, and Asp125. Notably, these interactions result in complete burial of this phosphate group (**Figure 3A**). By comparison, fewer interactions are generated between the other two phosphate groups and ApCarS: O1 of the β- phosphate group and O3’ of the sugar ring form a pair of hydrogen bonds with Arg114, whereas the α-phosphate group forms polar interactions with both Lys57 and Asp100 (**Figure 3B**). In addition to forming a hydrogen bond with Arg114, the sugar ring contributes to CTP binding by being sandwiched between Val32 and Lys57. Although Gly56 does not interact with CTP (separated by a minimum distance of~4.0 Å), this residue might play a role in CTP recognition through steric limitation. Meanwhile, the cytosine portion of CTP further inserts into the CTP-binding pocket, forming hydrogen bonds with Thr58 and packing against Pro31 and Lys57. An intramolecular hydrogen bond is thereby formed between O1 of the β-phosphate group and O3’ of the sugar ring of CTP, which likely plays a role in maintaining the curved conformation of the ApCarS-bound CTP (**Figure 3C**). Notably, each of these CTP-interacting residues is conserved among members of CDP-archaeol synthase (CarS) family, suggesting they employ a common mechanism for CTP recognition

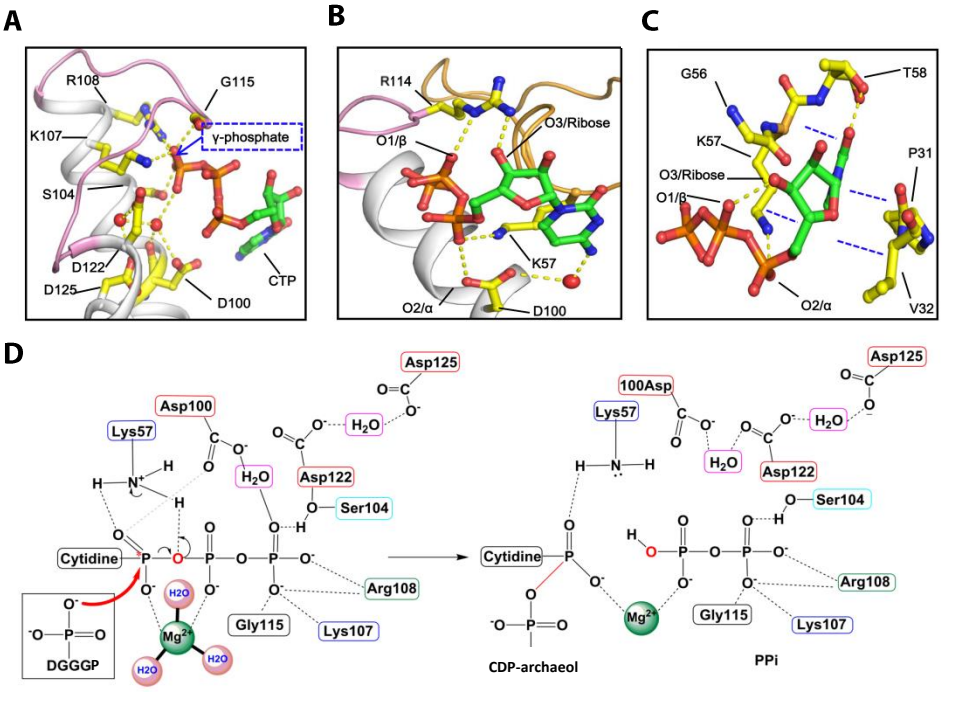


Figure 3| Recognition of CTP. (**A**) Network of interactions formed between the γ-phosphate group of CTP and residues within the polar pocket of the ApCarS active site. Amino acid residues are shown as sticks, water molecules are indicated by red balls, and CTP is shown as a green stick. (**B**) Residues K57, D100, and R114 participate in coordinating CTP within the active site via interactions with oxygen atoms from the α- and β-phosphate, ribose, and cytosine groups of CTP. (**C**) Interactions between CarS and the nucleotide portion of CTP. The blue dashed lines indicate hydrophobic packing. Hydrogen bonds are shown in yellow. CTP is shown as green stick. (**D**) A schematic represantion of the CTP recognition by CarS and proposed recognition by DGGGP at the active site. H bonds and salt bridges are marked by dotted lines. Red arrow refers to nucleophilic attack.

(**Supplementary Figure 1**). The elucidation of the CTP recongition mechanism by CarS allows prediction of the CTP interaction with the lipid substracte DGGGP. Herein, the phosphate group of DGGGP likely resides in the enzyme polar pocket where the hydroxyl activates the α-phosphate group of CTP, resulting in CDP-arhcaeol formation and in pyrophosphate realease. The latter is then stabilized by a network of positively charged residues (**Figure 3D**). Importantly, during the catalytic process, Mg^{2+}

interaction seems to be important for CTP stabilization and pyrophosphate realease. To this end, and to test this mechanism, site-directed mutagenesis and biochemical analysis were employed as descrbed in the following section.

Mg^{2+} is required for the enzymatic activity of ApCarS

Biochemical analyses demonstrated that Mg^{2+} significantly enhanced the CTP-binding and catalytic activities of ApCarS (**Figure 1B–D**). Notably, analysis of the crystal structure indicated that ApCarS contains an additional patch of electron density adjacent to the α- and β-phosphate groups (**Supplementary Figure 5B**). Modeling of Mg^{2+} into this electron dense region predicted coordination of the ion within 2.3 Å and 2.2 Å of the O1 and O2 atoms of the α-phosphate and β-phosphate groups, respectively. In addition to these two oxygen atoms, three water molecules are predicted to coordinate to the Mg^{2+} (**Figure 4A**). Notably, however, the Mg^{2+} ion does not interact directly with ApCarS. The important role of Mg^{2+} in the catalytic activity of ApCarS is reminiscent of the mechanism of Mg^{2+}-assisted transfer of the β- and γ-phosphate groups of ATP to substrates that is observed in kinases. In contrast, while Mg^{2+} coordinates to the β- and γ-phosphate groups of kinase-bound ATP molecules, this ion appears to coordinate to the α- and β-phosphate groups of ApCarS-bound CTP.

While the center of the periplasmic region encodes another electron-dense region, the density of this region is too strong for a water molecule. Modeling of Mg^{2+} into this region (**Figure 4B and Supplementary Figure 5C**) predicted that the ion might form coordination bonds with the carbonyl oxygen of Arg10 and the side chains of Asp8, Asp12, and Glu16. Additionally, two water molecules can be modeled into the density to facilitate the coordination of Mg^{2+} (**Supplementary Figure 5C**). This proposed Mg^{2+}-mediated interaction might function to stabilize the local conformation, allowing the N-terminal portion to fold back and interact with both TM4 and TM5; water could then mediate an interaction between Asn82, which is located in the short turn loop connecting TM2 and TM3, and Tyr139. Additionally, Tyr139 was predicted to pack against Trp9. These interactions could play a role in maintaining the stability of the periplasmic region. Consistent with this conclusion, substitution of the

amino acids at these positions with alanine residues greatly reduced the enzymatic activity of ApCarS (**Figure 4C,D**).

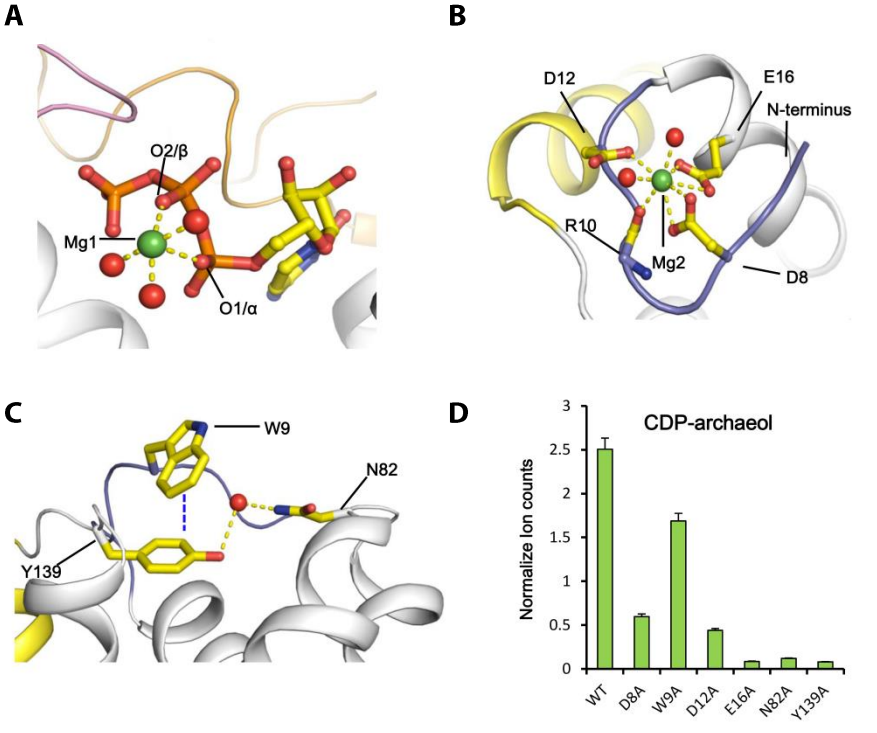


Figure 4| Mg²⁺ is required for the enzymatic activity of ApCarS. (**A**) The Mg²⁺ ion in the active site of ApCarS. Mg²⁺ is bridged by oxygen atoms from the α- and β-phosphate groups of CTP and water molecules. (**B**) The Mg²⁺ ion bound at the N-terminal region of ApCarS. (**C**) The interaction between Asn82 and Tyr139, bridged by a water molecule, which is proposed to enhance the stability of the structure. The π-π interaction is indicated with the blue dashed line. (**D**) CTP-transferase activities of ApCarS proteins with amino acid substitutions within the periplasmic region.

Biochemical analyses of ApCarS

To verify the importance of CTP-interacting residues identified during our crystal structure analysis, we generated a panel of ApCarS variants harboring amino acid substitutions within the CTP-binding site, and evaluated the CTP-binding activities and CDP-archaeol production levels of

these proteins (**Figure 5A,B and Supplementary Table 2**). Consistent with our structural observations, substitution of CTP-interacting residues with alanines resulted in a marked reduction or loss of CTP-binding activity, as indicated by ITC analysis. Likewise, the catalytic activity of these modified proteins was also significantly compromised. Specifically, modification of Lys57, Asp100, and Lys107 (K57A, D100A, and K107A, respectively), which form polar interactions with the phosphate groups of CTP, resulted in complete abrogation of the CTP-binding (**Figure 5A**) and catalytic activities (**Figure 5B**) of ApCarS. Similar results were obtained upon modification of Asn28 (N28A), which interacts with the α-phosphate group of CTP in a water- and Mg^{2+} -dependent manner (**Figure 5A,B**). While modification of Pro31 (P31A), which packs against the cytosine portion of CTP, disrupted CTP-binding, this variant retained appreciable catalytic activity at high CTP concentrations (**Figure 5B**); however, the reason for this discrepancy remains unclear.

Asp55, located within CL1, is not directly involved in the interaction with CTP. Instead, this residue forms a bifurcated salt bond with Arg51, thereby stabilizing the conformation of CL1. Notably, a D55A variant exhibited moderately reduced CTP-binding activity but strikingly compromised catalytic activity, indicating that proper CL1 conformation is essential for optimal ApCarS enzymatic activity (**Supplementary Figure 4A–C, Supplementary Figure 6 and Supplementary Table 2**).

Analysis of the glycerophospholipid binding site of ApCarS

While a DALI database search identified the *Thermotoga maritima* CDP-DAG synthetase (TmCds; PDB code 4q2e and 4q2g) [195] as the closest structural relative to ApCarS, with a score of 6.8, the two proteins share only12% sequence identity. The membrane domain of bacterial CDP-DAG synthetases is considerably larger than that of the archaeal CDP-archaeol synthases, comprising an active site formed by five transmembrane helices with a cytoplasmic region and additional N-terminal TM segments. The structural homology between the two proteins primarily lies within the portion of ApCarS encoding TM1–TM4 (**Figure 6A and Supplementary Figure 7A**), suggesting these domains might share a common evolutionary origin. Notably, there are striking differences in the amino acid sequences

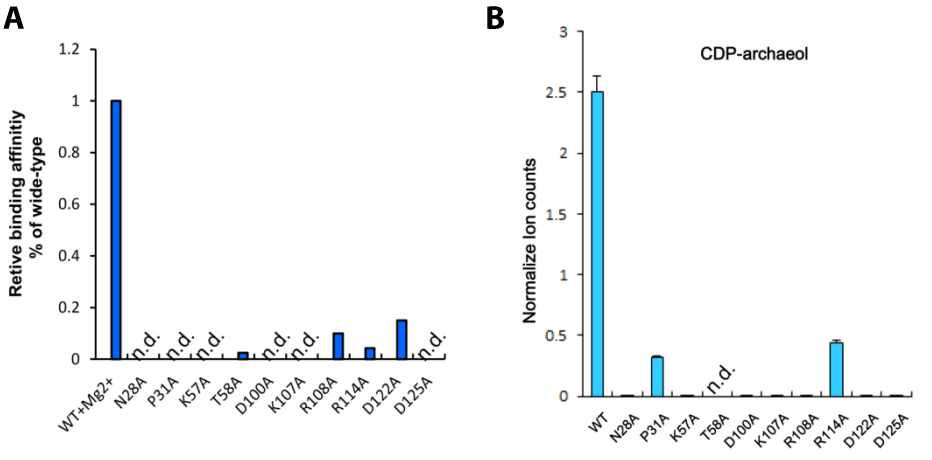


Figure 5| Biochemical analyses of ApCarS. (**A**) Relative CTP binding affinities of wild type (WT) and modified ApCarS proteins. The indicated amino acid residues were modified to alanine (A) residues, and CTP binding was analyzed in the presence of Mg^{2+}. The relative binding affinities correspond to ratios of the binding affinities measured for the WT and modified enzymes. Several variants exhibited such low binding affinities that the binding constants could not be determined (n.d.) by isothermal titration calorimetry (ITC). The original ITC data are presented in Supplementary **Figure 6** and **Supplementary Table 2**. (**B**) CTP-transferase activities of ApCarS variants. Amino acid residues around the active site have been modified to A [n.d., not detected; error bars, standard error of the mean (SEM) of the data from duplicate assays].

of the two proteins at the CTP-binding region, offering an explanation for the different substrate specificity of the two structurally related enzymes; however, the observed structural similarity at this region implies that TmCds might also use this site for CTP binding.

To date, extensive efforts to crystallize ApCarS and CTP in complex with 2,3-bis-O-geranylgeranyl sn-glycerol-phosphate have failed. As such, the glycerophospholipid binding site of this protein remains experimentally unidentified. However, the location of CTP and a structural comparison of ApCarS with TmCds allow prediction of the DGGGP binding site as supported by site-directed mutagenesis and biochemical analysis. Archaetidic acid contains two hydrophobic tails that can be mimicked by B-nonylglucoside (**Supplementary Figure 7A**). Notably, in ApCarS, TM5 loosely packs against TM1 and TM4, which form two fairly deep grooves

adjacent to the CTP-binding site. The structural flexibility of TM5 is consistent with structural findings, average B-factors of $TM5_{154-163}$ is higher than average of other helices (**Figure 6B, C**). The amino acid residues of TM5 that line the groove are largely hydrophobic, and are similar among CarS family members (**Figure 2B and Supplementary Figure 1**). These observations highlight the possibility that the two grooves act as a binding site for two long hydrophobic lipid tails [196–198]. Consistent with this hypothesis, structural superposition revealed that B-nonylglucoside binds to a similar site within TmCds (**Figure 6A and Supplementary Figure 7A**). The two hydrophobic linear hydrocarbon chains of DGGGP could therefore be expected to insert into these hydrophobic grooves, which would position the phosphate head groups adjacent to the phosphate group of CTP for catalysis (**Figure 6D and Supplementary Figure 7B**).

Discussion

In the current manuscript, we report the biochemical and structural characterization of the intramembrane enzyme ApCarS, which is involved in the biosynthesis of archaeal membrane lipids. Our analyses showed that ApCarS encodes CTP-transferase activity and is a functional homolog of AfCarS. Subsequently, analysis of the crystal structure generated in this study revealed the mechanism of CTP recognition by ApCarS. Notably, given the conservation of predicted CTP-interacting residues, this mechanism is likely conserved among CarS family members (**Supplementary Figure 1**).

Our data provide insight into the catalytic mechanism of ApCarS. Specifically, our biochemical data indicate that CTP binding results in stabilization of the ApCarS structure (**Figure 2B**), and that CTP-mediated stabilization of the CPD, in particular, is likely essential for the catalytic activity of this enzyme. Indeed, mutations that destabilized CL1 resulted in abrogation of the CTP-binding and catalytic activities of ApCarS (**Figure**

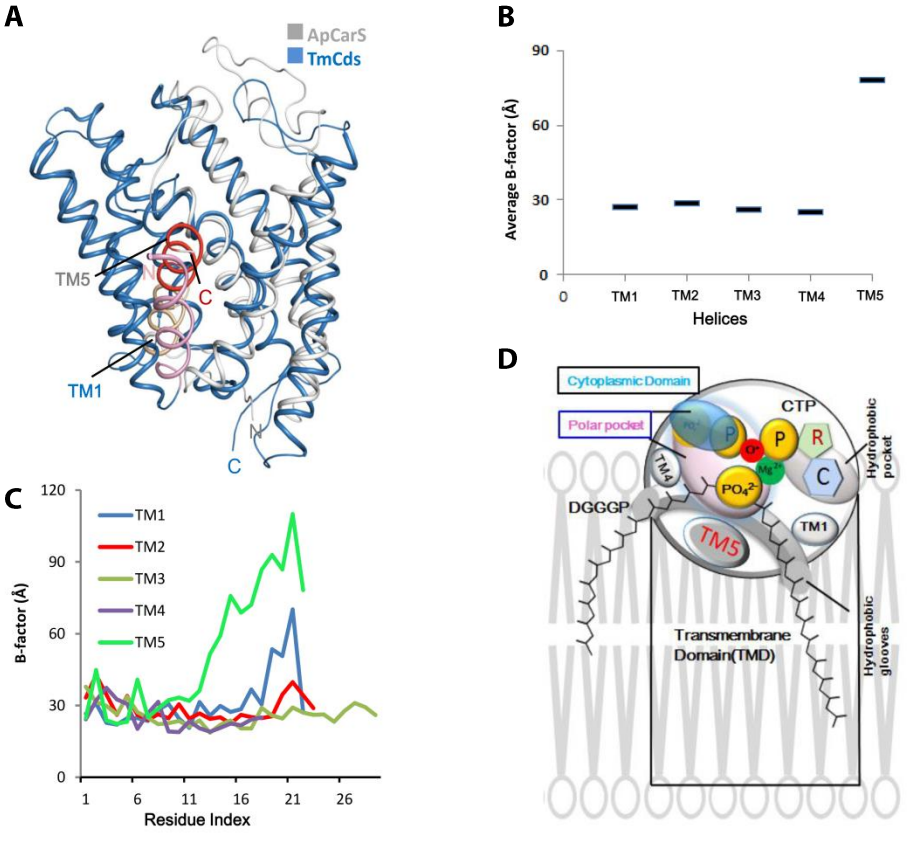


Figure 6| A working model of ApCarS-catalyzed reaction. (**A**) Structural superposition of ApCarS (gray) and the *Thermotoga maritima* CDP-DAG synthetase (TmCds; blue), which showed the highest similarity to that of ApCarS by DALI sever analysis [PDB code 4q2e [112]]. TM5 of ApCarS and TM1of TmCds are highlighted in red and pink, respectively which indicate structural flexibility. (**B**) The average b-factor of the helices from TM1 to TM5 of ApCarS. Residues positions are $TM1_{14\text{-}34}$, $TM2_{59\text{-}80}$, $TM3_{83\text{-}110}$, $TM4_{121\text{-}137}$ and $TM5_{154\text{-}164}$. The average b-factor of TM5 is rather higher than others, indicating the structure flexibility of TM5. (**C**) A plot of the averaged TMs b-factor of $TM1_{14\text{-}34}$, $TM2_{59\text{-}80}$, $TM3_{83\text{-}110}$, $TM4_{121\text{-}137}$ and $TM5_{154\text{-}164}$ showing the higher level of flexibility of $TM5_{154\text{-}164}$. (**D**) Cartoon depicting ApCarS-catalyzed reactions between CTP and DGGGP (2,3-bis-O-geranylgeranyl sn-glycerol-phosphate). The proposed DGGGP binding sites are shown as gray tubbiness. "P", "R", and "C" represent the phosphate groups, sugar ring, and cytosine of CTP, respectively. The phosphate group from DGGGP is indicated by "PO_4^{2-} ", whereas its hydrophobic tails are shown as folded lines (**Supplementary Figure 7B**).

5A,B and Supplementary Figure 4A–C). While binding to ApCarS resulted in burial of the γ- and β-phosphate groups of CTP/dCTP (**Figure2D, Figure 3A,B and Figure 6B**), the electron-withdrawing effect generated by Mg^{2+} and Lys57 rendered the α-phosphate group electrophilic, allowing it to react with the negatively charged phosphate head of the glycerophospholipid (**Figure 4A**). O1 of this α-phosphate group faces TM5, which together with TM1 and TM4, form a pair of hydrophobic grooves. As these grooves are clearly embedded within the plasma membrane, they should exhibit increased hydrophobicity, which would promote binding of the hydrophobic tails of glycerophospholipids. In turn, this interaction likely results in positioning of the phosphate head adjacent to the α-phosphate group of CTP/dCTP, enabling nucleophilic attack (**Figure 6B**).

Compared to the other domains of ApCarS, TM5 is less well defined, indicating its structural flexibility. This flexibility is consistent with a role for TM5 in glycerophospholipid substrate binding, and might also be associated with the temperature-dependent enzymatic activity of ApCarS. Indeed, increases in temperature are expected to promote, comparatively, more striking conformational changes in TM5 (**Figure 6B, C**), thus favoring substrate binding. Also, efficient dissociation of pyrophosphate and release of CDP-archaeol from the active site are likely important for optimal ApCarS catalytic activity. As such, structural flexibility of both TM5 and the cytoplasmic domain (CPD) may be required (**Figure 6D**).

Remarkably, although ApCarS and other CDP-archaeol synthases share only a low level of sequence identity with bacterial CDP-DAG synthase, while also being highly specific for their substrates, DGGGP and phosphatidic acid, respectively [41], the catalytic core domains of these proteins are structurally very similar. This finding supports an earlier proposal of the vertical inheritance of this gene from the cenancestor [8], but also suggests that this progenitor may have been less specific in regards to the lipophilic substrate.

In summary, we report the first crystal structure of CarS family. Our results provide critical new information for understanding the mechanism of ApCarS activity; however, future studies, including analysis of the

structure of ApCarS when bound to its two substrates, are required to fully characterize the catalytic pathway of this enzyme.

Experimental procedures

Protein expression and purification

The codon-optimized gene encoding CarS from *Aeropyrum pernix k1* (ApUbiA) was subcloned into the pET21b vector (Novagen, Madison, WI, USA). The 6×His-tagged protein was then expressed in *E. coli* BL21(DE3) cells, and cell membranes were harvested, solubilized in 1% DDM (dodecyl-β-D-maltopyranoside), and subsequently incubated with the nickel-NTA resin at 4 °C for 1–2 hours. The resin was gradient washed with 15, 30, and 50 mM imidazole in 15 mM Tris/HCl (pH 8.0) buffer containing 350 mM NaCl, 10% glycerol, and 0.03% DDM. Eluted protein in DDM buffer was then further purified by size-exclusion chromatography (Superdex-200 Increase 10/300; GE Healthcare) in 10 mM Tris/HCl (pH 8.0) buffer containing 0.02% DDM, 100 mM NaCl, and 10% glycerol. Peak ApCarS fractions were pooled and flash-frozen for crystallization.

Crystallization

Purified protein was concentrated to approximately 25 mg/ml. Crystals were grown using the lipid cubic phase method [199] at 20 °C in buffer containing 200 mM NaCl, 20% (w/v) PEG400, and 100 mM Tris/HCl (pH 8.0). The crystals (hexagonal 315 plates) grew to a size of 0.03×0.1×0.1mm within 3 weeks.

