Archaeal Lipid Biosynthesis: exploring the biosynthetic pathway towards complex membrane phospholipids

de Kok, Niels

DOI:
10.33612/diss.220142471

Publication date:
2022

Archaeal Lipid Biosynthesis: exploring the biosynthetic pathway towards complex membrane phospholipids

Niels Alexander Willem de Kok

university of groningen

Archaeal Lipid Biosynthesis: exploring the biosynthetic pathway towards complex membrane phospholipids

PhD thesis

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

This thesis will be defended in public on

Tuesday the 31st of May 2022 at 14:30 hours

by

Niels Alexander Willem de Kok

born on the 3rd of February 1992
in Helmond, the Netherlands

The research described in this thesis was carried out in the Department of Molecular Microbiology part of the Groningen Biomolecular sciences and Biotechnology institute. (GBB), University of Groningen, the Netherlands. It was financially supported by the Building Blocks of Life programme (737.016.006) of the Netherlands Organization for Scientific research (NWO).

Author: Niels de Kok

Cover: Guus Gijben - www.proefschrift-aio.nl

Layout: Guus Gijben - www.proefschrift-aio.nl

Supervisors

Prof. A.J.M. Driessen
Prof. A.J. Minnaard

Assessment Committee

Prof. T.J.G. Ettema
Prof. T.E.F. Quax
Prof. A. Guskov

Table of contents

Scope of the thesis

The goal of this thesis was to investigate archaeal phospholipid biosynthesis to ultimately unravel how complex phospholipids, such as cardiolipins and tetraether lipids, are synthesized. To accomplish this, a review was written to provide an overview of the phospholipid biosynthetic pathway in relation to enzyme structures in Archaea. This is followed by the description of the development of a new lipid analysis method which proved crucial for the subsequent characterization of two enzymes involved in these pathways. Namely, the study of di-geranylgeranylglycerol phosphate synthase from *Methanocaldococcus jannaschii* and the characterization of an archaeal cardiolipin synthase from *Methanospirillum hungatei*.

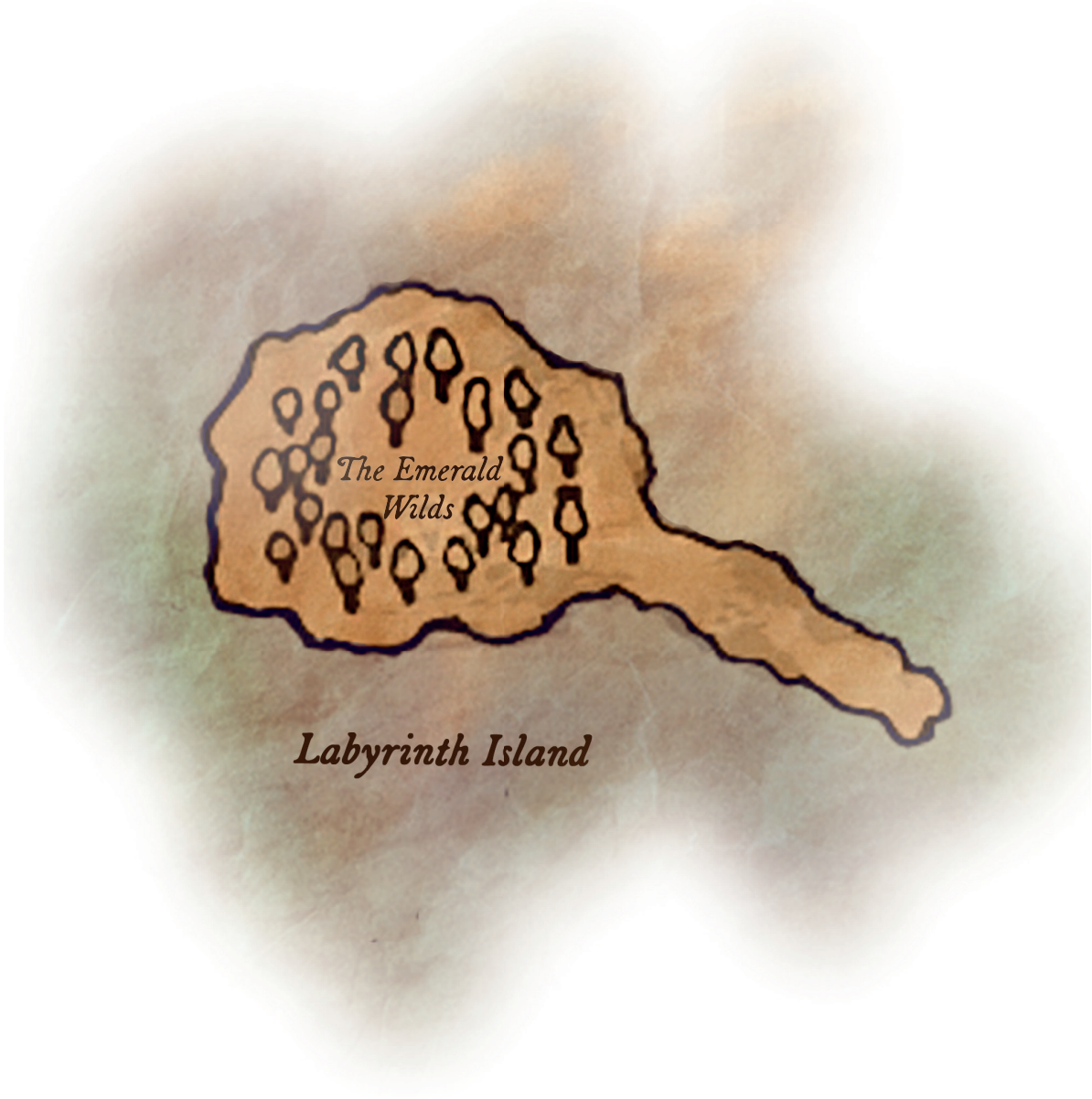
Chapter 1 reports on the enzymology of the archaeal phospholipid biosynthetic pathway and the structures of enzymes involved, leading to new insights pertaining to the "lipid-divide". Although archaeal ether-based phospholipids fulfil similar functions as bacterial ester-based phospholipid, the synthesis of the core phospholipid structure involves enzymatically and structurally unrelated enzymes indicating that the two pathways have evolved independently. In contrast, the pathways for activation and differentiation of lipid headgroups use very similar enzymes that are promiscuous in substrate specificity.

Chapter 2 describes the development of a 1-butanol-based ultra-high-performance-liquid-chromatography method for lipidomics analysis. Its performance is compared with more conventional mobile phase compositions, showing improved elution characteristics of more hydrophobic lipids. This enhances lipidomic analysis, crucial for reliable analysis of lipids as discussed in the other chapters of this thesis.

Chapter 3 focuses on the crystal structure of a key enzyme in archaeal phospholipid biosynthesis, di-geranylgeranyl phosphate synthase (DGGGPS). This work was the first to link the crystal structure of this enzyme to in vitro enzymatic activity, allowing structure-function analysis. Analysis of mutants resulted in the identification of several residues important for catalytic activity. A mutant of DGGGPS analogous to a mutant of human UbiAD1, implicated in the occurrence of Schnyder corneal dystrophy, showed reduced activity, not only providing insights in the molecular basis of disease but also showing that the study of distantly related organisms can lead to benefits for human health.

Chapter 4 discusses the first characterization of a phospholipase-D type archaeal cardiolipin synthase. This enzyme catalyzes the formation of cardiolipin from two archaetidylglycerol molecules. In addition, the reverse reaction using water results in hydrolysis of cardiolipin. In addition to water, the enzyme was shown to utilize a wide range of primary alcohols in the reversed reaction, resulting in novel diester/diether lipids and cardiolipin species. Moreover, this enzyme was shown to be promiscuous towards both archaeal and bacterial lipids yielding a novel hybrid archaeal-bacterial cardiolipin that has not been described before.

Chapter 5 summarizes the findings from the previous chapters and provides a future outlook on some of the topics discussed in this thesis.



Chapter 1

The catalytic and structural basis of archaeal glycerophospholipid biosynthesis

Extremophiles (2022)
In submission

Niels A.W. de Kok [a] and Arnold J.M. Driessen [a]*

[a] Department of Molecular Microbiology, Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, Groningen, Netherlands

* Corresponding author: *a.j.m.driessen@rug.nl*; Tel. + 31-50-3632164; Fax. +31-50-3632154.

Abstract

Archaeal glycerophospholipids are the main constituents of the cytoplasmic membrane in the archaeal domain of life and fundamentally differ in chemical composition compared to bacterial phospholipids. They consist of isoprenyl chains ether-bound to glycerol-1-phosphate. In contrast, bacterial glycerophospholipids are composed of fatty acyl chains ester bonded to glycerol-3-phosphate. This largely domain-distinguishing feature has been termed the “lipid-divide”. The chemical composition of archaeal membranes contributes to the ability of archaea to survive and thrive in extreme environments. However, ether-bonded glycerophospholipids are not only limited to extremophiles and found also in mesophilic archaea. Resolving the structural basis of glycerophospholipid biosynthesis is a key objective to provide insights in the early evolution of membrane formation and to deepen our understanding of the molecular basis of extremophilicity. Many of the glycerophospholipid enzymes are either integral membrane proteins or membrane-associated, and hence are intrinsically difficult to study structurally. However, in recent years the crystal structures of several key enzymes have been solved while unresolved enzymatic steps in the archaeal glycerophospholipid biosynthetic pathway have been clarified providing further insights in the lipid-divide and the evolution of early life.

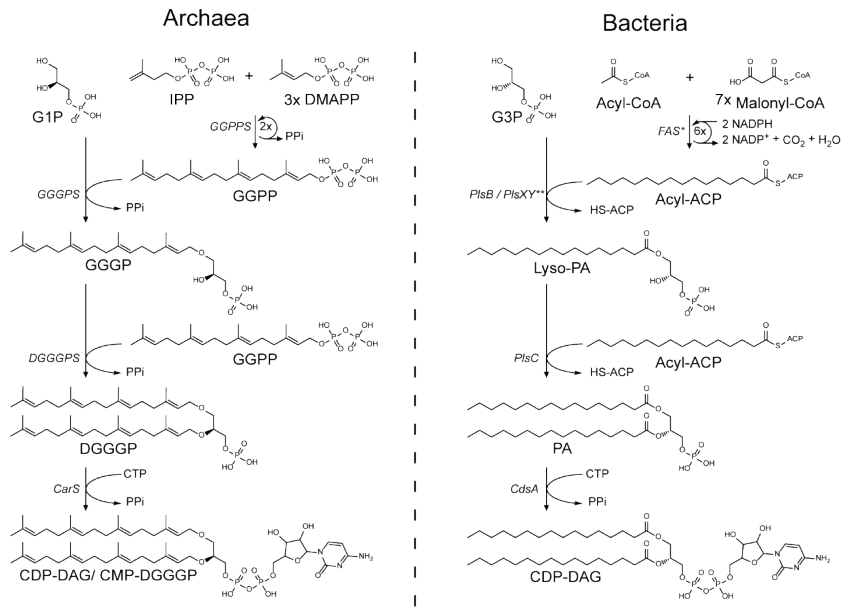


Figure 1: The biosynthesis pathway of the phospholipid core up to and including phospholipid headgroup activation. * In the initial fatty-acid biosynthesis reaction both substrates are transesterified from CoA to ACP. This also occurs during each malonyl-CoA addition cycle, but these ACP groups are regenerated and therefore not shown in this step. ** The PlsB of some organisms accepts acyl-CoA as well and the PlsXY pathway steps and acyl-phosphate intermediate are omitted for clarity.

Introduction

The cell membrane is an essential part of life. Membranes are required for compartmentalization of the cell, allowing for specialized reaction compartments and for maintenance of chemical gradients over the membrane; supporting processes such as transport, ATP synthesis and motility. On the basis of ribosomal RNA sequences, in 1977 Carl Woese proposed a phylogenetic tree of life containing 3 domains: the Eukarya, Bacteria and Archaea [1]. One of the fundamental features that distinguishes Archaea from Bacteria and Eukarya is the difference in phospholipid structure, the main constituent of cell membranes. Members of the domains of Bacteria and Eukarya generally synthesize phospholipids containing fatty acids esterified to glycerol-3-phosphate; whereas Archaea synthesize phospholipids containing isoprene moieties connected to glycerol-1-phosphate via ether bonds; these fundamental differences between archaeal lipids and those found in Bacteria and Eukarya are referred to as the "lipid-divide" [2–5]. While the nature of the hydrophobic radyl groups, the connection to the glycerol backbone, and the chirality of that backbone is markedly different; the core of the phospholipid biosynthesis pathways of Archaea and bacteria run in parallel with a remarkably similar overall organization. However, the enzymatic reactions for the formation of the ether bonds versus the ester bonds requires rather different enzymes and mechanisms (Figure 1). This is in stark contrast with the diversification pathways of the phospholipid polar headgroups, which for most modifications, are strikingly similar and involve similar enzymes.

Comparative studies on the structures of enzymes responsible for membrane lipid biosynthesis can provide insights in the molecular basis for the core differences between bacterial and archaeal lipids. Furthermore, these studies may deepen our understanding of the evolution of early life and what early life might have looked like. Typically, over long evolutionary periods, DNA sequences can be altered to a great extent by mutations, ultimately leading to low primary amino acid sequence homology. In the phospholipid biosynthetic enzymes, homology is primarily found in critical structural elements such as particular folds or substrate and co-factor binding sites. In particular sequences pertaining to hydrophobic features, which are often involved in lipid binding, amino acid identities tend to diverge more over time while amino acid properties are often more conserved. Thus, features could still be structurally conserved in the tertiary structure and be apparent in protein hydropathy profiles and crystal structures [6,7].

This review discusses the enzymatic basis of archaeal phospholipid biosynthesis, combining information from various works about the biochemical characteristics and structures of archaeal lipid biosynthesis enzymes to give an overview of the current state of knowledge. The combined knowledge of these fields revealed new insights which have implications for the continued characterization of the archaeal lipid biosynthesis pathway, and in extension, on the lipid-divide and evolution of early life.

Glycerophosphate lipid core biosynthesis

Mevalonate (MVA) pathway

Isoprenoids constitute a very diverse group of naturally occurring compounds. Several important Isoprenoid compounds include for example carotenoids, steroids, dolichols, plant terpenes and various prenylated compound groups such as archaeal phospholipids, quinones, chlorophylls and prenylated proteins. The essential building blocks of isoprenoids are isopentenyl pyrophosphate (IPP) and dimethylallyl pyrophosphate (DMAPP). Two non-homologous and unrelated pathways are known to exist for the synthesis of these compounds: the mevalonate (MVA) pathway and the unrelated non-mevalonate pathway (MEP/DOXP) [8]. Additionally, to date three alternate MVA pathways were discovered in archaea that are related to the canonical eukaryotic MVA pathway [9–11]. Even though there is no clear domain-related distribution of these pathways, Eukarya and members of the Sulfolobales tend to use the canonical MVA pathway while most other Archaea tend to use one of the three alternate MVA pathways. Bacteria generally synthesize isoprenoids through the MEP/DOXP pathway [8]. One of the three alternate MVA pathways found in archaea is the common "archaeal MVA pathway" named after its probable conservation in most archaea [11,12]. Archaea of the order Halobacteriales (and some members of the Chloroflexi bacteria) use the "haloarchaeal-type MVA pathway" [9,13]; while members of the Thermoplasmatales use the "Thermoplasma-type MVA pathway" [10,14,15].

The upper MVA pathway is shared between all organisms employing a MVA pathway and is responsible for the conversion of acetyl-CoA to MVA through 3 enzymatic reactions. The differences between the canonical eukaryotic- and alternate archaeal-type pathway variants are found in the later steps in the MVA pathway (Figure 2). The canonical eukaryotic MVA pathway converts MVA into IPP in 3 steps: First, MVA is phosphorylated to mevalonate-5-phosphate (MVA-5-P) which is phosphorylated a second time to yield mevalonate-(5)-pyrophosphate (MVA-5-PP) prior to decarboxylation to form IPP. In the haloarchaeal-type MVA

pathway, the second phosphorylation step and decarboxylation are swapped compared to the eukaryotic pathway: MVA is converted into IPP through MVA-5-P and isopentenyl phosphate (IP), a key intermediate for the archaeal pathway variants [9,13]. The Thermoplasma-type MVA pathway converts MVA into IPP using 4 steps through the mevalonate-3-phosphate (MVA-3-P), mevalonate-3,5-biphosphate (MVA-3,5-PP) and IP intermediaries into IPP [10,14,15]. The "archaeal" MVA pathway is similar to the haloarchaeal-type pathway but employs an extra intermediate. As with the canonical eukaryotic- and haloarchaeal-type MVA pathways, MVA is phosphorylated to yield MVA-5-P. MVA-5-P is dehydrated to form (tAMVA-5-P) which is subsequently decarboxylated to IP and phosphorylated to form IPP [11,12,16]. The last step of the lower MVA pathway is shared by all MVA pathway variants and involves the formation of DMAPP through the reversible isomerization of IPP.

Geranylgeranyl pyrophosphate synthesis

Isoprenyl pyrophosphate chains are formed by members of the isoprenyl pyrophosphate synthase (IPPS) family belonging to the prenyltransferase superfamily. Various IPPS are capable of synthesizing different chain lengths which are then used as precursors for terpene-based molecules such as quinones, carotenoids, steroids, hopanoids, dolichols, prenylated proteins or for archaeal phospholipid biosynthesis. Archaeal diether phospholipid biosynthesis typically requires C_{20} and/or C_{25} isoprene chains which are synthesized by all-trans, head-to-tail, short-chain IPPS; or more specifically geranylgeranyl pyrophosphate synthase (GGPPS) and geranylfarnesyl pyrophosphate synthases (GFPPS) respectively (Figure 1). While not used for glycerophospholipid biosynthesis in archaea, hexaprenyl pyrophosphate (HexPPS) and heptaprenyl pyrophosphate synthases (HepPPS) are also members of the isoprenyl pyrophosphate synthase family and are required for the biosynthesis of respiratory quinones [17,18]. The mechanism of action of these enzymes is conserved throughout the all trans-type IPPS family and follows a general "ionization-addition-elimination" mechanism [19]. First, DMAPP is ionized, losing its pyrophosphate group to form a carbocation allowing it to act as a prenyl donor. IPP then performs a nucleophilic attack with C-4 on the C-1 of the dimethylallyl carbocation ("tail-to-head") forming a carbon-carbon bond. This is followed by the elimination of H^+ from C-2 of the IPP moiety, forming the typical isoprenyl trans-double bond, resulting in the geranyl pyrophosphate (GPP, C_{10}) product [20–24]. In archaea, this reaction is iteratively repeated by GGPPS, consuming IPP, forming farnesyl pyrophosphate (FPP, C_{15}); and is repeated again to form geranylgeranyl pyrophosphate (GGPP, C_{20}) which can be used for the biosynthesis of archaeal phospholipids. Additionally, in some archaea, this iterative elongation reaction is repeated once more to form geranylfarnesyl pyrophosphate (GFPP, C_{25}) by GFPPS for the synthesis of C_{25} archaeal lipids [25].

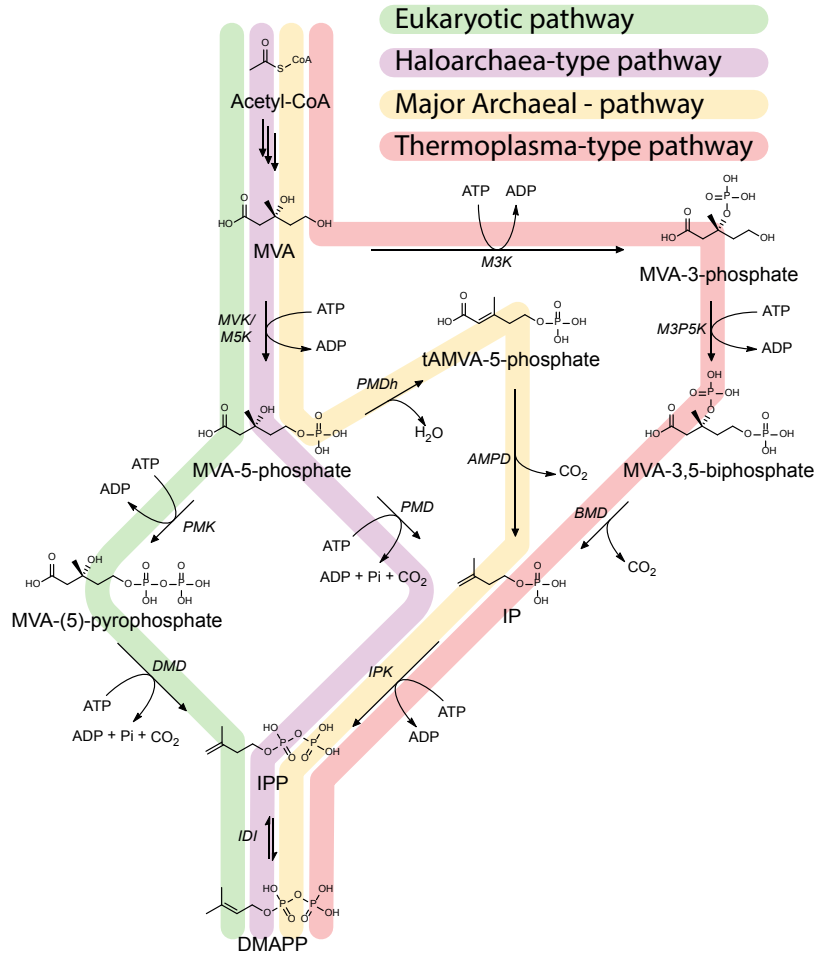


Figure 2: A schematic representation of the four known MVA pathways. The names of the enzymes are italicized. The pathways are marked based on their distribution. MVK, mevalonate kinase; M5K, mevalonate-5-kinase; PMK, phosphomevalonate kinase; DMD, diphosphomevalonate kinase; PMD, phosphomevalonate decarboxylase; PMDh, phosphomevalonate dehydratase; AMPD, anhydromevalonate phosphate decarboxylase; M3K, mevalonate-3 kinase; M3P5K, mevalonate-3-phosphate-5 kinase; BMD, bisphosphomevalonate decarboxylase; IDI, isopentenyl pyrophosphate:dimethylallyl pyrophosphate isomerase.

Remarkably, with some exceptions, class I all-trans head-to-tail isoprenyl pyrophosphate synthases such as FPPS and GGPPS are all structurally very similar proteins with the same basic protein folds and most forming homodimers [18,26–30] (Figure 3). They share substrates with the same basic structure, utilizing IPP as prenyl acceptor with an isoprenyl pyrophosphate donor such as DMAPP to form longer isoprenoid polymers of a specific length depending on the enzyme [31].

The first IPPS structure to be solved was that of avian FPPS [21]; and since various other IPPS structures have been determined. However, lacking archaeal representatives, the only GGPPS crystal structures solved to date are from human, *Plasmodium vivax, Saccharomyces cerevisiae, Oryza sativa* and *Arabidopsis thaliana* [32–36].

Sequence alignments revealed several highly conserved elements such as the first- and second-aspartate-rich motif (FARM and SARM, D-D-x(2)-D or D-D-x(4)-D) [37,38]. The FARM and SARM (along with other conserved residues) were found to be essential for catalysis and coordinate at least two Mg^{2+} ions [39–42]. Studies employing phylogenetic analysis with mutagenesis revealed three different types of GGPPS with a distinct phylogenetic distribution and molecular approach towards their product length specificity (See [31,43] and references therein).

Archaeal GGPPS (type-I) have bulky aromatic limiter residues on the fifth position N-terminal to the FARM without any inserted residues in the FARM. The region between the FARM and the fifth position N-terminal being coined the chain length determination region and is located on helix αD, one of the helices composing the product elongation cavity. Mutational experiments on this region revealed that product chain length could be significantly altered [25,30,44–47]. GGPPS from photosynthetic organisms such as plants, algae and cyanobacteria (type-II) have been less intensively studied and have smaller residues on the fourth and fifth position before the FARM, such as serine, alanine and methionine. Additionally these enzymes contain an insertion of two amino acid residues in the FARM, the first of which is a highly conserved proline residue [36,43]. GGPPS found in mammals and fungi (type-III) have recently attracted renewed interest as a pharmaceutical target in humans [33,42,48,49]. Type-III GGPPS contain similarly small residues at fourth and fifth positions N-terminal to the FARM compared to type-II GGPPS without any inserted residues in the FARM [50]. However, type-III GGPPS were found to employ a different region for chain-length determination and contain a conserved histidine at 2 positions N-terminal to the G(Q/E) motif located on helix αF which, like the FARM motif, is on one of the three helices that form the product elongation pocket [28,29]. Group-I, group-II and long-chain isoprenyl diphosphate synthases contain a variety of smaller residues at this position. Mutagenesis of the conserved histidine to alanine resulted in the longer GFPP product. Moreover, additional alanine mutations of residues located at the 2 and 3 position N-terminal to this location, that approximately corresponds to the pitch of an α-helix, resulted in the formation of even longer products [50]. Similar findings were observed with mutational studies employed on other closely related enzymes [21,46]. In essence, the positioning of larger residues along the product elongation cavity limits determines the final product chain length, reminiscent to the function of a hydrocarbon ruler [51].

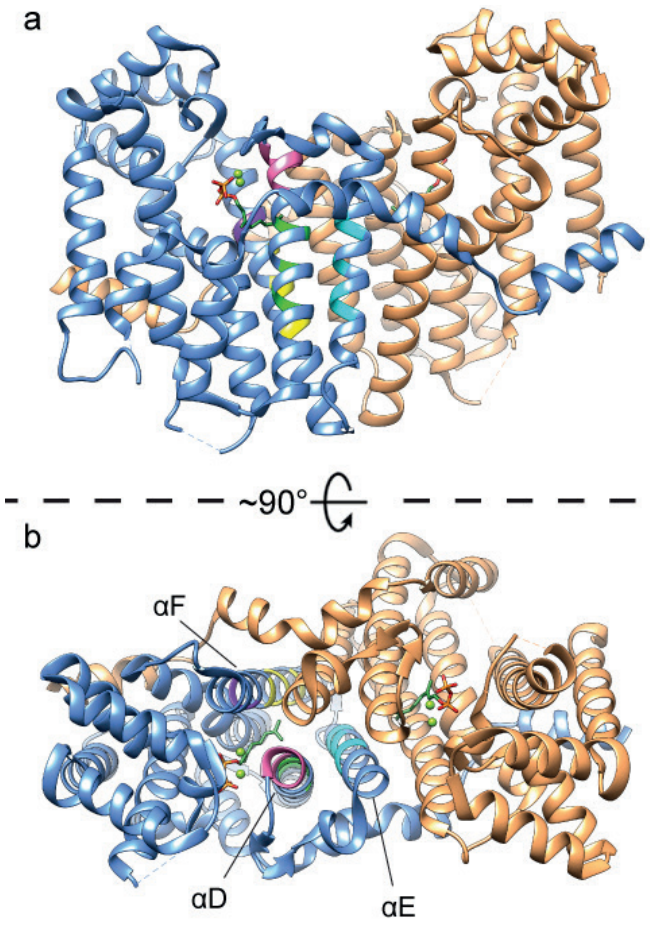


Figure 3: Side view (a) and top-down view (b) of the *Saccharomyces cerevisiae* GGPPS crystal structure with bound GPP (PDB: 2E8X, [32]). Both peptide chains of the crystal structure dimer are shown (Sandy brown for chain A and Cornflower blue for chain B). The FARM motif is highlighted in pink and the G(Q/E) motif in purple. The helices that form the product elongation cavity are annotated. Putative limiter residues based on the "three-floor" model are highlighted in yellow, green and cyan. Mg^{2+} ions are shown as green spheres.

Taken the mutagenesis data on the GGPPS types together, a study on type-II GGPPS proposed a unifying "three-floor" model which explains the molecular basis on final product chain length determination of short-chain IPPS [36]. The positioning of large- or medium-sized residues at positions corresponding to particular "floors" in any of the helices that make up the product elongation cavity (αD, αE or αF, figure 3b) limit the final product size. Large residues on the "first-floor" limit the enzyme to FPP as the final product, with medium residues on the "first-floor" or large residues on the "second-floor" limiting to GGPP and large residues on the "third-floor" resulting in a GFPP end-product [36,43].

Formation of the first ether bond to glycerol-1-phosphate

Prenyltransferases are responsible for the formation of the ether bonds in archaeal phospholipid biosynthesis. The first ether bond formation is catalyzed by geranylgeranylglycerol phosphate synthase (GGGPS) and involves the formation of an ether bond between glycerol-1-phosphate (G1P) and GGPP which results in geranylgeranylglycerol phosphate (GGGP) (Figure 1). This reaction marks the first committing step into the biosynthesis of archaeal phospholipids. Presumably, this biosynthesis step occurs in the cytosol as GGGPS enzymes do not contain transmembrane helices and are purified as soluble proteins from the soluble fraction of cell lysates [52–58].

In contrast, in bacteria, fatty acids are ester linked to glycerol-3-phosphate (G3P) (Figure 1). The first enzymes discovered performing this reaction were the enzyme of the PlsB/PlsC pathway. In this pathway the radyl moiety is attached through transesterification of acyl-ACP (or in some cases also acyl-CoA) to the C-1 position of G3P by the membrane-bound PlsB; resulting in the formation of lyso-phosphatidic acid (LPA) [59–62]. However, the increased availability of genome sequencing techniques revealed that the presence of PlsB is mostly limited to γ-proteobacteria [63]. Instead, most bacteria use the PlsX/PlsY/PlsC pathway, in which membrane-associated PlsX converts acyl-ACP to an acyl-phosphate intermediate, which subsequently is attached to C-1 of G3P by the integral membrane protein PlsY to form LPA [64–67]. Despite the similarities in overall organization of the bacterial and archaeal phospholipid biosynthetic pathways, the enzymatic reactions involved are very different. Hence, PlsB, PlsX and PlsY are not members of the prenyltransferase family and thus are also not structurally related to GGGPS.

Phylogenetic analysis of members of the GGGPS prenyltransferase sub-family revealed the presence of two distinct groups, group I and group II, which are further subdivided based on archaeal- (Ia and IIa) or bacterial origins (Ib and IIb) [57]. To date, the protein data bank (PDB) contains 6 reported GGGPS crystal structures.

Notably, most structures are from thermophilic archaea such as: *Archaeoglobus fulgidus* (AfGGGPS, group Ia, [52]), *Methanothermobacter thermoautotrophicus* (MtGGGPS, group IIa, [57]), *Thermoplasma acidophilum* (TaGGGPS, group IIa, [56]) and *Thermoplasma volcanium* (TvGGGPS, group IIa, [54]). The structure of a bacterial GGGPS (group IIb, [57]) from *Flavobacterium johnsoniae* was also reported (FjGGGPS). The GGGPS crystal structures all show a modified triose phosphate isomerase (TIM-) barrel structure in which an eight-stranded parallel β-barrel with a tightly packed hydrophobic core is surrounded by α-helices constituting a $(\beta\alpha)_8$-barrel structure (Figure 4a). Characteristic GGGPS TIM-barrel modifications include an additional helix α0 at the N-terminus. Helix α3 is often replaced by a string or strand without any secondary structure whereas in other cases is modelled as a short helix. Helix α5 is split in which helix α5′ is located on top of the barrel over the active site. Helix α3* has been suggested to act as a "swinging door" or part of a ratchet mechanism to aid in expelling the hydrophobic GGGP product from the hydrophobic GGPP binding site [52].

Archaea synthesize phospholipids with the G1P stereochemistry. Experiments with both purified enzyme and crude extracts showed high substrate specificity of GGGPS for G1P over G3P [68–70]. At high concentration, the *Methanococcus maripaludis* GGGPS can utilize both G1P and G3P, but exhibits a 8-fold higher affinity for G1P as compared to G3P [71]. The substrate specificity of GGGPS is a key point for the synthesis of lipid cores with the correct stereochemistry and therefore the lipid-divide. The positioning of the G1P phosphate group is facilitated by the TIM-barrel "standard phosphate-binding motif" [72,73]. This motif interacts with G1P through backbone amino groups and side-chains of loops βα6 (S169, G170), βα7 (G194, G195) and βα8 (V214, G215 and N216) (numbering according to AfGGGPS, group I) [52]. Notably, N216 is replaced by a conserved threonine in group II enzymes. It has been proposed that three strictly conserved tyrosine and glutamate residues are the main factors responsible for both substrate stereospecificity and catalytic activity in both group I and group II enzymes (Y124/Y165/E167 and Y134/Y178/E180 in AfGGGPS and MtGGGPS respectively) [52,57,74]. Of these three residues, the glutamate could facilitate prenyl acceptance by polarizing the G1P C-3 hydroxyl group [57].

The current proposed reaction mechanism for the GGGPS family involves the ionization of GGPP to form an allylic geranylgeranyl carbocation through removal of its pyrophosphate moiety by the Mg^{2+} ion and subsequent nucleophilic attack of the polarized C-3 hydroxyl group of G1P on C-1 of the geranylgeranyl carbocation forming the ether bond [54,57,74,75]. For enzymatic activity, the enzyme needs to bind the hydrophobic polyprenyl substrate of a specific length and the

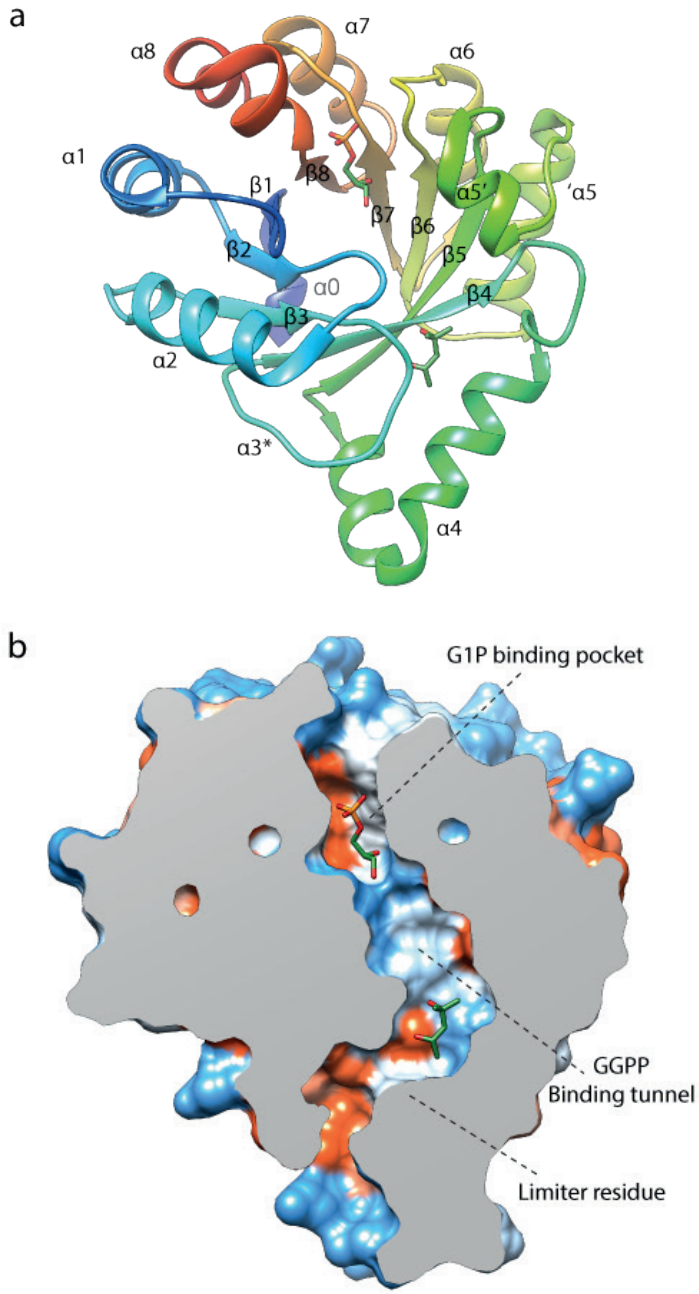


Figure 4: Ribbon view (a) and clipped solid-surface view (b) of the AfGGGPS crystal structure with bound G1P and 2-methyl-2,4-pentanediol (PDB: 2F6X, [52]). The surface view is colored according to the Kyte-Doolittle hydrophobicity scale. Red surfaces are hydrophobic, white is of mixed character and blue is hydrophilic.

diphosphate group should be oriented in such a way that it can be coupled to the C-3 hydroxyl oxygen of G1P. The diphosphate group of GGPP has been suggested to bind in complex with the Mg^{2+} ion which is coordinated by at least an essential conserved D13 and possibly by T37 in AfGGGPS (D25 and possibly E27/E28 and S54 in MtGGGPS) as shown by mutagenesis experiments and as seen in other prenyl transferases [52,57,74–76]. The GGPP binding pocket is varies considerably between different GGGPS crystal structures. It takes the form of a cleft [54,56], tunnel [52] or pocket [57], with a predominantly hydrophobic- or mixed-character. In AfGGGPS, the GGPP binding tunnel has an opening at the polar G1P binding site and runs over the top of the barrel towards the N-terminus of helix α4a and curves down over β3 and β4 between helix α4a and the loop between β3 and β4 (Helix α3*) [52,74] (Figure 4b). This tunnel is limited by a bulky limiter residue in helix α4a in group I enzymes (W99 in AfGGGPS) and bears some similarities to "hydrocarbon rulers" found in other acyl- and prenyl transferases such as GGPPS [27,51,52,77]. In group II enzymes, the limiter residue is positioned on helix α3' instead of helix α4a and is not as clearly defined as in group I since smaller residues seem to be involved (proposed to be I90 in FjGGGPS and V86 in MtGGGPS) [57].

The prenyltransferase family also contains enzyme subfamilies with different substrate specificity such as the bacterial heptaprenyltransferases (PcrB, "GGGPS" group Ib) which are homologous to archaeal GGGPS. Members of the PcrB subfamily catalyze the transfer of heptaprenyl diphosphate moieties (C_{35}) to G1P forming heptaprenylglycerol phosphate in contrast to the shorter tetraprenyl moieties (C_{20}) transferred by GGGPS [58,78]. Being homologs, the structures of these enzymes are quite similar. In addition, the shape and amino acid identity of the G1P binding pocket was found to be strongly conserved between *A. fulgidus* GGGPS and *Bacillus subtilis* PcrB [52,78]. The difference in substrate specificity between these two subfamilies was found to be caused by the bulky limiter residue at the bottom of the GGPP binding site in helix α4a of AfGGGPS (W99). For AfGGGPS the isoprenyl chain length is limited to C_{20} (GGPP), while in the related BsPcrB this residue is replaced with alanine (A100) and substrate limiting is facilitated by Y104, situated lower on helix α4a to allow for polyprenyl substrates up to C_{35} in length. Interestingly, haloarchaea producing C_{20}-C_{25} ether lipids were found to possess two paralogous GGGPS group I enzymes referred to as subgroups IaH1 and IaH2. Haloarchaeal GGGPS IaH1 enzymes contain the prototypical tryptophan limiter residue whereas IaH2 enzymes have a leucine which might allow for the binding of longer substrates [57]. These structural elements may form the molecular basis of the presence of C_{20}-C_{25} ether lipids in these organisms and supports a previous hypothesis by Boucher, et al. which stipulates that a few mutations could be enough to alter prenyl donor selectivity of GGGPS [79]. The crystal structure of

the more distantly related MoeO5 prenyltransferase has been solved as well [76]. MoeO5 shares many structural features with GGGPS and PcrB; such as a similar substrate binding site. However, the reaction catalyzed by MoeO5 is rather unusual as the prenyl group of FPP is transferred to 3-phosphoglycerate, during which the intramolecular trans-allylic bond of FPP gets converted into a cis-allylic bond.

GGGPS homologs adopt different oligomerization states, for example: AfGGGPS, TaGGGPS, TvGGGPS and FjGGGPS form dimers [52,54,57,69], whereas MtGGGPS, *Thermococcus kodakarensis* GGGPS and *Chitinophaga pinensis* GGGPS were found to form hexamers [57]. A correlation was found between similarity network clustering and the oligomerization state as revealed by selective light scattering and size exclusion chromatography. This showed that GGGPS enzymes belonging to group I form dimers; whereas almost all group II members are expected to form hexamers with the exception of bacterial GGGPS and GGGPS from Thermoplasmatales [53,55,57]. Most group II enzymes were found to have a conserved aromatic anchor residue (W, Y, F) on helix α5 participating in a cation-pi interaction which was shown to be essential for hexamerization [55,57]. Recently, Kropp, et al. [53] showed that most other group II GGGPS, which did not have the typically conserved W, Y or F aromatic anchor and thus were initially expected to form dimers, contained a histidine residue instead. This histidine was shown to be able to participate in cation-pi interactions and therefore those group II proteins are also expected to form hexamers [53]. The hexameric state of GGGPS has previously been described as a trimer of dimers and can be visualized as tilted dimers interlocking at a 120-degree angle forming a stacked "upper and lower" ring of three polypeptides of different dimers each (Figure 5). Thus, three different interfaces connect the subunits: the symmetrical dimer interface, which is found in natively dimeric GGGPS and is diagonal to the horizontal plane of the two-stacked-ring hexameric complex connecting the rings together; the ring interface, which contains the aromatic anchor residue and laterally connects the three polypeptides of three different dimers forming a ring; and the interconnection interface, which is also situated between three polypeptides of the three different dimers vertically connecting the two rings forming the stacked ring shape. Studies looking at residues which play a role in the formation of the hexameric complex found than an aromatic anchor residue in helix α5' to be essential to the formation of hexameric GGGPS complexes [53,55,57].

Oligomerization has been regarded as a factor enhancing thermostability [80,81]. Linde, et al investigated the effect of oligomerization on the thermostability of MtGGGPS and found that the temperature of the first denaturation transition step, which correlates with a loss of catalytic activity, increased with a higher order oligomeric state [55]. This indicates that the hexameric state confers increased thermostability. Furthermore, a dimeric MtGGGPS mutant had its catalytic efficiency drastically reduced (Km (G1P) value x50 higher), indicating hexamerization stabilizes the G1P binding pocket [53]. MD simulations at denaturing temperatures revealed that the flexibility of four regions was significantly higher than in other parts of the protein compared to non-denaturing temperature simulations. Three of these regions, Helix α4, loop βα6 and helix α5', are involved in substrate binding, explaining the loss of activity during the first denaturation transition. However, the region situated in helix α5' is of particular interest as this region normally covers the GGPP binding site but also contains the aromatic anchor residue which is part of the ring interface and would stabilize this region if the complex would be in the hexameric quaternary state. Thus, it has been suggested that hexamerization allows for a balance between more flexibility (possibly related to activity at lower temperatures) and thermostability. In contrast, not all hyperthermophilic GGGPS adopt hexameric oligomerization states, and thus it seems not a strictly essential feature for hyperthermophilic GGGPS enzymes. Furthermore, a recent ancestral sequence reconstruction (ASR) analysis study by Blank, et al. [54] which focused on the aromatic anchor residues (specifically W, Y and F only) showed that these residues are the result of convergent evolution without temperature being the primary driving factor for this development. This is in contrast with the study by Kropp, et al. [53] who also applied ASR analysis, and based on thermal denaturation experiments of ASR enzymes argues that hexamerization evolved as a general thermostability feature in group II GGGPS.

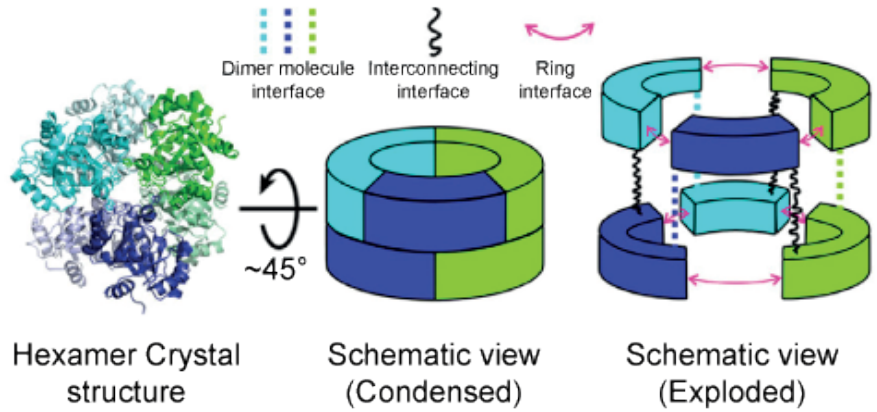


Figure 5: The hexameric configuration of MtGGGPS with a schematic representation of the different interfaces facilitating hexamerization. The dimers are colored green, cyan and purple. Adapted with permission from Linde, et al [55]. Copyright 2018 American Chemical Society.

Formation of the second ether bond to form the diether

The second ether bond formation between GGGP and GGPP results in di-geranylgeranylglycerol phosphate (DGGGP) and is catalyzed by di-geranylgeranylglycerol phosphate synthase (DGGGPS) (Figure 1), a member of the UbiA superfamily. Structural analysis shows that MjDGGGPS is an integral membrane protein with an N-terminal amphipathic helix, followed by 9 transmembrane helices forming a central cavity containing two hydrophobic lipid-binding tunnels (tunnel 1 and tunnel 2) and a cytoplasmic opening with a lateral portal into the membrane [82] (Figure 6). The cytoplasmic opening crowning the central cavity is referred to as the cytoplasmic domain and contains three loop/helix structures capping the cytoplasmic opening. The bottom of the lateral portal is surrounded by a pronounced "belt" of hydrophobic residues spanning ~30 Å in overall thickness covering the entire circumference of the enzyme and likely corresponds to the positioning of the phospholipid bilayer of the host organism *Methanocaldococcus jannaschii* (Figure 6b).

In bacteria, the analogous reaction to the formation of the second *ether* bond by DGGGPS, is the formation of the second *ester* bond facilitated by the membrane-associated protein PlsC [83–85]. PlsC facilitates the transesterification of acyl-ACP (sometimes acyl-CoA is accepted) with LPA resulting in the formation of PA. Similar to PlsB, PlsC belongs to family of phospholipid acyltransferases. Hence PlsC is not structurally or mechanistically related to DGGGPS. Hence, the early steps in the phospholipid biosynthetic pathways in Archaea and Bacteria must have evolved independently. Members of the UbiA superfamily generally transfer a polyprenyl group to hydrophobic acceptor molecules. However, the accepting groups and overall structure of acceptor molecules varies widely, including: An ethylene carbon in protoheme [86,87], a carbon as part of an aromatic ring for the synthesis of quinones or tocopheroles [88–91], the carboxyl carbon in a propionate side chain of chlorophyllides [92], or the C-2 of the glycerol backbone containing a hydroxyl group in GGGP [82,93,94]. Remarkably, phylogenetic analysis by Hemmi, et al. [94] revealed that UbiA enzymes cluster according to their prenyl acceptor substrate structure, and not prenyl donor substrate structure. This indicates that the conserved motifs are likely related to recognition of the prenyl donor and catalytic mechanism. Indeed, these structural elements in MjDGGGPS resemble that of soluble prenyl transferases such as GGPPS and FPPS [28,95].