Structure determination

X-ray diffraction data were collected at beamline BL19U1 of the Shanghai Synchrotron Radiation Facility. Data were integrated using autoPROC [200] and XDS [201] software, then reduced using the AIMLESS [202] program. Resolution cut-off was based on the default criteria of AIMLESS. The initial phases were solved with Arcimboldo Lite [203] through molecular replacement with combinations of polypeptide fragments and

density modification, and then a model for molecular replacement was built using ARP/wARP [204] software. Maximum likelihood-based refinement of the data was performed with Phenix [205], and the atomic model was fit using the Coot [206] program. The stereochemical quality of the final model was assessed with MolProbity [207].

In vitro activity assays and HPLC-MS analysis

The activity of wild type and variant ApCarS proteins was examined *in vitro* as described previously [41]. For these assays, purified ApCar (0.5 μM) was incubated in buffer containing 50 mM Tris-HCl (pH 7.5), 10 mM MgCl2, 75 mM NaCl, 0.1% DDM, 125 mM imidazole, and 5% glycerol, and enzymatic activity towards the chemically synthetized substrate DGGGP was evaluated in the presence of 2 mM CTP as a cofactor. The reaction mixtures were incubated for 1 hr at three different temperatures (37 °C, 60 °C and 90 °C) for wild type ApCarS, but only at 90 °C for the ApCarS variant proteins. Reaction products were analyzed by HPLC-MS, as described in our previous work [41].

ITC measurements

ITC measurements for the CTP-binding activity of wild type and variant ApCarS proteins were taken at 25 °C using an ITC200 microcalorimeter [MicroCal (Malvern Instruments), Malvern, UK]. Protein samples and substrates were prepared in buffer containing 20 mM TrisHCl (pH 8.0), 100 mM NaCl, and 0.02% DDM. Approximately 0.5 mM CTP was then titrated into buffer containing 0.15 mM ApCarS; the first injection was 0.5 μl in volume and the remainder were 2 μl in volume. The solution was stirred at a rate of 750 rpm. The data were fitted to a one site-binding model using MicroCal ORIGIN software. Mean thermodynamic parameters for triplicate titrations are shown in the **Supplementary Table 2** and **3**. All ITC measurements were performed in technical triplicate.

Limited proteolysis analysis of ApCarS

The stability of the structure of ApCarS following substrate binding was analyzed by limited proteolysis. ApCarS protein (3.5 mg/ml) was mixed with 0.03 mg/ml trypsin (Roche, Basel, Switzerland) in buffer containing

15 mM Tris/HCl (pH 8.0), 150 mM NaCl, 0.02% DDM, and 3 mM CTP in the presence or absence of 10 mM MgCl2. After 30 min of incubation on ice, proteins from the mixture were separated by SDS-PAGE and visualized by Coomassie blue staining.

Sequence alignment

Multiple sequence alignments were generated using ClustalW software, and were edited with the ESPript 3.0 program [208].

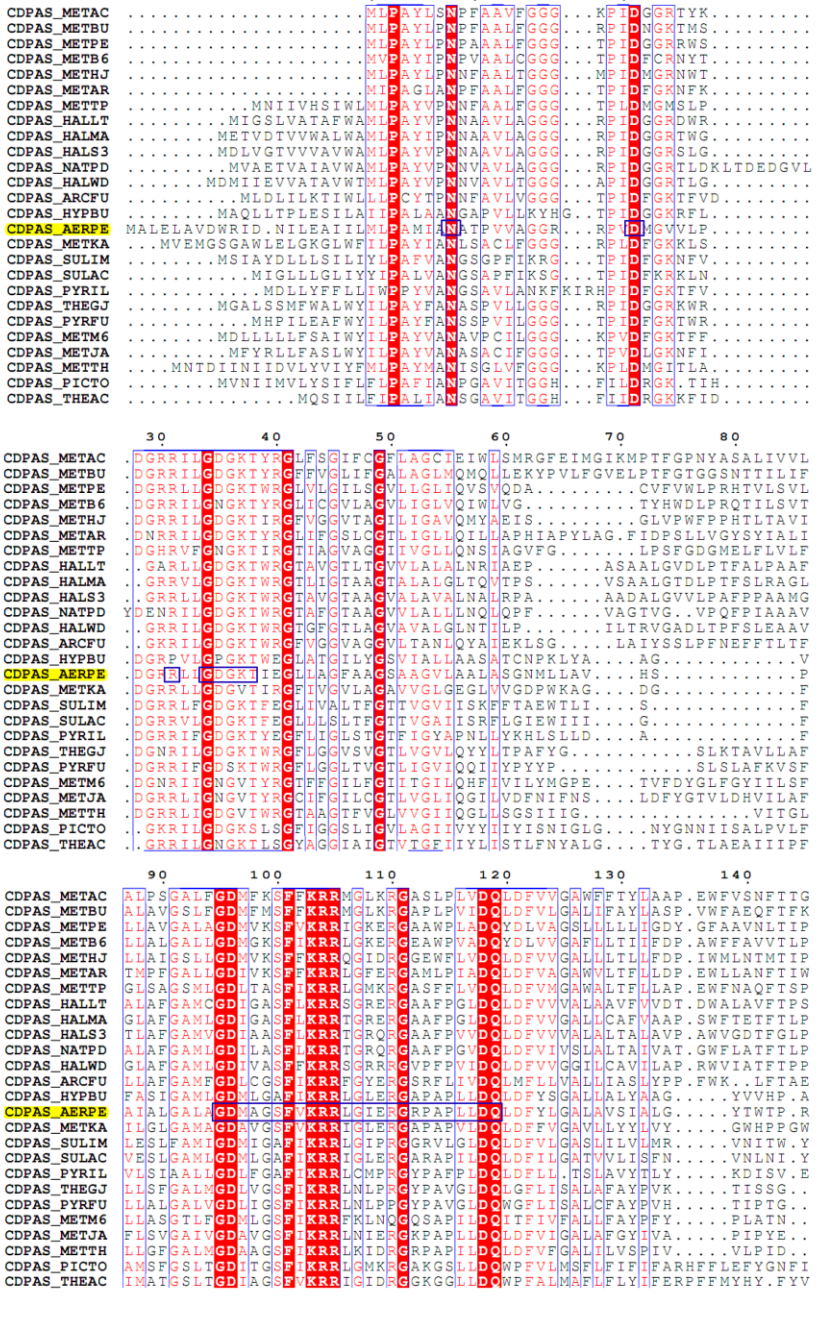
Acknowledgments

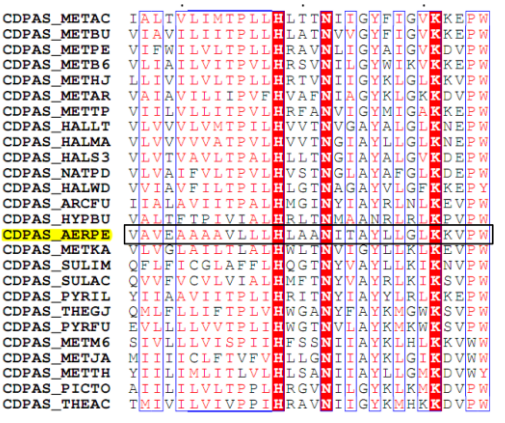
We thank J. Chai for critical comments. We thank the staffs from BL19U1 beamline of National Center for Protein Sciences Shanghai (NCPSS) at Shanghai Synchrotron Radiation Facility, for assistance during data collection. We thank Peter Fodran and Adriaan Minnaard for the synthesis of DGGGP. This work was funded by grants to W.C. from the National Natural Science Foundation of China (31570842), awards of National Young Thousand Talents Program and Sichuan Province Thousand Talents program in China. A.C. was supported by the research program of the biobased ecologically balanced sustainable industrial chemistry (BE-BASIC).

Author contributions

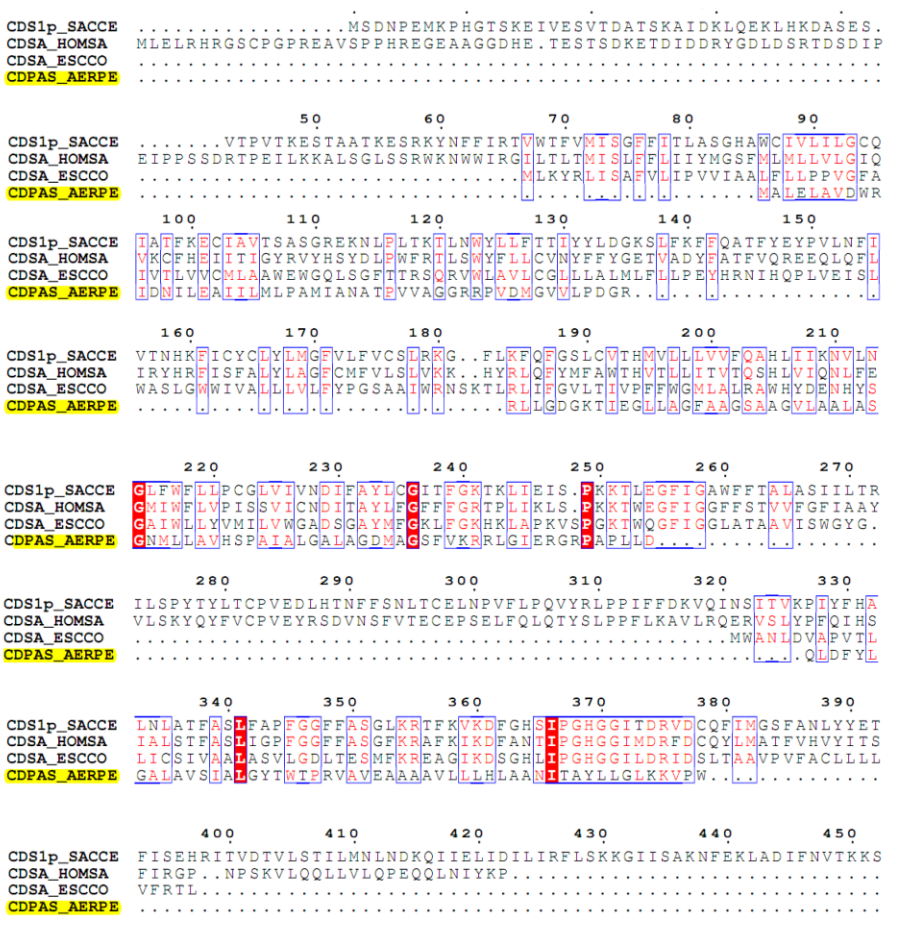
W.C. performed the research design; S.R. and J.W. made constructs; S.R.; Y.G. and Q.Y. purified the proteins; S.R. and B.S. grew and optimized the crystals; C.D. collected the data; F.Y. and X.Z. determined the structure; Y.G. and J.J. performed ITC; A. C. examined the enzyme activity and performed the mass spectrometry; W.C. and A.J.M.D analyzed the data; W.C. wrote manuscript. All authors participated in analyzing the data and writing the paper.

Supplemental Information

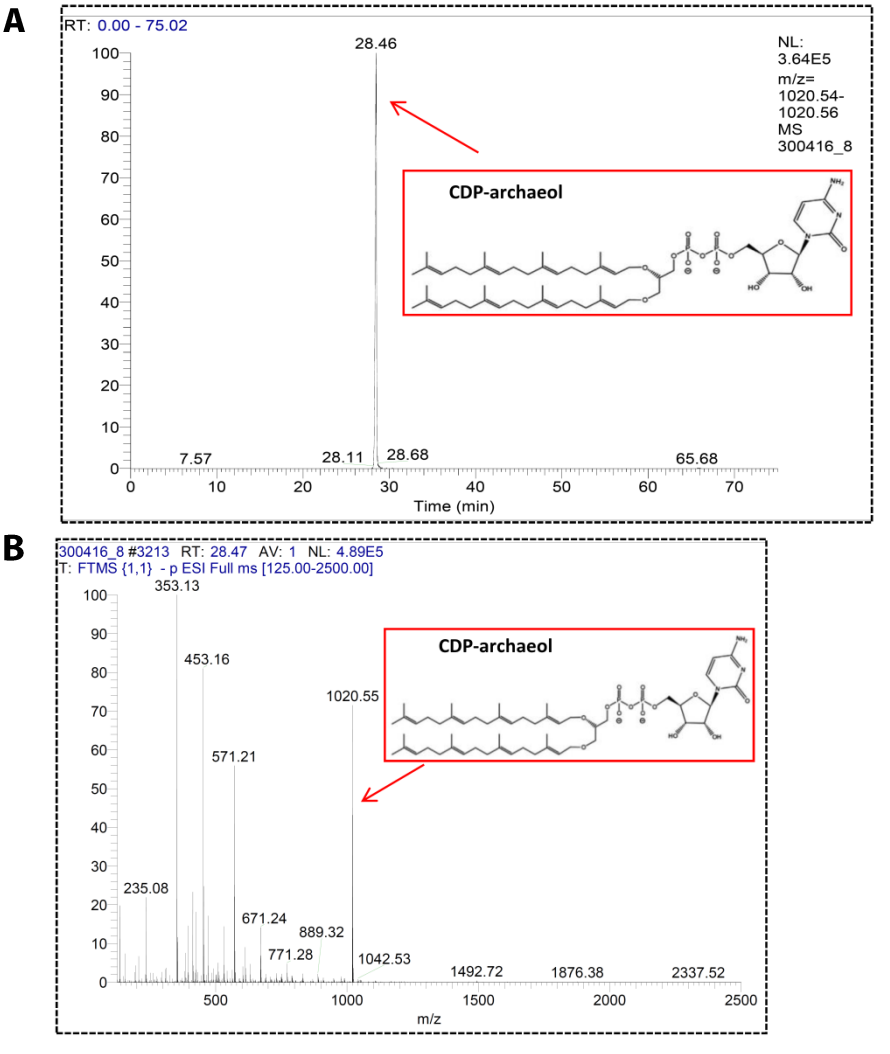




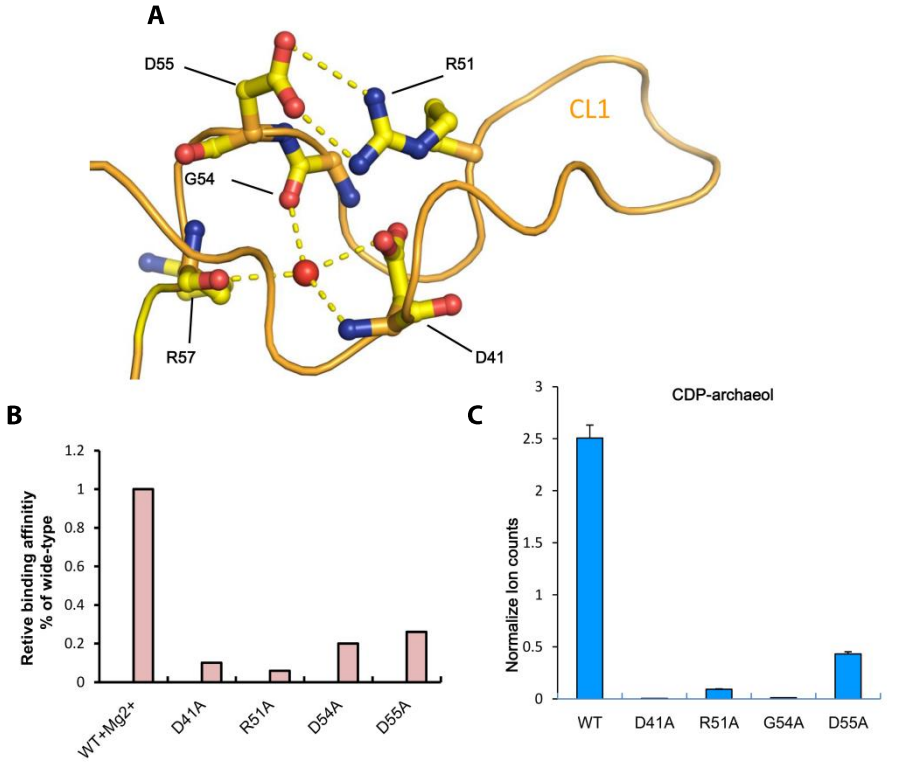
Supplementary Figure 1| Amino acid sequencealignment of CarS homologs among archaea. Pairwise sequence alignment of 26 CarS sequences, which are representative of the archaeal 3 CDP-archaeol synthetase family. The CarS of *Aeropyrum pernix*, which was analyzed in the current study, is highlighted in yellow. The conserved CTP binding motif is indicated with a blue box. Protein names and organisms are indicated as follows: CDPAS, CDP-archaeol synthetase; ARCFU, *Archaeoglobus fulgidus*. The alignment was generated using the MultAlin and ENDscript programs[206].



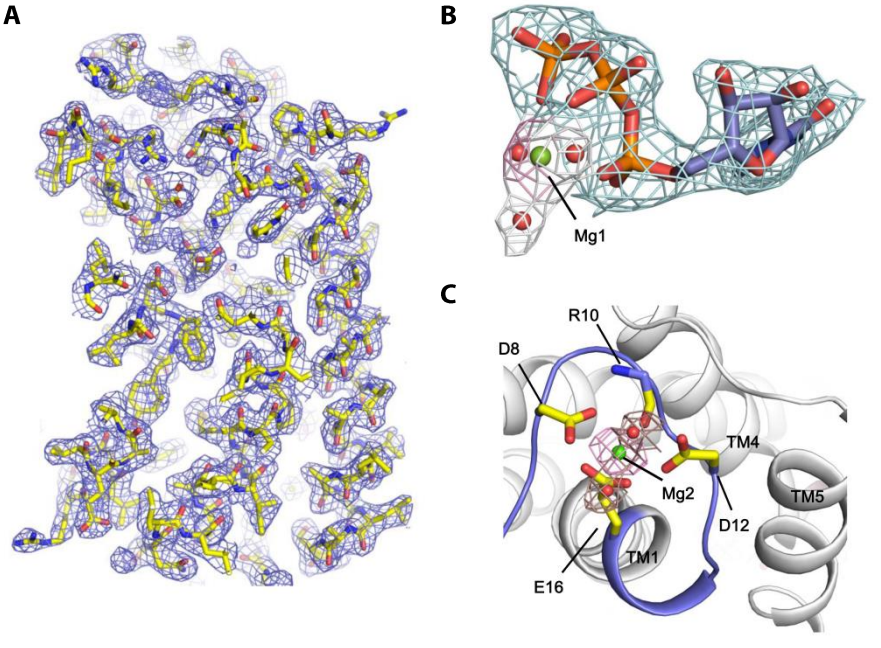
Supplementary Figure 2| Amino acid sequence alignment of members of the CTP-transferase superfamily from archaea, bacteria, and eukaryotes**.** Pairwise sequence alignment using four CarS sequences that are representative of archaea, bacteria, fungi, and mammals. Protein names and organisms are indicated as follows: CDPAS, CDP-archaeol synthetase; AERPE, *Aeropyrum pernix*.



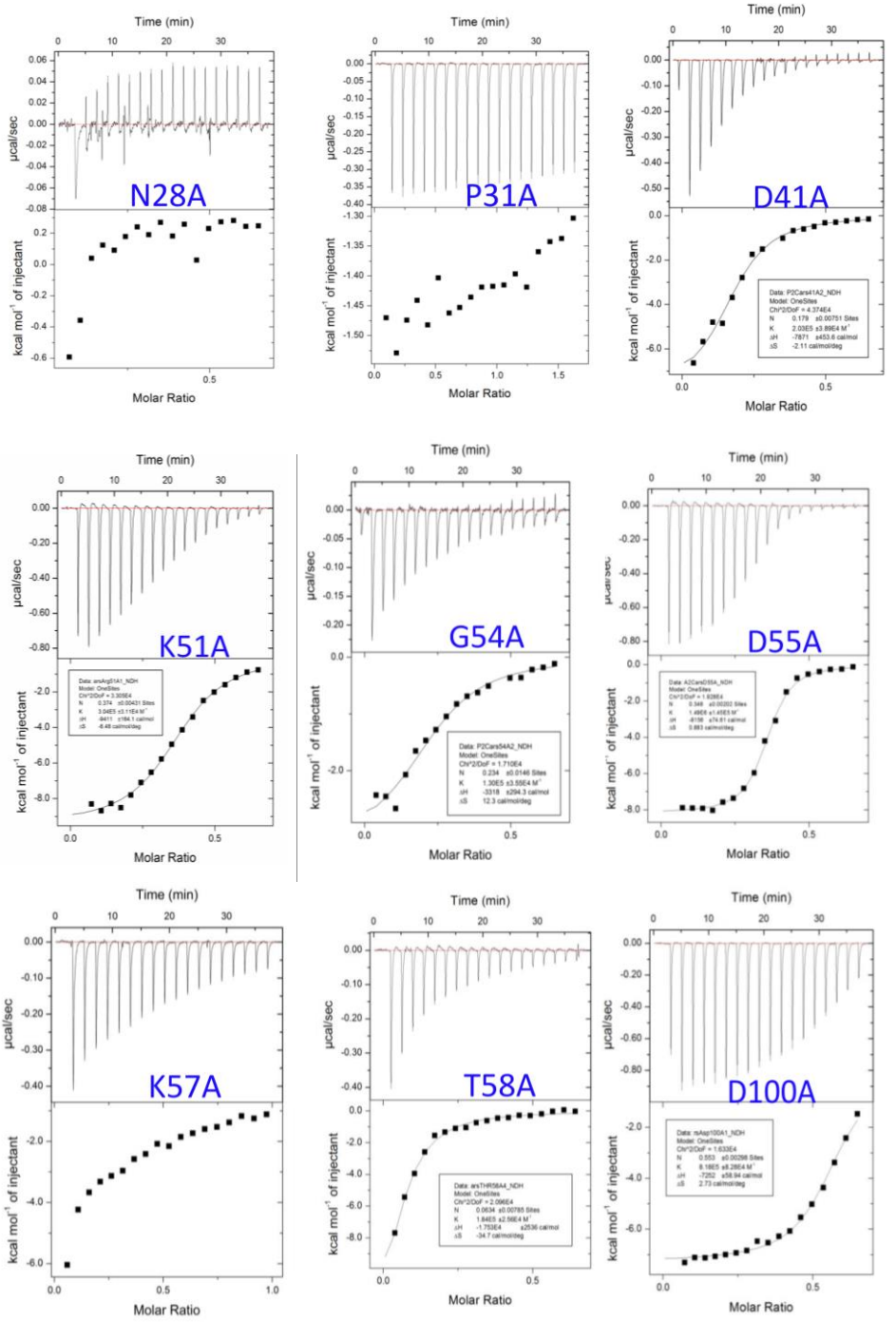
Supplementary Figure 3| Detection of products catalyzed by ApCarS using high performance liquid chromatography-massspectrometry (HPLC-MS). (**A**) Products from the ApCarS activity assay using DGGGP(2,3-bis-O-geranylgeranyl sn-glycerol-phosphate) and CTP as substrates, as depicted in **Figure 1D**, were separated by HPLC and analyzed by MS. The CDP-archaeol peak, with a m/z=1020.55 [M-H]**–**, is indicated by an arrow. (**B**)The total Ion Current Chromatogram filtered for the CDP-archaeol m/z value to detect the product.