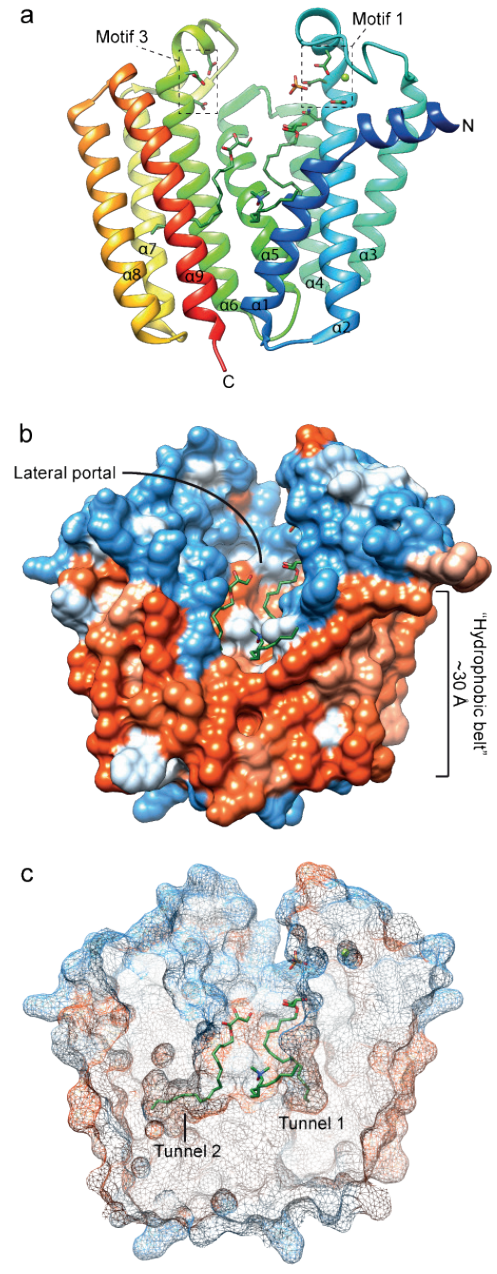


Figure 6: Ribbon view (a), solid-surface view (b) and clipped mesh-surface view of the MjDGGGPS crystal structure with bound [octadec-9-enyl]-2,3-hydroxy-propanoate and lauryldimethylamine oxide (LDAO) (PDB: 6M31, [82]). The surface views are colored according to the Kyte-Doolittle hydrophobicity scale. Red surfaces are hydrophobic, white is mixed character and blue is hydrophilic. Mg^{2+} ions are shown as green spheres. In panel A the aspartic acid residues of motif 1 and 3, and asparagine 62 are shown in liquorice style.

The molecular mechanism of the UbiA superfamily was explored by studying the activity of MjDGGGPS mutants [82]. Members of the DGGGPS family contain two conserved Asp-rich motifs with the D_{66}-x(3)-D_{70} (motif 1) and D_{183}-x(2)-D_{187}-x(3)-D_{190} (motif 3) motifs as well as a conserved Y_{125}-x(5)-K_{130} (motif 2) motif (MjDGGGPS numbering). Other members of the UbiA superfamily such as CoQ2, UbiA, MenA, Cox10 also contain similar motifs. These highly conserved motifs are each located on one of the capping loop/helix structures indicating that these structures are directly or indirectly involved in catalysis. This idea is further reinforced by the Mg^{2+} ion that is coordinated by motif 1 and the positioning of the pyrophosphate moiety of DMAPP, free phosphates and 2,3-hydroxy-propanoate "lipid backbone" of the co-crystallized [octadec-9-enyl]-2,3-hydroxy-propanoate. MjDGGGPS activity assays revealed that alanine mutants of D66, D70, D180, D183 and D187 strongly reduced MjDGGGPS activity, indicating that these residues indeed perform critical roles in catalysis. Moreover, in line with other prenyltransferases, MjDGGGPS prefers Mg^{2+} over other divalent cations while EDTA abolishes the activity, showing that divalent cations are essential for enzymatic activity. It was proposed that D66 and D70 are responsible for coordination of an Mg^{2+} ion whereas D180, D183 and D187 can either coordinate another Mg^{2+} ion or the pyrophosphate moiety by hydrogen bonding. In this context it is noteworthy that the previously reported *A. fulgidus* UbiA structures show two Mg^{2+} or Cd^{2+} ions bound in regions corresponding to motif 1 and motif 3 [96]. The structure of DGGGPS suggests that the orientation of the hydroxyl group on C-2 of GGGP could be important for catalysis as this is where the coupling of GGGP with the C-1 of GGPP happens. This is consistent with the chirality of the GGGP glycerol backbone produced by GGGPS in vivo. However, DGGGPS has been shown to accept G3P based GGGP as well [97]. Hence, it appears that this enzyme is less critical in defining the overall chirality of the phospholipid backbone compared to GGGPS.

Lipid binding experiments showed that the motif 1 (N62A, D66A, D70A) and motif 3 (D180A, D183A and D187A) mutations only had negligible effect on the binding of GGPP whereas GGGP binding was impaired to varying degrees indicating that these motifs are important for GGGP binding and catalysis. Mutations of residues lining "lipid-binding tunnel 2", I29, S171 and F148, abolished GGPP, but not GGGP binding. These data defined the substrate binding positions in the enzyme (GGPP in tunnel 2 and GGGP in tunnel 1) and indicates Mg^{2+} might not only have a catalytic role but might play a role in GGGP binding or positioning as well. This was further corroborated by MD simulations suggesting that GGGP binds less deep in the central cavity of the enzyme as compared to GGPP.

The N62A and Y125A mutants showed significantly decreased activity, without a large decrease in substrate affinity, indicating these residues do not play a major role in catalysis or substrate binding but might aid in properly orienting the substrates for catalysis. N62 and Y125 are expected to be located in close proximity of the glycerol backbone of GGGP hinting at a role in correct positioning of the prenyl acceptor (glycerol backbone C-2) to be accessible for the prenyl donor (GGPP, C-1). N62 is also located near the Mg^{2+} ion bound by motif 1, possibly aiding in Mg^{2+} positioning.

Interestingly, residue N62 in MjDGGGPS corresponds to N102 in the human UBIAD1 and is located along the central cavity and cytoplasmic opening (Figure 6a). Mutations in that residue and other nearby residues have been implicated in the occurrence of Schnyder corneal dystrophy (SCD) and other diseases caused by quinone deficiency in humans [98,99]. This association was further emphasized by the lack of activity of the corresponding mutants of homologously expressed CoQ2 (from *S. cerevisiae*) and heterologously expressed Human UBIAD1 [82]. The study of MjDGGGPS was the first study of an UbiA superfamily member that could directly couple enzymatic activity to a crystal structure.

The exact catalytic mechanism of DGGGPS remains to be confirmed. However, because of the similarity in conserved motifs, tertiary structure and prenyl donor substrates, DGGGPS is expected to employ a similar catalytic mechanism as that of soluble prenyltransferases and other members of the UbiA superfamily [82,96,100].

Polar head group activation and modification

The next step in archaeal phospholipid biosynthesis is the activation of DGGGP with CTP for polar headgroup attachment, yielding CMP-DGGGP (CDP-archaeol[1]) (Figure 1). CDP-archaeol is a key intermediate in the pathway and the precursor for phospholipid headgroup differentiation. The aforementioned reaction is achieved by the cytidyl-diphosphate-archaeol synthase enzyme CarS. The CTP-transferase reaction of CarS in Archaea is analogous to CDP-DAG formation in Bacteria but is achieved by a different, only distantly related enzyme (CdsA), although both enzymes belong to the CTP-transferase superfamily [101]. The activity of this enzyme was first demonstrated in crude lysates of *M. thermoautotrophicus* using synthetic substrates [102]. Activity assays using the crude membrane fraction of *M. thermoautotrophicus* showed a strong specificity towards the conversion of lipid substrates containing unsaturated geranylgeranyl groups, such as DGGGP over archaetidic acid (AA) or fatty acid based lipids. However, no strong selectivity was reported for DGGGP analogs containing ester instead of ether bonds or containing G3P- instead of G1P-backbone stereochemistry. Furthermore, catalytic studies with purified *Archaeoglobus fulgidus* CarS (AfCarS) confirm the high selectivity for DGGGP compared to bacterial PA; while the *Escherichia coli* CdsA shows significantly more activity on phosphatidic acids compared to DGGGP [101]. In addition, the archaeal AfCarS could not complement the conditional growth defect of an *E. coli cdsA* mutant strain. Taken together, these studies shows that CarS and CdsA effectively differentiate between archaeal and bacterial substrates and uphold the lipid-divide. However, the molecular basis of CarS substrate specificity between PA and DGGGP has not been studied *in vitro* in detail with purified enzyme instead of a crude membrane fraction [102].

To date, only the crystal structure of *Aeropyrum pernix* CarS (ApCarS) has been reported [103]. Five transmembrane helices provide the primary structural elements with two cytoplasmic loops forming a polar cytoplasmic domain (Figure 7a). The cytoplasmic domain contains a polar central cavity which in turn harbors a small hydrophobic cytidyl binding pocket on one side and a small polar phosphate binding pocket on the other side (Figure 7b). The central cavity is positioned in a cup shaped by the termini of transmembrane helices α1 to α4 where transmembrane helix α5 only seems to loosely cover the central cavity from the side. The rest of the central cavity is comprised by the two cytoplasmic loops forming the top half of the cavity. Despite being an integral membrane protein, ApCarS does not have hydrophobic tunnels to accommodate isoprenoid tails like DGGGPS but rather has two lipid binding grooves (LBG) split by helix α5. Moreover, the polar cavity is located around the top of the hydrophobic "belt" around the enzyme resulting in a somewhat thinner belt section directly adjacent to the active site (~18 Å, Figure 7b). The rear outer surface of the polar cavity is remarkably hydrophobic as well resulting in a ticker hydrophobic belt (~30 Å, Figure 7c). This might be an indication that the enzyme is able to tilt backwards into the membrane, possibly as a mechanism to bind CTP or discharge PPi (Figure 7c and 7d). Or, depending on perspective, the enzyme may tilt forward to bind DGGGP which would be present in the membrane. Alternatively, this large hydrophobic surface on the rear of the enzyme could serve as a surface for interaction with other proteins. The reaction mechanism of CarS is proposed to be comprised of 3 steps: Polarization of the CTP α-phosphate, transesterification by nucleophilic attack of the negatively charged phosphate oxygen of DGGGP on the CTP α-phosphate followed by product release [103].

ApCarS was co-crystallized with CTP that is bound in the polar cavity with the pyrimidine moiety pointing into the hydrophobic pocket and with the triphosphate moiety being coordinated by a Mg^{2+} and K^{+} ion [103]. This agrees with the observation that CarS is Mg^{2+} dependent, with its activity being abolished in the presence of EDTA. Furthermore, the presence of K^{+} ions considerably increases the activity of ApCarS [101,103]. The positioning of the triphosphate moiety is of particular interest as the γ- and β-phosphate are buried in a highly polarized pocket, whereas the α-phosphate is mostly solvent exposed and accessible for a nucleophilic attack by the DGGGP acceptor substrate.

DGGGP binds to CarS through hydrophobic interactions with two grooves lined with conserved hydrophobic residues and through hydrogen bonding with polar residues in the central cavity. These two lipid binding grooves are primarily formed by helix α5 loosely covering the central cavity containing the CTP and Mg^{2+} binding sites (Figure 7b). Interestingly, the *Thermotoga maritima* CdsA structure (TmCdsA, PDB: 4Q2E) around the active site aligns well with the structure of CarS. The largest structural difference seems to be caused by a different positioning of ApCarS helix α5 compared to the corresponding helix in TmCdsA (Helix α1), resulting in a different shape of the hydrophobic groove. The lipid binding grooves in ApCarS are overall wider compared to the groove in TmCdsA, rendering TmCdsA unable to properly accommodate the bulkier isoprenoid chains of archaeal phospholipids compared to the less-bulky bacterial fatty acyl based phospholipids [103]. This is likely the reason why *E. coli* CdsA does not readily accept DGGGP as a substrate [102].

1 In this work "archaeol" does not specifically refer to di-phytanyl-glycerol or di-sesterterpanyl-glycerol but rather as a general term for the archaeal diether glycerol core lipid regardless of its saturation state.

CdsA and CarS are both members of the CTP transferase family but only share a low degree of homology, highlighting that the enzymes are evolutionary distant and have diverged significantly from one another forming two separate CTP transferase subfamilies. Even though the overall structures of CarS and CdsA have diverged significantly, the enzyme catalytic core and structural elements involved in the binding of substrates have been conserved reasonably well.

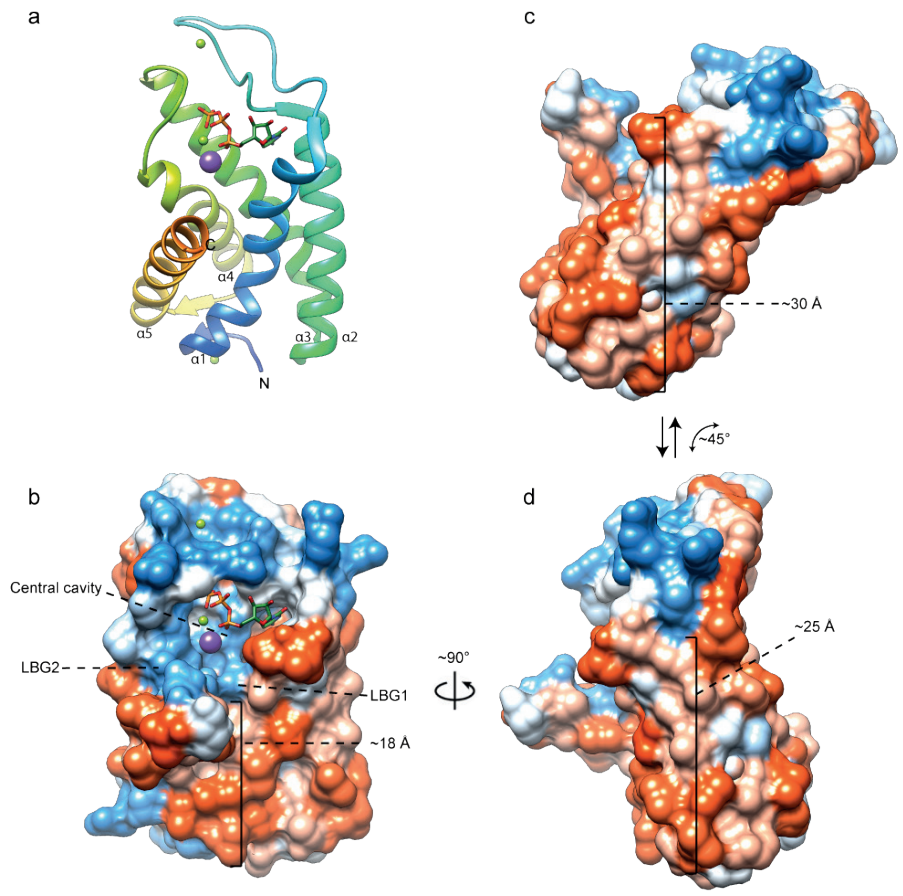


Figure 7: Ribbon view (a) and solid-surface view (b, c, d) of the ApCarS crystal structure with bound CTP (PDB: 5GUF, [103]). The surface views are colored according to the Kyte-Doolittle hydrophobicity scale. Panels (**c**) and (**d**) illustrate the suggested tilting motion where ApCarS could adopt a position that would favor CTP-binding/PPi release (**c**) and a position that would encourage lipid-substrate binding or product release (**d**). Red surfaces are hydrophobic, white is of mixed character and blue is hydrophilic. Mg^{2+} and K^{+} ions are shown as green and purple spheres respectively.

Phospholipid Polar Headgroup modification

The now-widespread use of liquid chromatography coupled to high-resolution MS (LC-MS) techniques has uncovered an unprecedented diversity of lipids in the domain of Archaea [104–113]. As such, there are many possible phospholipid headgroup modifications after the phosphate moiety of DGGGP has been activated by CarS. Thus headgroup modification is discussed from the perspective of well-studied bacteria such as *E. coli* or *B. subtilis* of which homologous enzymes have been identified in Archaea [114,115] (Figure 8).

The first enzymes involved in phospholipid headgroup modification after CDP-archaeol formation generally belong to the CDP-alcohol phosphatidyl transferase (CAPT) family. At this point the phospholipid biosynthesis pathway diversifies and the enzymes responsible for the synthesis of particular phospholipids becomes species specific. Members of this family catalyze the displacement of CMP from the CDP-alcohol with another alcohol, such as: glycerol-3-phosphate, *myo*-inositol phosphate, serine or ethanolamine. This enzymatic family is rather diverse in phylogenetic distribution and sequence length. However, they all share a hallmark CAPT-motif (D-x(2)-D-G-x(2)-A-R-x(7,8)-G-x(3)-D-x(3)-D, PROSITE PS00379) responsible for coordination of one or two catalytically important divalent metal ions [115–121]. Most members of the CAPT family employ an ordered sequential Bi-Bi mechanism for which both substrates have to bind before catalysis can take place and a product can be released [122–126]. Phylogenetic analysis revealed that bacterial and archaeal CAPTs cluster into groups according to their substrate specificity, for example: G3P, *myo*-inositol phosphate or serine [115,121,127,128]. Representatives are found for all three groups in bacteria, but only in the latter two groups are found in archaea. Due to the similarity in sequences between archaeal *myo*-inositol phosphate transferases and G3P transferases, the archaeal G3P transferases cluster into the *myo*-inositol-phosphate transferase group [115].

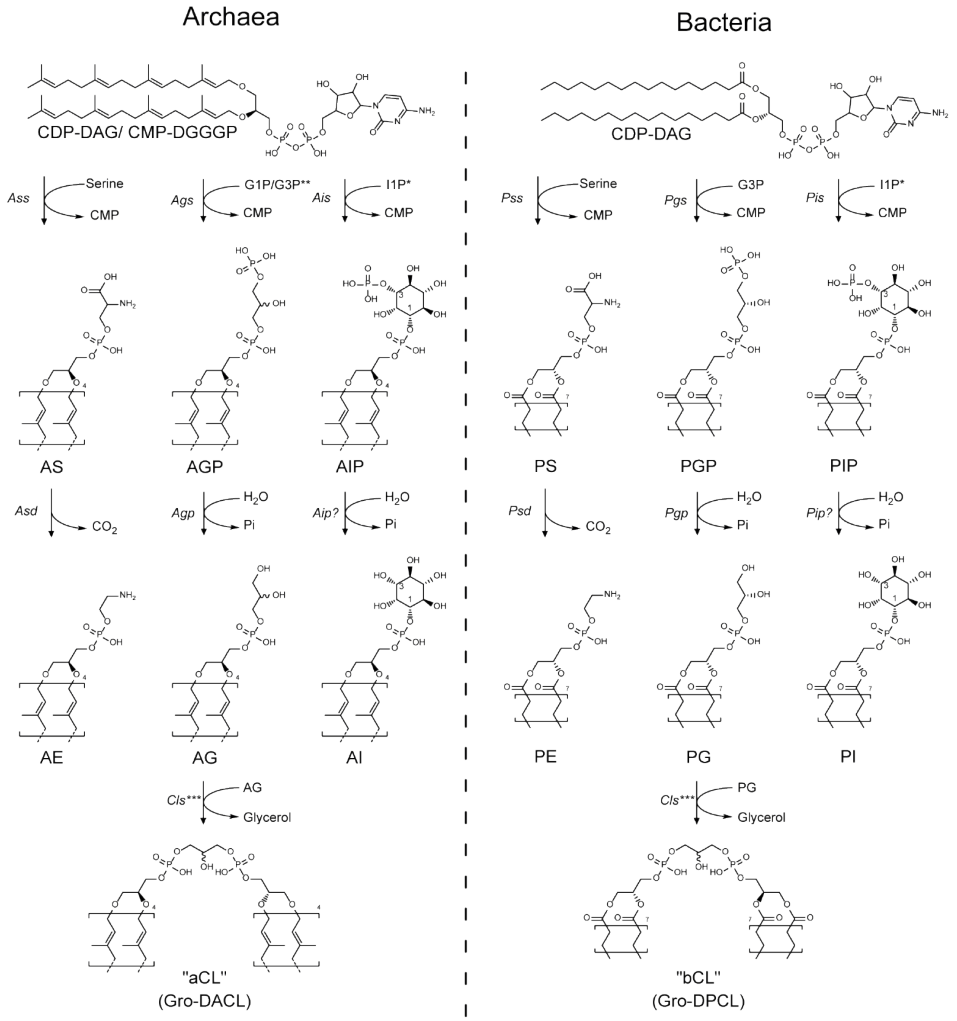


Figure 8: Phospholipid headgroup differentiation pathways of selected common phospholipid headgroups in Archaea and/or Bacteria. * 1L-*myo*-inositol-1-phosphate. The numbering changes once bound to the phosphatidyl phosphate. ** No information is available on whether archaeal Pgs (Ags) shows any specificity towards G1P or G3P. *** For clarity, only the prototypical addition of two AG/PG molecules to form a single archaeal- or bacterial-CL molecule (aCL/bCL) and glycerol is shown. It is not known whether the substrates for aCL formation are required to be fully saturated or whether saturation can still take place after aCL formation.

Phosphatidylethanolamine synthesis

Phosphatidylethanolamine (PE) forms a major portion of the phospholipids in the membranes of both *E. coli*, *B. subtilis* and many other bacteria [129]. In bacteria, PE is usually synthesized from CDP-DAG in a two-step pathway (Figure 8): The CMP moiety in CDP-DAG is exchanged for L-serine by phosphatidylserine synthase (Pss), typically a CAPT[2], to form phosphatidylserine (PS). PS is subsequently decarboxylated by phosphatidylserine decarboxylase (Psd) resulting in PE [130–132]. Homologs of these enzymes, archaetidylserine synthase (Ass) and archaetidylethanolamine decarboxylase (Asd), have been identified in some archaeal species, indicating that the biosynthetic pathway in these organisms is analogous to that of bacteria [114,115,128,133]. However, to date, no archaeal Ass or Asd structures have been published and most insights in this family of proteins have been obtained with the bacterial enzymes.

A structure of Pss from *Haemophilus influenzae* (PDB: 3HSI) has been deposited in PDB, but has not further been described. Recently, the structure of *E. coli* Psd has been solved [134,135]. Initially, EcPsd is transcribed as a single polypeptide. This pro-peptide undergoes an autocatalytic maturation event entailing proteolytic cleavage of a highly-conserved LGST motif of the pro-peptide into a smaller α-subunit and larger membrane-associated β-subunit [136,137]. Psds are unusual decarboxylases as they contain a unconventional pyruvoyl prosthetic group in the active site [138]. The serine-protease autoproteolytic mechanism leaves a dehydroalanine residue as the N-terminus of the α-subunit in lieu of the original serine of the pro-peptide LGST motif [139]. This dehydroalanine residue is rehydrated and consequently eliminates ammonia to form the pyruvoyl prosthetic group required for enzymatic activity [140,141].

The EcPsd α- and β-subunit together create one globular fold consisting of 7 α-helices and 18 β-strands in total. The structural core of the protein is formed by the main β-sheet-like structure consisting of 7 β-strands (Figure 9a). One broad side of the main sheet is covered by 4 small α-helices and 2 β-strands while the other broad side is covered by 8 β-strands of which 4 strands form a small β-sheet-like structure parallel to the main β-sheet. One of the short sides of the main β-sheet is crowned by the 3 large N-terminal α-helices with an amphipathic character arranged in a “U”-shape which harbors the substrate binding pocket and active site and likely peripherally associates the enzymes with the membrane and enables phospholipid substrates to diffuse from the membrane into the active site

2 *E. coli* PssA is atypical as it is a member of the phospholipase D family containing the prototypical HKD motif but capable of forming the same phosphodiester bond as members of the CAPT family.

[134] (Figure 9b and 9d). The pyruvoyl moiety is located on the N-terminus of the α-subunit which forms a β-strand which is securely stabilized as part of the main β-sheet-like structure (Figure 9a). The rest of the α-subunit forms 3 small β-strands that interact with the main β-sheet-like structure.

The crystal structures of EcPsd with bound PE show that the interactions of the lipid tail are mostly facilitated by non-specific hydrophobic interactions (Figure 9c and 9d). However, the phosphate moiety seems to engage in hydrogen bonding interactions with Y137 and the amide backbone of V167. In some structures the conserved S166 side chain also forms a hydrogen bond with PE [135]. The conserved residue H144 is in close proximity to the ethanolamine headgroup and is known to play an essential role in the decarboxylase activity; while D90 and H147 are not located in the active site and are likely only involved in protein maturation as suggested by mutagenesis experiments in which mutants of D90, H147 and S254 result in failure of protein maturation (and subsequent loss of activity). Regarding mutagenesis of H144, conflicting results have been reported. One study shows that H144 mutants of EcPsd are significantly impaired in pro-peptide maturation and lost almost all activity while in another study, apparent autoproteolytic maturation was demonstrated of mutants of H144 and H147 [134,135]. This could indicate that the conserved autoproteolytic catalytic residues (except for the conserved serine) might not be absolutely essential and other, non-conserved residues can fill-in and participate in autoproteolysis.

Several archaeal Pss homologs have been identified in archaea and two have been studied [128,133]. Notably, the Pss and Psd homologs of *Haloferax volcanii* (Ass and Asd) play a critical role in S-layer glycoprotein lipidation and the proteolytic processing of archaeosortase A (ArtA) substrates.

Traditionally, phospholipid biosynthesis was seen as homogenous process resulting in phospholipids that are randomly distributed over the membrane. The notion of heterogeneity in the phospholipid membrane, "a mosaic of phospholipids", was developed some decades ago, and has been further explored in various studies on phospholipid localization and lipid microdomains [131,142–146]. To this day, the heterogeneity of the archaeal phospholipid membrane and the enzymatic organization of phospholipid biosynthesis remains largely unexplored. The observation that Pss and Psd localize to the midcell in *H. volcanii* suggests that phospholipid biosynthesis is, at least to some extent, organized in space [131,133,147]. Further, despite not sharing the typical membrane-integral CAPT sequence and lacking transmembrane helices, *E. coli* PssA has been shown to be associated with the membrane where it forms complexes with acyl carrier

protein (ACP), PlsB and YbgC, suggesting that phospholipid biosynthesis is indeed organized in space [148]. With the observation in *H. volcanii*, it is not known whether these enzymes are specifically co-located, potentially forming a functional complex; or whether these proteins are associated due to an association with cell wall biosynthesis [131].

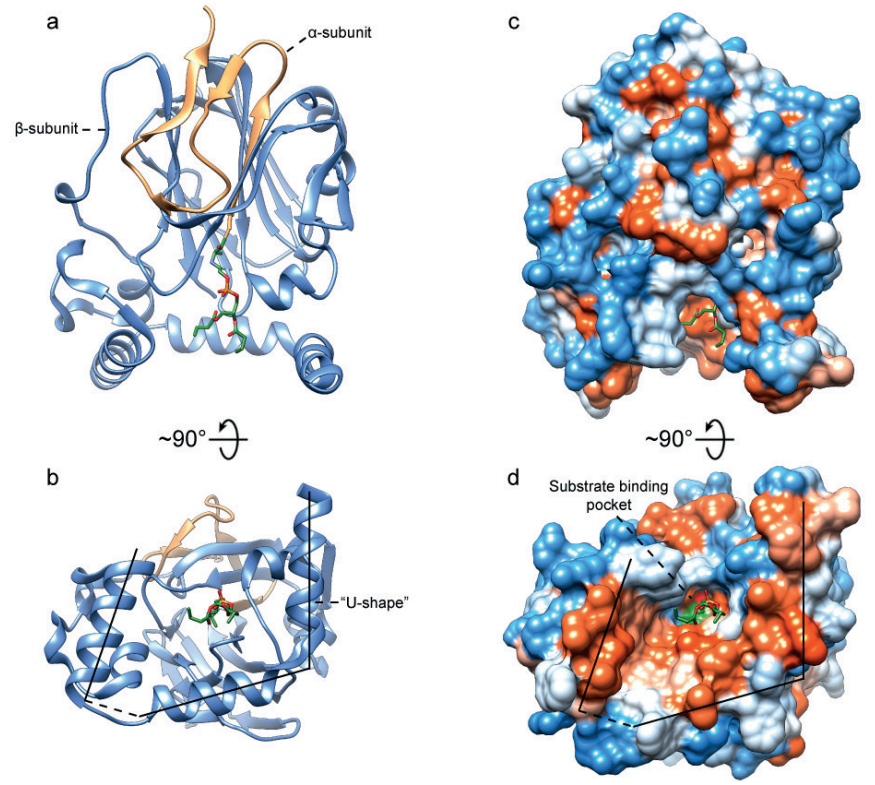


Figure 9: Ribbon view (a, b) and solid-surface view (c, d) of a monomer of the EcPsd crystal structure with bound PE (PDB: 6L07, [134]). The peptide chains corresponding to α- and β- subunits in the ribbon view are colored sandy brown and cornflower blue respectively. The surface views are colored according to the Kyte-Doolittle hydrophobicity scale where red surfaces are hydrophobic, white is of mixed character and blue is hydrophilic.

BLAST analysis shows that *H. volcanii* Pss is only distantly related to characterized homologs from *M. thermoautotrophicus* and *B. subtilis*. The active domains are conserved, but the evolutionary distance between these enzymes is quite large. Several motifs containing essential residues for the maturation of the enzyme have been identified [136,140]. Interestingly, HvPsd lacks one of the conserved two histidine residues (H147 in EcPsd, R89 in HvPsd and H198 in PkPsd), which has been shown to partake in the autoproteolytic triad and be essential for pro-peptide

maturation in the eukaryote *Plasmodium knowlesi* (PkPsd) [136]. However, this particular residue was not essential for the autoproteolytic processing of Psd1 in the eukaryote *Saccharomyces cerevisiae,* for which H345 was the essential residue (corresponding to H144 in EcPsd, H86 in HvPsd and H195 in PkPsd) [137].

EcPsd is capable of decarboxylating archaetidylserine (AS) to form archaetidylethanolamine (AE) [116]. However, earlier studies hitherto did not provide a strong structural basis for the substrate specificity of Psd enzymes and it is uncertain how well structures are conserved between bacterial Psd and archaeal Asd enzymes. Nevertheless, it appears that the enzyme is rather promiscuous, as it can accept both bacterial and archaeal phospholipid substrates; suggesting Psd has no, or only a mild preference for glycerol backbone chirality, or diester over diether phospholipids [116]. Also, there do not seem to be specific binding pockets for the radyl groups present in this enzyme. However, non-specific hydrophobic radyl group interactions are required for proper substrate binding and activity, as serine and phosphoserine do not compete with phosphatidylserine as substrate and glycerophosphoserine is not decarboxylated by EcPsd [149]. Instead, the specific interactions of EcPsd are centered around the domain-agnostic PS headgroup with the predominantly positively charged active site drawing in PS and stabilizing the phosphate backbone through Y137, S166 and potentially the backbone amide of V167. Unfortunately, residue Y137 is not conserved in archaeal Asds and is replaced with phenylalanine, while S166 is replaced with a proline or valine which all lack a polar group to stabilize the phosphate headgroup. These residues are of critical importance in EcPsd; therefore, the mode of substrate stabilization in archaeal Asd remains to be elucidated [134,135].

Phosphatidylglycerol and phosphatidylinositol synthesis

A very common phospholipid headgroup found in both Bacteria and Archaea is a phospholipid with a glycerol headgroup, phosphatidylglycerol (PG) and archaetidylglycerol (AG) respectively. While phosphatidyl *myo*-inositol (PI) is not as common in Bacteria, archaetidylinositol (AI) is a relatively common phospholipid in Archaea. PG and PI are synthesized in similar ways from CDP-DAG in two steps: In bacteria, the CMP headgroup is exchanged with G3P by phosphatidylglycerol phosphate synthase (Pgs) or with 1L-*myo*-inositol-1-phosphate[3] by phosphatidyl inositol phosphate synthase (Pis, occasionally also referred to as Pgs); both enzymes are CAPTs, forming phosphatidylglycerol phosphate (PGP) and phosphatidyl inositol phosphate (PIP) respectively [150] (Figure 8). The G3P and inositol

3 This work only discusses 1L-*myo*-inositol. From this point onwards "1L-*myo*-" is omitted for reading clarity.

phosphate headgroups of PGP and PIP are subsequently dephosphorylated by phosphatidylglycerol phosphatase (Pgp) and an unidentified phosphatidyl inositol phosphatase (Pip) to yield the plain glycerol- and inositol-headgroups of PG and PI respectively [127,151]. Although it is thought that, like bacterial Pgs, the archaeal homolog (Ags) uses G3P as a substrate [116]; to our knowledge, so far no experimental evidence has been published to confirm the stereospecificity of archaeal Ags for G1P or G3P.

E. coli Pgs (PgsA) contains the hallmark CAPT-motif and is predicted to form 6 transmembrane helices forming an integral membrane protein which is typical for members of the CAPT family [125,152]. Only recently, the first crystal structure of a bacterial PgsA has been reported for *Staphylococcus aureus* (SaPgsA) [153]. PI synthesis has been studied in more detail and three structures of bacterial Pis are deposited in PDB [121,151,154] (Figure 10). Several archaeal homologs of bacterial Pgs, Pis and Pgp (Ags, Ais and Agp) have been identified [115,155]. However, to date, no structures of archaeal Ags, Ais or AgpA have been reported, nor have these enzymes been studied in detail.

Both Pgs and PgpA from *E. coli* have been shown to be active on the archaeal analogs of CDP-DAG (CDP-archaeol) and phosphatidylglycerol phosphate (archaetidylglycerol phosphate, AGP) to form archaetidylglycerol [116]. *E. coli* contains two other Pgp homologs, PgpB and PgpC [156–158]. PgpB has a distinct substrate specificity from PgpA. While PgpA only dephosphorylates PGP, PgpB is also active on DAG pyrophosphate, PA and LPA [158,159].

The binding of inositol-1-phosphate (I1P) to Pis is dependent on the binding of CDP-DAG, suggesting an ordered sequential bi-bi rection mechanism as is common for most other members of the CAPT family [122,123,125,126]. It has been suggested that the reaction mechanism involves the deprotonation of an I1P hydroxyl group by a general base which enables the subsequent nucleophilic attack of I1P on the β-phosphate of CDP-DAG. The reaction presumably proceeds through a penta-coordinated intermediate before breaking the diphosphate bond and releasing the phosphatidylinositol phosphate and CMP products [121]. The activity of bacterial Pis on archaeal substrates compared to bacterial substrates is considerably lower, indicating some degree of substrate specificity [155]. Remarkably, the same study shows that while Ais accepted both substrates, it showed more activity on bacterial substrates compared to archaeal substrates. The latter may also be caused by non-optimal reaction conditions, as previously, it was shown that Ais activity in *M. thermoautotrophicus* membranes was greatly stimulated by increasing the detergent concentration [127]. Variations in used detergent concentration may give

rise to activity differences. Further, archaeal phospholipids have higher transition temperatures compared to bacterial phospholipids and thus the assay temperature might have a considerable impact on enzymatic activity [155].

The crystal structures of Pis from *Mycobacterium tuberculosis* (MtPis) and *Renibacterium Salmoninarum* shows these enzymes as dimers, possessing the typical CAPT structure with 6 transmembrane helices per monomer forming integral membrane proteins [121,154] (Figure 10). The MtPis crystal structure notably reveals the binding mode of two catalytically important Mg^{2+} cations which are coordinated by D68, D71, D89 and D93 of the conserved CAPT-motif (Figure 10b and 10c). Due to the complete loss of activity of D93 mutants it has been suggested that, in addition to coordinating one of the Mg^{2+} ions, D93 likely acts as a catalytic base for the deprotonation of the I1P hydroxyl group [119,121,124,154]. The cytidyl nucleotide moiety of CDP-DAG is bound into a pocket between TM helices α1-α3 (Figure 10a and 10c). The nucleotide binding pocket is lined by G72, A75, G85 and residues D31, T34 and T82 form weak hydrogen bonds with the nucleotide moiety. The diphosphate moiety of CDP-DAG interacts with at least one of the Mg^{2+} ions and the α-phosphate is additionally coordinated by G72 and R76. MtPis possesses a pronounced hydrophobic cleft running from the negatively charged diphosphate binding site containing the Mg^{2+} ions on the intracellular side of the enzyme towards the extracellular side of the enzyme [154]. Although the CDP-DAG radyl tails were poorly resolved and therefore truncated in the crystal structure, this hydrophobic cleft is presumed to be the radyl group binding site. The crystal structure of engineered MkPis shows the same general features but also shows the binding site of the phosphatidyl acceptor substrate I1P (residues are numbered according to MtPis) [151] (Figure 10d). Residues S132, K135, R155 and R195 were previously hypothesized to be involved in I1P binding and the MkPis crystal structure confirms this. Residue R94 and residues Y133 and R137 of the other peptide chain are positioned to interact with the phosphate moiety of I1P. Mutagenesis of R137 dramatically reduces enzyme activity, indicating the enzyme is functional as a dimer. Residue R152 is structurally strictly conserved in CAPT enzymes and is positioned to form the floor between the I1P and CDP-DAG binding pockets. Despite its essentiality, the function of R152 remains to be elucidated.

The structure of SaPgsA appears highly reminiscent of the MtPis crystal structure described above [121,153]. The phosphatidyl moiety of the co-crystallized PGP and CDP-DAG moieties is positioned in virtually the same way. Moreover, the cytidyl nucleotide headgroup of CDP-DAG is bound to the CDP binding pocket and oriented in the same way compared to the CDP-DAG in the MtPis crystal structure. Finally, the phosphatidyl acceptor binding site of SaPgsA (G3P) seems somewhat conserved.

Mainly lysine and arginine residues that are located on similar positions as I1P binding residues in MtPis interact with the G3P polar headgroup of co-crystallized PGP.

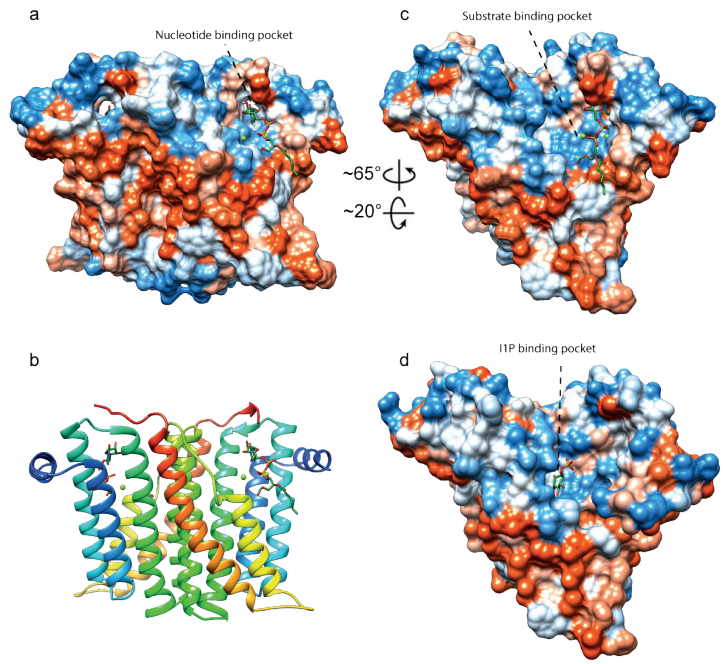


Figure 10: Solid-surface view (a, c, d) and ribbon view (b) of the MtPis dimer (a, b, c) and MkPis dimer (d) crystal structure with bound CDP-DAG and inositol-1-phosphate respectively (PDB: 6H59 and 6WMV, [121] and [151] respectively.). The surface views are colored according to the Kyte-Doolittle hydrophobicity scale where red surfaces are hydrophobic, white is of mixed character and blue is hydrophilic. Mg^{2+} ions are shown as green spheres.

Overall, archaea and bacteria use similar types of enzymes to produce the various phospholipids with common polar headgroups. These enzymes likely have a common origin which was possibly already present in the last universal common ancestor (LUCA) [2,4,114]. So far several studies showed that CAPTs can show some specificity, but ultimately tend to accept both archaeal and bacterial substrates [116,128,155]. However, CAPT substrate preference has not been studied in detail, while these enzymes are expected to be more likely to show some substrate specificity between archaeal and bacterial substrates when compared to downstream enzymes. CAPTs act on the shared CDP headgroup, they are integral membrane proteins and thus interact more closely with the lipid radyl groups of phospholipid

substrates compared to downstream enzymes [119–121]. Because of the diversity of downstream enzymes and limited structural insights, aspects of enzyme substrate specificity remain to be elucidated.

Cardiolipin synthesis

Cardiolipins (CL) are a class of phospholipids present in membranes in all three domains of life and were found to be involved in osmoregulation, membrane organization and is associated with bioenergetic proteins [160–173]. The structure of CL is rather unusual as it involves a phospholipid linked to a phospholipid or glycolipid through a polyol headgroup such as glycerol or a sugar moiety. The most commonly studied cardiolipin is 1,3-bis(*sn*-3'-phosphatidyl)-*sn*-glycerol, referred to as glycerol-di-phosphatidyl cardiolipin (Gro-DPCL) or simply cardiolipin. This phospholipid is usually a relatively minor component of the lipidome of *E. coli* or *B. subtilis* and has also been identified in Archaea [113,168,169,174–177]. Notably, one of the cardiolipin species identified in a halophilic archaeon, was a sulfated glyco-cardiolipin, found strongly associated with bacteriorhodopsin [169]. The headgroup of the sulfated glyco-cardiolipin bears a resemblance to phosphorylated- or sulfated hexoses such as di-*myo*-inositol phosphate or trehalose-sulfate which were both identified in archaea as osmolytes [178–180]. Moreover, the stimulation of *de novo* synthesis of these lipids upon osmotic shock, suggests that these lipids could play a role in osmoadaptation and stabilization of proteins [160,168,181].

Two main types of Cls have been identified. Typically, bacterial cardiolipin synthases (Cls) are members of the phospholipase-D (PLD) superfamily which also contains PS synthases (EcPssA), several endonucleases, poxvirus envelope proteins, a murine toxin from *Yersinia pestis* and the prototypical phospholipase D enzymes after which the superfamily is named [182]. PLD-type Cls synthesize CL through the reversible condensation of two phosphatidylglycerol moieties, forming CL and releasing a glycerol moiety in the process. In addition to cardiolipin synthase activity, PLD-type Cls have been shown to exhibit typical phospholipase D activity as well, irreversibly hydrolyzing the PG headgroup to yield PA and glycerol [183,184]. The second group of Cls belong to the CAPT family and have been identified in eukaryotes; forming CL by linking CDP-DAG to the terminal hydroxyl group of the glycerol headgroup on PG, releasing a CMP moiety in the process [173,182]. CAPT-type Cls have only been found only in Eukaryotes and Actinobacteria. Hence this work will only discuss PLD-type Cls from this point onwards as these are also present in Archaea.

PLD-type Cls have been subdivided into 4 subtypes based on various properties: ClsA-, ClsB- and ClsC-type and putative halophilic archaeal Cls. ClsA-types contain a conserved hydrophobic N-terminal domain not found in ClsB or ClsC [185]. ClsC clearly forms a different group on the basis of sequence similarity to ClsA and ClsB and EcClsC has been shown to be capable of utilizing PE in contrast to EcClsA and EcClsB [186]. The putative haloarchaeal Cls form a separate phylogenetic group, distinct from ClsA, ClsB, ClsC and known PLDs [184] (Figure 11a). Interestingly, phylogenetic analysis revealed at least 5 fairly distinct groups. Based on the phylogenetic tree one could argue that the ClsA group could be subdivided into ClsA1 and ClsA2; with both groups containing members which possess the conserved N-terminal region and ClsA1 mostly containing members from Gram-negative bacteria and the ClsA2 group methanogenic archaea and Gram-positive bacteria (Figure 11a).

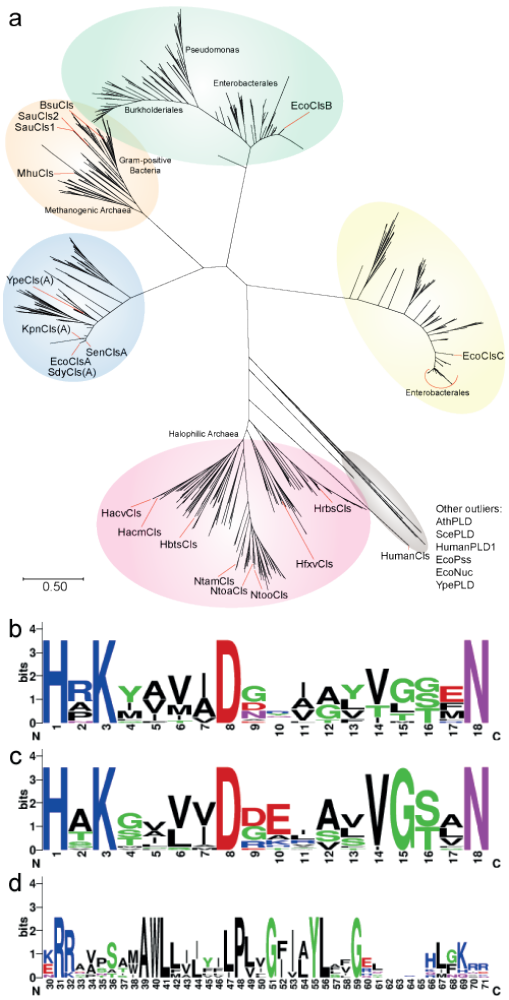


Figure 11: (a) A maximum likelihood phylogenetic tree showing the phylogenetic distribution of different types of Cls. Weblogo sequence of the first- **(b)** and second HKD motif **(c)** from a selection of in total 18 Cls from well-known bacteria, MhCls and several haloarchaeal PLD-type Cls homologs [187]. The weblogo of the second N-terminal hydrophobic region **(d)** was generated from only bacterial ClsA-type enzymes and MhCls. The phylogenetic tree was constructed from sequences obtained by BLAST analysis on archaea and bacteria with a maximum of 5000 hits using EcClsA, B, C, MhCls and Cls homologs from *H. volcanii* and *Halobacterium salinarum*. Prior to maximum likelihood phylogenetic tree construction using the JTT model in MEGA-X [188,189], EcClsA, B and C hit redundancy was reduced to 80% sequence similarity and for the other queries 99%. The top-200 hits of each redundancy reduced query were then pooled, duplicates removed, and redundancy was again reduced to 80% sequence similarity which yielded a final list of 798 hits. Cls homologs (18) of selected organisms and outliers (7) were added before alignment using Clustal Omega [190].

Some bacteria possess multiple Cls homologs. *Staphylococcus aureus* is known to possess at least two closely related Cls homologues, of which one is thought to be active during low-pH stress conditions in which the other Cls is not active [191]. *E. coli* possess three cardiolipin synthases, ClsA, ClsB and ClsC, of which ClsA has been the most thoroughly studied [185,186]. Archaea typically contain a single Cls homolog but their distribution is mostly restricted to halophiles and methanogens. In *E. coli,* the expression of the Cls homologues differs during the various stages of culture growth and stress conditions. The exact purpose of multiple Cls is not completely clear. Aside from differential expression during different growth phases or stress conditions, different Cls subtypes have a different substrate specificity allowing them to either use different phospholipid and/or alcohol substrates [186,192,193]. However, this has not been exhaustively studied with comparable *in vitro* conditions for all different Cls subtypes from a single organism or close homologs. Perhaps, Cls enzymes are involved in phospholipid remodeling.