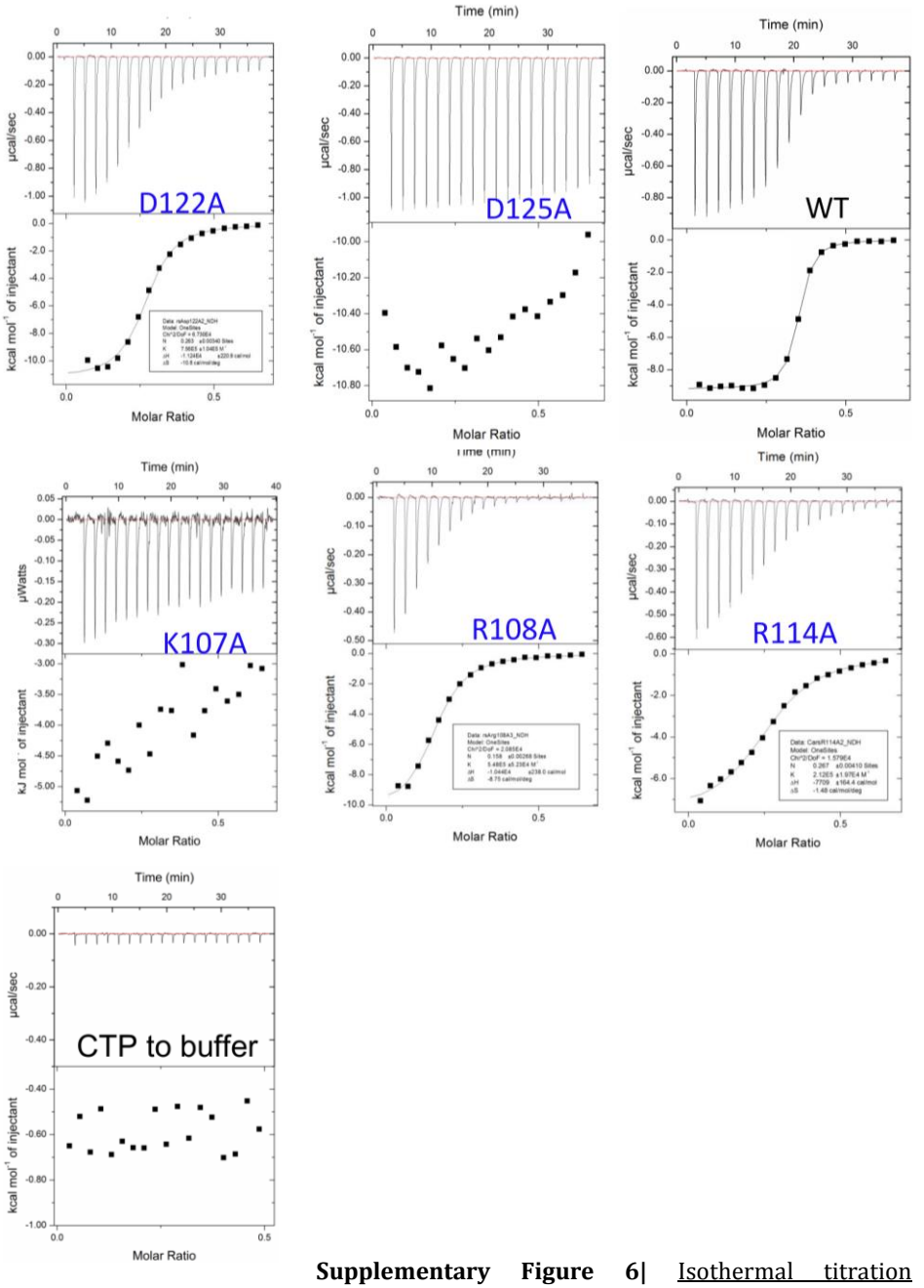


Supplementary Figure 4| Stabilization of cytoplasmic loop 1 (CL1) is essential for the CTP-binding and catalytic activity of the ApCarS. (**A**) Interaction networks within CL1. (**B**) Relative CTP-binding affinities of ApCarS proteins encoding single amino acid substitutions in CL1. The original isothermal titration calorimetry (ITC) data are presented in **Supplementary Figure 6** and in **Supplementary Table 2**. Amino acid residues located within CL1 (D41, K51, G54 and D55) were modified to alanine.(**C**) Enzymatic activity of ApCarS proteins encoding single amino acid substitutions within CL1. The assays were performed as described in the legend for **Figure 1B,D**.

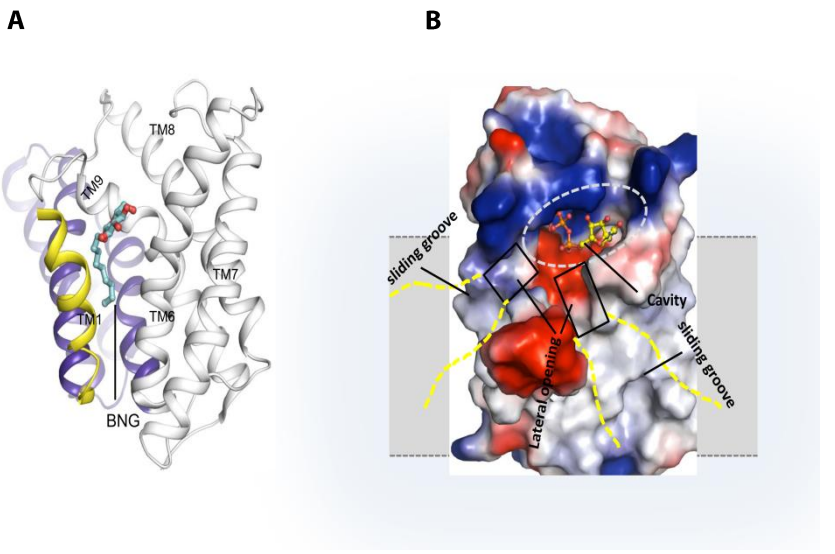


Supplementary Figure 5| Electron density maps for the ApCarS bound to Mg^{2+} ions. (**A**) Sigma A-weighted 2Fo–Fc electron density map of the finally refined model contoured at 2.2σ. (**B**) Mg1 at the active site of ApCarS. Representative 2Fo-Fc density maps of Mg^{2+} (in pink) and water molecules (in gray) contoured at 2.2σ and 1.5σ respectively. (**C**) Coordination of Mg2 at the N-terminal loop of ApCarS. Representative 2Fo-Fc density maps of Mg^{2+} (pink) and water molecules (violet) contoured at 2.5σ and 1.5σ respectively.


Time (min)
µcal/sec
kcal mol⁻¹ of injectant
Molar Ratio
N28A
P31A
D41A
K51A
G54A
D55A
K57A
T58A
D100A



Supplementary Figure 6| Isothermal titration calorimetry (ITC) measurements of the CTP-binding activity of ApCarS proteins encoding amino acid substitutions in the cytoplasmic domain (CPD). Residues adjacent to the active site were substituted with alanines (Ala).The representative raw data and binding isotherms for CTP-binding are shown here.



Supplementary Figure 7| Structural basis for dual-substrate binding**.** (**A**) Crystal structure of *Thermotoga maritima* CDP-DAG synthetase (TmCds)in complex with B-nonylglucoside (PDB ID: 4q2g). (**B**) Overall electrostatic surface representation of ApCarS (**Figure 6B**).

Supplementary Table 1. Data collection and refinement statistics.

Data collect and refinement statistics	
CTP-bound (PDB ID:)	
Data collect	
Space group	C2221
Cell dimensions	
a, b, c (Å)	56.505, 89.249, 98.876
α, β, γ (°)	90, 90,90
Resolution (Å)	47.741(2.397)
R_{merge} (%)	21
R_{pim} (%)	6.3
I/σ (I)	24.4
Completeness (%)	98.8
Redundancy	10.2
Refinement	
No. reflections	9961
R_{work}/R_{free}	20.71/27.06
No. atoms	
Protein	1141
Ligand/ion	8
B-factors	
Protein	163
Ligand/ion	4
R.m.s deviations	
Bond length (Å)	0.007
Bond angles (°)	1.009
Ramachandran(%)	
Favoured	96.73
Allowed	3.2
disallowed	0

Value in parentheses is for the highest-resolution shell.

Supplementary Table 2. Affinity constants and binding parameters demonstrating the effect of amino acid substitutions within the cytoplasmic domain (CPD) of ApCarS on CTP-binding.

	K_D (μM)	N (sites)	ΔH (kcal/mol)	ΔS (cal/mol/deg)
WT+Mg^{2+}	0.2±0.014	0.36	-9490±37.09	-1.19
WT -Mg^{2+}	1.6±0.18	0.31	-11340±168.1	-11.6
N28A	n.d.			
P31A	n.d.			
D41A	2.1±0.16	0.18	-7753±348.76	-1.43
R51A	3.4±0.34	0.37±0.004	-9476±165.9	-6.75
D54A	1.05±0.11	0.16±0.008	-8153.5±537.3	4.95
D55A	0.76±0.07	0.36±0.002	-8446.5±78.76	-0.3785
K57A	n.d.			
T58A	8.85±1.41	0.157±0.012	-6747±710.2	0.485
D100A	n.d.			
K107A	n.d.			
R108A	1.95±0.25	0.173±0.004	-9551±294.9	-5.915
R114A	4.6±3.39	0.257±0.00323	-8042±137.35	-2.555
D122A	1.4±0.07	0.26±0.002	-11460±193.5	-11.65
D125A	n.d.			

Data are shown as mean and s.e.m. of three technical replicates, n.d. refer to not detect. ApCarS protein concentration is about 0.15mM, CTP is 0.5mM, CTP as ligand titrated into protein ApCarS. Notably, here N (sites) is not about 1, because CTP associated with ApCarS during expression and remained bound throughout all phases of purification and crystallization, which was supported by our experimental results, including structure model, trypsin limited proteolysis, and mass spectrometry.

Chapter 6

Saturation of the ether lipids double bonds: Geranylgeranyl Reductase

Antonella Caforio[1] and Arnold J. M. Driessen[1]

[1]Department of Molecular Microbiology, Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, 9747 AG Groningen, The Netherlands; The Zernike Institute for Advanced Materials, University of Groningen, 9747 AG Groningen, The Netherlands

Abstract

The presence of fully saturated isoprenoid chains is one of the distinct features of the archaeal membrane lipids compared to phospholipids of eukarya and bacteria. Hydrogenation of the double bonds is achieved via a reduction reaction which is catalyzed by enzymes belonging to the geranylgeranyl reductase (GGR) family. However, it is still a matter of debate at which step of the biosynthetic pathway archaetidic acid is reduced. GGR and ferredoxin of two different archaeal organisms were tested for the formation of the saturated species of the archaeal lipids archaetidylglycerol (AG) and archaetidylethanolamine (AE) in an *E. coli* strain harboring the ether lipid biosynthetic pathway.

Introduction

All living cells are surrounded by a barrier called cytoplasmic membrane that separates the inside from the outside of the cell. The membrane is involved in many important cellular processes. The lipid composition of cytoplasmic membrane is an important feature that distinguish organisms from each other, in particular with the differentiation between Archaea and Bacteria and Eukarya [88,126]. Archaeal lipids consist of fully saturated isoprenoid chain ether linked to the glycerophosphate backbone compared to the phosphatidic acid chain linked via ester bonds to the enantiomeric glycerophosphate form found in bacteria and eukarya. One of the striking features of archaeal lipids is the high chemical stability of their hydrocarbon chains which contribute to the survival in extreme environments [3,71]. This property is partially conferred by an important step in ether lipid biosynthesis which is the saturation of the double bonds on the isoprenoid chain. The archaeal ether lipid biosynthetic pathway has been extensively described in the thesis chapters, but the exact enzymology of saturation remains unclear.

One mechanism proposed is that the hydrogenation of the double bonds of the isoprenoid chain to produce the final unsaturated archaeal lipids is catalyzed by a geranylgeranyl reductase (GGR) enzyme. It belongs to the GGR family which includes GGR members from bacteria [209] and plants [210] mainly involved in photosynthesis and carotenoid production [51]. Among archaea many GGR homologous are found that contain a highly conserved FAD binding site and a certain degree of conservation in the substrate binding domain. In particular the conserved motif sequence YxWxFPx7-8GxG is important for keeping the substrate in a FAD parallel position optimal for double bond reduction [50]. Several GGR enzymes from different archaeal organisms have been investigated and characterized in the past years. Reductase activity was initially observed in cell free extracts of the hyperthermophilic archaea associated protein [211] for which the crystal structure was solved in a FAD bound complex [50]. The proposed hydrogenation mechanism requires the presence of NADH or other reducing agents which reduce the flavin cofactor of the enzyme, followed by the hydrogenation of the lipid substrate double bonds

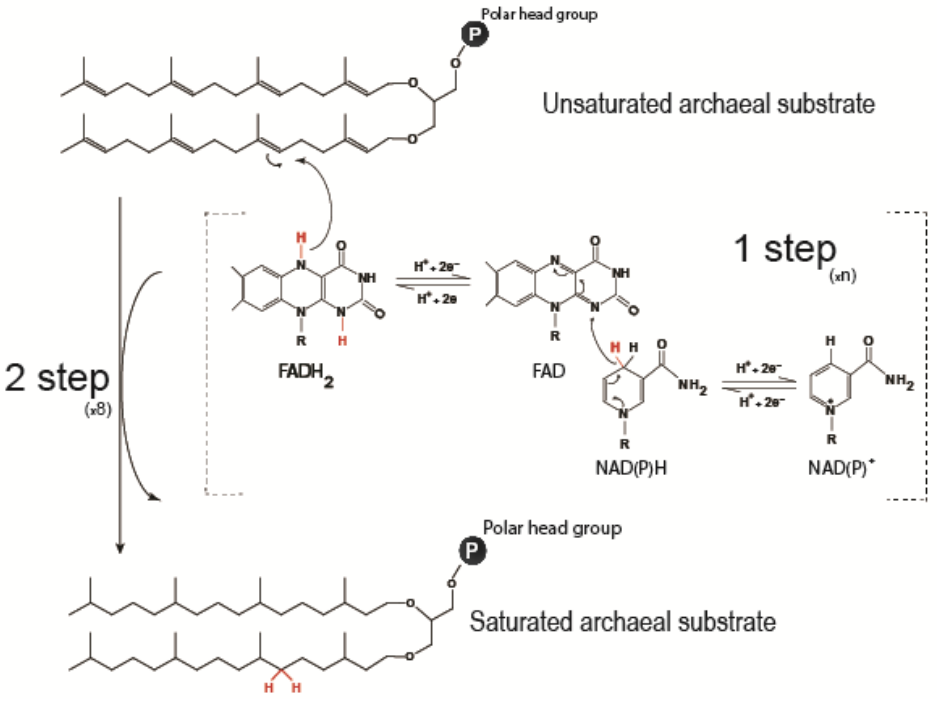


Figure 1 | Schematic representation of the geranylgeranyl reductase putative mechanism. The proposed mechanism for the hydrogenation of double bonds requires an initial step in which the flavin cofactor bound to the enzyme is reduced by NAD(P)H or other reducing agent. The reduced flavin cofactor transfers the received hydrogens to the substrate to saturate the double bonds on the archaeal lipid isoprenoid chain. The transfer of the two hydrogens (colored in red) are highlighted with arrows. The two steps are repeated to obtain full saturation of all the double bonds.

by the reduced flavin (**Figure 1**). The GGR from *Sulfolobus acidocaldarius* has a similar structure as the GGR from *T. acidophilum* in the FAD binding and active sites [52]. The enzyme clearly shows a substrate preference for DGGGP, GGGP and to a lesser extent GGPP [212]. A recent study investigated the GGR from the mesophilic archaea *Methanoscarcina acetivorans*. The corresponding gene was expressed in a engineered *E. coli* strain containing four genes of the archaeal ether lipid biosynthetic pathway. Upon the expression of the reductase enzyme the saturated form of DGGGP could be detected [53]. Interestingly, the hydrogenation activity was enhanced when the ferredoxin gene, localized upstream the GGR gene

in *M. acetivorans* genome, was coexpressed in the bacterial strain. This suggested the possibly role of ferredoxin as biological electron donor. However, the overall saturation of double bonds was weak leaving high levels of the archaeal phospholipids unsaturated.

In the present section the GGR and ferredoxin from *M. acetivorans* and *Methanococcus maripaludis* were used to characterized the double bonds hydrogenation step in the ether lipid biosynthetic pathway. The expression of these genes was realized in an engineered *E. coli* strain containing up to seven ether lipid genes [163] aiming to *in vivo* produce the two saturated form of archaetidylglycerol (AG) and archaetidylethanolamine (AE).

Results

In vivo saturation of AG and AE by GGR and ferredoxin of *M. acetovirans*

On the basis of the ability of the GGR enzyme and ferredoxin from the archaeon *M. acetivorans* to generate fully saturated archaeal lipid species when coexpressed with four ether lipid biosynthesis genes in *E. coli* [53], the activity of GGR and ferredoxin was tested in the *E. coli* strain containing up to seven ether lipids enzymes in order to produce saturated AG and AE [163]. Herein, the codon optimized genes *MA1484* and *MA1485* encoding for the GGR and ferredoxin from *M. acetivorans*, respectively, were added to the set of four compatible vectors listed in **Table 1**. Aerobically growing *E. coli* strains were induced with 0.25 mM IPTG for 3.5 hours, and alternatively cells were grown aerobically until mid-exponential phase (OD_{600} = 0.6) and then induced under oxygen depriving condition [213]. Total lipid analysis was performed by LC-MS using the previously described method [163]. Three different engineered *E. coli* strains were compared containing a various combination of ether lipid enzymes: 1 - seven ether lipid enzymes (*E. coli* Idi, *E. coli* mutant IspA, *B. subtilis* AraM, *M. maripaludis* GGGPS, *A. fulgidus* DGGGPS, *A. fulgidus* CarS and *B. subtilis* PssA), II eight ether lipid genes as I but with *M. acetivorans* GGR and III – nine ether lipid enzymes as I with *M. acetivorans* GGR and *M. acetivorans* Ferredoxin. The LC-MS data revealed no differences in the

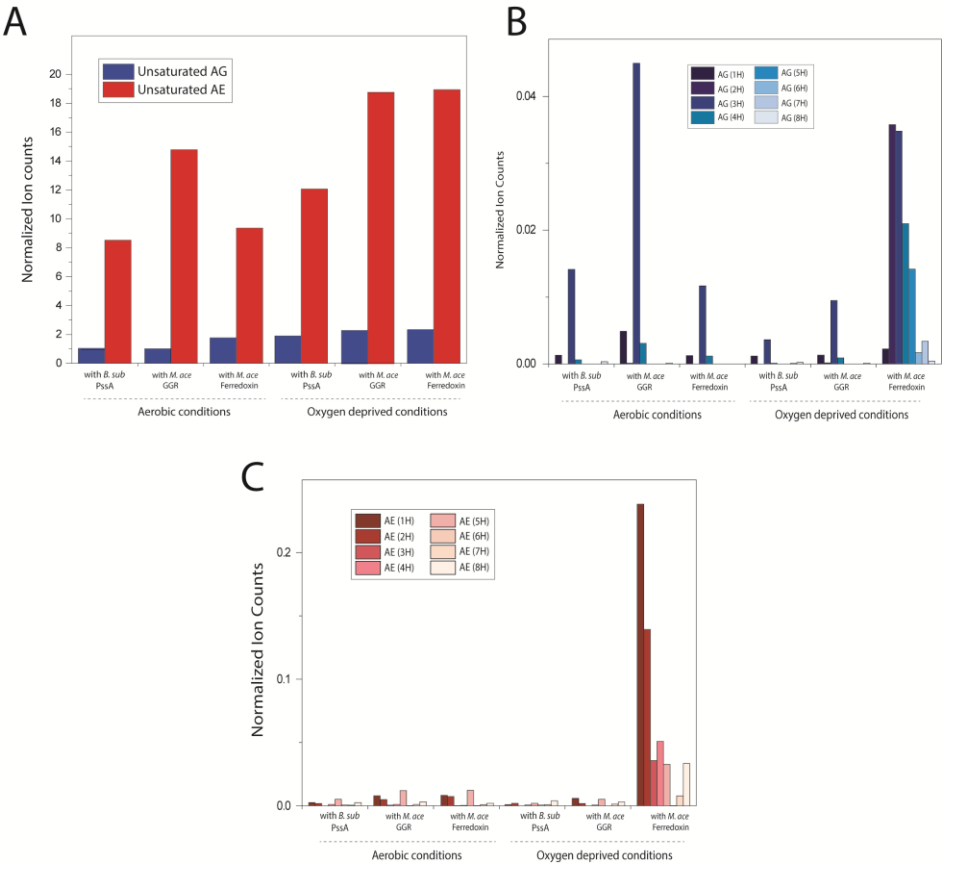


Figure 2 | *In vivo* saturation of the archaeal AG and AE lipids by GGR and ferredoxin from *M. acetivorans.* Total lipid analysis of three engineered *E. coli* strains harboring different ether lipids enzymes combinations. Growth and induction condition are indicated. (**A**) Two unsaturated AG and AE archaeal lipids. (**B**) Saturated AG species. (**C**) Saturated AE species. The H+ in the legend indicates the amount of hydrogens added to the substrate and therefore the number of the reduced double bonds. The LC-MS ion counts were normalized using Eicosane as internal standard.

amount of the two unsaturated archaeal species AG and AE in the two different tested induction conditions (**Figure 2A**). A detailed analysis on the saturated AG and AE species (**Figure 2B and C**) showed the presence of the two archaeal lipids with a different degree of double bonds saturation only in the *E. coli* strains containing all the ether lipids enzymes, including GGR and ferredoxin induced in an oxygen deprived condition.

However, incompletely saturated AG and AE were also observed in the other analyzed strains while the amounts of saturated lipid species were very low.

In vivo optimization of the saturation activity

The poor saturation activity of the GGR and Ferredoxin from *M. acetivorans* was optimized in order to produce higher amount of saturated AG and AE species in the engineered *E. coli* strains. To this end, two different IPTG concentration, 0.1 and 0.25 mM, were used to induce the two bacterial strains (with GGR and with GGR/Ferredoxin) at different growth phases (OD_{600} = 0.0, 0.2 and 0.5). Total lipid analysis of the membrane fraction of the tested strains revealed that induction of the *E. coli* strain harboring nine ether lipid enzymes (with GGR/Ferredoxin) with 0.1 mM of ITPG and at the beginning of growth yielded a lower amount of unsaturated AG and AE (**Figure 3A**) than the *E. coli* strain containing eight ether lipid genes (with GGR) induced under the same conditions. Induction with 0.25 mM of IPTG did not improve the saturation. When the LC-MS data was screened for the saturated AG and AE species, none of the conditions resulted in higher amounts of these lipid species (**Figure 3B and C**).

In vivo saturation of AG and AE by GGR and ferredoxin of *M. maripaludis*

Given the lower saturation activity of the GGR from the thermophilic archaea *M. acetovirans*, the same enzyme from the mesophilic archaea *M. maripaludis* was examined as mesophilic enzymes may have a higher activity at 37 °C, temperature at which *E. coli* grows optimally. The gene *MMARC5_RS06325* encoding for GGR and the upstream gene *MMARC5_RS06320* encoding for the ferredoxin protein were codon optimized for expression in the *E. coli* cells. The genes were cloned into a system of four compatible vectors and the *E. coli* strains harboring different combinations of ether lipid enzymes were grown and induced using the above described procedure. Induction at oxygen deprived

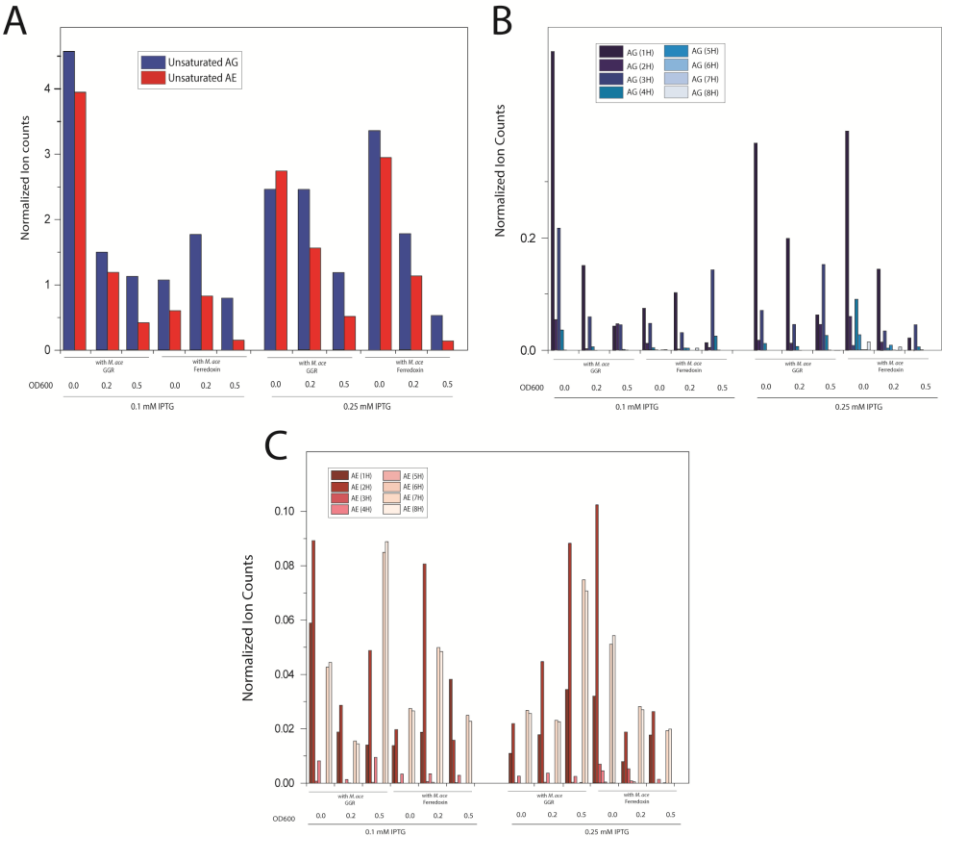


Figure 3 | *In vivo* optimization of the archaetidyl phospholipid saturation by GGR and ferredoxin from *M. acetivorans.* Total lipid analysis of engineered *E. coli* strains harboring all the ether lipid enzymes with GGR or with GGR and Ferredoxin. Different IPTG concentrations and growth phases of induction are indicated. (**A**) Unsaturated AG and AE archaeal lipids. (**B**) Saturated AG species. (**C**) Saturated AE species. The H+ in the legend indicates the amount of hydrogens added to the substrate and therefore the number of the reduced double bonds. The LC-MS ion counts were normalized using Eicosane as internal standard.

conditions lead to higher levels of unsaturated AG and AE production compared to the aerobic conditions (**Figure 4A**). In the presence of GGR and ferredoxin, a remarkable decrease of the two unsaturated AG and AE species was noted but this was not accompanied with an increase in the saturated AG and AE species (**Figure 4B and C**).

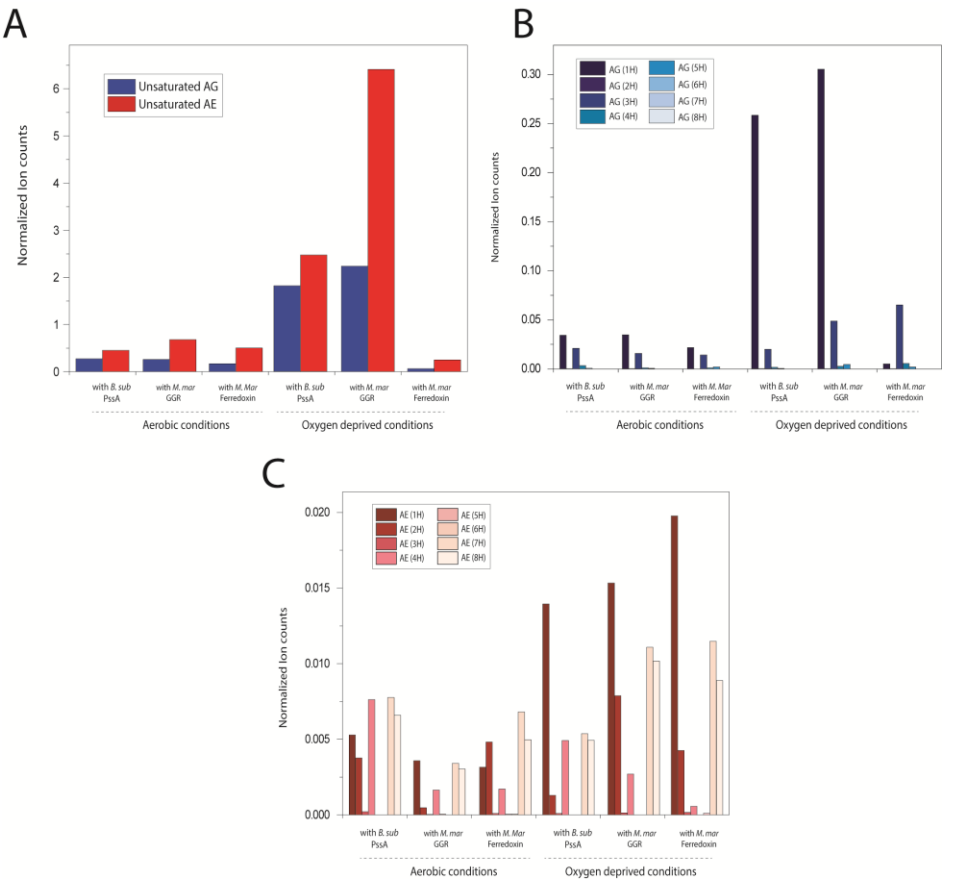


Figure 4 | *In vivo* saturation of the archaeal AG and AE lipids by GGR and ferredoxin from *M. maripaludis.* Total lipid analysis of three engineered *E. coli* strains harboring different ether lipid enzymes combinations. Growth and induction condition are indicated. (**A**) Unsaturated AG and AE archaeal lipids. (**B**) Saturated AG species. (**C**) Saturated AE species. The H+ in the legend indicates the amount of hydrogens added to the substrate and therefore the number of the reduced double bonds. The LC-MS ion counts were normalized using Eicosane as internal standard.

Discussion

Archaeal intermediates so far characterized in the heterologously expressed ether lipid biosynthetic pathway all contain double bonds in their hydrocarbon side chain. However, the mature archaeal lipid consists of fully saturated isoprenoid chains which is brought about by a hydrogenation step involving the GGR enzymes. Crystallographic studies

on various GGR enzymes [50,52,214] has shed some light on the interaction of the enzyme with the FAD cofactor and the putative substrate providing insights into the putative saturation mechanism. Different reducing agents were also tested in order to understand the enzymatic requirement of this biosynthetic step. Here, we aimed to characterize double bond saturation of the ether lipids synthesis in order to place this biosynthetic step in the archaeal lipid pathway and to generate fully saturated lipid species.

In vitro characterization using the purified GGR enzyme from *M. acetivorans* and *M. maripaludis* employing the chemically synthetized substrate DGGGP, GGPP and different reducing agents, such as NAD(P)H, sodium dithionite, purified Ferredoxin however failed to show any reductase activity towards DGGGP and GGPP (data not shown). Therefore, the enzymes were expressed *in vivo* to exam saturation of AG and AE. By the introduction of the GGR and ferredoxin enzymes in the system of seven archaeal and bacterial enzymes [163], double bonds saturation could be demonstrated in *E. coli*. The lipid analysis with the GGR and ferredoxin from the hyperthermophilic archaea *M. acetivorans* revealed a low saturation activity of AG and AE. Despite the low activity, the presence of the archaeal AG and AE with different number of saturated double bonds could be detected by LC-MS. The abundance of these species was remarkably higher when ferredoxin was also present in the *E. coli* strain compared to GGR alone. Thus, this result underlines the biological role of ferredoxin as reducing agent in the double bonds saturation mechanism, confirming previous studies [53]. GGR and Ferredoxin from the mesophilic archaea *M. maripaludis* were also investigate to attempt to a higher enzymatic activity. Again, only partially saturated archaeal lipids were detected. Further studies are required to determine how GGR saturates its substrates and to assess at what stage of lipid biosynthesis isoprenoid chain saturation occurs.

Experimental procedures

Cloning procedures and bacterial strain

E. coli BL21 was used for the expression of the entire ether lipid biosynthetic pathway. The primers and plasmids used to clone the investigated enzymes are listed in the **Table 1** and **2**.

Table 1. Expression vectors used in the present study.

Plasmid	Description	Reference
pRSF-Duet-1	Cloning and expression vector (Kan^R), T7 promoter	Novagen
pET-Duet-1	Cloning and expression vector (Amp^R), T7 promoter	Novagen
pCDF-Duet-1	Cloning and expression vector (Str^R), T7 promoter	Novagen
pACYC-Duet-1	Cloning and expression vector (Cm^R), T7 promoter	Novagen
pSJ135	*ispA* gene with a double mutation Y79H and S140T, the *idi* gene, both with a N-terminal His-tag from *E. coli K12* cloned into pCDF-Duet vector using the primers 62, 63, 24 and 57	[163]
pSJ138	Synthetic gene encoding codon optimized GGGP synthase from *M. maripaludis* with N-terminal His-tag, and the *B. subtilis araM* with C-terminus His-tag cloned into pET-Duet vector using the primers 70, 71, 11 and 12	[163]
pSJ140	Synthetic gene encoding codon optimized DGGGP synthase from *A. fulgidus* with N-terminal His-tag and redesigned ribosome binding site AGGACGTTAACAT, and a synthetic gene encoding codon optimized CDP-archaeol synthase from *A. fulgidus* with a C-terminus His-tag cloned into pRSF-Duet vector using the primers 32, 20, 84 and 86	[163]
pAC004	*B. subtilis pss* gene with C-terminal His-tag cloned into pACYC-Duet vector using the primers 89 and 90	[163]

pAC021	Synthetic gene encoding codon optimized GGR from *M. acetivorans* with N-terminal His-tag cloned into pAC004 vector using the primers 564 and 565	This study
pAC022	Synthetic gene encoding codon optimized ferredoxin from *M. acetivorans* with N-terminal His-tag cloned into pAC021 vector using the primers 547 and 548	This study
pAC032	Synthetic gene encoding codon optimized GGR from *M. maripaludis* with N-terminal His-tag cloned into pAC004 vector using the primers 591 and 592	This study
pAC033	Synthetic gene encoding codon optimized ferredoxin from *M. maripaludis* with N-terminal His-tag cloned into pAC032 vector using the primers 593 and 594	This study

Table 2. Oligonucleotide primers used in the study.

Primers name	Primer sequence 5' ⟶ 3'	Restriction site
11	GCGCGAATTCATGCATCACCACCACC	EcoRI
12	GCGCAAGCTTTCATTTTTTGGACAGC	HindIII
20	GCGCCTCGAGGACAGGTTTCCCGACTGGAAAG	XhoI
24	GATATACCATGGGCAGCCATCACCATC	NcoI
32	GGCGCCATATGCTGGATCTGATTCTGAA	NdeI
57	GCGCGAATTCTTATTTATTACGCTGGATGATGTAG	EcoRI
62	CACTCATTAATTCATGATGATTTACCGGCAATGG	blunt
63	AGCGTGGATACACTCAACGGC	blunt
70	GCGCCATATGAATCGTATCGCAGCTGAC	NdeI
71	GCGCCTCGAGTTAGTGATGATGGTGGTGATGTTCATATAGACCA TGATCAGCG	XhoI
84	GCCGCCATGGGTAGTCATCATCACCACCATC	NcoI
86	GCGCGAATTCTTAGAATGCACCGGCGA	EcoRI
89	GCGCCATATGAATTACATCCCCTGTATGATTACG	NdeI
90	GCGCCTCGAGTTAGTGATGGTGATGGTGGTGATGATGATTCCAT AGACTCCAG	XhoI
564	CGCGGAGCTCATGAAAGACATTTACGACGTGCTGG	SacI
565	CGCGGTCGACTTAGTGATGGTGGTGATGGTGGTGATGCG	SalI
547	CGCGCCATGGTCTGGCACTACAC	NcoI

548	CGCGGGATCCTCAGTGATGGTGGTGATG	BamHI
591	CGCGGAGCTCATGCGTGCACTGAATAACAG	SacI
592	CGCGGTCGACTTAGTGATGGTGATGGTGGTGATG	SalI
593	CGCGGCGGCCGCATGAAAGTGAACTATAACAAATGC	NotI
594	CGCGCTTAAGTTAATGATGGTGATGGTGGTGATG	AflII

Bacterial growth and lipid analysis

Engineered *E. coli* strains were grown at 37 °C in LB medium, supplemented with the required antibiotics: ampicillin (50 µg/ml), kanamycin (50 µg/ml), streptomycin (50 µg/ml) and chloramphenicol (34 µg/ml), 0.2% glucose and 1 mM $NiCl_2$ were added when necessary. Cells were aerobically grown and induced with 0.25 mM of IPTG for 3.5 hours [163]. When specified, the strains were aerobically grown until exponential phase (OD_{600} = 0.6) and induced under oxygen deprived condition. Membrane isolation and lipid extraction was performed as reported previously [163]. Lipids were analyzed by LC-MS using an Accela1250 HPLC system coupled with an ESI-MS Orbitrap Exactive (Thermo Fisher Scientific) as described [163].

Chapter 7

Summary

Samenvatting

Sommario

Archaea are a group of single-celled prokaryotic organisms that represent one of the three domains of life together with Bacteria and Eukarya. These microorganisms were originally discovered in extreme environments such as hot and acidic springs, high salt and pressure conditions [3,215]. One of the important features that provides archaeal organisms with the capability to survive harsh conditions is their unique structure of the cellular membrane lipids. Archaeal lipids differ from bacterial and eukaryotic counterparts in three main aspects. Phospholipids in archaea are composed of highly methylated-branched isoprenoid chains linked via ether bonds to glycerol-1-phosphate as the glycerophosphate backbone. This differs from the bacterial and eukaryal phospholipids which consist of straight fatty acid chain esters linked to the enantiomeric glycerol-3-phosphate. Within the class of archaeal lipids there is great structure diversity with respect to the length, composition and configuration of the isoprenoid side chains. The most common core lipid in archaeal is *sn*-2,3-diphytaniylglycerol diether, generally called archaeol. This molecule can undergo several modifications, including hydroxylation and condensation. In fact a tail-to-tail condensation of two diether lipid molecules is one of the most frequent and functionally important structural variations that leads to the glycerol-dialkylglycerol tetraether lipid, known as caldarchaeol. This core lipid can exist with several modification such as the presence of penta- or hexacyclic rings depending on the archaeal species [72,73]. **Chapter 1** provides an introduction to the thesis and discusses the structure of the archaeal membrane lipids, their biosynthesis and changes in composition in response to environmental stress. A different degree of cyclopentane rings or an accurate tuning of the diether, macrocyclic diether and tetraether lipids ratio are some of the mechanisms that archaea use to maintain the proper membrane fluidity in response to temperature changes [76,77]. The described structural variability found in archaeal lipids is underlined by the existence of different biosynthetic genes which are functionally flexible, while the entire biosynthetic pathway can be considered as a complex multienzymatic and partially membrane-bound system.

In the past years, the archaeal lipid biosynthetic pathway has been extensively described and characterized [2]. In **Chapter 1,** the six

biochemical steps that lead to the final saturated archaeal lipid from isoprenoid building blocks are described in details. In archaea, isopentenyl pyrophosphate (IPP) and dimethylallyl pyrophosphate (DMAPP) synthesis proceeds through the mevalonate pathway and their sequential condensation via a characteristic polyprenyl synthase, geranylgeranyldiphosphate (GGPP) synthase, yielding the C-20 GGPP compound. The archaeal glycerol-1-phosphate (G1P) configuration is then conferred by an archaeal specific enzyme, G1P dehydrogenase, which is not evolutionally and functionally related to the bacterial G3P dehydrogenase enzyme. The two synthetized compounds, GGPP and G1P, are then linked together via two ether bonds, steps catalyzed by a cytoplasmic protein, geranylgeranylglycerol phosphate (GGGP) synthase and a membrane associated protein di-geranylgeranylglycerol phosphate (DGGGP) synthase. The following biosynthetic step concerns the formation of a CDP-activated precursor via a CTP-transferring reaction brought about by an archaeal enzymes named CDP-archaeol synthase (CarS).

Chapter 2 in detail describes the synthesis of CDP-archaeol, a central step in biosynthesis as it is the precursor for polar head group attachment. In Bacteria and Eukarya similar reactions take place where the CDP-diacylglycerol synthase (CdsA) links CTP to phosphatidic acid leading to the formation of CDP-diacylglycerol (CDP-DAG) [100,101]. Bioinformatics analysis revealed a remote homology of a previously uncharacterized archaeal enzyme with the bacterial CdsA [4]. This protein, now termed CarS, is universally conserved in archaea. A biochemical characterization of the archaeal CarS protein was performed. The corresponding gene from *Archaeoglobus fulgiduds* was expressed in *E. coli*, purified and tested in *in vitro* reactions towards the chemically synthetized substrate DGGGP. Only in presence of CTP and divalent ions (Mg^{2+}), CDP-archaeol formation could be detected establishing the role of CarS as a CDP-archaeol synthase. The discovery of the key enzyme enabled the reconstitution of the entire archaeal lipid biosynthesis pathway in a cell free system. By the use of a combination of purified bacterial and archaeal enzymes, CDP-archaeol could be synthetized *in vitro* from the simple isoprenoid building blocks.

CDP-archaeol is an important intermediate in the lipid biosynthesis pathway. In the next biosynthetic step, the CDP group is replaced by a

polar head group such as L-serine, ethanolamine, glycerol or *myo*-inositol. Enzymes that catalyze polar head group attachment are describe in details in **Chapter 3**. The enzymes responsible for these catalytic steps exhibit a high sequence homology and share a common mode of action in all three domain of life, being members of the same protein superfamily, CDP-alcohol phosphatidyltransferase [7,43]. In bacteria the replacement of the CDP group with glycerol is catalyzed by the two enzymes PgsA and PgpA leading to the formation of phosphatidyglycerol (PG). Because of the functional homology the substrate specificity of the two bacterial enzymes was tested towards archaeal substrates. In the presence of CDP-archaeol the archaeal lipid archaetidylglycerol (AG) was formed demonstrating a promiscuity towards bacterial and archaeal lipid substrates. In contrast to AG formation, archaetidylethanolamine (AE) formation required the presence of a PssA enzyme from a different bacteria than *E. coli*. PssA enzymes are classified in two groups: group-I contains the enzymes from Gram-negative bacteria (*E. coli*) and group-II includes the PssA enzymes from Gram-positive bacteria (*B. subtilis*), yeast and archaea [148]. Previous studies demonstrated that PssA from *E. coli* is highly selective for CDP-diacylglycerol while PssA from *B. subtilis* likely also recognizes archaeal-like substrates [47]. Purified PssA from *B. subtilis* was biochemically tested together with the bacterial phosphatidylserine decarboxylase, Psd, for the synthesis of the archaeal lipid AE. The biochemical characterization of these enzymes enabled the reconstitution of the entire archaeal lipid synthesis *in vitro* starting from the two isoprenoid building blocks until AG and AE formation. Furthermore, conditions were defined for AG and AE production *in vivo* using the bacterium *E. coli* as a host. By introduction of seven genes, indeed AG and AE could be synthesized but at very low levels (less than 1% of total lipid).

The aforementioned characterization of the diether phospholipid pathway enzymes, provided a novel design for the engineering of *E. coli* for high level ether lipid production and is presented in **Chapter 4**. To generate higher levels of isoprenoid building blocks in *E. coli*, the endogenous MEP-DOXP pathway was upregulated via chromosomal integration of the MEP-DOXP operon in the bacterial genome. This led to a higher production of IPP and DMAPP as evidenced by the production of

lycopene. Upon the co-expression of the heterologous ether lipid enzymes the engineered *E. coli* strain produced up to 32% of AG. This meant a nearly complete replacement of the endogenous PG pool by the archaeal AG demonstrated the functional integration of the archaeal lipid biosynthetic pathway into the bacterial host yielding a viable cell with a mixed heterochiral membrane. The proper stereoconfiguration of the archaeal AG was validated by chiral-selective NMR, but even in the absence of the G1P dehydrogenase, chiral correct AG was produced. Remarkably, this must mean that *E. coli* is capable of G1P formation, a process that has not been studied or realized before. We also noticed a weak substrate selectivity of the GGGP synthase for G1P over G3P, thus potentially GGGP might be synthesized with an incorrect chirality but this was not observed *in vivo*, likely because of the overall selectivity of the pathway. The observation that GGGP synthase is less selective than expected sheds some new light on the lipid divide and the segregation between bacteria and archaea. While the analogous reaction in bacteria is catalyzed by the highly selective PlsB enzyme, the poorer chiral selectivity of GGGPS suggests that primordial presence of non-stereoselective enzymes in the common ancestor from which bacterial organisms evolved and diverged upon the appearance of higher stereoselective enzymes as PlsB.

The presence of archaeal lipids altered the cell growth, cell morphology and conferred a certain degree of resistance to environmental stress. The cells with a mixed heterochiral membrane show a long lag phase before growth commenced with nearly the same growth rate as wild-type *E. coli*. This suggests that the cells undergo an adaptation process to deal with the altered membrane composition. However, genome sequencing did not reveal any significant change or mutation where this property was not stable as upon the transfer to fresh medium, growth commenced only after a long lag phase. The majority of the engineered cells show elongated and thinner cell shape, and with higher induction levels of the pathway also the presence of appendages and extrusions. These structures also contain DNA and they appear to arise from an aberrant division mechanism that does no longer occur at mid-cell. This phenomenon seems to be associated with a high level production of lipids, and is not apparent in the cells in which the highest level of AG formation was seen, that occurs at low inducer

concentration. Finally, the presence of archaeal lipids conferred to the bacterial cells an increased tolerance to heat and freezing exposure and resistance against butanol treatment. Although the latter features are subtle, they are significant and apparently overcome possible increased susceptibility by the expression of the many foreign enzymes.

In **Chapter 5** the catalytic mechanism of one of the key enzymes of the archaeal lipid biosynthetic pathway is investigated. To this end, the crystal structure of the CarS protein from the archaeon *Aeropyrum pernix k1* was solved in a CTP- and Mg^{2+}-bound complex at 2.38 Å. Structural characterization revealed an unusual fold of five transmembrane segments forming a hydrophilic pocket at the cytoplasmic side of the membrane which is defined as the CTP-binding site. Mutagenesis of the amino acid residues involved in the CTP interaction resulted in a loss in CTP binding and CDP-archaeol formation. The crystal structure further identifies a certain flexibility of the transmembrane segment 5 consistent with its predicted role in glycerophospholipid binding. Thus, given the high conservation of the CTP-binding domain among the members of the CTP-transferase family, the characterization of CarS provided insights into the CTP recognition mechanism. However, further studies are required to completely understand the CarS enzymatic mechanism, and in particular glycerophospholipid binding and product release.

Finally, the possible mechanism of double bond hydrogenation of the archaeal lipid with an unsaturated isoprenoid chain is described in the **Chapter 6**. The mature archaeal lipids consist of fully saturated isoprenoid chain and saturation is catalyzed by the enzyme geranylgeranyl reductase (GGR). Despite the availability of crystal structures of few archaeal GGRs [52,211], the exact reductase mechanism and the stage at which double bond reduction occurs are unclear. Two GGRs from two different archaea were tested for their reductase activity in conjunction with an associated ferredoxin. However, the *in vivo* reduction of the archaeal phospholipids produced in the engineered *E. coli* cell was poor and thus non-conclusive with respect to the stage at which saturation occurs. The requirement for ferredoxin as biological cofactor was instead confirmed. Further study are required to elucidate the GGR saturation mechanism and to determine at

what stage of the ether lipid pathway the hydrogenation of the isoprenoid chain occurs.

To summarize, the presented thesis describes the reconstitution of the complete ether lipid biosynthetic pathway both *in vitro* and *in vivo*. The study identified a novel enzyme in the archaeal lipids biosynthesis, namely the CDP-archaeol synthase and revealed a high promiscuity of bacterial enzymes for the attachment of the polar head groups. The work also sheds some light on the significance of some evolutionary theories. The obtainment of a bacterial strain with a heterochiral mixed membrane showed the coexistence of the two different lipid species contrasting the previous hypothesis of chemical instability of such membrane which pushed the evolution of bacterial and archaeal organisms towards a more stable homochiral membranes. Thus, the coexistence of bacterial and archaeal lipids in a living cell suggested the presence of a common ancestor with a mixed membrane from which organisms evolved. Further, the different substrate specificity found in the archaeal GGGPS and in the bacterial PlsB raises new possibilities about the initial insurgence of archaeal organisms, followed by the differentiation into bacteria. In this way the trigger that caused the segregation from the primordial cell can be seen as the appearance of stereoselective enzymes such as PlsB which pushed the differentiation of bacterial organisms from the ancient cell. The latter one then further evolved towards an archaeal organism with the conserved capability to live in extreme environments but with an evolved and specific membrane lipid composition.