PLD-type Cls, like typical PLD enzymes, contain a duplicated PLD domain which consists of a set of four structural regions. The third region of the set contains a highly conserved HKD motif (H-x-K-x(4)-D) for a total of two HKD motifs per Cls enzyme [194,195]. Some sources cite the HKD motif to be longer, including residues from the fourth structural region (H-x-K-x(4)-D-x(6)-G-[GST]-x-N) while others refer to these extra residues as a separate motif [185,196,197]. Notably, the degree of conservation between the first and second (extended) HDK motif among Cls is different. With the Cls from several well-known bacteria, including the Cls from *Methanospirillum hungatei* (MhCls) and several putative haloarchaeal PLD-type Cls, the first extended HKD motif is less conserved and rather resembles "H-x-K-x(4)-D-x(6)-[GLT]-[GST]-x-N" whereas the second HKD motif is more defined and resembles the cited motif "H-x-K-x(4)-D-x(5)-[VI]-G-[ST]-x-N" (Figure 11b and 11c). Additionally, *E. coli* ClsA (EcClsA) contains an N-terminal domain of approximately 60-141 residues in length which seems to be present in ClsA homologs from other bacteria (and including MhCls). It is especially prevalent in Gram-negative bacteria such as members of the Enterobacterales [198]. The regions of residues 7-29 and 33-64 in the N-terminal domain of EcClsA contain stretches of hydrophobic residues flanked C-terminally by positively charged residues and are predicted to each form a transmembrane helix. The second hydrophobic region contains a motif of unknown function and is conserved among ClsA of some well-known bacteria and MhCls [184] (Figure 11d). Protein mass fingerprinting and Edman degradation data suggests that EcClsA undergoes a post-translational modification at its N-terminus resulting in the loss of a region from residue 1 up to residue 32 that harbors the first hydrophobic region (unpublished data and [171,199]).

Previous data indicated that ClsA from *E. coli* accepts substrates other than glycerol for the reverse "alcoholysis" reaction of CL, such as simple primary alcohols and D-mannitol [183,200]. The ClsA-type homolog from the methanogenic archaeon *M. hungatei* (MhCls) has been characterized and shown to exhibit a high degree of promiscuity. In contrast to the EcClsA, the enzyme does not seem to undergo the same posttranslational modification and could be purified as a full-length protein. It was found to accept various alcohols other than glycerol in the reverse reaction. This resulted in the synthesis of natural- and various non-natural phospholipids *in vitro.* Furthermore, MhCls was found to be active on bacterial G3P-based ester phospholipids in addition to the archaeal G1P-based ether phospholipids [184]. Remarkably, the enzyme is capable of generating a hybrid phospholipid species, not before seen in nature, where PG is coupled to AG. Notably, both EcClsA and MhCls are unable to utilize trehalose as an alcohol acceptor (unpublished results) unlike the *Salmonella typhi* ClsB that is responsible for the formation of phosphatidyl trehalose and di-phosphatidyl trehalose cardiolipin [192].

No PLD-type Cls crystal structures have been published to date. However, crystal structures of other PLD superfamily members from eukaryal- and bacterial sources have been reported [196,201,202]. The EcClsA structure has been modeled based on a bacterial PLD family member structure using the Phyre2 one-to-one threading function, resulting in a saddle- or bilobed structure typical for PLD structures [171,201–203] (Figure 12a). The modeled EcClsA structure lacks a pronounced hydrophobic surface, and thus the mechanism of membrane association remains unclear if in the native enzyme both predicted N-terminal transmembrane helices are indeed cleaved off [198]. The conserved histidine and lysine from both HKD motifs are located in close proximity of one another in a pocket at the interface of the PLD domain lobes (Figure 12b, 12c). The catalytic pocket is further lined with the glycine (G238, G418), serine (S239), threonine (T419), asparagine (N241, N421) residues of the extended HKD motif and seems to be larger than the binding pocket of the template PLD (*Streptomyces sp.* PMF, PMFPLD, PDB: 1F0I, [202]). Although, most of the difference in the size of the catalytic pocket seems to be caused by the "external" conformation of a HKD motif histidine (H448) in the PMFPLD structure compared to the EcClsA model where both HKD motif histidines are in the "internal" conformation and thus can be attributed to a modeling- or crystallization artefact [202,204] (Figure 12b). The PMFPLD crystal structure shows three loops (Loop A, B and C) near the catalytic pocket which could be functionally significant, potentially inhibiting the binding of larger phosphatidyl acceptor substrates compared to EcClsA (Figure 12c). Loop A, B and C in PMFPLD are formed from three sequences inserted between predicted structurally conserved elements in a region between structural region II and III exclusively in the C-terminal PLD domain (D336-R350,

A373-L389 and F423-Y437 in PMFPLD corresponding to Y343-D349, R370-L376 and F398-G401 in EcClsA resp). Particularly loop B in PMFPLD (A373-L389) is situated near the C-terminal histidine residue which activates the acceptor alcohol. Loop B might provide a structural basis for PMFPLDs preference for phosphodiester hydrolysis instead of transphosphatidylation by reducing space in- or access to the active site. However, as models only have limited accuracy; the orientation of residues might be different in an actual crystal structure of EcClsA and enzymatic conformation could be different in an aquatic environment or in presence of a phospholipid bilayer. Moreover, PLDs are thought to undergo some conformational shifting around the active site upon the binding of substrates or inhibitors and thus no strong conclusions can be drawn from this modeled result [201,204,205].

Even though the (extended) HKD motif is conserved among PLD family members, they show different substrate and/or product specificities and different ratios of hydrolysis to transphosphatidylation [206–208]. The latter of which can also be interpreted as a difference in substrate specificity of water versus an acceptor alcohol. These observations indicate that other residues are likely responsible for substrate specificity. However, despite several solved crystal structures of PLD members, the structural basis for their substrate specificity is not known.

Several studies have attempted to clarify the structural basis of PLD-member substrate specificity. It was shown that a substitution of serine to threonine, or glycine to serine, in the "GG/GS" motif (considered to be part of the extended HKD motif) in the C-terminal structural region 4 of PLD, strongly reduced overall activity or resulted in a significant increase of transphosphatidylase activity respectively [197,205]. Notably, Cls often have a threonine instead of a serine in that position as well. Mutational analysis of various PLDs in a region near the active pocket [209] or in loop C revealed residues that are reported to impact phospholipid affinity, enzyme stability, substrate specificity and transphosphatidylase- or hydrolase activity [209–211] (Figure 12d). Importantly, many of these mutated residues are not conserved among PLDs. And even though the mutations of residues in loop C significantly alters the characteristics of the enzyme; the exact role of the mutated residues and loop C itself remains to be determined. Overall, while several residues affecting substrate specificity were identified in PLDs, no study has provided overarching conclusive answers.

The catalytic mechanism of PLD enzymes is conserved among the entire superfamily. All reactions involve the transesterification or hydrolysis of a glycerophosphodiester lipid with either an organic molecule containing a hydroxyl group or water, respectively. The first step of the reaction involves the formation

of a phosphatidyl-enzyme intermediate through a nucleophilic attack on the phosphatidyl phosphorus atom by the histidine from the N-terminal HKD motif which is activated by a nearby conserved aspartic or glutamic acid from the C-terminal PLD domain (H224/E432 in EcClsA or H167/D465 in PMFPLD numbering, respectively (Figure 12d)) [202,209,212]. In the second step, the histidine from the C-terminal HKD motif which is activated by an aspartic acid residue from the N-terminal PLD domain protonates the headgroup of the substate, forming the alcohol leaving group. In the third and final step, the histidine of the C-terminal HKD-motif then activates a primary alcohol or water molecule for a nucleophilic attack on the phosphatidyl-histidine intermediate, forming the free phospholipid with the new alcohol headgroup or PA respectively.

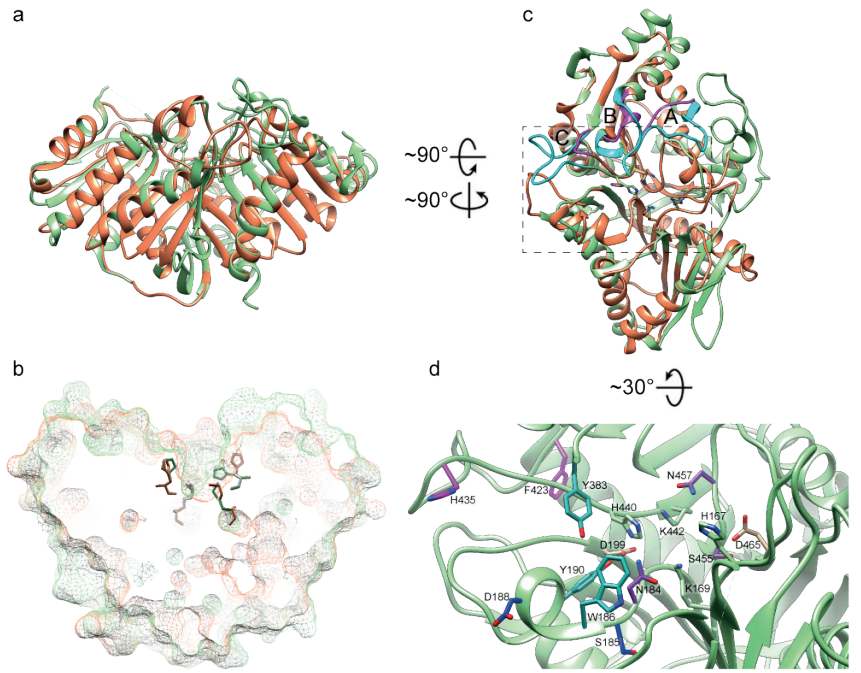


Figure 12: Ribbon view (a, c, d) and mesh-surface view (d) of the PMFPLD crystal structure (PDB: 1F0I, [202]). (a, c) The EcClsA structure (orange) is modelled onto the PMFPLD crystal structure (light green) where loops A B and C are colored magenta and cyan for the EcClsA model and PMFPLD structure respectively. **(b)** Shows a mesh surface section of the EcClsA model and PMFPLD structure with the histidine and lysine residues of the HKD motifs shown as sticks. **(d)** A tilted and zoomed-in view of panel c (visible area marked) in which histidines and lysines of the HKD motif and residues discussed in the text are shown as sticks and are colored according to cited studies in which they, or corresponding residues in other PLDs are mutated (blue [209], magenta [210], dark-cyan [211]). Or purple for residues

of interest other than glycine in the extended HKD motif [197,205], or tan for two conserved aspartic acid residues involved in catalysis. In accordance with the catalytic mechanism discussed above, residues around the C-terminal HKD motif histidine have been shown to affect acceptor alcohol substrate specificity. Therefore, it is likely that residues around the N-terminal HKD motif histidine could influence phosphatidyl donor substrate specificity. Ultimately, enzymes of the PLD family could be more efficiently designed to reduce hydrolysis and accept larger or different alcohol substrates when cardiolipin synthases are used as design template to learn more about substrate selectivity mechanisms in PLD family members. Alternatively, Cls could be engineered to accept new donor or acceptor substrates in order to generate synthetic phospholipids with customized properties. A cardiolipin synthase crystal structure would further aid this research and possibly further clarify alcohol acceptor substrate recognition.

Lipid core modification

Isoprene saturation

In most archaea, the isoprene chains of mature membrane glycerophospholipids are fully saturated. Saturation is assumed to occur after polar headgroup activation due to the specificity of CarS for unsaturated archaetidic acid (DGGGP) [102]. Prenyl reductases are flavoenzymes responsible for the saturation of isoprenyl groups. Specifically, in the context of archaeal phospholipid biosynthesis, geranylgeranyl reductase (GGR) fulfils this function. The search for prenyl reductases in archaea led to the identification of a membrane-associated GGR from *T. acidophilum* (TaGGR) which is active on DGGGP and partially unsaturated archaeols (di-phytyl) with a polar head group such as AE and AG [213–215]. The degree of saturation of bacterial fatty acyl chains is determined during fatty acid synthesis by FabA (anaerobic) or fatty acids are aerobically desaturated by desaturases [216]. For example in *E. coli*, trans-unsaturated bonds are anaerobically introduced and subsequently cis-isomerized during fatty acid synthesis in intermediates by FabA and directed back into the main fatty acid synthesis cycle by FabB [217–220]. As part of the fatty acid synthesis cycle, FabZ introduces unsaturated bonds as well; however, these are subsequently saturated by FabI [218]. The saturation of fatty acid biosynthesis intermediates by FabI and introduction of unsaturated bonds by FabA (and FabZ) or aerobic desaturases occur through entirely different mechanisms compared to archaeal isoprenyl synthesis and isoprenyl saturation. Hence, bacterial fatty acid saturation is not further discussed in this work.

The only two crystal structures of archaeal GGRs are those of *Sulfolobus acidocaldarius* GGR (SaGGR) and TaGGR [215,221,222]. SaGGR is a monomer composed of two functional domains resembling TaGGR: a FAD-binding Rossman-type fold domain and a ligand binding domain (Figure 13a). The substrate binding cavity is situated between the two established functional domains and contains a catalytic- (Pocket A) as well as a non-catalytic (Pocket B) hydrophobic binding pocket to accommodate substrates containing two lipid tails (Figure 13b and 13c). Interestingly, despite the high similarity between the SaGGR and TaGGR sequences and structures, SaGGR is purified from the cytosolic fraction whereas TaGGR is associated with the cell membrane fraction. Compared to TaGGR, SaGGR contains 60 extra amino acids at the C-terminus forming three α-helices that are considered part of the ligand-binding domain. A DALI search revealed similarities with members of the *p*-hydroxy-benzoate hydroxylase (PHBH) flavin monooxygenase family which adopt a similar two-domain organization and structure, but lacking the 60 C-terminal amino acids found in SaGGR [215,221,223]. Notably, the PHBH family is known to have members with three-domain structures that are approximately 50-70 amino acids longer than the two-domain enzyme sequences (See [221] and references therein). However, the three-domain PHBH sequences were not reported to correspond more closely to SaGGR or its C-terminal sequence.

The secondary structures of the FAD-binding and ligand binding domains are highly conserved between PHBH and GGR with some noteworthy exceptions: Helix αK at the domain interface leading into the C-terminus of the ligand binding domain is kinked, tilting the C-terminal portion of the ligand binding domain away from the substrate binding pocket opening and attributing to the creation of a larger internal cavity compared to PHBH enzymes [215] (Figure 13a). This larger cavity accommodates the lipid tail which is bound in the non-catalytic pocket (Pocket B) of a diether phospholipid substrate such as DGGGP.

The two hydrophobic binding pockets for the accommodation of substrate lipid tails are separated by Y215 (SaGGR numbering) and a loop from residues 290-301 with Y215 forming a hydrogen bond with the backbone carbonyl of H297 [222] (Figure 13b and 13c). While co-crystallized PG from the expression host in the TaGGR structure did not form many specific interactions, the phosphate headgroup moiety of PG in the SaGGR crystal structure formed salt-bridges with residues H55, H297 and hydrogen bonds with residues Y340 and N294. This is remarkable as the PG from the expression host has a G3P based backbone whereas archaeal phospholipids as found in *S. acidocaldarius* have a G1P based backbone. This indicates that SaGGR might be glycerol backbone agnostic, despite having fairly deep substrate binding pockets as the different chiral center of G3P vs G1P only

affects the positioning of the secondary lipid tail in the non-catalytic position. However, this would depend on whether the bulkier isoprenyl-ether substrate can also fit in pocket B without hindering the positioning of the primary isoprenyl lipid tail in the catalytic pocket when bound to a G3P-based backbone.

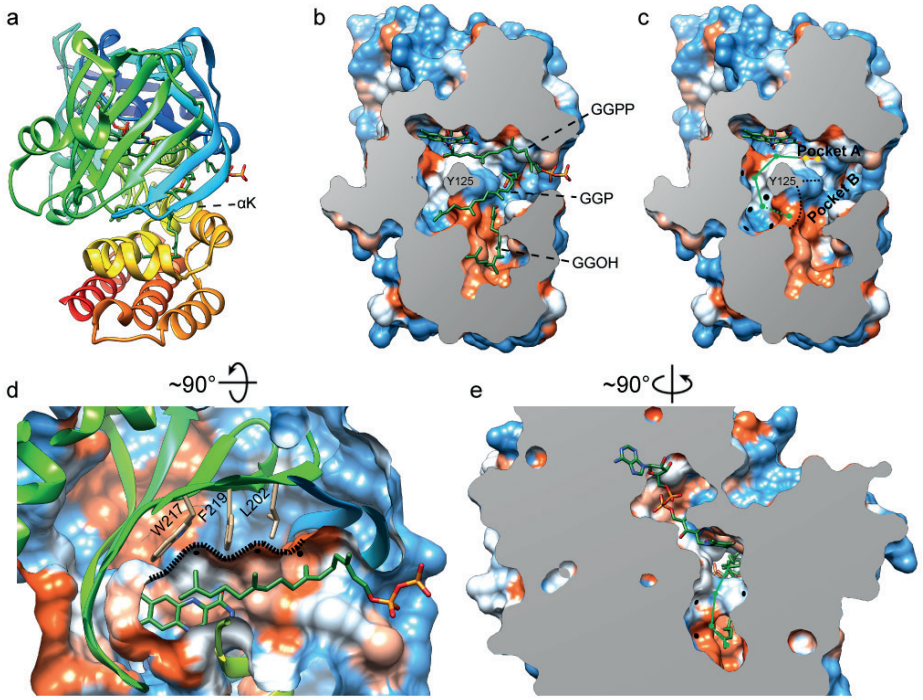


Figure 13: Ribbon view (a, d) and solid-surface view (b, c, d, e) of the SaGGR crystal structure with bound FAD, GGPP and its derivatives (PDB: 4OPD, [222]). The surface views are colored according to the Kyte-Doolittle hydrophobicity scale where red surfaces are hydrophobic, white is of mixed character and blue is hydrophilic. Possible positions of tri-substituted carbon atoms and the distance between them for each isoprenyl moiety in GGPP are schematically illustrated in a 2D-plane as circles and lines respectively with colors representing different elements (**c**, **e**). Pockets which could potentially harbor (**c**, **e**), or have been shown to be able to harbor isoprene methyl moieties are marked with black dots (**d**).

When co-crystallized with GGPP, three moieties corresponding to GGPP with different phosphorylation states were found in both hydrophobic binding pockets of SaGGR [222] (Figure 13b). Only one GGPP moiety was bound in a catalytic manner (1, GGPP) in pocket A, while the other two moieties were found in pocket B (2, GGP and 3, GG-OH). The GGPP phosphate moiety was found to bind in a different manner compared to the bacterial PG in TaGGR or SaGGR and only forming a hydrogen bond with the β-phosphate to the backbone carbonyl of N90 [215,222]. The terminal double bond

of GGPP (Δ^{14}) is in a suitable position for reduction, directly facing the N-5 nitrogen of FAD (Figure 13b and 13d). Notably, the methyl group directly adjacent to this double bond is situated between two conserved residues (W217 and F219) which likely act as a biological resistive detent mechanism or stabilization feature to position the double bond properly for catalysis in concert with Y215 [222] (Figure 13d). Moreover, particularly pocket A seems to contain more sub-pockets that occur with a spacing that might be suitable to accommodate isoprene-methyl moieties, suggesting that these pockets could harbor the lipid methyl groups for when the isoprenoid substrate is bound differently for the reduction of double bonds other than the terminal bond (Figure 13c, 13d and 13e). While the GGP moiety is bound in a non-catalytic manner, its phosphate headgroup engages in similar interactions to PG and forms salt-bridges with H55, H297 and additionally K343, and hydrogen bonds with Y340 and N294; indicating that these indeed play a role in binding anionic moieties such as the phosphate moiety in glycerophospholipids such as GGPP, GGGP, DGGGP, unsaturated CDP-archaeol or headgroup modified derivatives thereof [222,224].

Sato et al [225,226] showed that purified GGR from *S. acidocaldarius* (SaGGR) was more active than a GGR isolated from *A. fulgidus* (AfGGR). Interestingly, it was also observed that SaGGR exhibited a different reduction pattern depending on the substrate, fully reducing diether-committed geranylgeranyl moieties such as GGGP and only partially reducing GGPP, leaving the proximal double bond intact and forming Phytyl-PP [226]. In recent years, two studies have shown that GGRs have a surprisingly broad substrate specificity, as in addition to GGPP, GGGP and DGGGP, AfGGR and SaGGR are also active on terpene acids, free terpene alcohols and terpene alcohol derivatives with headgroups other than pyrophosphate, such as hemiester, ester and aromatic headgroups [224,227]. In contrast to earlier observations, SaGGR was shown to be able to fully reduce a portion of both GGPP, geranylgeraniol and farnesol substrates [224,226,227]. It should be stressed that full reduction of GGPP, forms phytanyl-PP, which is no longer suitable as a prenyl donor and therefore a dead-end product. Saturation of the proximal double bond prevents stabilization of the carbocation intermediate, as seen in GGPP during catalysis by GGGPS, IPPS and likely DGGGPS [54,70]. The activity of GGR on terpene acids is remarkable, as this indicates that the carbonyl oxygen of bacterial acyl chains should not prevent substrate recognition by GGR. However, unsaturated acyl chains typically lack the methyl groups of isoprene chains which could potentially prevent efficient catalysis as these methyl groups, as in GGPP, are thought to play a role in substrate stabilization in the active site [222]. Both AfGGR and SaGGR were found to saturate synthetic isoprenoids with a succinate ester headgroup closer to completion compared to free isoprenols or isoprenoids with an acetate ester or benzyl ether headgroup; suggesting that the distance between charged (or polar) headgroup atoms to the proximal double

bond is important for reduction of (or the lack thereof) the proximal bond [224]. By extension indicating that the difference in product specificity between AfGGR and SaGGR is caused by a longer distance from the active site to the phosphate binding site in SaGGR.

The native biological reducing agent for some GGRs remains elusive. The activity of identified GGRs *in vitro* is dependent on the abiotic reducing agent sodium dithionite; although activity of TaGGR has been observed in the presence of NADH and a low level of activity of SaGGR was observed with extremely high concentrations of NADH [213,221]. SaGGR and MaGGR are only poorly active in the *E. coli* cells used to express the protein, indicating these GGRs are incompatible with the *E. coli* electron transport system components and therefore the non-covalently bound FAD cofactor of GGR cannot be effectively reduced [228]. This can concern specific co-factors or specific ferredoxins that can couple particular enzymatic reactions to the electron transport system. A ferredoxin-like protein encoded downstream to the GGR gene in the *M. acetivorans* genome was able to reduce MaGGR, leading to GGR activity in *E. coli* cells [228]. A SyntTax synteny conservation search revealed that the GGR-ferredoxin synteny is largely conserved among selected Euryarchaeotes with the exception of members of the Halobacteriales and notably, *A. fulgidus* and *M. mazei* [229]. However, the GGR-ferredoxin synteny found in *M. acetivorans* is not conserved in any Crenarchaeote in the SyntTax database (unpublished results).

Glycerol dialkyl glycerol tetraether (GDGT) biosynthesis

Many archaea and some bacteria synthesize macrocylic glycerol dialkyl glycerol tetraether lipids (GDGTs) which resemble two tail-to-tail dimerized dialkyl glycerol diether lipids (DGDs) [230–234] (Figure 14). In some archaea, such as *S. acidocaldarius,* GDGTs are the main constituent of the cell membrane forming a predominantly monolayer membrane [235–238]. In Crenarchaeota and Euryarchaeota, GDGTs can contain up to eight cyclopentane rings in the dialkyl moieties of the caldarchaeol lipid core; while in Thaumarchaeota up to four cyclopentane rings can be found, with an extra cyclohexane ring to form the crenarchaeol lipid core unique to this clade [110,239,240]. GDGTs can be found modified in various other ways, such the additional methyl groups, bridging the alkyl moieties, or the use of a butane- or pentane-triol backbone instead of glycerol, resulting in a diverse spectrum of GDGT derivatives [106,108,109]. The proportion of GDGTs in the membrane is subject to homeoviscous adaptation and tends to increase at higher cultivation temperatures [241,242]. Additionally, the degree of cyclization in tetraethers also increases at higher temperatures and tetraether containing fossil remains are used in historical temperature proxies [243–246]. However, the relation of the degree of GDGT pentacyclization and the pH is still unclear [231,243,247]. It has been suggested that the introduction of pentacyclic rings allows for more dense lipid packing, which aids the growth of hyperthermophilic organisms by stabilizing the membrane structure at high temperatures and likely decreases proton permeability [248–250]. In contrast, the hexane ring in crenarchaeol is thought to serve the opposite function, destabilizing lipid packing to form a more fluid membrane, for instance to allow for growth of Thaumarchaeota at lower temperatures [240].