Further, the work has potential biotechnological implications as the re-programming of the lipid composition of *E. coli* resulted in some increase in robustness. Thus, when further tuned, this may be seen as a general methodology for other microorganisms of industrial relevance when aspects of solvent tolerance are addressed. It would not be possible to inflict such radical alteration of the lipidome by mutagenesis and selection. However, still some important steps of the archaeal lipid biosynthetic pathway, such as double bond reduction, tetraethers formation and sugar attachment remain unclear and need to be studied in order to elucidate the mechanisms of archaeal lipid synthesis.

Archaea behoren tot de groep van prokaryoten organismen die, gezamenlijk met de Bacteriën en Eukaryoten, de drie domeinen van het leven representeren. Deze eencellige micro-organismen zijn oorspronkelijk ontdekt op plaatsen waar extreme omstandigheden heersen, zoals in de diepzee bij hoge druk, hoge zoutconcentraties en in hete en zure of alkalische bronnen [3,215]. Een van de belangrijkste eigenschappen die de overleving van archaea onder zulke moeilijke omstandigheden mogelijk maakt is de unieke structuur van de lipiden in hun cellulaire membraan. De lipiden van Archaea onderscheiden zich op tenminste drie vlakken van hun bacteriële en eukaryote analogen. Fosfolipiden in archaea bevatten methyl-vertakte isoprenoide ketens, die via een ether-binding gekoppeld zijn aan een glycerolfosfaat ruggengraat waarbij glycerol-1-fosfaat gebruikt wordt als precursor tijdens de biosynthese. Dit verschilt van de bacteriële en eukaryote fosfolipiden die uit rechte vetzuurstaart esters bestaan die gekoppeld zijn aan het enantiomere glycerol-3-fosfaat. Binnen de groep van archaeale lipiden is er een grote structurele verscheidenheid met betrekking tot de lengte, samenstelling en configuratie van de isoprenoide ketens. Het meest voorkomende kern lipide in archaea is sn-2,3-diphytanylglycerol di-ether, beter bekend als archaeol. Dit molecuul kan verschillende modificaties ondergaan, waaronder hydroxylatie en condensatie. Een tail-to-tail condensatie van twee di-ether lipide moleculen is een van de meest voorkomende en functioneel belangrijke structurele variaties die leidt tot het glycerol-dialkylglycerol tetra-ether lipide, beter bekend als caldarchaeol. Afhankelijk van de archaea soort komt dit lipide voor met verscheidene modificaties, waaronder de aanwezigheid van penta- of hexacyclische ringen en heeft het de bijzondere eigenschap dat het membraanomspannend is [72,73]. **Hoofdstuk 1** begint met een introductie van het proefschrift en beschrijft de structuur van de membraanlipiden in Archaea, het mechanisme van hun biosynthese en veranderingen in de samenstelling in het membraan als gevolg van veranderende omgevingscondities. Bijvoorbeeld, bij toenemende temperstuur, neemt het aantal cyclopentaan ringen in de isoprenoide ketens toe of verandert de ratio van di-ether, macro-cyclische di-ether en tetra-ether lipiden. Dit zijn een aantal mechanismen die archaea gebruiken

om de juiste membraan vloeibaarheid te behouden bij veranderende omgevingstemperatuur [76,77]. De gehele biosynthese route kan worden beschouwd als een complex meervoudig enzymatisch en gedeeltelijk membraan gebonden systeem.

Gedurende de afgelopen 20 jaar is er veel onderzoek gedaan aan het mechanisme van lipide biosynthese in archaea en is deze route uitgebreid beschreven en gekarakteriseerd [2]. In **Hoofdstuk 1** worden de zes enzymatische stappen die leiden tot de uiteindelijke vorming van verzadigde lipiden uitgaande van simpele isoprenoide bouwstenen in detail beschreven. Isopentenyl pyrofosfaat en dimethylallyl pyrofosfaat zijn de precursors voor de synthese van de isoprene keten die aangemaakt worden via de mevalonaat route. De condensatie van deze precursors vindt plaats via het karakteristieke polyprenyl synthase, geranylgeranyldifosfaat synthase, en dit resulteert uiteindelijk in de productie van C-20 GGPP. Glycerol-1-fosfaat wordt aangemaakt door een archaea specifiek enzym, het G1P dehydrogenase, dat evolutionair en functioneel niet gerelateerd is aan de familie van bacteriële G3P dehydrogenases. Twee moleculen GGPP en G1P worden daarna aan elkaar gekoppeld via 2 ether-verbindingen, stappen die achtereenvolgend gekatalyseerd worden door het cytoplasmatisch eiwit geranylgeranylglycerol fosfaat synthase en een membraan geassocieerd eiwit di-geranylgeranylglycerol fosfaat synthase. De volgende stap in de biosynthese betreft de vorming van een CDP-geactiveerde precursor waarbij er CTP overdracht plaatsvindt, een reactie die wordt uitgevoerd door het veronderstelde CDP-archaeol synthase. De identiteit van dit enzym was bij aanvang van dit promotieonderzoek nog niet bekend. Echter uit eerder onderzoek was al wel bekend dat deze activiteit bestaat daar deze werd eerder aangetoond middels biochemische studies gebruikmakende van cellulaire extracten van Archaea.

Hoofdstuk 2 beschrijft de identificatie en de karakterisatie van het CDP-archaeol synthase (CarS) dat noodzakelijk is voor de synthese van CDP-archaeol. Dit is een centrale stap in de biosynthese van ether lipiden, doordat het de precursor voor kopgroep-aanhechting betreft. In Bacteria en Eukaryoten vinden vergelijkbare reacties plaats, waarbij het CDP-diacylglycerol synthase (CdsA) CTP koppelt aan fosfaatzuur, wat leidt tot

de vorming van CDP-diacylglycerol [100,101]. Een bio-informatische analyse onthulde een verre homologie van een nog niet gekarakteriseerd archaeal enzym met het bacteriële CdsA [4]. Dit eiwit, vanaf nu CarS genoemd, is universeel geconserveerd in archaea. Het CarS eiwit werd biochemisch gekarakteriseerd. Het overeenkomende gen van *Archaeoglobus fulgidus* werd tot expressie gebracht in *E. coli,* vervolgens gezuiverd en middels chemisch gesynthetiseerde substraat DGGGP in een *in vitro* reactie getest voor de veronderstelde activiteit. Alleen in de aanwezigheid van CTP en divalente ionen (Mg^{2+}), kon de vorming van het CDP-archaeol gedetecteerd worden. Met dit experiment kon eenduidig worden vastgesteld dat CarS als CDP-archaeol synthase functioneert. De ontdekking van dit sleutelenzym maakte de weg vrij om de gehele archaeale lipide biosynthese route te reconstitueren in een cel vrij systeem. Door een combinatie van gezuiverd bacteriële en archaeale enzymen te gebruiken, kon CDP-archaeol worden gesynthetiseerd vanuit de eerder genoemde simpele isoprenoide bouwstenen.

CDP-archaeol is een belangrijk tussenproduct in de ether lipide biosynthese route. In de volgende stappen wordt de CDP groep vervangen voor een polaire kopgroep, zoals serine, glycerol, ethanolamine of myo-inositol. Enzymen die deze polaire kopgroep koppeling katalyseren worden in detail beschreven in **Hoofdstuk 3**. Deze enzymen beschikken over een zeer hoge sequentie homologie en ze delen een gemeenschappelijke modus operandi in alle domeinen van het leven. Ze behoren tot dezelfde eiwit superfamilie, de CDP-alcohol fosfatidyltransferases [7,43]. In bacteriën wordt de substitutie van de CDP groep door glycerol gekatalyseerd door de twee enzymen PgsA en PgpA, waarbij de vorming van fosfatidylglycerol (PG) verloopt via het tussenproduct fosfatidylglycerolfosfaat (PGP). Vanwege de hoge mate van homologie tussen de bacteriële en archaeale enzymen, is de substraat specificiteit van de twee bacteriële enzymen getest met CDP-archaeol. In deze reacties werd het archaeale lipide archaeatidylglycerol (AG) gevormd, wat erop duidt dat de bacteriële enzymen weinig specifiek zijn voor het type lipide. In tegenstelling tot de vorming van AG, bleek het *E. coli* PssA enzym dat betrokken is bij de vorming van fosfatidylethanolamine niet in staat om CDP-archaeol te herkennen als substraat. Voor de vorming van

archaeatidylethanolamine (EA) was daarom een PssA eiwit uit een andere bacterie nodig. PssA enzymen vormen fosfatidylserine uit diacylglycerol en serine en ze worden geclassificeerd in twee groepen: Groep-1 omvat enzymen van Gram-negatieve bacteriën (zoals *E. coli)* en groep-II omvat de PssA enzymen van Gram-positieve bacteriën (zoals *B. subtilis*), gisten en archaea [148]. Zodoende bestaat er de mogelijkheid dat het PssA van *B. subtilis* ook CDP-archaeol herkent [47]. Gezuiverd PssA van *B. subtilis* in combinatie met het *E. coli* fosfatidylserine decarboxylase, Psd, was inderdaad in staat om AE uit CDP-archaeol en serine te vormen. Deze waarneming maakte het mogelijk de gehele archaeale lipide synthese *in vitro* na te bootsen beginnende met de twee isoprenoide bouwstenen tot aan de vorming van AG en AE. Tevens kon nu de minimale set van enzymen worden gedefinieerd waarmee AG en AE gevormd kunnen worden in levende *E. coli* cellen. Door zeven nieuwe genen in *E. coli* te introduceren kon daadwerkelijk de productie van kleine hoeveelheden AG en AE worden aangetoond (minder dan 1% van het totale lipide gehalte).

De eerdergenoemde di-ether fosfolipide route enzymen zijn tevens tot expressie gebracht in een *E. coli* stam met een verhoogde isoprenoide productie zoals beschreven in **Hoofdstuk 4**. Om meer isoprenoide bouwstenen te generen in *E. coli*, is een additionele kopie van de endogene MEP-DOXP route via chromosomale integratie in het *E. coli* genoom geplaatst. Dit resulteerde in een verhoogde productie van IPP en DMAPP, wat mede bleek uit een verhoogde productie van het pigment lycopeen. Door de co-expressie van de heterologe ether lipide enzymen bleek het mogelijk om in deze *E. coli* stam veel AG te maken en wel tot op 32% van het totale fosfolipide. Dit betekende een vrijwel complete vervanging van de endogene PG-pool voor het archaeale AG. Door de functionele integratie van de archaeale lipide biosynthese route in de bacteriele gastheer te realiseren was het voor het eerst mogelijk om het effect van een gemengd heterochirale membraan op de levensvatbaarheid van een bacteriële cel te onderzoeken en daarmee van evolutionaire hypothese te testen die verondersteld dat archaea en bacteriën voortgekomen zijn uit een gemeenschappelijke voorouder (“last universal common ancestor” - LUCA) met een gemengd heterochirale membraan. Deze hypothese verondersteld dat een gemengd heterochirale membraan niet stabiel is wat vervolgens

heeft geresulteerd in de 'lipid divide'; de splitsing in bacteriën en archaea elk met een homochirale membraan. Alhoewel een dergelijke evolutionaire hypothese moeilijk op experimentele wijze te testen, kan het in vivo model wel een antwoord gegeven op de vraag in hoeverre een gemengde heterochirale membraan inderdaad instabiel is. De juiste stereo-configuratie van het AG in de *E. coli* cellen kon gevalideerd worden middels chiraal-selectieve NMR, maar zelfs zonder de inbreng van het G1P dehyrdogenase, werd chiraal correcte AG geproduceerd. Dit is een opmerkelijke waarneming en houd in dat *E. coli* in staat is G1P aan te maken, een proces dat nog niet eerder is bestudeerd. Daarnaast is ook een zwakke substraat specificiteit van het GGGP synthase waargenomen voor G1P ten opzichte van G3P maar dit heeft geen effect gehad op de biosynthese omdat alleen de correcte chiraliteit is waargenomen. Dit laatste is mogelijk het gevolg van de algehele selectiviteit van de biosynthese route waardoor het niet mogelijk is G3P-gebasseerd AG te synthetiseren in de latere biosynthese stappen. De waarneming dat GGGP synthase minder selectief voor het glycerolfosfaat is opmerkelijk en geeft nieuw inzichten in de moleculaire basis van de lipide splitsing en scheiding tussen bacteriën en archaea. Waar de analoge reactie in bacteriën wordt gekatalyseerd door het zeer selectieve PlsB enzyme, suggereert de minder chirale selectiviteit van GGGPS dat de gemeenschappelijke voorouder van de bacteriën en archaea LUCA over niet-stereoselectieve enzymen beschikte en dus mogelijk in staat was meerdere chiraliteiten te realiseren van de glycerolfosfaat gebaseerde lipiden. Vervolgens is de bacteriële lijn verder geëvolueerd en afgesplitst en in dit proces zijn er meer stereoselectieve enzymen ontstaan zoals het eerder genoemde PlsB eiwit.

De aanwezigheid van archaeale lipiden verandert de celgroei en morfologie en draagt bij aan een versterkte weerstand ten opzichte van omgevingsfactoren. Cellen met een gemengd heterochirale membraan vertonen eerst een lange lag-fase waarna groei wordt ingezet met vrijwel dezelfde groeisnelheid als de wilde type cellen. Dit suggereert dat de cellen een aanpassingsproces ondergaan mogelijk om de niveaus van de ingebrachte enzymen aan te passen en/of om met de veranderde membraansamenstelling overweg te kunnen. Echter, uit de genoomsequentie bleek dat er geen significante veranderingen of mutaties

plaats hebben gevonden gedurende de lag fase. Ook na overplaatsing naar een nieuwe medium wordt er opnieuw, zij het een verkorte, lag-fase in de groei waargenomen. De AG-bevattende cellen vertonen ook morfologische veranderingen. De cellen zijn met name langer terwijl bij een sterkere inductie van de ether lipide biosyntheses, ook uitstulpingen en afsplitsingen worden waargenomen. Deze laatstgenoemde structuren bevatten zelfs DNA en ze lijken te ontstaan ten gevolge van een afwijkend celdelingsmechanisme. Deling vindt niet langer plaats in het midden van de cel maar op willekeurige plaatsen. Daar dit alleen in significante mate optreedt na een sterk verhoogde lipidenproductie en het niet zichtbaar is in de cellen met de grootste hoeveelheid AG, lijkt het celdelingsdefect veroorzaakt te worden door de hoge lipiden overproductie. Mogelijk is dit proces dus niet specifiek en wordt een soortgelijk fenomeen waargenomen indien de bacteriële lipiden worden overgeproduceerd. Immers normaliter zijn groei, lipide biosynthese en celdeling gekoppelde processen. Door de introductie van de ether lipid biosynthese genen en de hoge inductie van lipide biosynthese, wordt dit proces mogelijk ontkoppeld van celdeling en vindt er een verstoorde deling plaats. Tot slot, de aanwezigheid van de ether lipiden draagt bij aan een verhoogde tolerantie van de bacteriële cellen voor blootstelling aan hitte en vrieskou, alsmede een verhoogde weerstand tegen butanol behandeling. Alhoewel de verbetering van deze eigenschappen subtiel lijken, zijn deze waarneming wel degelijk significant en doen ze klaarblijkelijk mogelijk negatieve effecten van de hoge expressie van de soortvreemde enzymen teniet.

In **Hoofdstuk 5** is de structuur en het katalytische mechanisme van CarS onderzocht. Het CarS eiwit van *Aeropyrum pernix k1* is gezuiverd en gekristalliseerd in de CTP- en Mg^{2+} gebonden toestand. De structuur kon worden opgehelderd met een resolutie van 2.38 Å. Het eiwit bestaat uit vijf membraanomspannende α–helices met een hydrofiele groeve aan de cytoplasmatische zijde van het membraan, die fungeert als CTP-bindende plaats. Mutagenese van de aminozuurresiduen betrokken bij de binding van CTP resulteerde in een verlies in CTP binding en de daar aan gekoppelde enzymatische activiteit gemeten als de vorming van CDP-archaeol. Het membraan deel is op een unieke wijze gevouwen met een flexibel membraan helix 5 dat gepositioneerd is tegen de veronderstelde

glycerofosfolipide bindingsplaats. Afgaande op de hoge conservering van het CTP-bindingsdomein in de CTP-transferase familie, heeft de verdere biochemische karakterisering van het CarS verdere inzichten in het CTP-herkenningsmechanisme opgeleverd. Opmerkelijk is de waarneming dat CarS structureel zeer verwant is met een gedeelte van het bacteriële CDP-DAG synthetase CdsA terwijl deze enzymen op basis van aminozuurvolgorde slechts beperkt verwant zijn. Er zal echter meer onderzoek nodig zijn om het complete enzymatische mechanisme van CarS en met name de binding van glycerofosfolipide en het loskoppelen van CDP-archaeol te begrijpen.

Tot slot, beschrijft in **Hoofdstuk 6** een mogelijke mechanisme van de verzadiging van de dubbele bindingen in de onverzadigde isoprenoid ketens van de archaeale lipiden. Deze lipiden worden normaliter volledig verzadigd in een proces dat wordt gekatalyseerd door het enzym geranylgeranyl reductase (GGR). Ondanks de beschikbaarheid van kristalstructuren van een aantal van deze archaeale GGR enzymen [52,211] is het exacte reductase mechanisme nog niet bekend. Ook is niet bekend gedurende welke stap in de biosynthese route de reductie van de dubbele bindingen plaatsvindt. Twee GRR enzymen uit verschillende archaea zijn getest op hun reductase activiteit in aanwezigheid van het bij de desbetreffende genen geassocieerde ferredoxin. Echter, de *in vivo* reductie van de ether lipiden gevormd in *E. coli* lijkt verrassend inefficiënt waardoor het niet mogelijk was om vast te stellen wanneer de verzadiging plaatsvindt. De experimenten bevestigde echter wel de noodzaak voor het ferredoxin als biologische co-factor. Er zal echter meer onderzoek nodig zijn om het GGR verzadigings mechanismen op te helderen en een meer efficiënte conversie te realiseren om deze enzymen te kunnen toepassen voor de verzadiging van ether lipiden in *E. coli*.

Samenvattend, dit proefschrift beschrijft de reconstitutie van de complete ether-lipide biosynthese route met gezuiverde enzymen *in vitro* alsmede de productie van deze lipiden *in vivo* in een bacteriele gastheer. Dit onderzoek heeft ondermeer geleid tot de identificatie van een nieuw enzym in de lipide biosynthese in archaea, het CDP-archaeol synthase, en toonde aan dat bacteriële enzymen in staat zijn diverse polaire kopgroepen aan het CDP-archaeol te koppelen wat een beperkte

selectiviteit suggereert. Het werk geeft ook nieuwe inzichten in bepaalde evolutionaire theorieën. Het werk toont eenduidig aan dat bacteriële stam met een gemengd heterochirale membraan levensvatbaar is en dat de twee chiraal verschillende lipide soorten naast elkaar kunnen bestaan. Dit maakt een eerdere hypothese minder waarschijnlijk die stelt dat chemische instabiliteit van dergelijke membranen heeft geleid tot het ontstaan van bacteriën en archaea met de veronderstelde stabielere homochirale membranen. Onze waarnemingen ondersteunen ook *in vitro* waarnemingen dat heterochirale membranen stabiel zijn. Het werk is echter in overeenstemming met het idee dat er een gemeenschappelijke voorouder moet zijn geweest met een gemengde membraan van waaruit bacteriën en archaea zijn geëvolueerd. De waargenomen beperkte stereoselectiviteit van het archaeale GGPS enzym voor glycerolfosfaat en de exclusieve selectiviteit van het bacteriële PlsB voor glycerol 3-fosfaat ondersteund de hypothese dat LUCA beschikte over beperkte stereoselectieve lipide biosynthese enzymen wat heeft bijgedragen aan de vorming van een gemengd heterochiral lipide membraan. Gedurende de daaropvolgende differentiatie naar bacteriën, is de stereoselectiviteit van de (ester fosfolipiden) route toegenomen. Dit proces heeft mogelijk ook bijgedragen aan de lipide splitsing. Tijdens de evolutie van archaea hebben mogelijk andere factoren bijgedragen zoals het kunnen overleven onder extreme omstandigheden wat specifieke eisen stelt aan de stabiliteit van de membraan.

Het werk heeft potentiele biotechnologische implicaties, doordat de herprogrammering van de lipide samenstelling van *E. coli* geresulteerd heeft in een verhoogde robuustheid. Alhoewel dit verder ontwikkelt zal moeten worden, kan deze methode mogelijk generiek gebruikt worden voor micro-organismen van industriële relevantie bijvoorbeeld om de tolerantie tegen organische oplosmiddelen te verhogen. Het is niet mogelijk om de in dit proefschrift radicale verandering van het lipidoom te bewerkstelligen via klassieke methoden van mutagenese en selectie en daarom heeft kan deze methodiek mogelijk radicalere verbeteringen opleveren dan de klassieke methodieken. Echter alvorens deze methode kan worden toegepast zullen een aantal nog onopgehelderde vragen moeten worden beantwoord zoals het mechanisme van reductie van

dubbele bindingen, de vorming van tetraether lipiden en de mechanismen van suiker koppeling als polaire kopgroep.

Il termine Archea indica un gruppo di organismi procariotici unicellulari che rappresentano uno dei tre regni della vita insieme a Batteri e Eucarioti. Tali microorganismi vennero inizialmente scoperti in ambienti con condizioni estreme di temperatura, aciditá, concentrazioni saline e pressione [3,215]. Una delle caratteristiche principali che fornisce agli archea la capacità di sopravvivere in condizioni così ostili è data dalla struttura unica dei lipidi che compongono la loro membrana cellulare. In archea i lipidi differiscono dai quelli presenti in batteri e eucarioti in tre aspetti principali. Essi sono infatti composti da catene isoprenoidi altamente ramificate con gruppi metilici, collegate tramite legami eterei alla struttura glicofosfatica, nota come glicerolo-1-fosfato. Tale struttura si differenzia da quella dei fosfolipidi batterici e eucariotici in quanto questi ultimi sono costituiti da catene lineari di acidi grassi, legate tramite legami esterei alla forma enantiomerica glicerolo-3-fosfato. Rispetto alla struttura lipidica di base, in archea è presente un grandissima varietà di lipidi in themrini di lunghezza, composizione e configurazione delle catene isoprenoidi laterali. Il nucleo lipidico più comune trovato negli organismi archea è *sn*-2,3-difitanilglicerolo, comunemente chiamato archeolo. Tale molecola può subire ulteriori modificazioni, inclusi processi di idrossilazione e condensazione. Infatti, una condensazione coda-coda di due molecole archeolo è una delle variazioni strutturali più frequenti e funzionalmente più importanti che comporta la formazione di una struttura lipidica glicerolo-dialchilglicerolo tetraetere, conosciuta come caldarcheolo. Tale nucleo lipidico può esistere in una varietà di strutture modificate dalla presenza, ad esempio, di anelli penta- o esaciclici dipendentemente dalla specie in cuisi trova [72,73]. Il **Capitolo 1** fornisce, pertanto, l'introduzione alla tesi, discutendone la struttura dei lipidi di membrane degli archea, la loro biosintesi e come la loro composizione all'interno della membrana viene cambiata in risposta agli stress ambientali. Un diverso numero di strutture pentacicliche o un'accurata regolazione del rapporto tra archeolo/archeolo macrociclico e caldarcheolo, come strutture lipidiche, rappresentano alcuni dei meccanismi che gli archea utilizzano per mantenere la giusta fluidità di membrana in risposta ai cambiamenti di temperatura [76,77]. La descritta varietà strutturale presente nei lipidi di questi organismi è enfatizzata

dall'esistenza di diversi geni biosintetici funzionalmente flessibili, mentre l'intera via biosintetico può essere considerata come un complesso multi-enzimatico e un sistema parzialmente associato a membrana.