Recently, two GDGT (cyclopentane) ring synthases (GrsA and GrsB) and a tetraether (GDGT) synthase (Tes) have been identified in *S. acidocaldarius* and two methanogens *(Methanococcus aeolicus* and *M. acetivorans),* respectively [251,252]. These enzymes are part of the radical SAM superfamily of enzymes which use a [4Fe-4S]$^{+}$ cluster to cleave S-adenosylmethionine to typically generate a 5′-deoxyadenosyl radical. This radical can be used to perform a large variety of reactions involving relatively unreactive substrates, including the formation of carbon-carbon bonds. This function finds use in enzymes that are involved in various critical cellular processes, such as co-factor synthesis, DNA methylation, enzyme activation and other protein posttranslational modifications, but also in lipid-headgroup and lipid-tail modification [251–254] (Figure 14).

The exact biosynthetic route of GDGTs has remained elusive since the discovery of their structure more than four decades ago [257]. Several studies have touched upon this topic with varying hypotheses and mixed results. Pulse-labelling experiments by Langworthy revealed that *Thermoplasma* cells initially incorporate [^{14}C]MVA in the archaeol core lipid; with the signal in caldarchaeol core lipid increasing after the signal in archaeol reached its maximum intensity [258]. This suggested that archaeol lipids (DGD) are a direct precursor to the caldarchaeol lipids (GDGT). Later experiments by Nemoto et al. [259] also used pulse-chase labelling techniques and verified previous observations by Langworthy. Furthermore, a study showed that terbinafine, a squalene-epoxidase inhibitor, also inhibits GDGT synthesis resulting in DGD accumulation in *T. acidophilum* and that particular GDGT-exclusive headgroups are attached only after synthesis of GDGT from DGD [259–262]. A feeding experiment with [^{14}C] radiolabeled saturated archaeol or caldarchaeol core lipids for incorporation in the membranes of *M. hungatei* did not show interconversion of radioactivity from DGD the fraction into the GDGT fraction of the lipid extracts [255]. Moreover, the same study showed that double bonds in geranylgeraniol are important for the incorporation of a [^{3}H] radiolabel in both DGD and GDGT lipids, providing a possible explanation as to why the radioactivity of the radiolabeled saturated archaeol and caldarchaeol core lipids is not converted from DGDs to GDGTs [255] (Figure 14). Another feeding study on *M. hungatei* with [^{32}P]PO_4^{3-} revealed that radiolabeled phosphate is readily incorporated in DGD phospholipids and that the incorporation

into GDGT phospholipids started later and occurred more slowly [263]. This suggests that GDGT phospholipids have a larger precursor pool than DGD phospholipids supporting the notion that DGDs are potentially precursors of GDGTs in *M. hungatei*.

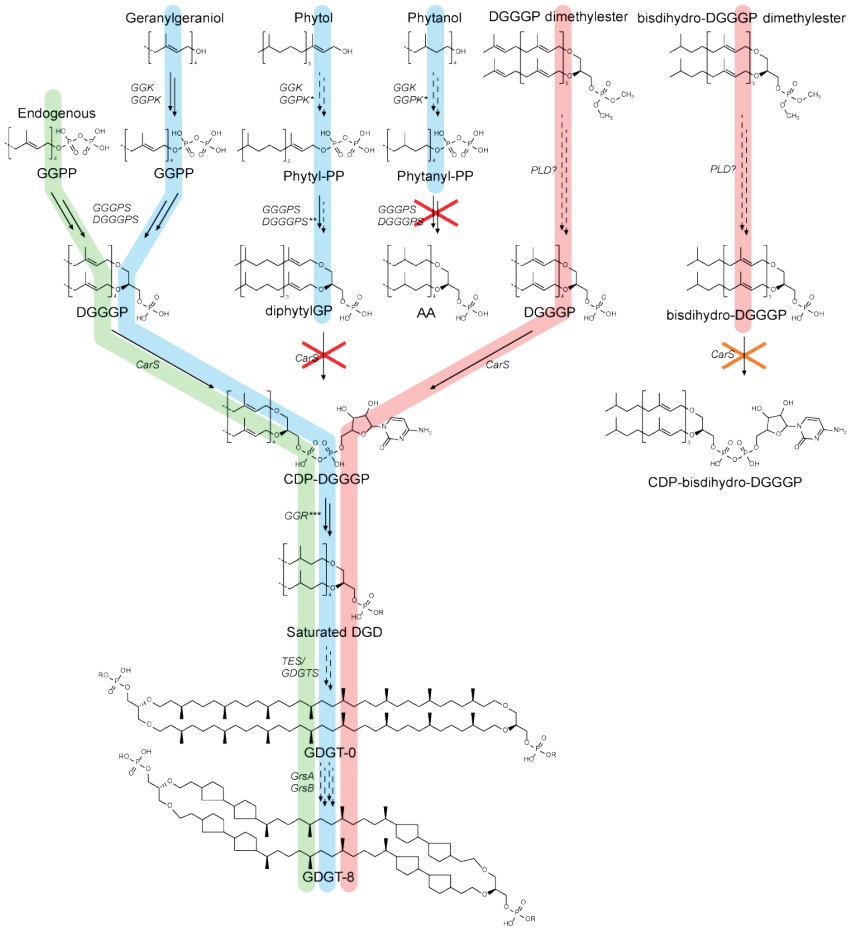


Figure 14: Proposed biosynthetic pathway with GGGPS, DGGGPS and CarS acting as (isotopically labeled) feeding substrate gatekeepers. The pathway for endogenous GGPP is colored green, feeding substrates from Poulter, et al. in blue [255] and feeding substrates from Eguchi, et al. in red [256]. Dashed arrows indicate uncharacterized reactions. Crosses indicate reactions that are not possible (red) or thought impossible on the basis of enzyme characterization experiments or the feeding study result (orange). * GGK and GGPK have not been shown to be active on (partially) saturated isoprenoids. ** While GGGPS can utilize phytyl-PP as a prenyl donor, this has not been demonstrated for DGGGPS. *** Due to the substrate promiscuity of GGR (towards headgroups) and headgroup diversification enzymes (towards the lipid backbone and radyl tails), isoprenyl saturation and headgroup modification could occur in parallel.

In a recent study, a tetraether (GDGT) synthase (Tes) was identified [252]. GDGT synthase is a radical SAM enzyme, strongly reinforcing the notion that GDGTs are formed from the dimerization of diether lipids. However, the question whether the terminal double bonds of the diether substrate are required for catalysis remains unanswered. In 2000, Eguchi et al. [264,265] presented dissimilar ^{1}H NMR spectra between multiple-deuterated DGD lipids obtained by lipid extraction and purification from D_9-mevalonolactone fed cultures of *Haloarcula japonica, M. thermoautotrophicus* and *M. jannaschii*. The dissimilar deuterium labeling was attributed to an unusual double-bond migration and linked with the presence or absence of GDGTs in the lipidome. It was suggested that this unusual double-bond migration leads to the formation of a terminal methylene group; enabling the synthesis of GDGTs from unsaturated diether lipids with a reductive coupling by an oxidoreductase at the step of double bond saturation [265]. However, this hypothesis was later retracted [256]. Moreover, our structural analysis of the GGR active site strongly suggested that GGR is not able to perform the previously proposed reductive coupling as the product would be unable to leave the active site of the enzyme the moment the macrocycle would be formed. Furthermore, it was shown that the terminal carbon atom retained both deuterium atoms upon formation of GDGT in *M. jannaschii* [265]. This observation is highly similar to a study on diabolic acid synthesis in *Butyrivibrio fibrisolvens* fed with C-16 $^{2}H_3$ palmitic acid where the resulting diabolic acid was shown to retain all 6 deuterium atoms [266]. This excludes the formation of higher oxidation states of the terminal carbon atom during carbon-carbon bond formation in both cases and has been proposed to involve a radical mechanism. This radical reaction mechanism may be similar to that of GDGT formation, indicating that the terminal double bonds in GDGT are not required for catalysis.

A later study by Eguchi et al. [256] involved the feeding of cultures of *M. thermoautotrophicus* with DGGGP di-methyl esters with various saturation patterns. Fully unsaturated DGGGP di-methyl ester and DGGGP di-methyl ester with the terminal bonds saturated were found to be incorporated into GDGTs and diethers and only diethers respectively (Figure 14). These data suggest that DGGGP or an equivalent molecule is the substrate for the GDGT synthase and that the terminal double bond is essential for GDGT synthase activity on DGGGPS or its equivalent in *M. thermoautotrophicus*. Studies with membrane fractions of *M. thermoautotrophicus* suggest that the CarS prefers unsaturated substrates [102], which may imply that lipids are typically saturated *after* headgroup activation. Thus, the DGGGP di-methyl ester with the saturated terminal double bonds might not be accepted by CarS and therefore not further processed into GDGTs (Figure 14). Counterintuitively, the substrate specificity of CarS could prevent the incorporation

of this molecule into GDGTs if the GDGT synthase is specific for saturated substrates or substrates with a particular headgroup. However, the proposed specificity of CarS for unsaturated substrates [102] needs to be verified with the purified enzyme in assays with a better kinetic resolution using both saturated and unsaturated substrates.

Possibly, the mechanism to form GDGTs is indeed similar to the formation of diabolic acid which has also been proposed to involve a radical mechanism [266]. This mechanism is advantageous as the terminal isoprenyl moiety might not need to remain unsaturated for GDGT synthesis. Considering GGRs substrate promiscuity, this would be much simpler to regulate. This would mean that in *M. thermoautotrophicus*, most GDGT precursors (GGPP, GGGP and DGGGP) would remain unsaturated until the headgroup of DGGGP has been activated by CarS, forming CMP-DGGGP (CDP-archaeol). GGR could then reduce the double bonds of various diether lipids formed after headgroup activation by CarS [213], which then would allow GDGT formation.

Reports on the directionality of the progressive nature of double bond saturation are conflicting [222,227]. However, progressive saturation starting at the proximal double bonds towards to terminal double bonds would be an effective mechanism to aid in preventing the formation of unsaturated GDGTs, regardless of whether the GDGT synthase requires saturated or unsaturated terminal isoprenyl moieties. If the GDGT synthase requires the terminal isoprenyl moiety to be saturated, and recognizes a headgroup, this mechanism would avoid the potential GDGT saturation problem and explain the observations of the feeding studies published so far. If the GDGT synthase requires unsaturated terminal bonds, it has to recognize the state of the other bonds in addition to the terminal double bond to prevent the formation of unsaturated GDGTs. Alternatively, the formation of unsaturated GDGTs would require the existence of an unidentified reductase specific for tetraethers as GDGTs are likely not suitable as substrates for GGR. Overall, it seems more likely that the GDGT synthase is specific towards saturated double bonds and employs a radical reaction mechanism similar to that of diabolic acid formation [266].

Intuitively, the tetraether synthase can be expected to be a membrane protein as the substrate is likely embedded in the membrane. However, the primary amino acid sequence of the identified GDGT synthases are not predicted to contain any transmembrane helices or domains (TMHMM V2, [267]) [252]. Also, the recently identified GDGT cyclopentane ring synthases are not integral membrane proteins [251], whereas they are thought to catalyze a reaction in the lipid tails of GDGTs which are buried in the membrane. Interestingly, other enzymes in the lipid biosynthesis pathway are also not integrally associated with the membrane despite the nature of their substrates. For instance, GGR enzymes are not integral membrane proteins either, but can still be found associated with the membrane [213,225–227]. This presents a practical issue, as the enzyme needs to reach its substrate (or *vice versa)*. A possible solution can be found in the mechanism of Psd and several other enzymes; these enzymes bind to the polar surface of the membrane and provide a diffusion pathway to 'extract' the lipid from the membrane bilayer into the active site [268,269]. However, it should be noted that the enzymes described in this fashion typically catalyze reactions at the acyl-ester bond (proximal terminus) or at the lipid headgroup, so far none have been described to catalyze reactions at the distal terminus of the radyl group.

Most studies suggest a mechanism involving the tail-to-tail dimerization of diethers to form tetraethers (Figure 14), which is further strengthened by the recent discovery of the tetraether synthase [252]. This is also substantiated by the observation that in some organisms tetraether lipids are found where only one of the carbon-carbon bonds appears to have been formed (referred to as glycerol trialkyl glycerol tetraethers (GTGTs)) [239,270–272] which is considered an intermediate in the tail-to-tail condensation process [252,273] or terminal products caused by a quirk in the pathway where GDGT synthases have altered substrate specificity; or possibly where GDGT synthases are specific for unsaturated terminal prenyl moieties and GGR saturates one of the terminal prenyl units before the GDGT synthase performs the second coupling. Villanueva, et al. [274] proposed an alternative lipid biosynthetic pathway in which IPPS are able to accept IPP intermediates containing pentacyclic rings and suggesting that the C_{20} isoprenoid-pyrophosphates dimerize to form C_{40} di-pyrophosphate isoprenoids with and without rings. These precursors would then be ether bonded, at both ends, to G1P by GGGP and DGGGP forming GDGTs. However, the discovery of the GDGT synthase nullifies the hypothetical role for IPPS in C_{20} isoprenoid dimerization [252]. Moreover, this would require an unprecedented level of functional plasticity of the phospholipid biosynthetic enzymes, or alternatively, involve sofar undiscovered enzymes. Notably, GGPPS, GGGPS and DGGGPS must be able to accept $C_{>20}$-isoprenoid substrates that can also contain rings. In this respect, the hydrophobic substrate binding grooves or tunnels of AfGGGPS and MjDGGGPS do not appear deep enough to accept C_{40}-isoprenoids and are unlikely to be wide enough to accommodate pentacyclic rings [52,82] (Figures 4b and 6b) despite the fact that *A. fulgidus* and *M. jannaschii* are both GDGT-synthesizing organisms [241,275]. Lastly, the proposed pathway suggests that isoprenyl chain saturation occurs after addition of the second glycerol moiety. As described above the GGR substrate binding cavity is split in two pockets [215,221] (Figure 13c). For an isoprenyl

moiety to bind in the active site, the isoprenyl tails of a typical diether isoprenoid lipid substrate are wedged apart. Due to the macrocyclic structure of GDGTs, the isoprenyl chains can no longer be wedged apart, and it is difficult to envisage how an unsaturated GDGT isoprenyl moiety would reach the active site.

Biphytane diols containing ring moieties have been detected in environmental samples and glycerol dialkanol diethers in both environmental samples and archaeal cultures of methanogens [276–278]. Biphytane diols could be produced from GGPP if the GDGT synthase and GrsA/GrsB accepts Phytanyl-PP and biphytanyl-PP, respectively; as GGRs can accept GGPP (See GGR section). The proposition of glycerol dialkanols as precursors of GDGTs would be mechanistically quite challenging as both glycerol dialkanol tails would need to attach to the same glycerol moiety, and in the correct position, to form the GDGT macrocycle. Reported glycerol dialkanol diethers are fully saturated and lack a phosphate headgroup [276,278]. As described above, known prenyltransferases require the prenyl donor to have a pyrophosphate "head-" group and the proximal isoprenyl moiety to be unsaturated. Thus, the hydroxyl groups characteristic of glycerol dialkanols would need to be phosphorylated by enzymes related to geranylgeraniol kinase (GGK) and GGP kinase (GGPK) that would need to accept longer isoprenoid substrates [279]. Additionally, an unidentified enzyme would be required to create the macrocycle by the attachment of the second glycerol moiety. Moreover, lipids are synthesized with a phosphate headgroup originating from G1P, which plays an important role in substrate binding by both GGGPS and DGGGPS. This phosphate headgroup is either retained for phospholipid synthesis, or is likely eliminated after the diether phospholipid core has formed for- or during the synthesis of other lipid types such as glycolipids [253,280]. As such, a more feasible explanation would be that these molecules are breakdown products of GDGTs.

The intermediates of the proposed alternative biosynthetic pathway would be expected to be widely present among GDGT synthesizing archaea. So far, reports detecting these suggested intermediates in lipidomes are scarce and evidence for the unprecedented enzyme functional plasticity is lacking. Thus, tail-to-tail dimerization of diether lipids by GDGT synthase remains the most likely mechanism for tetraether lipid formation.

Concluding remarks

The glycerophospholipid core biosynthesis pathway is fairly well characterized in both Archaea and Bacteria, but a particularly interesting group of organisms remains. In recent years, the new "Asgard" superphylum of archaea has been described. Members of this superphylum include the Thor-, Odin-, Heimdall- and Lokiarchaea, and are of particular interest as these archaea are the most closely related members of the Archaea to Eukaryotes [281–283]. The first Lokiarchaeal metagenome was found to not contain a complete archaeal phospholipid biosynthesis pathway, lacking G1P dehydrogenase (G1PDH) [3,284]. Surprisingly, Lokiarchaea were found to possess homologs to PlsC and only PlsY of the PlsXY pathway which are both typically only found in bacteria, leading to the proposition that Lokiarchaea might synthesize archaeal phospholipids with a G3P backbone or chimeric phospholipids containing both archaeal and bacterial radyl groups. Later metagenome sequencing studies resulted in more Lokiarchaeal genomes which revealed the presence of a G1PDH, suggesting that Lokiarchaea are capable of synthesizing complete archetypical archaeal phospholipids [285]. Indeed, isoprenoid based lipids have been found in the only Lokiarchaeote that could be cultivated so far [286]. However, the function of the PlsY and PlsC homologs remains to be elucidated. For this reason, Lokiarchaea remain an interesting group of organisms to study for the structural basis for archaeal versus bacterial phospholipid biosynthesis and the lipid-divide.

The synthesis of the glycerophospholipid cores in Archaea and Bacteria occurs along similar lines while employing different molecular mechanisms (Figure 1). For example, the utilization of pyrophosphate groups for the formation of the ether bonds in Archaea compared to the transesterification of acyl-ACP to glycerol in Bacteria. The reactions involving CarS and CAPT enzymes for headgroup activation and diversification are saliently similar between Archaea and Bacteria (Figure 1 and 8). However, despite the common presence of CAPT enzymes in both pathways, only relatively little structural and biochemical information has been reported. For instance, Ags is thought to attach G3P as a headgroup, but this enzyme has not been characterized [116]. In bacteria, G3P is utilized both for the phospholipid backbone and the phosphatidylglycerol polar headgroup. Archaea utilize G1P for the phospholipid backbones; but as not all archaea seem to possess typical G3P synthesis-related enzymes, it remains to be determined whether the lipid-divide extends to the headgroup of AG and which chiral form of glycerol phosphate is used by Ags [3].

The substrate specificities of CarS and GGR are two key points that have insufficiently been addressed. While a particular group of GGRs seems to use ferredoxin as a biological reducing agent, the biological reducing agent of many GGRs is still not known. *In vitro* studies using sodium dithionite as reducing agent suggest that GGR is a rather promiscuous enzyme as it accepts free isoprenols and synthetic isoprenoid esters, GGPP, GGGP, DGGGP and perhaps even CDP-lipids and mature phospholipids with diverse polar headgroups. Experiments with crude lysates of *M. thermoautotrophicus* suggest a specificity of MtCarS for unsaturated phospholipids which would suggest that phospholipid saturation by GGR occurs after headgroup activation by CarS. However, this needs to be further examined with the purified enzyme in kinetic assays. Does the *in vitro* activity of GGR demonstrate the true substrate specificity of GGR? Or is the biosynthetic pathway more organized in space where GGR acts only on CDP-archaeol and downstream products? On CarS substrate specificity, GGGPS, and likely DGGGPS, are specific towards substrates with the G1P stereochemistry. On the basis of the CarS crystal structure, CarS is expected to be specific for G1P based substrates as well. However, this has not been tested *in vitro* with a purified enzyme; hence it remains unclear whether CarS contributes to the lipid-divide on the basis of the glycerol backbone stereochemistry or radyl ester vs ether bonding or the nature of the radyl groups themselves. Thus, thorough biochemical characterization of GGR and CarS, and protein-protein interactions of lipid biosynthesis enzymes, are of prime interest to further our knowledge on lipid biosynthesis in Archaea.

Recently, two GDGT ring synthases (GrsA and GrsB [251]) and the tetraether (GDGT) synthase were identified [252]. However, the enzymes have not been biochemically characterized. For instance, are GrsA and GrsB only active on tetraethers and not on diethers or substrates with a single radyl group? While the GDGT synthase is expected to tail-to-tail couple diether lipids, its substrate specificity has not been studied either. A key question remains as to whether the GDGT synthase is active on saturated- or unsaturated terminal isoprenyl groups. Moreover, it remains to be determined how substrate recognition is accomplished for all of these enzymes; as GrsA, GrsB and the GDGT synthase are not predicted to be integral membrane proteins while their substrates with their sites of catalysis are thought to be buried in the membrane. The fact that these enzymes are radical SAM enzymes complicates their biochemical characterization. However, this is key towards a better understanding of the biochemical basis of GDGT formation and the formation of related lipids such as cyclic diethers and other tetraethers such as H-bridged lipids, crenarchaeol, bacterial bridged-GDGTs and potentially other lipids with unusual carbon structures.

Funding information

The work was supported by the Building Blocks of Life program (737.016.006), and the 'BaSyC – Building a Synthetic Cell' Gravitation grant (024.003.019), both programs are subsidized by NWO. Neither program influenced the development or submission of this work.

Author contributions

Literature research: Niels A.W. de Kok, Arnold J.M. Driessen; Writing - original draft preparation, revisioning and editing: Niels A.W. de Kok; Visualization: Niels A.W. de Kok; Writing – critical reviewing: Arnold J.M. Driessen; Funding acquisition: Arnold J.M. Driessen; Supervision: Arnold J.M. Driessen

Competing interests

The authors have no competing interests to declare.