In passato, la biosintesi lipidica in archea è stata estensivamente descritta e caratterizzata [2]. Nel **Capitolo 1**, sei passaggi biosintetici che portano alla formazione di lipidi completamente maturi e saturi a partire dai semplici costituenti di partenza, vengono descritti in dettaglio. In archea, la sintesi di isopentenil pirofosfato (IPP) e dimetilallil pirofosfato (DMAPP) procede attraverso la via del mevalonato. Tali unitá di base vengono condensate in modo sequenziale attraverso una caratteristica poliprenile sintasi, determinando la formazione di un composto a 20 atomi di carbonio, noto come GGPP. La configurazione tipica dei lipidi di archea, aventi glicerolo-1-fosfato (G1P), è conferita da un enzima archea-specifico, G1P deidrogenasi, il quale non è né evolutivamente né funzionalmente correlato all'enzima batterico G3P deidrogenasi. Pertanto, le due molecole sopradescritte, GGPP e G1P, una volta sintetizzate, vengono unite da due legami eterei formati da una proteina citoplasmatica, geranilgeranil-glicerolfosfato (GGGP) sintasi e da una proteina associata a nmembrana, di-geranilgeranilglicerolfosfato (DGGGP) sintasi. Il passaggio biosintetico successivo riguarda la formazione di un importante precursore metabolico, il quale viene attivato attraverso una reazione che trasferisce il gruppo CTP ad opera di un enzima noto come CDP-archeolo sintasi (CarS).

Il **Capitolo 2** descrive in dettaglio la sintesi di CDP-archeolo, un passaggio centrale nella biosintesi dato che tale metabolita rappresenta il precursore per l'attaccamento dei gruppi polari. In batteri e eucarioti è possibile trovare una reazione molto simile in cui l'enzima CDP-diacilglicerolo sintasi (CdsA) trasferisce CTP sull'acido fosfatico con la conseguente formazione di CDP-diacilglicerolo (CDP-DAG) [100,101]. Analisi bioinformatiche hanno messo in evidenza la presenza di una remota omologia tra l'enzima batterico CdsA e un enzima archea non precedentemente caratterizzato [4]. Tale proteina, ora definita come CarS, è universalmente conservata nel regno archea e per poter caratterizzarla biochimicamente, il gene corrispondente, derivante dall'archea *Archaeoglobus fulgidus*, è stato espresso in *E. coli*. La proteina codificata da tale gene è stata purificata e testata attraverso delle reazioni *in vitro*

usando come substrato la molecola chimicamente sintetizzata DGGGP. Solo in presenza di CTP e di ioni divalenti (Mg^{2+}), è stato possibile osservare la formazione del prodotto CDP-archeolo, stabilendo così la funzione di CarS come CDP-archeolo sintasi. La scoperta di tale enzima chiave ha consentito pertanto la ricostituzione dell'intera via biosintetica per la sintesi di lipidi archea in un sistema indipendente dalla cellula. Attraverso l'uso di una combinazione di enzimi batterici e archea purificati, è stato possibile ottenere la sintesi di CDP-archeolo *in vitro* a partire dalle semplici e iniziali unità isoprenoidi.

CDP-archeolo rappresenta un importante metabolita intermedio nella biosintesi dei lipidi. Il passaggio biosintetico successivo riguarda infatti la sostituzione del gruppo CDP con uno dei gruppi polari come L-serina, etanolammina, glicerolo o mio-inositolo. Gli enzimi che catalizzano l'attaccamento di tali gruppi polari sono descritti in dettaglio nel **Capitolo 3**. Gli enzimi responsabili di tali passaggi catalitici esibiscono un'alta omologia di sequenza e condividono un simile meccanismo enzimatico in tutti e tre i domini della vita, essendo membri della stessa superfamiglia proteica, CDP-alcol fosfatidiltrasferasi [7,43]. Nei batteri, la sostituzione del gruppo CDP con glicerolo è catalizzata da due enzimi, PgsA e PgpA, che determinano la formazione di fosfatidilglicerolo (PG) passando dal metabolita intermedio fosfatidilglicerolo fosfato (PGP). Data l'omologia funzionale esibita da tali enzimi batterici, la loro specificità di substrato è stata testata verso substrati basati sulla struttura base dei lipidi archea. In presenza di CDP-archeolo, il lipide archea archeatidilglicerolo (AG) è stato prodotto dimostrando una promiscuità verso substrati lipidici batterici e archea. A differenza della formazione di AG, la formazione di archetidiletanolammina (AE) richiede la presenza dell'enzima PssA proveniente da un batterio diverso da *E. coli*. Gli enzimi PssA sono responsabili della formazione di fosfatidiliserina a partire da CDP-DAG e L-serina e sono classificati in due gruppi principali: gruppo-I contiene gli enzimi presenti in batteri Gram-negativi (*E. coli*) e il gruppo- II include gli enzimi PssA di batteri Gram-positivi (*B. subtilis*), lieviti e archea [148]. Studi precedenti hanno dimostrato che l'enzima PssA di *E. coli* è altamente selettivo per CDP-diacilglicerolo mentre PssA di *B. subtilis* probabilmente riconosce anche substrati aventi proprietà archea [47]. L'enzima purificato

PssA da *B. subtilis* è stato quindi testato biochimicamente insieme all'altro enzima batterico, fosfatidilserina decarbossilasi Psd, per la sintesi del lipide archea AE. La caratterizzazione biochimica di tali enzimi ha permesso la ricostituzione dell'intera via biosintetica per lipidi archea *in vitro* a partire dalle due unità isoprenoidi fino alla formazione di AG e AE. Per di più sono state definite le condizione per la sintesi di AG e AE *in vivo* utilizzando il batterio *E. coli* come ospite. Attraverso l'introduzione di sette geni, di fatto la sintesi dei due lipidi AG e AE è stata ottenuta, anche se a livelli molto bassi (inferiori all'1% rispetto al totale lipidico batterico).

La caratterizzazione degli enzimi coinvolti nella via metabolica per la sintesi di fosfolipidi dieterei sopradescritta, ha fornito le basi per un progetto innovativo riguardante l'ingegnerizzazione del batterio *E. coli* per ottenere alti livelli di produzione di lipidi eterei ed è presentato nel **Capitolo 4**. Per generare alti livelli di unità isoprenoidi in *E. coli*, il pathway MEP-DOXP, endogeno al batterio usato, è stato sovraespresso attraverso l'integrazione cromosomica dell'operone MEP-DOXP nel genoma batterico. Tale procedura ha determinato un'elevata sintesi di IPP e DMAPP come evidenziato dalla produzione di licopene. In seguito alla co-espressione degli enzimi eterologhi per la sintesi di lipidi eterei, il ceppo ingegnerizzato di *E. coli* ha prodotto fino a 32% di AG. Questo ha significato un quasi completo rimpiazzo della quantità di PG endogena con l'AG archea dimostrando l'integrazione funzionale dell'intera via biosintetica per la sintesi di lipidi archea nell'ospite batterico determinando così la formazione di una cellula vitale con una membrana mista ed etero-chirale. La propria configurazione stereochimica del lipide archea AG è stata validata tramite una tecnica NMR selettiva per la chiralità, la quale ha dimostrato che anche in assenza dell'enzima G1P deidrogenasi, la produzione di AG è avvenuta con la corretta chiralità. Degno di nota, tale risultato indica la capacità di *E. coli* di sintetizzare G1P, un processo che non è stato studiato o identificato prima. Inoltre, una debole selettività di substrato dell'enzima GGGP sintasi verso G1P rispetto a G3P è stata osservata. Pertanto, potenzialmente GGGP potrebbe essere sintetizzato con un'incorretta chiralità anche se ciò non è stato osservato *in vivo*, molto probabilmente in seguito alla generale selettività del processo metabolico. L'osservazione che l'enzima GGGP sintasi è meno selettivo rispetto a

quanto aspettato ha gettato una nuova luce sulle ipotesi che tentano di spiegare come la segregazione tra batteri e archea sia avvenuta. Al contrario, in batteri la reazione analoga è catalizzata dall' enzima altamente selettivo PlsB. Pertanto, la scarsa selettività chirale di GGGPS suggerisce la primordiale presenza di enzimi non-stereoselettivi nell'antenato comune dal quale gli organismi batterici si sono evoluti e differenziati in seguito all'apparizione di enzimi altamente stereoselettivi come PlsB.

La presenza di lipidi archea ha causato un'alterazione della crescita cellulare, della morfologia cellulare e ha conferito un certo grado di resistenza agli stress ambientali. Le cellule con una membrane mista ed eterochirale mostrano infatti una lunga fase lag prima che la crescita si normalizzi a quasi la stessa velocità di crescita del ceppo wild-type di *E. coli*. Questo suggerisce che le cellule subiscono un processo di adattamento alla presenza di una composizione alterata della membrana. Comunque, il sequenziamento del genoma batterico non ha rilevato nessun cambiamento o mutazione significativi dato che il comportamento cellulare osservato non è risultato stabile e costante, in quanto in seguito al trasferimento in un terreno di coltura fresco, la crescita cellulare iniziava sempre solo dopo una lunga fase lag. La maggior parte delle cellule ingegnerizzate mostra inoltre una forma cellulare allungata e sottile e, in presenza di alti livelli di induzione del processo metabolico, anche la presenza di appendici ed estrusioni. Queste strutture contengono DNA e sembrano originarsi da un aberrante meccanismo di divisione che non avviene più nel centro cellulare. Questo fenomeno sembra essere associato ad una elevata produzione di lipidi e non risulta evidente nelle cellule in cui il più alto livello di sintesi di AG è stato osservato, il che avviene a basse concentrazioni di induttore. Infine, la presenza di lipidi archea ha conferito alle cellule batteriche un'aumentata tolleranza alle esposizioni di calore e di congelamento e una resistenza contro trattamenti di butanolo.

Nel **Capitolo 5** la struttura e il meccanismo catalitico di CarS sono stati investigati. A tale scopo, la struttura cristallografica della proteina CarS derivante dall'archea *Aeropyrum pernix k1* è stata risolta in un complesso con CTP e Mg^{2+} a 2.38 Å. La caratterizzazione strutturale ha rivelato un'organizzazione della struttura secondaria della proteina inusuale, dove

i cinque segmenti transmembrana sono arrangiati in modo da formare una tasca idrofilica al lato citoplasmatico della membrana che è stata definita come il sito di legame di CTP. La mutagenesi di residui amminoacidici coinvolti nell'interazione con CTP è risultata nella perdita del legame con CTP e nella formazione di CDP-archeolo. La struttura cristallografica ha inoltre identificato una certa flessibilità nel segmento transmembrana cinque consistente con il suo ruolo previsto nel legame con il glicerolfosfolipide. Pertanto, data l'elevata conservazione del dominio di legame del CTP tra i membri della famiglia CTP-trasferasi, la caratterizzazione di CarS ha fornito ulteriori approfondimenti nel meccanismo di riconoscimento di CTP. Comunque, ulteriori studi sono necessari per capire completamente il meccanismo enzimatico di CarS e in particolare il legame con glicerofosfolipidi e il rilascio del prodotto sintetizzato.

Infine, il possibile meccanismo di idrogenazione dei doppi legami presenti nei lipidi archea aventi una catena isoprenoide insatura è descritto nel **Capitolo 6**. I lipidi di archea, quando maturi, consistono di catene isoprenoidi completamente sature e tale saturazione è catalizzata dall'enzima geranilgeranil reduttasi (GGR). Nonostante la disponibilità di strutture cristallografiche di alcune GGR di archea [52,211], l'esatto meccanismo di riduzione e a che livello avviene la riduzione dei doppi legami non è chiaro. Due GGR provenienti da due diversi archea sono state testate per la loro attività enzimatica insieme all'associata proteina ferredoxina. Comunque, la riduzione *in vivo* dei fosfolipidi archea prodotti nelle cellule ingegnerizzate di *E. coli* è stata povera e pertanto non conclusiva in rispetto a che livello biosintetico avviene la saturazione. La necessità di ferredoxina come cofattore biologico è stata invece confermata. Ulteriori studi sono richiesti per elucidare il meccanismo di saturazione da parte di GGR e per determinare a che livello nella via metabolica per la sintesi di lipidi eterei avviene l'idrogenazione delle catene isoprenoidi.

Per riassumere, la tesi presentata descrive la ricostituzione dell'intero e completo processo biosintetico per la produzione di lipidi archea sia *in vitro* che *in vivo*. Lo studio ha indentificato un nuovo enzima nella biosintesi di lipidi archea, cioè CDP-archeolo sintasi e ha rivelato l'elevata

promiscuità degli enzimi batterici coinvolti nell'attaccamento dei gruppi polari. Il lavoro ha anche gettato una nuova luce sull'importanza di alcune teorie evoluzionistiche. Il nostro lavoro nell'ingegnerizzare un ceppo batterico con una membrana mista ed eterochirale mostra la co-esistenza di due specie lipidiche differenti, chiralmente distinte, contrastando le precedenti ipotesi su una instabilità chimica di tale membrana, la quale ha spinto l'evoluzione degli organismi batterici e archea verso una maggiormente stabile membrana omochirale. Pertanto, le nostre osservazioni supportano l'idea che ci deve essere stato un comune antenato con una membrana mista dal quale batteri e archea si sono evoluti. Inoltre, la differenza in specificità di substrato trovata nell'enzima di archea GGGPS e in quello batterico PlsB pone nuove possibilità sull'iniziale insorgenza degli organismi archea, seguiti dalla differenziazione in batteri. In tale prospettiva l'innesco che ha causato la segregazione dalla cellula primordiale può essere visto come l'apparizione di enzimi stereo-selettivi come PlsB i quali hanno spinto la differenziazione degli organismi batterici dalla cellula antenata. Questa poi si sarebbe evoluta in un organismo archea mantenendo la capacità di vivere in ambienti estremi ma con una composizione lipidica della membrana evoluta e specifica.

Infine, il lavoro ha potenziali implicazioni biotecnologiche in quanto la riprogrammazione della composizione lipidica di *E. coli* ha determinato un'aumentata robustezza. Pertanto, quando finemente regolata, ciò potrebbe essere visto come una metodologia generale per altri micro-organismi di rilevanza industriale nel momento in cui una certa tolleranza ai solventi è richiesta. Non sarebbe comunque possibile ottenere tale alterazione radicale del contenuto lipidico tramite mutagenesi e selezione. Comunque, alcuni importanti passaggi del processo metabolico per la sintesi di lipidi archea, come la riduzione dei doppi legami, la formazione delle strutture tetraeteree e l'attaccamento di zuccheri rimangono non chiari e necessitano di ulteriori studi per elucidare completamente il meccanismo di sintesi dei lipidi di archea.

Acknowledgements

It was March 2012 when the first time I came to Groningen and I started my master project in the Molecular Microbiology department.
It has been more that 4 years and now I came at the end of my PhD. It was a long journey full of emotions, fun, difficulties and rewards. I would not be the person who I am now without all the persons that I met and were part of my life in this period of time!

First of all I would like to thank my supervisor professor Arnold Driessen for giving me the opportunity to do my PhD projet in your group. Thanks for being always available for suggestions and advises during our meetings, for all the freedom that you gave me to handle the project. Thanks for your support and guidance during all these years that helped to bring my PhD work to a successful end. Thanks for being always on my side in some difficult moments and let me express all my opinions and feelings. Thank you a lot!!

Dear Samta, then I would like to thank you for being my daily supervisor during my first year. You taught me really a lot and gave me the knowledge to carry on the project by myself. Thanks for all the motivation, the useful tips and nice advises which graetly helped me to come at the end of this project. It was a pleasure to work with you!!

My dear Amalina, my paranymph, my collegue but most of all my close friend. We started together this chapter of our life. We were both master students sharing the same lab, being tip boxes-filling parterns. We both came back and started the PhD, sharing the same office. Thanks for all the time spent together, for the support that you gave me in difficult moments, for the nice chatting, laughing and confidences. You made the days at work always coloured with your smile and enthusiasm. Never change this aspect of you!!

I would like to thank my other paranymph, Marten. You started to be part of “lipid-group” and soon you could experience by yourself what it means use the LCMS for the project! Thanks for all the sharing about the projects, the frustration about the LCMS problems and all the jokes. Thanks for listening to all my complains about the Dutch weather and, at times, Dutch

people. It was nice to have someone to actually share project results and have fruitful discussion with. Thanks for all the support and for being always available everytime I needed help. Dank je wel mjin lieve maneer!!

I would like to thank also the collaborators in Wageningen University, Melvin, John van der Oost and Serve Kengen for the fruitful discussions and for the joint work to reach the final goal of the project. I would like to thank prof. Wei Cheng, and its group, from SiChuan University for its enthusiasm in resolving the structure of CarS and for bringing new knowledge and challenging aspects to the project.

The Molecular Microbiology group..what to say.. It has been always a pleasure to work in this group. The people around (those who left and the new who started) were always friendly and smiling, making an enjoyable work place. In particular thanks to Janny, for all your patience and your availability to solve any kind of problems. Thanks for reminding me rules of the lab and keep it always tidy and clean. Thanks to Greetje, Jeroen and Susan to have always an answer whenever I had questions and for all the technical help during these years.
I would like to thank Juke and Drik-Jan for their useful advises and suggestions during the work discussions. It was always great to receive comments about the project and results seen by a different point of view. Thank you!

Thanks also to Sasha and Annebart. Sasha, you were always there to help me everytime I dropped in your office asking for LCMS assistance. With you smile and kindness you taught me really a lot about how to proper use the machine and without all your experience in solving MS breaking down I would not reach the end of my PhD. Annebart, I would like to thank you for all the help and suggestions that you gave me during these years in making graphs, thesis cover, colony counting software. You were always available to help me in any moment I asked you. Thanks so much for that and for all the jokes and funny moments.

Thanks to Bea, Manon and Anmara for being always available in helping me with burocratic university matters. I would like to thanks Jan Peter and Iuliia with who I shared the office in the first part of my Phd. It was always

nice to talk with you. Thanks to Reto for introducing me in his group of friends at the beginning of my PhD. You improved my social life and I will never forget all the fun I had in Groningen and all the crazy nights out! I also thanks Ana for being always smiling and funny during all these years!!
Then thanks to all the people of old MolMic group: Ilja, Jeanine, Stefan, Intan, Alexej, Yang and of the new MolMic group: Sabrina, Zsofi, Irfan, Ciprian, Riccardo, Annarita, Carsten, Fernando, Ewa, Danae, Elke, Gosja, Hyon, Ming and all the other members.
I would like also to thank my Italian friends here in Groningen, Giulia, Barbara and Roberta. It is always nice to speak Italian and to share feelings and emotions with you. Thanks also to my friend far away, Stefania, Sarah, Serena, Simonetta.. Thanks for being always present!

Finally I would like to thanks my family.
Grazie mamma e papá per avermi sempre sostenuto in tutti questi anni di studio. Sono passati ormai quasi dieci anni da quando sono andata via da casa per iniziare questo percorso che mi ha portato ad ottenere il titolo di dottore. E se sono giunta fino a qui è soprattutto grazie a voi che mi avete sempre incoraggiato nelle mie decisioni e lasciato che io seguissi la mia strada, anche se questa mi ha portato a migliaia di km di distanza. Grazie per essere sempre presenti in ogni momento di felicitá e di difficoltá. Le stesse parole le rivolgo a mia sorella. La mia confidente, la mia amica piú stretta e sincera e una parte di me. Abbiamo condiviso quasi tutto nella vita e vorrei condividere anche questo importante traguardo con te! La mia forza piú grande è quella di sapere che potró sempre contare su di voi e che ci sarete sempre per me. VI AMO!!

L'ultimo ringraziamento va a te amore mio. In questi due anni hai rimpieto le mie giornate di amore e felicitá. Nonostante la distanza sei sempre stato accanto a me in ogni momento a sostenermi nei momenti peggiori e a condividere con me le piú belle soddisfazioni, annullando tutti i km che ci tenevano separati. Sei la mia metá, la mia anima gemella e ti ringrazio per aver fatto si che io possa averti al mio fianco. TI AMO! !

References

1 Woese, C. R. and Fox, G. E. (1977) Phylogenetic structure of the prokaryotic domain: the primary kingdoms. Proc. Natl. Acad. Sci. U. S. A. **74**, 5088–90.

2 Koga, Y. and Morii, H. (2007) Biosynthesis of ether-type polar lipids in archaea and evolutionary considerations. Microbiol. Mol. Biol. Rev. **71**, 97–120.

3 Koga, Y. (2012) Thermal adaptation of the archaeal and bacterial lipid membranes. Archaea **2012**, 789652.

4 Lombard, J., López-García, P. and Moreira, D. (2012) The early evolution of lipid membranes and the three domains of life. Nat. Rev. Microbiol. **10**, 507–15.

5 Boucher, Y., Kamekura, M. and Doolittle, W. F. (2004) Origins and evolution of isoprenoid lipid biosynthesis in archaea. Mol. Microbiol. **52**, 515–27.

6 Villanueva, L., Damsté, J. S. S. and Schouten, S. (2014) A re-evaluation of the archaeal membrane lipid biosynthetic pathway. Nat. Rev. Microbiol., Nature Publishing Group.

7 Daiyasu, H., Kuma, K.-I., Yokoi, T., Morii, H., Koga, Y. and Toh, H. (2005) A study of archaeal enzymes involved in polar lipid synthesis linking amino acid sequence information, genomic contexts and lipid composition. Archaea **1**, 399–410.

8 Lombard, J., López-García, P. and Moreira, D. (2012) Phylogenomic investigation of phospholipid synthesis in archaea. Archaea **2012**, 630910.

9 Lombard, J. and Moreira, D. (2011) Origins and early evolution of the mevalonate pathway of isoprenoid biosynthesis in the three domains of life. Mol. Biol. Evol. **28**, 87–99.

10 Matsumi, R., Atomi, H., Driessen, A. J. M. and van der Oost, J. (2011) Isoprenoid biosynthesis in Archaea--biochemical and evolutionary implications. Res. Microbiol., Elsevier Masson SAS **162**, 39–52.

11 Xue, J. and Ahring, B. K. (2011) Enhancing isoprene production by genetic modification of the 1-deoxy-d-xylulose-5-phosphate pathway in Bacillus subtilis. Appl. Environ. Microbiol. **77**, 2399–405.

12 Grochowski, L. L., Xu, H. and White, R. H. (2006) Methanocaldococcus jannaschii uses a modified mevalonate pathway for biosynthesis of isopentenyl diphosphate. J. Bacteriol. **188**, 3192–8.

13 Dellas, N. and Noel, J. P. (2010) Mutation of archaeal isopentenyl phosphate kinase highlights mechanism and guides phosphorylation of additional isoprenoid monophosphates. ACS Chem. Biol., American Chemical Society **5**, 589–601.

14 Chen, M. and Poulter, C. D. (2010) Characterization of thermophilic archaeal isopentenyl phosphate kinases. Biochemistry, American Chemical Society **49**, 207–17.

15 Dellas, N., Thomas, S. T., Manning, G. and Noel, J. P. (2013) Discovery of a metabolic alternative to the classical mevalonate pathway. Elife **2**, e00672.

16 Wang, K. and Ohnuma, S. (1999) Chain-length determination mechanism of

isoprenyl diphosphate synthases and implications for molecular evolution. Trends Biochem. Sci. **24**, 445–51.

17 Vandermoten, S., Haubruge, E. and Cusson, M. (2009) New insights into short-chain prenyltransferases: structural features, evolutionary history and potential for selective inhibition. Cell. Mol. Life Sci. **66**, 3685–95.

18 Hemmi, H., Ikejiri, S., Yamashita, S. and Nishino, T. (2002) Novel Medium-Chain Prenyl Diphosphate Synthase from the Thermoacidophilic Archaeon Sulfolobus solfataricus. J. Bacteriol. **184**, 615–620.

19 Tachibana, A., Yano, Y., Otani, S., Nomura, N., Sako, Y. and Taniguchi, M. (2000) Novel prenyltransferase gene encoding farnesylgeranyl diphosphate synthase from a hyperthermophilic archaeon, Aeropyrum pernix. Molecularevolution with alteration in product specificity. Eur. J. Biochem. **267**, 321–8.

20 Ohnuma, S., Suzuki, M. and Nishino, T. (1994) Archaebacterial ether-linked lipid biosynthetic gene. Expression cloning, sequencing, and characterization of geranylgeranyl-diphosphate synthase. J. Biol. Chem. **269**, 14792–7.