List of abbreviations

AA	Archaetidic acid
aCL	Archaeal CL
ACP	Acyl carrier protein
AE	Archaetidylethanolamine
AG	Archaetidylglycerol
AGP	Archaetidylglycerol phosphate
Agp	Archaetidylglycerol phosphatase
Ags	Archaetidylglycerol synthase
AI	Archaetidylinositol
AIP	Archaetidylinositol phosphate
Aip	Archaetidylinositol phosphatase
Ais	Archaetidylinositol synthase
AMPD	Anhydromevalonate phosphate decarboxylase
ArtA	Archaeosortase A
AS	Archaetidylserine
Asd	Archaetidylserine decarboxylase
ASR	Ancestral structure reconstruction
Ass	Archaetidylserine synthase
bCL	Bacterial CL
BLAST	Basic Local Alignment Search Tool
BMD	Bisphosphomevalonate decarboxylase
CAPT	CDP-alcohol phosphatidyl transferase
CarS	CDP-archaeol synthase
CDP	Cytidine diphosphate
CdsA	CDP-DAG synthase
CL	Cardiolipin
Cls	Cardiolipin synthase
CMP	Cytidine monophosphate
CoA	Coenzyme A
COG	Cluster of orthologous genes
CoQ2	4-hydroxybenzoate polyprenyltransferase
CoX10	Protoheme IX farnesyltransferase
CTP	Cytidine triphosphate
DAG	Diacylglycerol
DGD	Dialkyl glycerol diether
DGGGP	Di-geranylgeranylglycerol phosphate
DGGGPS	Di-geranylgeranylglycerol phosphate synthase
DMAPP	Dimethylallyl pyrophosphate
DMD	Diphosphomevalonate kinase
EDTA	Ethylenediaminetetraacetic acid
FabA	3-hydroxydecanoyl-[acyl-carrier-protein] dehydratase
FabB	3-oxoacyl-[acyl-carrier-protein] synthase
FabI	Enoyl-[acyl-carrier-protein] reductase
FabZ	3-hydroxyacyl-[acyl-carrier-protein] dehydratase
FAD	Flavin adenine dinucleotide
FARM	First aspartate-rich motif
FAS	Fatty acid synthase
FPP	Farnesyl pyrophosphate
FPPS	Farnesyl pyrophosphate synthase
G1P	Glycerol-1-phosphate
G1PDH	G1P dehydrogenase
G3P	Glycerol-3-phosphate
GDGT	Glycerol dialkyl glycerol tetraether
GDGTS	GDGT synthase
GFPP	Geranylfarnesyl pyrophosphate
GFPPS	Geranylfarnesyl pyrophosphate synthase
GGGP	Geranylgeranylglycerol phosphate
GGGPS	Geranylgeranylglycerol phosphate synthase
GGK	Geranylgeraniol kinase
GG-OH	Geranylgeraniol
GGPK	Geranylgeranyl phosphate kinase
GGPP	Geranylgeranyl pyrophosphate
GGPPS	Geranylgeranyl pyrophosphate synthase
GGR	Geranylgeranyl reductase
GPP	Geranyl pyrophosphate
Gro-DACL	Glycerol-di-archaetidyl-CL
Gro-DPCL	Glycerol-di-phosphatidyl-CL
Grs	GDGT ring synthase
HepPPS	Heptaprenyl pyrophosphate synthase
HexPPS	Hexaprenyl pyrophosphate
I1P	(1L-*myo*-) Inositol-1-phosphate
IDI	Isopentenyl pyrophosphate:dimethylallyl pyrophosphate isomerase
IP	Isopentenyl phosphate
IPP	Isopentenyl pyrophosphate
IPPS	Isoprenyl pyrophosphate synthase
LBG	Lipid binding groove
LC-MS	Liquid chromatography – mass spectrometry
LDAO	Lauryldimethylamine oxide
LPA	Lyso-phosphatidic acid
LUCA	Last universal common ancestor
M3K	Mevalonate-3 kinase
M3P5K	Mevalonate-3-phosphate-5 kinase
M5K	Mevalonate-5-kinase
MD	Molecular dynamics
MenA	1,4-dihydroxy-2-naphthoate octaprenyltransferase
MVA	Mevalonate
MVA-3,5-PP	Mevalonate-3,5-biphosphate
MVA-3-P	Mevalonate-3-phosphate
MVA-5-P	Mevalonate-5-phosphate
MVA-5-PP	Mevalonate-5-pyrophosphate
MVK	Mevalonate kinase
NADPH	Nicotinamide adenine dinucleotide phosphate
PcrB	Heptaprenylglycerol phosphate synthase
PDB	Protein data bank
PE	Phosphatidylethanolamine

PG	Phosphatidylglycerol
PGP	Phosphatidylglycerol phosphate
Pgp	Phosphatidylglycerol phosphatase
Pgs	Phosphatidylglycerol synthase
PHBH	p-hydroxy-benzoate hydroxylase
PI	Phosphatidylinositol
PIP	Phosphatidylinositol phosphate
Pip	Phosphatidylinositol phosphatase
Pis	Phosphatidylinositol synthase
PLD	Phospholipase D
PlsB	Glycerol-3-phosphate acyltransferase
PlsC	1-acyl-*sn*-glycerol-3-phosphate acyltransferase
PlsX	Phosphate acyl transferase
PlsY	Glycerol-3-phosphate acyltransferase
PMD	Phosphomevalonate decarboxylase
PMDh	Phosphomevalonate dehydratase
PMFPLD	Streptomyces sp. PMF PLD
PMK	Phosphomevalonate kinase
PPi	Inorganic pyrophosphate
PS	Phosphatidylserine
Psd	Phosphatidylserine decarboxylase
Pss	Phosphatidylserine synthase
SARM	Second aspartate-rich-motif
SCD	Schnyder corneal dystrophy
tAMVA-5-P	trans-anhydromevalonate-5-phosphate
TES	Tetraether synthase
TIM-barrel	Triose phosphate isomerase barrel
TM	Transmembrane
UbiA	4-hydroxybenzoate octaprenyltransferase

Chapter 2

A versatile method to separate complex lipid mixtures using 1-butanol as eluent in a reverse-phase UHPLC-ESI-MS system

Chemistry and Physics of Lipids (2021)
volume 240, 105125

https://doi.org/10.1016/j.chemphyslip.2021.105125

Niels A.W. de Kok [a], Marten Exterkate [a], Ruben L.H. Andringa [b], Adriaan J. Minnaard [b], Arnold J.M. Driessen [a]*

[a] Department of Molecular Microbiology, Groningen Biomolecular Sciences and Biotechnology Institute and Zernike Institute for Advanced Materials, University of Groningen, Nijenborgh 7, 9747 AG, Groningen, The Netherlands

[b] Department of Chemical Biology, Stratingh Institute for Chemistry, University of Groningen, Nijenborgh 7, 9747 AG Groningen, The Netherlands

* Corresponding author: *a.j.m.driessen@rug.nl*; Tel. + 31-50-3632164; Fax. +31-50-3632154.

Abstract

Simple, robust and versatile LC-MS based methods add to the rapid assessment of the lipidome of biological cells. Here we present a versatile RP-UHPLC-MS method using 1-butanol as the eluent, specifically designed to separate different highly hydrophobic lipids. This method is capable of separating different lipid classes of glycerophospholipid standards, in addition to phospholipids of the same class with a different acyl chain composition. The versatility of this method was demonstrated through analysis of lipid extracts of the bacterium *Escherichia coli* and the archaeon *Sulfolobus acidocaldarius.* In contrast to 2-PrOH-based methods, the 1-butanol-based mobile phase is capable of eluting highly hydrophobic analytes such as cardiolipins, tetraether lipids and mycolic acids during the gradient instead of the isocratic purge phase, resulting in an enhanced separation of cardiolipins and extending the analytical range for RPLC.

Introduction

Cellular membranes fulfil the essential function of compartmentalization, separating cellular contents from the extracellular environment and enable the formation and maintenance of specialized reaction compartments. Membranes are predominantly formed by glycerophospholipids that due to their amphipathic characteristics can self-assemble into bilayer structures. Glycerophospholipids that constitute the membrane of Archaea are markedly different from those found in Bacteria and Eukarya. One of the defining features of archaeal glycerophospholipids is the presence of ether-bound isoprenoid-based lipid tails instead of ester-bound acyl lipid tails found in bacteria and eukaryotes. This feature makes archaeal lipids more hydrophobic compared to their fatty-acid based counterparts. Various archaea also synthesize more complex glycerophospholipids such as cardiolipins (CLs) or membrane spanning tetraether lipids that resemble two tail-to-tail dimerized diether lipids forming a macrocyclic structure [169,181,231]. These lipids in particular are strongly hydrophobic and therefore can complicate chromatographic analysis using traditional reversed-phase LC methods. The study of archaeal lipid membranes and the mechanisms of biosynthesis conveys information about what ancient cell membranes might have looked like and provide the basis for the Tex86 historical temperature proxy that is extensively used to predict surface temperatures in the Cenozoic era [246].

Due to high sample complexity, contemporary lipidomics often requires multi-dimensional analysis [287]. This is typically facilitated by methods such as LC-MS, a powerful tool allowing for detailed sample analysis [288]. Shotgun lipidomics by direct infusion MS has been documented as well and has the advantage of short analysis times facilitating high-throughput studies. Although, the lack of analyte separation can cause ion suppression, interfering with the identification of low abundance lipids [289]. Furthermore, the lack of analyte separation can interfere with absolute quantification and requires the use of a mass filter prior to MS/MS analysis for structural characterization [104,236,290,291]. The low resolution of such mass filters (~0.4 Da) leaves analysis methods that have no analyte separation, and therefore rely on "MS-only" information, susceptible to interference of isobaric lipids. Thus these methods are generally less suitable for detailed structural characterization of lipids in highly complex samples containing isobaric or isomeric lipid species [289,292,293].

Methods of lipid separation are usually based on normal-phase (NP) or reversed-phase (RP) LC [105]. Other methods employ 2-dimensional separation for more comprehensive analysis. These methods combine various separation

techniques and select these techniques on the basis of desired separation selectivity, compatibility and throughput [294–297]. NPLC is often used to analyze environmental samples and generally uses relatively apolar mobile phase solvent mixtures consisting of 2-propanol (2-PrOH), hexane, chloroform and small amounts of water. This is used in combination with bare silica or diol-columns to separate lipids into different lipid classes, mostly based on polar head group hydrophilicity [298–302]. The separation of lipids into classes is particularly useful for biomarker studies, hence NPLC methods are more commonly used in environmental biology. An NPLC-based technique, hydrophilic interaction liquid chromatography (HILIC), is commonly employed for lipidomic analyses [303–307]. HILIC uses a hydrophilic stationary phase, like NPLC; with solvents that are less hydrophobic, often employing an acetonitrile (MeCN)-water gradient, and thus are more comparable to RPLC mobile phase compositions. As with NPLC, HILIC methods tend to separate lipids according to their class. However, lipids often elute in a relatively narrow time window, resulting in a single chromatographic peak per lipid class in which lipid species can only be distinguished by mass. This property greatly benefits quantification of lipid species, as a single internal lipid standard per class is sufficient to correct for a difference in response factor for that class [308].

In contrast to NPLC-based techniques, RPLC uses hydrophobic stationary phases, most commonly C18-modified silica combined with more polar solvent mixtures such as water, MeCN and 2-PrOH. This difference in solvents and column material shifts the selectivity of RPLC methods to separate lipids mostly based on the hydrophobic properties of the lipid tail. Separation based on polar headgroup characteristics still takes place, but to a much lower degree compared to NPLC-based techniques [105,287,309–314]. While most RPLC methods are suitable for the separation of fatty acids, lyso-glycerophospholipids and diester glycerophospholipids; the separation of highly hydrophobic lipids, such as cardiolipins or archaeal tetraether lipids can be challenging for water/MeCN- and water/2-PrOH-based methods. These highly hydrophobic lipids are difficult to elute from C18 columns, which often results in co-elution, peak broadening and in more extreme cases even leading to an absence of signal. These issues limit analytical coverage of RPLC and can be particularly problematic for the analysis of complex biological samples. As these samples contain many different lipids, which can have very different concentrations to one another; co-elution can obscure low abundance lipids, or interfere with reliable quantification due to ion suppression effects. Another common problem in the study of lipids through LC-MS is the co-elution of isomeric lipids [287,292,311,315,316]. For instance, cis-trans isomeric lipids cannot be distinguished by common mass spectrometry techniques and thus isomer analysis largely relies on chromatographic separation techniques [292,317,318]. However, while not commonplace at this time, ion mobility spectrometry (IMS) techniques are gaining in popularity for lipidomic analyses. IMS offers an extra orthogonal source of information in addition to chromatographic separation and MS [319,320]. This technique is able to distinguish and separate isomeric lipids based on their conformational shape, despite often having identical retention times, providing more information and confidence for lipid identification and lipidomic characterization [321–324].

Classic normal-phase or HILIC methods do not always have the desired resolution and selectivity that RPLC methods have to offer for a given set of analytes; for example, lipidomic analyses requiring in-depth structural characterization, for which a large variety of acyl tails or isoprene chains of varying length or saturation needs to be detected and separated. In addition, most LC systems are configured to run under RPLC conditions and switching to an NPLC configuration can involve significant downtime. Despite these drawbacks, normal-phase chromatographic separation excels in the separation of similar ether-lipid cores and has the advantage of allowing stronger injection solvents [230,231,239,243,325]. To increase the eluotropic range of RPLC with the aim to separate more hydrophobic lipids, we increased the eluotropic strength of the secondary mobile phase compared to established methods by using 1-butanol (1-BuOH) as a solvent. Here we present a versatile RP-UHPLC-ESI-MS method capable of separating complex biological lipid mixtures, including especially hydrophobic molecules such as the isoprene-based cardiolipin molecules and tetraether lipid species that are often found in archaeal lipid membranes.

Materials and methods

Growth conditions and cell harvesting

Escherichia coli DH5α was cultured under aerobic conditions at 37°C with 200 RPM shaking. A pre-heated 4 L Erlenmeyer flask containing 1 L of LB medium was inoculated to OD_{600} 0.01 and incubated for 16 h (OD_{600} 2.5). *Sulfolobus acidocaldarius* MW001 was cultured under aerobic conditions at 76°C with 200 RPM shaking in Brock medium [326,327]. A pre-heated 1 L Erlenmeyer flask containing 500 mL of Brock medium pH 3 supplemented with 10 µg/mL uracil was inoculated to OD_{600} 0.03 and incubated to OD_{600} 1.5 over 70 h. Cells were harvested by centrifugation at 7,500 x g for 15 min at 4°C and washed with 15 mL buffer (50 mM Tris-HCl pH 8, 100 mM KCl). Washed cells were pelleted (4,000 x g, 15 min at 4°C) and flash frozen in liquid N_2. Frozen cell pellets were placed in a pre-cooled freeze dryer set to -55°C at 0.01 mbar pressure and left to dry for 3 days.

Total lipid extraction and sample preparation

Lipids were extracted from 10 mg freeze-dried cells from *E. coli* DH5α or *S. acidocaldarius* MW001 pellets using an adapted acidic Bligh and Dyer method employing 0.1 M HCl instead of 2 M HCl or 5% trichloroacetic acid as described elsewhere [328,329]. The crude chloroform fraction was dried and the lipid film was three times re-extracted using 400 µL chloroform. The resulting fraction was dried, three times re-extracted with 400 µL chloroform - methanol (1:2), dried again and finally three times re-extracted with 400 µL methanol. The resulting lipid film was dissolved to 5 mg/mL in methanol and filtered through a 0.2 µm PTFE syringe filter prior to analysis.

All pure lipid stocks and the *E. coli* cardiolipin extract (Avanti polar lipids, cat: 841199) were dissolved in chloroform. To prepare these samples for analysis, chloroform containing a measured amount of lipids (or a mix thereof) was evaporated under a N_2-stream at 37°C to form a lipid film. This film was then re-dissolved in methanol (to 0.25 mg/mL for pure lipids or 0.1 mg/mL for lipid standard mixtures) for UHPLC-ESI-MS analysis.

Reverse-phase UHPLC method development

For the separation of lipids, Waters C18 charged surface hybrid (CSH) and bridged ethylene hybrid (BEH) (150 x 2.1 mm internal diameter) columns with a 1.7 µm particle size were selected [105,312]. Typically, RPLC methods for lipid analysis employ mixtures of MeCN, methanol and ultrapure water (MQ) as eluent A. 2-PrOH typically forms the basis of eluent B. To increase the eluotropic strength of eluent B, in-house distilled 1-BuOH (Sigma Aldrich, cat#: 901351, supplemental methods) was selected as solvent for its stronger interaction with lipids compared to 2-PrOH. Ammonium formate (5 mM) was employed as buffering agent to promote negative ion formation; as lipids such as phosphatidic acid (PA) and phosphatidylethanolamine (PE) tend to ionize more easily in negative mode. The addition of aqueous ammonia to the mobile phase of a similar method has been reported to improve peak shape [105]. However, aqueous ammonia was not included to avoid retention time increases, lipid-class specific ionization efficiency alterations and reduced column lifetime. A mixture of MeCN and MQ was selected for the composition of eluent A. To provide sufficient retention for less hydrophobic analytes, such as lyso-phospholipids, up to 40% of MQ was added. This was sufficient to allow for separation of dodecyl maltoside (DDM) and lyso-phosphatidic acid (LPA) and lyso-phosphatidyl glycerol (LPG) without causing miscibility issues when combined with eluent B. In turn, eluent B was modified to contain 0.5% MQ to allow for buffer solubilization and 10% MeCN to prevent miscibility problems with eluent A. Ultimately, the analytes in a standard mixture (DDM; 18:1/18:1 PG, PC, PE, PA; 18:1/18:1/18:1/18:1 CL and AG) responded virtually identical to both the CSH and BEH columns. The intended purpose of this method was to separate many different lipids in complex samples, such as total lipid extracts. As these samples contain a large variety of analytes, the C18 CSH column was selected to further develop the method to reduce the potential risk of peak shape issues for compounds not included in our standard samples. At 60°C the peak width of the standard compounds eluting from the C18 CSH column was found to be between 1.5 and 2 times smaller compared to the peak width at 40°C. In favor of column lifetime, we decided to use 55°C for a better balance between manufacturer recommended operating parameters and lower backpressure, shorter run time and sharper peaks.

UHPLC-MS analysis

UHPLC-MS analyses were performed using an Accela1250 UHPLC system (Thermo Fisher Scientific) coupled to a Thermo Exactive Orbitrap mass spectrometer (Thermo Fisher Scientific) equipped with an ESI ion-source in negative ionization mode. The spray voltage was set at 3.0 kV whereas the capillary, tube lens and skimmer voltage were set to -75 V, -190 V and -46 V respectively. The capillary temperature was set at 300°C. Sheath- and aux-gas flow rate was set at 60 and 5 units respectively. Generally, a scan range of *m/z* 125-2500 was used in full-MS mode and *m/z* 77.5-1550 for MS/MS mode. The Thermo Exactive Orbitrap mass spectrometer model used in this study lacks a quadrupole mass filter and therefore was not capable of parent ion mass filtering before ion fragmentation. As a result, all ions were fragmented together in MS/MS mode.

Compounds were separated using an Acquity UPLC CSH C18 column (2.1 x 150 mm, 1.7 µm; Waters, cat#: 186005298) coupled to an Acquity UPLC CSH VanGuard pre-column (2.1 x 5 mm, 1.7 µm; Waters, cat#: 186005303). The column assembly was maintained at 55°C with an eluent flow rate of 300 µL/min to allow for shorter analysis times. The injection volume was 5 µL. Eluent A consisted of MQ:MeCN (40:60) containing 5 mM of ammonium formate, and eluent B consisted of MQ:MeCN:1-BuOH (0.5:10:90) also containing 5 mM ammonium formate. A linear gradient elution was applied as follows: 5% eluent B for 2.5 min; a gradient from 5% to 90% eluent B over 36.5 min; holding for 3 min; returning to 5% eluent B in 0.5 min; and holding for 8 min. The narrow gradient window method is the same with a 45% eluent B baseline and that the gradient runs from 45% to 90% eluent B over 19.5 min maintaining the same gradient slope (2.31% B / min).

Lipid headgroups, the total amount of carbon atoms in the ether- or acyl tails carbon atoms and number of double bonds therein were identified by full-MS and MS/MS using an in-house calculated *in-silico* lipid database and comparison to lipid standards of the same class where possible. Analyte retention time trends were also used as an additional indicator for identification. Lipid acyl tail isomers were identified using full-MS/MS in a similar fashion as described elsewhere [311,330,331].

Results

Separation of a Lipid standard

To improve upon the ability of current 2-PrOH-based gradient RPLC methods to separate highly hydrophobic lipids, the eluotropic strength of the eluents was increased by using 1-BuOH instead of 2-PrOH, and the elution profile was adjusted compared to previously described methods [105,312]. Thus, lipids were loaded on the column with eluent A consisting of MQ:MeCN (40:60) and 5 mM of ammonium formate whereas eluent B consisted of MQ:MeCN:1-BuOH (0.5:10:90) and 5 mM ammonium formate. After sample injection, a linear elution gradient was applied (See Materials and Methods). Standard solutions of lipids belonging to several different classes were injected on a CSH C18 UPLC column to establish retention times of commonly observed, well-characterized lipids and internal standards for lipid extractions. All lipid species (Table 1) were detected in the negative ionization mode predominantly as $[M-H]^-$ ions with the exception of phosphatidyl choline (PC 18:1/18:1). For this lipid, a formate adduct ($[M+CHO_2^-]^-$) was the most abundant ion, presumably caused by the positive charge of PC promoting formate ion attachment. The separation of these species is demonstrated in Figure 1.

Table 1: List of the compounds injected as lipid standards with peak numbering, corresponding retention time (tR) and observed mass. The peak numbers correspond to the chromatogram in Figure 1. AG, archaetidyl glycerol. *: Likely *sn2* and *sn1* lyso-lipid isomer separation respectively. **: The observed and theoretical mass of PC refers to the $[M+CHO_2^-]^-$ adduct as PC preferentially forms formate adducts.

Peak #	t_R	Compound	observed $[M-H]^-$	Theoretical $[M-H]^-$	ppm error
1	2.59	DDM	509.29707	509.29674	0.648
2	3.54/3.89*	LPG (18:1)	509.28871	509.28849	0.432
3	3.92/4.27*	LPA (18:1)	435.25208	435.25171	0.850
4	5.09	PG (10:0/10:0)	553.31512	553.31471	0.741
5	13.8	Oleic acid (18:1)	281.24881	281.24860	0.747
6	21.5	PG (18:1/18:1)	773.53443	773.53381	0.802
7	22.12	PA (18:1/18:1)	699.49795	699.49703	1.315
8	23.01	PG (18:0-18:1)	775.55019	775.54946	0.941
9	24.37	PC (18:1/18:1)	830.59298**	830.59166**	1.589
10	24.87	PE (18:1/18:1)	742.54022	742.53923	1.333
11	26.83	AG (20:0/20:0)	805.67018	805.66918	1.241
12	33.08	CL (4x18:1)	1456.02867	1456.02755	0.769

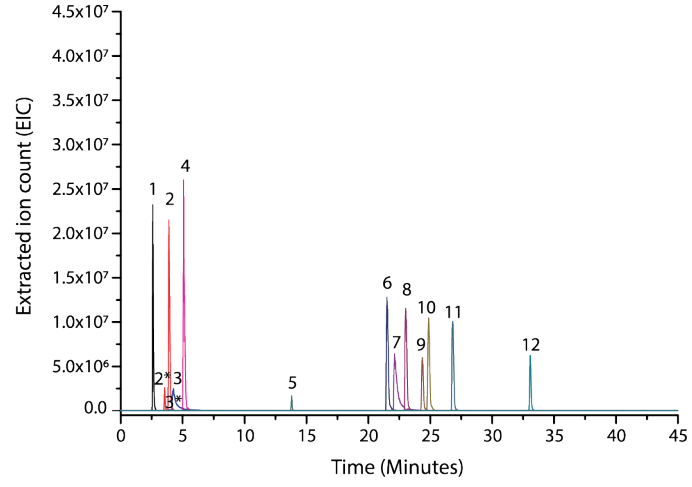


Figure 1: LC-MS extracted ion chromatograms showing separation of defined lipid standards containing a variety of lipid classes using the full gradient method. The peak numbers correspond to the compounds listed in table 1.

Additionally, mobile phase performance was compared to the same mobile phase with 2-PrOH instead of 1-BuOH and two 2-PrOH-based mobile phases described in literature [312,332] (Supplemental Figure 1 and Supplemental Tables 1-3).

With the employed elution gradient slope using the 1-BuOH-based mobile phase (See Materials and Methods), all lipid species separate well and are visible as separate peaks, including some different classes with the same lipid core (i.e., di-oleoyl-ester glycerol phospholipids). In general, lipids are mainly separated based on their hydrophobicity, meaning lipids with more carbon atoms on the acyl chains generally elute later, whereas the introduction of double bonds in the acyl chain causes analytes to elute earlier. These results are consistent with previous findings using 2-PrOH based methods [304,312]. In addition, the presence of ether bonds instead of ester bonds likely contributes to the later elution of archaeal diether lipids compared to their corresponding bacterial diester lipids. Both LPG and LPA show a smaller peak eluting just before the main peak, which can be attributed to the *sn1/sn2* acyl migration equilibrium of 90% *sn1* and 10% *sn2* isomer [333].

Noteworthy, previously used RP methods often suffered from peak tailing problems with lipids containing amine and ammonium groups such as phosphatidyl ethanolamine (PE) and PC, respectively. This peak tailing is suggested to be caused by interactions between the amine and ammonium groups on the lipid headgroup with free silanol groups of the column packing [334]. However, no significant PE or PC peak tailing was observed with either the CSH or BEH technologies of the columns employed in this work [312]. On the contrary, in the lipid standard where the PE peak shows slight tailing similar to most other analytes, the PC peak shows slight fronting. PA and phosphatidyl serine (PS) often exhibited excessive peak tailing when analyzed using RPLC methods in the past as well [335]. However, with the current method, only moderate peak tailing of PA (and LPA) was observed in the lipid standard (Figure 1).