21 Gunsalus, R. P., Lai, D., Lluncor, B., Schro, I., Liao, J. C. and Monbouquette, H. G. (2009) Reconstruction of the archaeal isoprenoid ether lipid biosynthesis pathway in Escherichia coli through digeranylgeranylglyceryl phosphate **11**, 184–191.

22 Fujiwara, S., Yamanaka, A., Hirooka, K., Kobayashi, A., Imanaka, T. and Fukusaki, E.-I. (2004) Temperature-dependent modulation of farnesyl diphosphate/geranylgeranyl diphosphate synthase from hyperthermophilic archaea. Biochem. Biophys. Res. Commun. **325**, 1066–74.

23 Chen, A. and Poulter, C. D. (1993) Purification and characterization of farnesyl diphosphate/geranylgeranyl diphosphate synthase. A thermostable bifunctional enzyme from Methanobacterium thermoautotrophicum. J. Biol. Chem. **268**, 11002–7.

24 Ohnuma, S., Hirooka, K., Hemmi, H., Ishida, C., Ohto, C. and Nishino, T. (1996) Conversion of product specificity of archaebacterial geranylgeranyl-diphosphate synthase. Identification of essential amino acid residues for chain length determination of prenyltransferase reaction. J. Biol. Chem. **271**, 18831–7.

25 Lee, P. C., Petri, R., Mijts, B. N., Watts, K. T. and Schmidt-Dannert, C. (2005) Directed evolution of Escherichia coli farnesyl diphosphate synthase (IspA) reveals novel structural determinants of chain length specificity. Metab. Eng. **7**, 18–26.

26 Han, J.-S. and Ishikawa, K. (2005) Active site of Zn(2+)-dependent sn-glycerol-1-phosphate dehydrogenase from Aeropyrum pernix K1. Archaea **1**, 311–7.

27 Koga, Y., Sone, N., Noguchi, S. and Morii, H. (2003) Transfer of pro-R hydrogen from NADH to dihydroxyacetonephosphate by sn-glycerol-1-phosphate dehydrogenase from the archaeon Methanothermobacter thermautotrophicus. Biosci. Biotechnol. Biochem. **67**, 1605–8.

28 Nishihara, M. and Koga, Y. (1997) Purification and properties of sn-glycerol-1-phosphate dehydrogenase from Methanobacterium thermoautotrophicum: characterization of the biosynthetic enzyme for the enantiomeric glycerophosphate backbone of ether polar lipids of Archaea. J. Biochem. **122**, 572–6.

29 Han, J.-S., Kosugi, Y., Ishida, H. and Ishikawa, K. (2002) Kinetic study of sn-glycerol-1-phosphate dehydrogenase from the aerobic hyperthermophilic archaeon, Aeropyrum pernix K1. Eur. J. Biochem. **269**, 969–76.

30 Koga, Y., Ohga, M., Tsujimura, M., Morii, H. and Kawarabayasi, Y. (2006) Identification of sn-glycerol-1-phosphate dehydrogenase activity from genomic information on a hyperthermophilic archaeon, Sulfolobus tokodaii strain 7. Biosci. Biotechnol. Biochem. **70**, 282–5.

31 Guldan, H., Sterner, R. and Babinger, P. (2008) Identification and characterization of a bacterial glycerol-1-phosphate dehydrogenase: Ni(2+)-dependent AraM from Bacillus subtilis. Biochemistry **47**, 7376–84.

32 Guldan, H., Matysik, F.-M., Bocola, M., Sterner, R. and Babinger, P. (2011) Functional assignment of an enzyme that catalyzes the synthesis of an archaea-type ether lipid in bacteria. Angew. Chem. Int. Ed. Engl. **50**, 8188–91.

33 Podar, M., Makarova, K. S., Graham, D. E., Wolf, Y. I., Koonin, E. V and Reysenbach, A.-L. (2013) Insights into archaeal evolution and symbiosis from the genomes of a nanoarchaeon and its inferred crenarchaeal host from Obsidian Pool, Yellowstone National Park. Biol. Direct **8**, 9.

34 Ren, F., Feng, X., Ko, T.-P., Huang, C.-H., Hu, Y., Chan, H.-C., Liu, Y.-L., Wang, K., Chen, C.-C., Pang, X., et al. (2013) Insights into TIM-barrel prenyl transferase mechanisms: crystal structures of PcrB from Bacillus subtilis and Staphylococcus aureus. Chembiochem **14**, 195–9.

35 Peterhoff, D., Beer, B., Rajendran, C., Kumpula, E.-P., Kapetaniou, E., Guldan, H., Wierenga, R. K., Sterner, R. and Babinger, P. (2014) A comprehensive analysis of the geranylgeranylglyceryl phosphate synthase enzyme family identifies novel members and reveals mechanisms of substrate specificity and quaternary structure organization. Mol. Microbiol. **92**, 885–99.

36 Payandeh, J., Fujihashi, M., Gillon, W. and Pai, E. F. (2006) The crystal structure of (S)-3-O-geranylgeranylglyceryl phosphate synthase reveals an ancient fold for an ancient enzyme. J. Biol. Chem. **281**, 6070–8.

37 Zhang, D. and Poulter, C. D. (2002) Biosynthesis of archaebacterial ether lipids. Formation of ether linkages by prenyltransferases, American Chemical Society **2**, 1270–1277.

38 Hemmi, H., Shibuya, K., Takahashi, Y., Nakayama, T. and Nishino, T. (2004) (S)-2,3-Di-O-geranylgeranylglyceryl phosphate synthase from the thermoacidophilic archaeon Sulfolobus solfataricus. Molecular cloning and characterization of a membrane-intrinsic prenyltransferase involved in the biosynthesis of archaeal ether-linked memb. J. Biol. Chem. **279**, 50197–203.

39 Zhang, H., Shibuya, K., Hemmi, H., Nishino, T. and Prestwich, G. D. (2006) Total synthesis of geranylgeranylglyceryl phosphate enantiomers:

substrates for characterization of 2,3-O-digeranylgeranylglyceryl phosphate synthase. Org. Lett. **8**, 943–946.

40 Yokoi, T., Isobe, K., Yoshimura, T. and Hemmi, H. (2012) Archaeal phospholipid biosynthetic pathway reconstructed in Escherichia coli. Archaea **2012**, 438931.

41 Jain, S., Caforio, A., Fodran, P., Lolkema, J. S., Minnaard, A. J. and Driessen, A. J. M. (2014) Identification of CDP-Archaeol Synthase, a Missing Link of Ether Lipid Biosynthesis in Archaea. Chem. Biol.

42 Morii, H., Nishihara, M. and Koga, Y. (2000) CTP:2,3-di-O-geranylgeranyl-sn-glycero-1-phosphate cytidyltransferase in the methanogenic archaeon Methanothermobacter thermoautotrophicus. J. Biol. Chem. **275**, 36568–74.

43 Koga, Y. (2011) Early evolution of membrane lipids: how did the lipid divide occur? J. Mol. Evol. **72**, 274–82.

44 Tachibana, A., Tanaka, T., Taniguchi, M. and Oi, S. (1993) Potassium-stimulating mechanism of geranylgeranyl diphosphate synthase of Methanobacterium thermoformicicum SF-4. J. Biochem. **114**, 389–92.

45 Soderberg, T., Chen, A. and Poulter, C. D. (2001) Geranylgeranylglyceryl phosphate synthase. Characterization of the recombinant enzyme from Methanobacterium thermoautotrophicum. Biochemistry **40**, 14847–54.

46 Nemoto, N. (2003) Purification and Characterization of Geranylgeranylglyceryl Phosphate Synthase from a Thermoacidophilic Archaeon, Thermoplasma acidophilum. J. Biochem. **133**, 651–657.

47 Morii, H. and Koga, Y. (2003) CDP-2,3-Di-O-Geranylgeranyl-sn-Glycerol:L-Serine O-Archaetidyltransferase (Archaetidylserine Synthase) in the Methanogenic Archaeon Methanothermobacter thermautotrophicus. J. Bacteriol. **185**, 1181–1189.

48 Morii, H., Kiyonari, S., Ishino, Y. and Koga, Y. (2009) A novel biosynthetic pathway of archaetidyl-myo-inositol via archaetidyl-myo-inositol phosphate from CDP-archaeol and D-glucose 6-phosphate in methanoarchaeon Methanothermobacter thermautotrophicus cells. J. Biol. Chem. **284**, 30766–74.

49 Morii, H., Ogawa, M., Fukuda, K. and Taniguchi, H. (2013) Ubiquitous distribution of phosphatidylinositol phosphate synthase and archaetidylinositol phosphate synthase in Bacteria and Archaea, which contain inositol phospholipid. Biochem. Biophys. Res. Commun.

50 Xu, Q., Eguchi, T., Mathews, I. I., Rife, C. L., Chiu, H.-J., Farr, C. L., Feuerhelm, J., Jaroszewski, L., Klock, H. E., Knuth, M. W., et al. (2010) Insights into substrate specificity of geranylgeranyl reductases revealed by the structure of digeranylgeranylglycerophospholipid reductase, an essential enzyme in the biosynthesis of archaeal membrane lipids. J. Mol. Biol. **404**, 403–17.

51 Ogawa, T., Isobe, K., Mori, T., Asakawa, S., Yoshimura, T. and Hemmi, H. (2014) A novel geranylgeranyl reductase from the methanogenic archaeon Methanosarcina acetivorans displays unique regiospecificity. FEBS J. **281**, 3165–76.

52 Sasaki, D., Fujihashi, M., Iwata, Y., Murakami, M., Yoshimura, T., Hemmi, H.

and Miki, K. (2011) Structure and mutation analysis of archaeal geranylgeranyl reductase. J. Mol. Biol. **409**, 543–57.

53 Isobe, K., Ogawa, T., Hirose, K., Yokoi, T., Yoshimura, T. and Hemmi, H. (2014) Geranylgeranyl reductase and ferredoxin from Methanosarcina acetivorans are required for the synthesis of fully reduced archaeal membrane lipid in Escherichia coli cells. J. Bacteriol. **196**, 417–23.

54 Nemoto, N., Shida, Y., Shimada, H., Oshima, T. and Yamagishi, A. (2003) Characterization of the precursor of tetraether lipid biosynthesis in the thermoacidophilic archaeon Thermoplasma acidophilum. Extremophiles **7**, 235–43.

55 Eguchi, T., Takyo, H., Morita, M., Kakinuma, K. and Koga, Y. (2000) Unusual double-bond migration as a plausible key reaction in the biosynthesis of the isoprenoidal membrane lipids of methanogenic archaea. Chem. Commun., The Royal Society of Chemistry 1545–1546.

56 Poulter, C. D., Aoki, T. and Daniels, L. (1988) Biosynthesis of isoprenoid membranes in the methanogenic archaebacterium Methanospirillum hungatei. J. Am. Chem. Soc., American Chemical Society **110**, 2620–2624.

57 Kamekura, M. and Kates, M. (1999) Structural diversity of membrane lipids in members of Halobacteriaceae. Biosci. Biotechnol. Biochem. **63**, 969–72.

58 Koga, Y., Nishihara, M., Morii, H. and Akagawa-Matsushita, M. (1993) Ether polar lipids of methanogenic bacteria: structures, comparative aspects, and biosyntheses. Microbiol. Rev. **57**, 164–82.

59 Macalady, J. L., Vestling, M. M., Baumler, D., Boekelheide, N., Kaspar, C. W. and Banfield, J. F. (2004) Tetraether-linked membrane monolayers in Ferroplasma spp: a key to survival in acid. Extremophiles **8**, 411–9.

60 UDA, I., SUGAI, A., ITOH, Y. H. and ITOH, T. (2004) Variation in Molecular Species of Core Lipids from the Order Thermoplasmales Strains Depends on the Growth Temperature. J. Oleo Sci. **53**, 399–404.

61 Schouten, S., Hopmans, E. C. and Sinninghe Damsté, J. S. (2013) The organic geochemistry of glycerol dialkyl glycerol tetraether lipids: A review. Org. Geochem. **54**, 19–61.

62 Damsté, J. S. S., Rijpstra, W. I. C., Hopmans, E. C., Jung, M.-Y., Kim, J.-G., Rhee, S.-K., Stieglmeier, M. and Schleper, C. (2012) Intact polar and core glycerol dibiphytanyl glycerol tetraether lipids of group I.1a and I.1b thaumarchaeota in soil. Appl. Environ. Microbiol. **78**, 6866–74.

63 Pitcher, A., Rychlik, N., Hopmans, E. C., Spieck, E., Rijpstra, W. I. C., Ossebaar, J., Schouten, S., Wagner, M. and Damsté, J. S. S. (2010) Crenarchaeol dominates the membrane lipids of Candidatus Nitrososphaera gargensis, a thermophilic group I.1b Archaeon. ISME J. **4**, 542–52.

64 Koga, Y. and Morii, H. (2005) Recent advances in structural research on ether lipids from archaea including comparative and physiological aspects. Biosci. Biotechnol. Biochem. **69**, 2019–34.

65 Untersteller, É., Fritz, B., Bliériot, Y. and Sinaÿ, P. (1999) The structure of calditol isolated from the thermoacidophilic archaebacterium Sulfolobus acidocaldarius. Comptes Rendus l'Académie des Sci. - Ser. IIC - Chem. **2**,

429–433.

66 Gambacorta, A., Caracciolo, G., Trabasso, D., Izzo, I., Spinella, A. and Sodano, G. (2002) Biosynthesis of calditol, the cyclopentanoid containing moiety of the membrane lipids of the archaeon Sulfolobus solfataricus. Tetrahedron Lett. **43**, 451–453.

67 Sugai, A., Sakuma, R., Fukuda, I., Kurosawa, N., Itoh, Y. H., Kon, K., Ando, S. and Itoh, T. (1995) The structure of the core polyol of the ether lipids from Sulfolobus acidocaldarius. Lipids **30**, 339–44.

68 Gambacorta, A., Gliozzi, A. and De Rosa, M. (1995) Archaeal lipids and their biotechnological applications. World J. Microbiol. Biotechnol., Kluwer Academic Publishers **11**, 115–131.

69 Ulrih, N. P., Gmajner, D. and Raspor, P. (2009) Structural and physicochemical properties of polar lipids from thermophilic archaea. Appl. Microbiol. Biotechnol. **84**, 249–60.

70 Zhang, Y.-M. and Rock, C. O. (2008) Membrane lipid homeostasis in bacteria. Nat. Rev. Microbiol., Nature Publishing Group **6**, 222–33.

71 Oger, P. M. and Cario, A. (2013) Adaptation of the membrane in Archaea. Biophys. Chem. **183**, 42–56.

72 Sprott, G. D., Meloche, M. and Richards, J. C. (1991) Proportions of diether, macrocyclic diether, and tetraether lipids in Methanococcus jannaschii grown at different temperatures. J. Bacteriol. **173**, 3907–10.

73 Jacquemet, A., Barbeau, J., Lemiègre, L. and Benvegnu, T. (2009) Archaeal tetraether bipolar lipids: Structures, functions and applications. Biochimie **91**, 711–717.

74 Koga, Y. and Nakano, M. (2008) A dendrogram of archaea based on lipid component parts composition and its relationship to rRNA phylogeny. Syst. Appl. Microbiol. **31**, 169–82.

75 Kates, M. (1992) Archaebacterial lipids: structure, biosynthesis and function. Biochem. Soc. Symp. **58**, 51–72.

76 De Rosa, M., Esposito, E., Gambacorta, A., Nicolaus, B. and Bu'Lock, J. D. (1980) Effects of temperature on ether lipid composition of Caldariella acidophila. Phytochemistry **19**, 827–831.

77 Benvegnu, T., Lemiègre, L. and Cammas-Marion, S. (2008) Archaeal Lipids: Innovative Materials for Biotechnological Applications. European J. Org. Chem. **2008**, 4725–4744.

78 Tenchov, B., Vescio, E. M., Sprott, G. D., Zeidel, M. L. and Mathai, J. C. (2006) Salt tolerance of archaeal extremely halophilic lipid membranes. J. Biol. Chem. **281**, 10016–23.

79 Russell, N. J. and Nichols, D. S. (1999) Polyunsaturated fatty acids in marine bacteria--a dogma rewritten. Microbiology **145 (Pt 4**, 767–79.

80 Chang, E. L. (1994) Unusual thermal stability of liposomes made from bipolar tetraether lipids. Biochem. Biophys. Res. Commun. **202**, 673–9.

81 Elferink, M. G. L., de Wit, J. G., Driessen, A. J. M. and Konings, W. N. (1994) Stability and proton-permeability of liposomes composed of archaeal tetraether lipids. Biochim. Biophys. Acta - Biomembr. **1193**, 247–254.

82 Brown, D. A., Venegas, B., Cooke, P. H., English, V. and Chong, P. L.-G. (2009)

Bipolar tetraether archaeosomes exhibit unusual stability against autoclaving as studied by dynamic light scattering and electron microscopy. Chem. Phys. Lipids **159**, 95–103.

83 Komatsu, H. and Chong, P. L. (1998) Low permeability of liposomal membranes composed of bipolar tetraether lipids from thermoacidophilic archaebacterium Sulfolobus acidocaldarius. Biochemistry, American Chemical Society **37**, 107–15.

84 Shimada, H., Nemoto, N., Shida, Y., Oshima, T. and Yamagishi, A. (2008) Effects of pH and temperature on the composition of polar lipids in Thermoplasma acidophilum HO-62. J. Bacteriol. **190**, 5404–11.

85 Choquet, C. G., Patel, G. B., Sprott, G. D. and Beveridge, T. J. (1994) Stability of pressure-extruded liposomes made from archaeobacterial ether lipids. Appl. Microbiol. Biotechnol. **42**, 375–384.

86 Dannenmuller, O., Arakawa, K., Eguchi, T., Kakinuma, K., Blanc, S., Albrecht, A. M., Schmutz, M., Nakatani, Y. and Ourisson, G. (2000) Membrane properties of archaeal macrocyclic diether phospholipids. Chemistry **6**, 645–54.

87 Koga, Y., Kyuragi, T., Nishihara, M. and Sone, N. (1998) Did archaeal and bacterial cells arise independently from noncellular Precursors? A hypothesis stating that the advent of membrane phospholipid with enantiomeric glycerophosphate backbones caused the separation of the two lines of descent. J. Mol. Evol. **47**, 631.

88 Martin, W. and Russell, M. J. (2003) On the origins of cells: a hypothesis for the evolutionary transitions from abiotic geochemistry to chemoautotrophic prokaryotes, and from prokaryotes to nucleated cells. Philos. Trans. R. Soc. Lond. B. Biol. Sci. **358**, 59–83; discussion 83–5.

89 Wächtershäuser, G. (2003) From pre-cells to Eukarya--a tale of two lipids. Mol. Microbiol. **47**, 13–22.

90 Fan, Q., Relini, A., Cassinadri, D., Gambacorta, A. and Gliozzi, A. (1995) Stability against temperature and external agents of vesicles composed of archael bolaform lipids and egg PC. Biochim. Biophys. Acta **1240**, 83–8.

91 Shimada, H. and Yamagishi, A. (2011) Stability of heterochiral hybrid membrane made of bacterial sn-G3P lipids and archaeal sn-G1P lipids. Biochemistry **50**, 4114–20.

92 Valentine, D. L. (2007) Adaptations to energy stress dictate the ecology and evolution of the Archaea. Nat. Rev. Microbiol. **5**, 316–23.

93 Koga, Y. (2014) From Promiscuity to the Lipid Divide : On the Evolution of Distinct Membranes in Archaea and Bacteria 234–242.

94 Yao, J. and Rock, C. O. (2013) Phosphatidic acid synthesis in bacteria. Biochim. Biophys. Acta **1831**, 495–502.

95 Wang, C. W., Oh, M. K. and Liao, J. C. (1999) Engineered isoprenoid pathway enhances astaxanthin production in Escherichia coli. Biotechnol. Bioeng. **62**, 235–41.

96 Lorenzen, W., Ahrendt, T., Bozhüyük, K. A. J. and Bode, H. B. (2014) A multifunctional enzyme is involved in bacterial ether lipid biosynthesis. Nat. Chem. Biol. **10**, 425–7.

97 van de Vossenberg, J. L., Driessen, A. J. and Konings, W. N. (1998) The essence of being extremophilic: the role of the unique archaeal membrane lipids. Extremophiles **2**, 163–70.
98 Nowicka, B. and Kruk, J. (2010) Occurrence, biosynthesis and function of isoprenoid quinones. Biochim. Biophys. Acta - Bioenerg. **1797**, 1587–1605.
99 Liang, C. and Sigel, H. (1992) Metal Ion Binding Properties of Dihydroxyacetone Phosphate and Glycerol 1 -Phosphate **3**, 7780–7785.
100 Kelley, M. J. and Carman, G. M. (1987) Purification and characterization of CDP-diacylglycerol synthase from Saccharomyces cerevisiae. J. Biol. Chem. **262**, 14563–70.
101 Sparrow, C. P. and Raetz, C. R. (1985) Purification and properties of the membrane-bound CDP-diglyceride synthetase from Escherichia coli. J. Biol. Chem., American Society for Biochemistry and Molecular Biology **260**, 12084–91.
102 Dowhan, W. (1997) CDP-diacylglycerol synthase of microorganisms. Biochim. Biophys. Acta **1348**, 157–65.
103 Dowhan, W. (2013) A retrospective: use of Escherichia coli as a vehicle to study phospholipid synthesis and function. Biochim. Biophys. Acta **1831**, 471–94.
104 Heacock, A. M. and Agranoff, B. W. (1997) CDP-diacylglycerol synthase from mammalian tissues. Biochim. Biophys. Acta **1348**, 166–72.
105 Chang, Y.-F. and Carman, G. M. (2008) CTP synthetase and its role in phospholipid synthesis in the yeast Saccharomyces cerevisiae. Prog. Lipid Res. **47**, 333–9.
106 Carman, G. M. and Han, G.-S. (2011) Regulation of phospholipid synthesis in the yeast Saccharomyces cerevisiae. Annu. Rev. Biochem. **80**, 859–83.
107 Martin, D., Gannoun-Zaki, L., Bonnefoy, S., Eldin, P., Wengelnik, K. and Vial, H. (2000) Characterization of Plasmodium falciparum CDP-diacylglycerol synthase, a proteolytically cleaved enzyme. Mol. Biochem. Parasitol. **110**, 93–105.
108 Lolkema, J. S. and Slotboom, D. J. (1998) Hydropathy profile alignment: a tool to search for structural homologues of membrane proteins. FEMS Microbiol. Rev. **22**, 305–22.
109 Crooks, G. E., Hon, G., Chandonia, J.-M. and Brenner, S. E. (2004) WebLogo: a sequence logo generator. Genome Res. **14**, 1188–90.
110 Shen, H., Heacock, P. N., Clancey, C. J. and Dowhan, W. (1996) The CDS1 gene encoding CDP-diacylglycerol synthase in Saccharomyces cerevisiae is essential for cell growth. J. Biol. Chem. **271**, 789–795.
111 Ganong, B. R. and Raetz, C. R. (1982) Massive accumulation of phosphatidic acid in conditionally lethal CDP-diglyceride synthetase mutants and cytidine auxotrophs of Escherichia coli. J. Biol. Chem., ASBMB **257**, 389–94.
112 Wu, L., Niemeyer, B., Colley, N., Socolich, M. and Zuker, C. S. (1995) Regulation of PLC-mediated signalling in vivo by CDP-diacylglycerol synthase. Nature **373**, 216–22.
113 Salis, H. M., Mirsky, E. A. and Voigt, C. A. (2009) Automated design of synthetic ribosome binding sites to control protein expression. Nat.

Biotechnol. **27**, 946–50.

114 de Crécy-Lagard, V., Phillips, G., Grochowski, L. L., El Yacoubi, B., Jenney, F., Adams, M. W. W., Murzin, A. G. and White, R. H. (2012) Comparative genomics guided discovery of two missing archaeal enzyme families involved in the biosynthesis of the pterin moiety of tetrahydromethanopterin and tetrahydrofolate. ACS Chem. Biol. **7**, 1807–16.

115 Shastri, S., Zeeman, A.-M., Berry, L., Verburgh, R. J., Braun-Breton, C., Thomas, A. W., Gannoun-Zaki, L., Kocken, C. H. M. and Vial, H. J. (2010) Plasmodium CDP-DAG synthase: an atypical gene with an essential N-terminal extension. Int. J. Parasitol. **40**, 1257–68.

116 Deems, R. A. (1995) Lipid Signaling Enzymes and Surface Dilution Kinetics. J. Biol. Chem. **270**, 18711–18714.

117 Kaufmann, A., Manting, E. H., Veenendaal, A. K., Driessen, A. J. and van der Does, C. (1999) Cysteine-directed cross-linking demonstrates that helix 3 of SecE is close to helix 2 of SecY and helix 3 of a neighboring SecE. Biochemistry **38**, 9115–25.

118 Koga, Y. and Morii, H. (2006) Special methods for the analysis of ether lipid structure and metabolism in archaea. Anal. Biochem. **348**, 1–14.

119 Wycuff, D. R. and Matthews, K. S. (2000) Generation of an AraC-araBAD promoter-regulated T7 expression system. Anal. Biochem. **277**, 67–73.

120 Caruthers, M. H., Barone, A. D., Beaucage, S. L., Dodds, D. R., Fisher, E. F., McBride, L. J., Matteucci, M., Stabinsky, Z. and Tang, J. Y. (1987) Chemical synthesis of deoxyoligonucleotides by the phosphoramidite method. Methods Enzymol. **154**, 287–313.

121 Watanabe, Y., Nakamura, T. and Mitsumoto, H. (1997) Protection of phosphate with the 9-fluorenylmethyl group. Synthesis of unsaturated-acyl phosphatidylinositol 4,5-bisphosphate. Tetrahedron Lett. **38**, 7407–7410.

122 Wagner, S., Klepsch, M. M., Schlegel, S., Appel, A., Draheim, R., Tarry, M., Högbom, M., van Wijk, K. J., Slotboom, D. J., Persson, J. O., et al. (2008) Tuning Escherichia coli for membrane protein overexpression. Proc. Natl. Acad. Sci. U. S. A. **105**, 14371–6.

123 Howell, S. J., Spencer, N. and Philp, D. (2001) Recognition-mediated regiocontrol of a dipolar cycloaddition reaction. Tetrahedron **57**, 4945–4954.

124 Alcaraz, M.-L., Peng, L., Klotz, P. and Goeldner, M. (1996) Synthesis and Properties of Photoactivatable Phospholipid Derivatives Designed To Probe the Membrane-Associate Domains of Proteins. J. Org. Chem., American Chemical Society **61**, 192–201.

125 Yinghua Jin, F. G. and Roberts, and R. M. C. (2007) Stereoselective isoprenoid chain extension with acetoacetate dianion. Org. Synth. **84**, 43–57.

126 Dowhan, W. (1997) Molecular basis for membrane phospholipid diversity: why are there so many lipids? Annu. Rev. Biochem. **66**, 199–232.

127 Jain, S., Caforio, A. and Driessen, A. J. M. (2014) Biosynthesis of archaeal

membrane ether lipids. Front. Microbiol. **5**, 641.

128 Warner, T. G. and Dennis, E. A. (1975) Phosphatidylserine decarboxylase: analysis of its action towards unsaturated and saturated phosphatidylserine and the effect of Triton X-100 on activity. Arch. Biochem. Biophys. **167**, 761–8.

129 Kent, C. (2003) Eukaryotic Phospholipid Biosynthesis, Annual Reviews 4139 El Camino Way, P.O. Box 10139, Palo Alto, CA 94303-0139, USA.

130 Cronan, J. E. (2003) Bacterial membrane lipids: where do we stand? Annu. Rev. Microbiol. **57**, 203–24.

131 Raetz, C. R. and Dowhan, W. (1990) Biosynthesis and function of phospholipids in Escherichia coli. J. Biol. Chem. **265**, 1235–8.

132 Yung, B. Y. and Kornberg, A. (1988) Membrane attachment activates dnaA protein, the initiation protein of chromosome replication in Escherichia coli. Proc. Natl. Acad. Sci. U. S. A. **85**, 7202–5.

133 Sekimizu, K., Yung, B. Y. and Kornberg, A. (1988) The dnaA protein of Escherichia coli. Abundance, improved purification, and membrane binding. J. Biol. Chem. **263**, 7136–40.

134 Xia, W. and Dowhan, W. (1995) In vivo evidence for the involvement of anionic phospholipids in initiation of DNA replication in Escherichia coli. Proc. Natl. Acad. Sci. U. S. A. **92**, 783–7.

135 Vrije, T. de, Swart, R. de, Dowhan, W., Tommassen, J. and Kruijff, B. de. (1988) Phosphatidylglycerol is involved in protein translocation across Escherichia coli inner membranes. Nat. Int. Wkly. J. Sci. **334**, 173–175.

136 Lill, R., Dowhan, W. and Wickner, W. (1990) The ATPase activity of SecA is regulated by acidic phospholipids, SecY, and the leader and mature domains of precursor proteins. Cell **60**, 271–80.

137 Phoenix, D. A., Kusters, R., Hikita, C., Mizushima, S. and de Kruijff, B. (1993) OmpF-Lpp signal sequence mutants with varying charge hydrophobicity ratios provide evidence for a phosphatidylglycerol-signal sequence interaction during protein translocation across the Escherichia coli inner membrane. J. Biol. Chem. **268**, 17069–73.

138 Nishijima, M. and Raetz, C. R. (1979) Membrane lipid biogenesis in Escherichia coli: identification of genetic loci for phosphatidylglycerophosphate synthetase and construction of mutants lacking phosphatidylglycerol. J. Biol. Chem. **254**, 7837–44.

139 Heacock, P. N. and Dowhan, W. (1987) Construction of a lethal mutation in the synthesis of the major acidic phospholipids of Escherichia coli. J. Biol. Chem. **262**, 13044–9.

140 Gopalakrishnan, A. S., Chen, Y. C., Temkin, M. and Dowhan, W. (1986) Structure and expression of the gene locus encoding the phosphatidylglycerophosphate synthase of Escherichia coli. J. Biol. Chem. **261**, 1329–38.

141 Lu, Y.-H., Guan, Z., Zhao, J. and Raetz, C. R. H. (2011) Three phosphatidylglycerol-phosphate phosphatases in the inner membrane of Escherichia coli. J. Biol. Chem. **286**, 5506–18.

142 Icho, T. and Raetz, C. R. (1983) Multiple genes for membrane-bound

phosphatases in Escherichia coli and their action on phospholipid precursors. J. Bacteriol. **153**, 722–30.

143 Dillon, D. A., Wu, W. I., Riedel, B., Wissing, J. B., Dowhan, W. and Carman, G. M. (1996) The Escherichia coli pgpB gene encodes for a diacylglycerol pyrophosphate phosphatase activity. J. Biol. Chem. **271**, 30548–53.

144 Icho, T. (1988) Membrane-bound phosphatases in Escherichia coli: sequence of the pgpB gene and dual subcellular localization of the pgpB product. J. Bacteriol. **170**, 5117–5124.

145 Funk, C. R., Zimniak, L. and Dowhan, W. (1992) The pgpA and pgpB genes of Escherichia coli are not essential: evidence for a third phosphatidylglycerophosphate phosphatase. J. Bacteriol. **174**, 205–13.

146 KANFER, J. N. and KENNEDY, E. P. (1962) Synthesis of phosphatidylserine by Escherichia coli. J. Biol. Chem. **237**, PC270–PC271.

147 Bae-Lee, M. and Carman, G. (1984) Phosphatidylserine synthesis in Saccharomyces cerevisiae. Purification and characterization of membrane-associated phosphatidylserine synthase. J. Biol. Chem. **259**, 10857–10862.

148 Matsumoto, K. (1997) Phosphatidylserine synthase from bacteria. Biochim. Biophys. Acta - Lipids Lipid Metab. **1348**, 214–227.

149 Sciara, G., Clarke, O. B., Tomasek, D., Kloss, B., Tabuso, S., Byfield, R., Cohn, R., Banerjee, S., Rajashankar, K. R., Slavkovic, V., et al. (2014) Structural basis for catalysis in a CDP-alcohol phosphotransferase. Nat. Commun. **5**, 4068.

150 Li, Q. X. and Dowhan, W. (1990) Studies on the mechanism of formation of the pyruvate prosthetic group of phosphatidylserine decarboxylase from Escherichia coli. J. Biol. Chem. **265**, 4111–4115.

151 Li, Q. X. and Dowhan, W. (1988) Structural characterization of Escherichia coli phosphatidylserine decarboxylase. J. Biol. Chem. **263**, 11516–22.

152 Schuiki, I. and Daum, G. (2009) Phosphatidylserine decarboxylases, key enzymes of lipid metabolism. IUBMB Life **61**, 151–62.

153 Bligh, E. G. and Dyer, W. J. (1959) A RAPID METHOD OF TOTAL LIPID EXTRACTION AND PURIFICATION. Can. J. Biochem. Physiol., NRC Research Press Ottawa, Canada **37**, 911–917.

154 Farmer, W. R. and Liao, J. C. (2000) Improving lycopene production in Escherichia coli by engineering metabolic control. Nat. Biotechnol. **18**, 533–7.

155 Martin, V. J. J., Pitera, D. J., Withers, S. T., Newman, J. D. and Keasling, J. D. (2003) Engineering a mevalonate pathway in Escherichia coli for production of terpenoids. Nat. Biotechnol. **21**, 796–802.

156 Reiling, K. K., Yoshikuni, Y., Martin, V. J. J., Newman, J., Bohlmann, J. and Keasling, J. D. (2004) Mono and diterpene production in Escherichia coli. Biotechnol. Bioeng. **87**, 200–12.

157 Damen, C. W. N., Isaac, G., Langridge, J., Hankemeier, T. and Vreeken, R. J. (2014) Enhanced lipid isomer separation in human plasma using reversed-phase UPLC with ion-mobility/high-resolution MS detection. J. Lipid Res. **55**, 1772–1783.

158 Szostak, J. W., Bartel, D. P. and Luisi, P. L. (2001) Synthesizing life. Nature

409, 387–90.

159 Gogarten, J. P., Kibak, H., Dittrich, P., Taiz, L., Bowman, E. J., Bowman, B. J., Manolson, M. F., Poole, R. J., Date, T. and Oshima, T. (1989) Evolution of the vacuolar H+-ATPase: implications for the origin of eukaryotes. Proc. Natl. Acad. Sci. U. S. A. **86**, 6661–5.

160 Baymann, F., Lebrun, E., Brugna, M., Schoepp-Cothenet, B., Giudici-Orticoni, M.-T. and Nitschke, W. (2003) The redox protein construction kit: pre-last universal common ancestor evolution of energy-conserving enzymes. Philos. Trans. R. Soc. Lond. B. Biol. Sci. **358**, 267–74.

161 Cao, T. B. and Saier, M. H. (2003) The general protein secretory pathway: phylogenetic analyses leading to evolutionary conclusions. Biochim. Biophys. Acta **1609**, 115–25.

162 de Rosa, M., de Rosa, S., Gambacorta, A., Minale, L. and Bu'lock, J. D. (1977) Chemical structure of the ether lipids of thermophilic acidophilic bacteria of the Caldariella group. Phytochemistry, Pergamon **16**, 1961–1965.

163 Caforio, A., Jain, S., Fodran, P., Siliakus, M., Minnaard, A. J., van der Oost, J. and Driessen, A. J. M. (2015) Formation of the ether lipids archaetidylglycerol and archaetidylethanolamine in Escherichia coli. Biochem. J. **470**, 343–55.

164 Datsenko, K. A. and Wanner, B. L. (2000) One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. Proc. Natl. Acad. Sci. U. S. A. **97**, 6640–5.

165 Baba, T., Ara, T., Hasegawa, M., Takai, Y., Okumura, Y., Baba, M., Datsenko, K. A., Tomita, M., Wanner, B. L. and Mori, H. (2006) Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. Mol. Syst. Biol. **2**, 2006.0008.

166 Yoon, S.-H., Kim, J.-E., Lee, S.-H., Park, H.-M., Choi, M.-S., Kim, J.-Y., Lee, S.-H., Shin, Y.-C., Keasling, J. D. and Kim, S.-W. (2007) Engineering the lycopene synthetic pathway in E. coli by comparison of the carotenoid genes of Pantoea agglomerans and Pantoea ananatis. Appl. Microbiol. Biotechnol. **74**, 131–9.

167 Albermann, C. (2011) High versus low level expression of the lycopene biosynthesis genes from Pantoea ananatis in Escherichia coli. Biotechnol. Lett. **33**, 313–9.

168 Tazuya-Murayama, K., Aramaki, H., Mishima, M., Saito, K., Ishida, S. and Yamada, K. (2006) Effect of L-serine on the biosynthesis of aromatic amino acids in Escherichia coli. J. Nutr. Sci. Vitaminol. (Tokyo). **52**, 256–60.

169 Saha, S. K., Furukawa, Y., Matsuzaki, H., Shibuya, I. and Matsumoto, K. (1996) Directed mutagenesis, Ser-56 to Pro, of Bacillus subtilis phosphatidylserine synthase drastically lowers enzymatic activity and relieves amplification toxicity in Escherichia coli. Biosci. Biotechnol. Biochem. **60**, 630–3.

170 Zhang, C., Chen, X., Zou, R., Zhou, K., Stephanopoulos, G. and Too, H. P. (2013) Combining Genotype Improvement and Statistical Media Optimization for Isoprenoid Production in E. coli. PLoS One **8**.

171 Green, P. R., Merrill, A. H. and Bell, R. M. (1981) Membrane phospholipid

synthesis in Escherichia coli. Purification, reconstitution, and characterization of sn-glycerol-3-phosphate acyltransferase. J. Biol. Chem. **256**, 11151–9.

172 Kameda, K. and Nunn, W. D. (1981) Purification and characterization of acyl coenzyme A synthetase from Escherichia coli. J. Biol. Chem. **256**, 5702–7.

173 Black, P. N. and DiRusso, C. C. (2003) Transmembrane movement of exogenous long-chain fatty acids: proteins, enzymes, and vectorial esterification. Microbiol. Mol. Biol. Rev. **67**, 454–72, table of contents.

174 Yamagishi, M. and Kakinuma, K. (1989) Confirmation of the absolute stereochemistry of s/7-2,3-di-O-phytanyl glycerol,the unit lipid of the cell membrane of halophilic archaebacteria Halobacterim halobium. Agric. Biol. Chem. **53**, 867–868.

175 Ajikumar, P. K., Xiao, W.-H., Tyo, K. E. J., Wang, Y., Simeon, F., Leonard, E., Mucha, O., Phon, T. H., Pfeifer, B. and Stephanopoulos, G. (2010) Isoprenoid pathway optimization for Taxol precursor overproduction in Escherichia coli. Science **330**, 70–4.

176 Mileykovskaya, E. and Dowhan, W. (2005) Role of membrane lipids in bacterial division-site selection. Curr. Opin. Microbiol. **8**, 135–142.

177 Atilla-Gokcumen, G. E., Muro, E., Relat-Goberna, J., Sasse, S., Bedigian, A., Coughlin, M. L., Garcia-Manyes, S. and Eggert, U. S. (2014) Dividing cells regulate their lipid composition and localization. Cell, Elsevier **156**, 428–39.

178 Nichols, D. S., Miller, M. R., Davies, N. W., Goodchild, A., Raftery, M. and Cavicchioli, R. (2004) Cold Adaptation in the Antarctic Archaeon Methanococcoides burtonii Involves Membrane Lipid Unsaturation. J. Bacteriol., American Society for Microbiology **186**, 8508–8515.

179 Sakamoto, T. and Murata, N. (2002) Regulation of the desaturation of fatty acids and its role in tolerance to cold and salt stress. Curr. Opin. Microbiol. **5**, 206–210.

180 DITTMER, J. C. and LESTER, R. L. (1964) A SIMPLE, SPECIFIC SPRAY FOR THE DETECTION OF PHOSPHOLIPIDS ON THIN-LAYER CHROMATOGRAMS. J. Lipid Res., American Society for Biochemistry and Molecular Biology **5**, 126–7.

181 Stewart, J. C. (1980) Colorimetric determination of phospholipids with ammonium ferrothiocyanate. Anal. Biochem. **104**, 10–4.

182 Armarego, W. L. F. and Chai, C. L. L. (2013) Purification of Laboratory Chemicals 7th edition, Butterworth-Heinemann.

183 Fishov, I. and Woldringh, C. L. (1999) Visualization of membrane domains in Escherichia coli. Mol. Microbiol. **32**, 1166–72.

184 Król, E., de Sousa Borges, A., da Silva, I., Polaquini, C. R., Regasini, L. O., Ferreira, H. and Scheffers, D.-J. (2015) Antibacterial activity of alkyl gallates is a combination of direct targeting of FtsZ and permeabilization of bacterial membranes. Front. Microbiol. **6**, 390.

185 Mouzin, G., Cousse, H., Rieu, J.-P. and Duflos, A. (1983) A Convenient One-Step Synthesis of Glycidyl Ethers. Synthesis (Stuttg). **1983**, 117–119.

186 Fodran, P. (2015) Stereoselective Synthesis of Glycerol-based Lipids.
187 Schneiter, R. and Toulmay, A. (2007) The role of lipids in the biogenesis of integral membrane proteins. Appl. Microbiol. Biotechnol. **73**, 1224–32.
188 Bogdanov, M., Xie, J. and Dowhan, W. (2009) Lipid-protein interactions drive membrane protein topogenesis in accordance with the positive inside rule. J. Biol. Chem., American Society for Biochemistry and Molecular Biology **284**, 9637–41.
189 Dowhan, W. and Bogdanov, M. (2009) Lipid-dependent membrane protein topogenesis. Annu. Rev. Biochem. **78**, 515–40.
190 Di Paolo, G. and De Camilli, P. (2006) Phosphoinositides in cell regulation and membrane dynamics. Nature, Nature Publishing Group **443**, 651–657.
191 Wymann, M. P. and Schneiter, R. (2008) Lipid signalling in disease. Nat. Rev. Mol. Cell Biol. **9**, 162–76.
192 Chaurio, R. A., Janko, C., Muñoz, L. E., Frey, B., Herrmann, M. and Gaipl, U. S. (2009) Phospholipids: Key Players in Apoptosis and Immune Regulation. Molecules, Molecular Diversity Preservation International **14**, 4892–4914.
193 Icho, T., Sparrow, C. P. and Raetz, C. R. (1985) Molecular cloning and sequencing of the gene for CDP-diglyceride synthetase of Escherichia coli. J. Biol. Chem., American Society for Biochemistry and Molecular Biology **260**, 12078–83.
194 Langley, K. E. and Kennedy, E. P. (1978) Partial purification and properties of CTP:phosphatidic acid cytidylyltransferase from membranes of Escherichia coli. J. Bacteriol., American Society for Microbiology **136**, 85–95.
195 Liu, X., Yin, Y., Wu, J. and Liu, Z. (2014) Structure and mechanism of an intramembrane liponucleotide synthetase central for phospholipid biosynthesis. Nat. Commun. **5**, 4244.
196 Cheng, W., Li, W., Forsgren, M., Attersand, A., Lake, S., Grünler, J., Swiezewska, E., Dallner, G., Climent, I., Nakagawa, K., et al. (2014) Structural insights into ubiquinone biosynthesis in membranes. Science, American Association for the Advancement of Science **343**, 878–81.
197 Huang, H., Levin, E. J., Liu, S., Bai, Y., Lockless, S. W., Zhou, M., Hirota, Y., Tsugawa, N., Nakagawa, K., Suhara, Y., et al. (2014) Structure of a Membrane-Embedded Prenyltransferase Homologous to UBIAD1. PLoS Biol. (Dutzler, R., ed.), Public Library of Science **12**, e1001911.
198 Chung, B. C., Zhao, J., Gillespie, R. A., Kwon, D.-Y., Guan, Z., Hong, J., Zhou, P. and Lee, S.-Y. (2013) Crystal structure of MraY, an essential membrane enzyme for bacterial cell wall synthesis. Science **341**, 1012–6.
199 Caffrey, M. (2009) Crystallizing membrane proteins for structure determination: use of lipidic mesophases. Annu. Rev. Biophys. **38**, 29–51.
200 Vonrhein, C., Flensburg, C., Keller, P., Sharff, A., Smart, O., Paciorek, W., Womack, T. and Bricogne, G. (2011) Data processing and analysis with the autoPROC toolbox. Acta Crystallogr. D. Biol. Crystallogr. **67**, 293–302.
201 Kabsch, W. (2010) XDS. Acta Crystallogr. D. Biol. Crystallogr. **66**, 125–32.
202 Evans, P. R. and Murshudov, G. N. (2013) How good are my data and what is the resolution? Acta Crystallogr. D. Biol. Crystallogr. **69**, 1204–14.

203 Sammito, M., Millán, C., Frieske, D., Rodríguez-Freire, E., Borges, R. J. and Usón, I. (2015) ARCIMBOLDO_LITE: single-workstation implementation and use. Acta Crystallogr. D. Biol. Crystallogr. **71**, 1921–30.

204 Langer, G., Cohen, S. X., Lamzin, V. S. and Perrakis, A. (2008) Automated macromolecular model building for X-ray crystallography using ARP/wARP version 7. Nat. Protoc. **3**, 1171–9.

205 Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L.-W., Kapral, G. J., Grosse-Kunstleve, R. W., et al. (2010) *PHENIX*: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr. Sect. D Biol. Crystallogr., International Union of Crystallography **66**, 213–221.

206 Emsley, P., Lohkamp, B., Scott, W. G., Cowtan, K., IUCr, W., G.-K. R., L.-W., H., R., I. T., J., M. A., W., M. N., et al. (2010) Features and development of *Coot*. Acta Crystallogr. Sect. D Biol. Crystallogr., International Union of Crystallography **66**, 486–501.

207 Chen, V. B., Arendall, W. B., Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S. and Richardson, D. C. (2010) MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallogr. D. Biol. Crystallogr. **66**, 12–21.

208 Robert, X. and Gouet, P. (2014) Deciphering key features in protein structures with the new ENDscript server. Nucleic Acids Res. **42**, W320–4.

209 Gomez Maqueo Chew, A., Frigaard, N.-U. and Bryant, D. A. (2008) Identification of the bchP gene, encoding geranylgeranyl reductase in Chlorobaculum tepidum. J. Bacteriol. **190**, 747–9.

210 Giannino, D., Condello, E., Bruno, L., Testone, G., Tartarini, A., Cozza, R., Innocenti, A. M., Bitonti, M. B. and Mariotti, D. (2004) The gene geranylgeranyl reductase of peach (Prunus persica [L.] Batsch) is regulated during leaf development and responds differentially to distinct stress factors. J. Exp. Bot. **55**, 2063–73.

211 Nishimura, Y. and Eguchi, T. (2006) Biosynthesis of archaeal membrane lipids: digeranylgeranylglycerophospholipid reductase of the thermoacidophilic archaeon Thermoplasma acidophilum. J. Biochem. **139**, 1073–81.

212 Sato, S., Murakami, M., Yoshimura, T. and Hemmi, H. (2008) Specific partial reduction of geranylgeranyl diphosphate by an enzyme from the thermoacidophilic archaeon Sulfolobus acidocaldarius yields a reactive prenyl donor, not a dead-end product. J. Bacteriol. **190**, 3923–9.

213 Demir, V. and Dincturk, H. B. (2006) Semi-anaerobic growth conditions are favoured by some Escherichia coli strains during heterologous expression of some archaeal proteins. Mol. Biol. Rep. **33**, 59–63.

214 Murakami, M., Shibuya, K., Nakayama, T., Nishino, T., Yoshimura, T. and Hemmi, H. (2007) Geranylgeranyl reductase involved in the biosynthesis of archaeal membrane lipids in the hyperthermophilic archaeon Archaeoglobus fulgidus. FEBS J. **274**, 805–14.

215 Schiraldi, C., Giuliano, M. and De Rosa, M. (2002) Perspectives on biotechnological applications of archaea. Archaea **1**, 75–86.