Overall, from the elution profile of the standard lipids, it appears that the 1-BuOH-based method effectively separates a wide range of lipids, including the more hydrophobic cardiolipin and archaeal diether lipid species. The overall retention time is reduced compared to 2-PrOH-based methods, leaving chromatographic space for the analysis of more hydrophobic compounds (Supplemental Figure 1). However, this comes at the cost of decreased ionization efficiency of less hydrophobic lipids in the lipid standard mix using the same ionization conditions.

Separation and analysis of an *E. coli* cardiolipin extract

An *E. coli* cardiolipin extract was analyzed to validate the efficacy on a lipid mixture that is difficult to separate and strongly retains on C18 columns. For this purpose, a narrower gradient window was applied to optimize separation of more hydrophobic analytes (see Materials and Methods). The MS was set to negative mode while alternating between full-MS mode and CID MS/MS (80 eV) each scan. The lipid mixture eluted in 5 peak clusters (Figure 2A). However, a cardiolipin database mass search revealed the presence of a 6th cluster containing very small amounts of lipids not visible in the TIC chromatogram (Supplemental Table 4). In addition, the mobile phase performance was compared to the same mobile phase with 2-PrOH instead of 1-BuOH and two previously published 2-PrOH-based mobile phases [312,332] (Supplemental Figure 2 and Supplemental Tables 1-2).

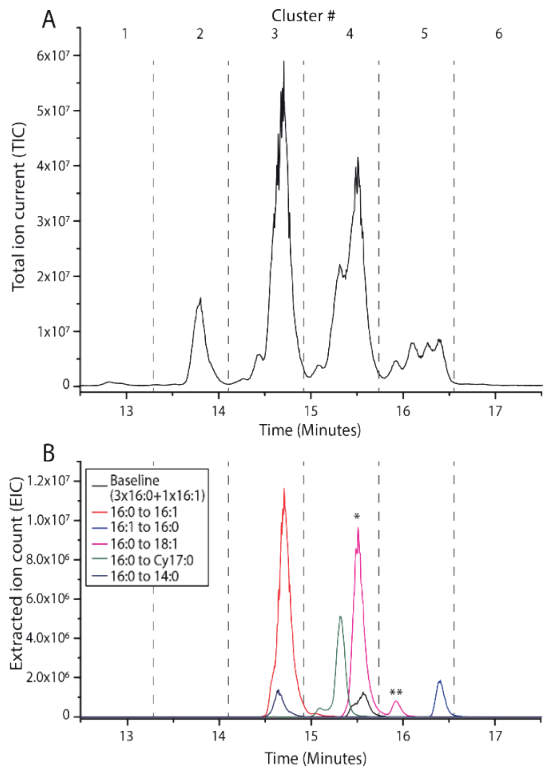


Figure 2: LC-MS chromatograms showing **(A)** the elution of various cardiolipin species in 6 peak clusters in a total ion chromatogram using the narrow gradient window method. **(B)** Extracted ion chromatograms of selected CL molecule masses that are closely related to CL 16:0/16:1/16:0/16:0, differing by a single acyl substitution as listed in the legend. Peak * represents CL 16:0/16:1/18:1/16:0 whereas peak ** represents the isomeric CL 16:0/Cy17:0/Cy17:0/16:0.

In order to investigate the effect of small differences in acyl tail composition, the mass of CL 16:0/16:1/16:0/16:0 and of closely related CLs (CL molecules with a single different acyl chain) were extracted from the total ion chromatogram and plotted in Figure 2B. This revealed that the cluster elution interval corresponds to the loss or gain of a double bond or a C_2H_4 moiety (Figure 2B and Supplemental Figure 3). The peaks in a cluster were incompletely separated based on other features such as the presence of cyclopropyl groups or, for example, double bond isomers. Noteworthy, the separation observed here is similar to that of positional isomers as reported previously for PC diester lipids [336]. However, no positional isomers were identified with sufficient confidence to confirm the capability to separate acyl positional isomers of CL. The separation data collected does allow for a calculation of approximate retention time shifts for a given modification of a base lipid. For example, CL 16:0/16:1/**16:1**/16:0 (t_R 14.71) elutes 0.86 min earlier, CL 16:0/16:1/**Cy17:0**/16:0 (t_R 15.32) elutes 0.25 min earlier and CL 16:0/16:1/**14:0**/16:0 (t_R 14.64) elutes 0.93 min earlier than CL 16:0/16:1/**16:0**/16:0 (t_R 15.57). Thus, based on these molecules, a retention time increase of approximately ~0.47 min per introduced CH_2 moiety in the fatty acyl chain is observed. Alternatively, the introduction of a double bond or cyclopropyl moiety resulted in a retention time decrease of 0.86 and 0.72 min respectively. This effect is amplified when the sample is analyzed using two in-line coupled CSH UPLC columns (result not shown). As shown in Figure 2B, some lipids elute in multiple peaks. Slightly different retention properties of different lipid modifications can cause isomeric lipid species to elute at a different time. For example, CL 16:0/Cy17:0/Cy17:0/16:0 elutes slightly later than its isomer CL 16:0/16:1/18:1/16:0 (Figure 2B and Supplemental Figure 5).

The MS/MS spectra yielded strong signals of acyl carboxylate ions allowing for easy identification of acyl chain moieties. As a result of ketene and fatty acid elimination, $[M\text{-}H\text{-}RCHCO_{sn1/2}]^-$ and $[M\text{-}H\text{-}RCH_2COOH_{sn1/2}\text{-}H_2O]^-$ ions were also detected. Due to preferred elimination of the *sn2* over the *sn1* ketene/fatty acid moiety, the MS/MS data was used to determine the acyl chain substituent position of several of the more abundant cardiolipin species as originally reported for *Lactobacillus acidophillus* and *E. coli* diester lipids [311,330,331] (Supplemental Table 4, Supplemental Figures 4 and 5). Using this method, our analysis showed that there is a preference for 16:0 acyl chains on the outer chain positions (*sn1*) while there is a preference for alternative chain substituents on the inner chain positions (*sn2*) of cardiolipin molecules present in this *E. coli* cardiolipin extract.

Overall, this analysis demonstrates that this method is capable of differentiating between similar and large lipids containing more than 50 carbon atoms in the lipid tails. In comparison with conventional 2-PrOH-based mobile phases, our method showed improved separation of these hydrophobic lipid species. In general, a larger retention time difference between peak clusters was obtained, which resulted in the appearance of some extra total ion current peaks. Moreover, using the 1-BuOH-based mobile phase resulted in a significantly better peak shape for lyso-cardiolipins.

Separation and analysis of an *E. coli* total lipid extract

To illustrate that the RP-UHPLC method can be used for conventional full-scale lipidomics, an *E. coli* total lipid extract was injected using the full-gradient method (5-90% B) in order to also adequately retain and separate less hydrophobic lipids. The *E. coli* lipidome was then analyzed for the presence of the major phospholipid class species. Masses corresponding to at least 33 distinct phosphatidyl glycerol (PG) and 40 distinct PE diester glycerophospholipids were identified in the extract (Supplemental Table 5). Diesters with the same core, but a different head group are well separated from one another as in the standard lipid mixture. Additionally, masses corresponding to 36 different cardiolipins are well separated from diester lipids (Figure 3).

The wide variety of lipid species separated in this complex sample shows that the 1-BuOH-based method is versatile. At the same time, it allows for very similar lipid species with only minor structural rearrangements to be separated as well. This is exemplified by the extracted ion chromatogram of one of the more abundant diester lipids, PG (33:1). This phospholipid was selected as it contains acyl chains with an even and uneven number of carbon atoms in length, where one of the acyl chains either contains a cyclopropyl moiety or a double bond. This theoretically allows for many combinations of acyl chains with slightly different properties and possibly resulting in multiple chromatographic peaks for the same mass. In a similar fashion to the cardiolipin extracted ion chromatograms (Figure 2B), the extracted ion chromatogram of PG (33:1) reveals multiple peaks, two main peaks with the major peak showing a shoulder (Supplemental figure 6). The retention time differences between the peaks suggests that these are isomers containing cyclopropyl moieties and/or double bonds. The MS/MS data indicates that the major peak represents PG (16:0/Cy17:0). However, due to the lack of a mass filter in the instrument used, the acyl tail identity of the other isomeric PG lipid species could not be confidently established using the lyso- and acyl fragments. Multiple other co-eluting PG lipid species rendered it impossible to confidently assign specific acyl tails to the *sn1*- or *sn2*-position of particular parent ions.

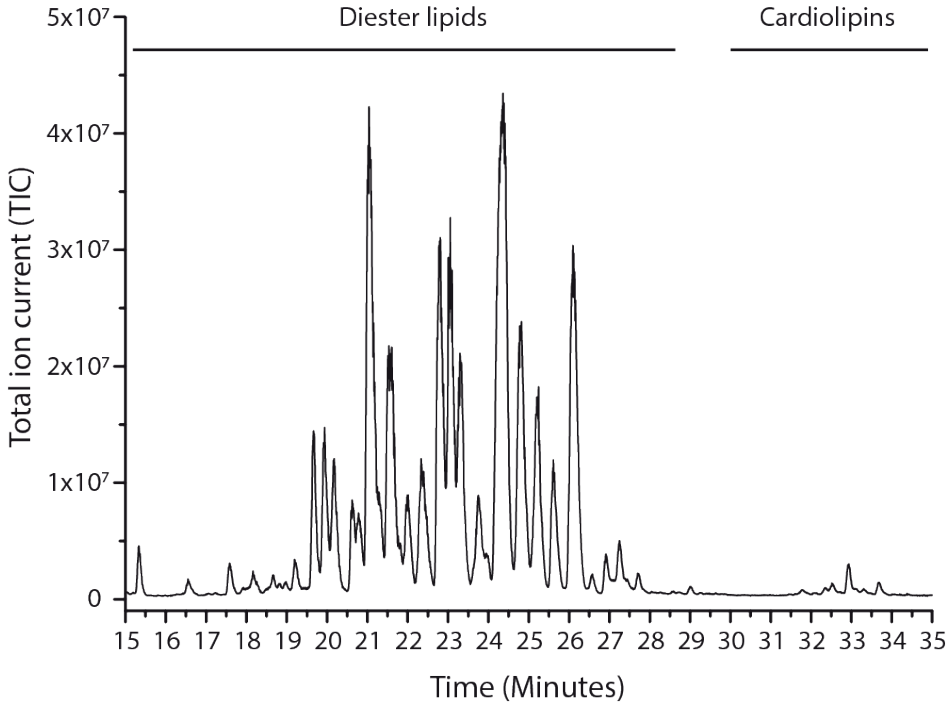


Figure 3: A total ion LC-MS chromatogram showing the elution of an *E. coli* DH5α total lipid extract using the full gradient method.

Separation and analysis of a *S. acidocaldarius* total lipid extract

Next, our RP-UHPLC method was employed to analyze archaeal lipid extracts containing tetraether lipids. A total lipid extract of the crenarchaeon *S. acidocaldarius* MW001 was injected on the column using the same full-gradient as used for the *E. coli* lipid extract. Four major elution peaks were found with masses corresponding to an inositol phosphate dialkyl glycerol diether lipid (IP-DGD, 1) and three main glycerol dialkyl glycerol tetraether (GDGT) peaks identified as inositol phosphate-di-hexose (IP-DH-GDGT, 3), inositol phosphate (IP-GDGT, 4), di-hexose (DH-GDGT, 5) and a smaller caldarchaeol (GDGT, 6) peak (Figure 4A). It appears that *S. acidocaldarius* tetraether lipids elute later (35-40 min, Figure 4A) with 76-88% eluent B, as compared to the cardiolipins from *E. coli* (30-35 min, Figure 3 and Supplemental Table 5) with 64-76% eluent B, illustrating the separation of more hydrophobic compounds.

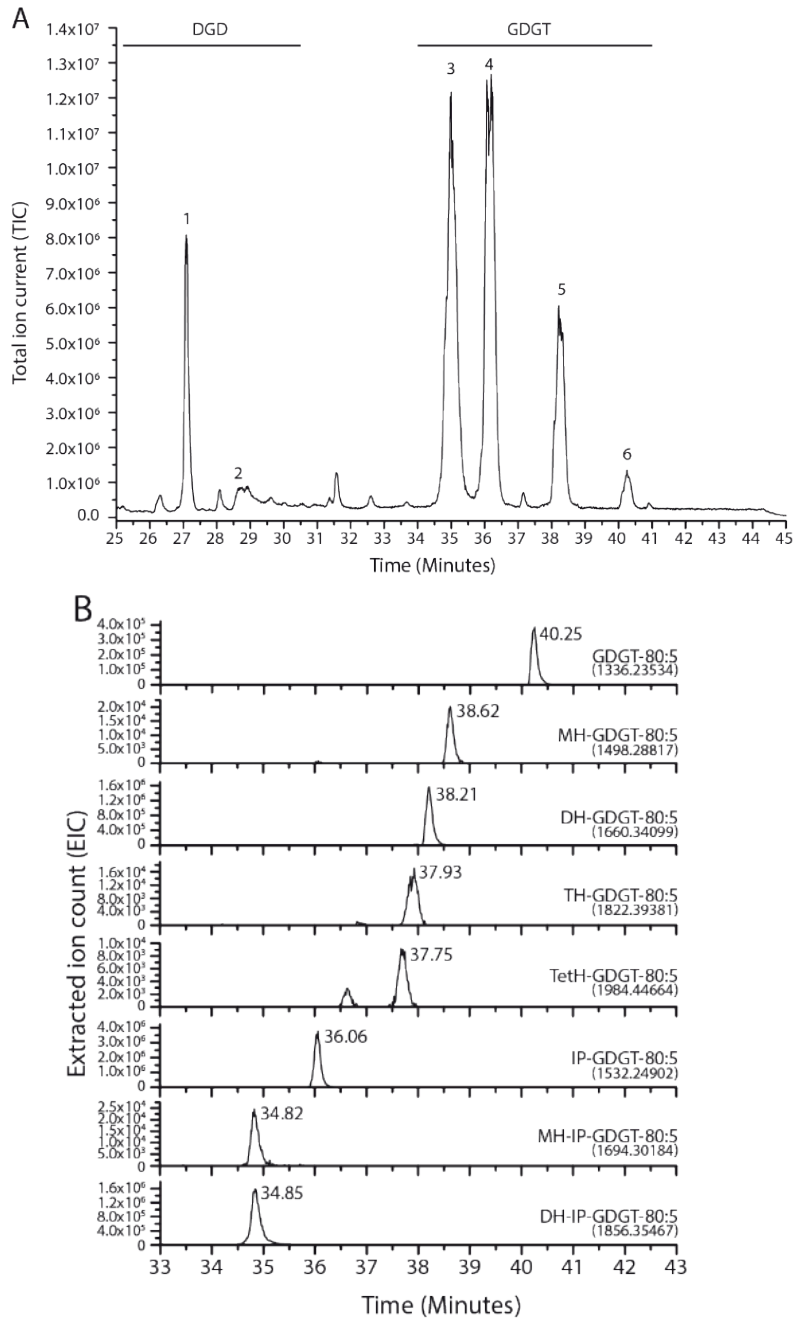


Figure 4: (A) Total ion LC-MS chromatogram showing the elution of a *S. acidocaldarius* MW001 total lipid extract using the full gradient method. **(B)** Extracted ion chromatograms of GDGT-80:5 species and their +1 isotopologue containing varying numbers of hexose units. GDGT-80:5 indicates a tetraether with 80 carbon atoms in the alkyl chains containing 5 cyclopentane rings. The theoretical mass for each observed formate adduct ion ($[M+CHO_2^-]^-$) is listed below the name.

The addition of a hexose to a tetraether lipid containing no hexoses usually results in a significant retention time decrease (e.g., GDGT to MH-GDGT and IP-GDGT to MH-IP-GDGT, Figure 4B). However, the retention time decrease for additional hexoses suffers from diminishing returns, resulting in a smaller retention time decrease for each hexose added. Possibly, hexose addition on the less polar glycerol backbone shields a part of the hydrophobic tetraether core from interactions with the C18 column, resulting in a reduced retention time. Thus, GDGTs with two hexoses on the same end (DH-GDGT) should have a different retention time from GDGTs with a single hexose on both ends (MH-MH-GDGT) due to a difference in lipid core shielding. In accordance with literature, we only observe a single chromatographic peak for the mass corresponding to MH-MH-GDGT and DH-GDGT indicating the presence of only DH-GDGT. MH-MH-GDGT or other tetraether lipids with two hexose headgroups have not been described for the *S. acidocaldarius* lipidome, but these lipids have for instance been detected in sediment cores taken from the northern Arabian Sea [337,338].

Normal phase HPLC methods are particularly well suited to separate tetraethers with different numbers of cyclopentane rings in the lipid core [231,236,243,251]. Therefore, the capability of our RP-UHPLC method to separate tetraether lipids containing different amounts of rings was investigated using the narrow LC gradient (Figure 5, Supplemental Table 6). Separation was apparent from tetraether lipid species containing 5 rings or more and improved with a further increase in the number of rings. Thus, the method is capable of tetraether lipid separation of species with different numbers of rings to some extent.

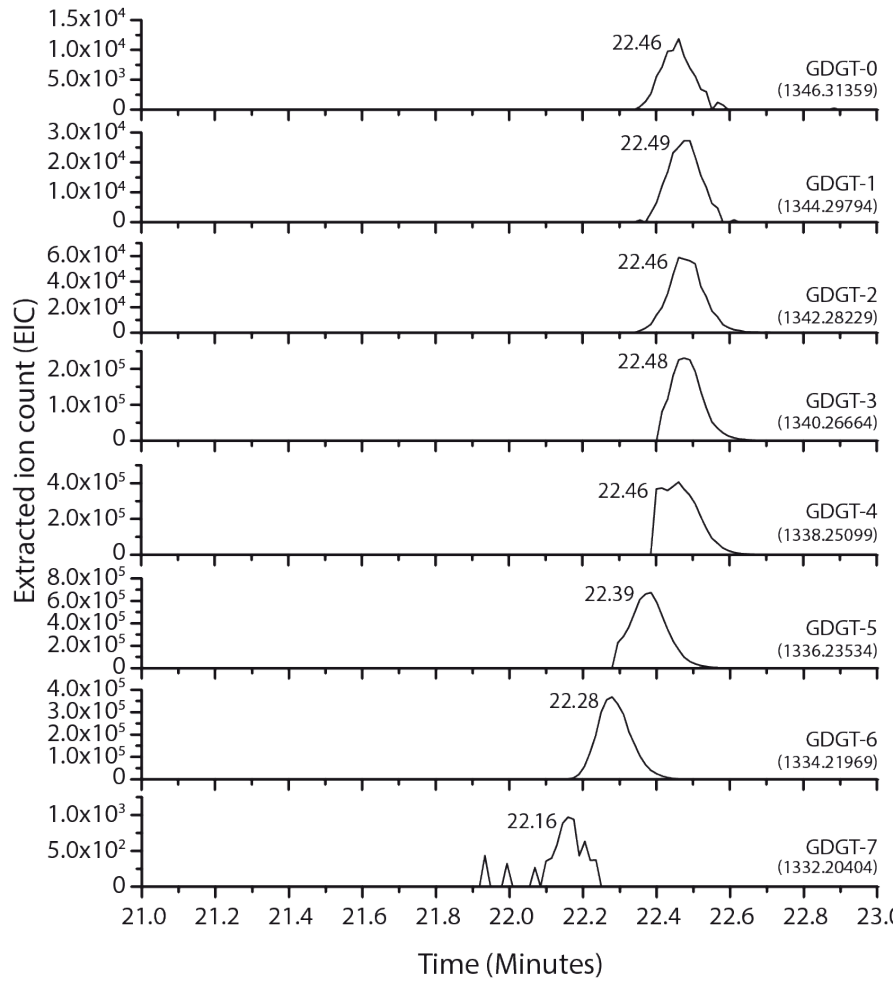


Figure 5: Extracted ion count chromatograms of masses corresponding to formate adducts of GDGT core lipids ([GDGT+CHO_2^-]$^-$) with different numbers of rings from a *S. acidocaldarius* MW001 total lipid extract eluted with the narrow gradient window method. The name format used is "GDGT-X", where "X" denotes the number of rings that lipid species has with the theoretical formate adduct ion mass for that lipid species listed below the name.

Discussion & conclusions

1-Butanol as a strong eluent in RPLC

Here we report on a versatile RP-UHPLC-ESI-MS method that can be used to separate and analyze complex biological lipid mixtures. The method depends on a routinely used C18 column but employs 1-BuOH instead of 2-PrOH as a solvent in the secondary mobile phase. While different hydrophobic alcohols, including 1-BuOH, have been used in the past as organic modifiers to alter the chromatographic properties of mobile phases, 1-BuOH is currently not widely used as primary eluent in RPLC separations [339,340]. The use of this stronger eluent facilitates earlier elution of common lipid analytes and allows for the elution of highly hydrophobic lipid species (e.g., cardiolipins, tetraethers and mycolic acids) before the end of the elution gradient, enhancing separation.

Analytical coverage is one of the barriers for further advancement in the field of lipidomics [292]. The use of stronger eluents, such as 1-BuOH in RPLC, pushes that barrier forward and allows for more method flexibility for the aforementioned application. Overall, RPLC methods, such as this one, are suited to analyze eukaryotic or bacterial lipidomes which often have a large diversity of acyl lipid tail species in a particular class. Although, most archaeal organisms lack the lipid tail diversity of bacteria or eukaryotes, this method is also of interest for archaeal lipidomics. As it showed an enhanced ability to elute highly hydrophobic lipid species such as tetraether lipids, which are often found in archaeal membranes.

Separation of different lipid classes

Besides the separation of strongly hydrophobic lipid species, different classes (e.g.: PG, PA, PC and PE) of less hydrophobic lipids such as lyso-, di-ester and di-ether lipids can also be separated based on their headgroup. Sufficient separation of these lipid classes reduces the chance to observe ion suppression effects and aids in detailed structural characterization. Separation effects based on lipid headgroup hydrophilicity are also apparent for highly hydrophobic species such as tetraether lipids. However, additional modifications that further increase headgroup hydrophilicity, such as adding additional hexose moieties, seem to suffer from diminishing returns. Although some GDGTs containing different numbers of cyclopentane rings could be separated, normal phase methods remain far superior for this purpose.

LPA and PA peak tailing

The peaks corresponding to LPA and PA did show moderate tailing in the lipid standard. PA is known to exhibit excessive tailing in RPLC separations and has an affinity to metals. To address this issue, Ogiso and coworkers [335] previously showed that an increase in buffering agent concentration, a decrease in pH 6.7 to pH 4 and the addition of 5 µM phosphoric acid to the mobile phase, suppressed the unidentified PA column interactions. The addition of phosphoric acid possibly resulted in more competing anion interactions from the mobile phase, thereby suppressing the unidentified PA analyte interactions, but at the same time may harm MS equipment. Additionally, the increase in buffer concentration and change of pH can cause retention changes potentially requiring method redevelopment specifically for LPA/PA analysis. Later, Cífková and coworker [341] also described the tailing behavior of PA during HILIC method optimization and managed to reduce PS and PA tailing. This was accomplished by employing a Hydride-modified silica HILIC column (replacing Si-OH with Si-H bonds) and using a relatively high concentration (40 mM) of ammonium formate pH 4 in the mobile phase. However, with our RPLC method, LPA and PA tailing was found to be greatly attenuated when using a column with a passivating coating on the inner metal surface of the column (Supplemental figure 7). This suggests that phosphate-metal interactions with metal column surfaces, not only free silanol group interactions, are a major factor leading to the observed peak tailing of phosphatidic acids such as LPA and PA using the RPLC method described in this work.

Adaptations for more specific applications

Alternative techniques for the analysis of highly hydrophobic compounds exist; for example, non-aqueous reversed-phase (NARP) LC. NARPLC is often used to analyze triacylglycerol lipids, mycolic acids and other highly hydrophobic compounds [287,295,342–345]. Typically, NARPLC uses RPLC columns but requires organic solvents that are often incompatible with RPLC system configurations. Although, the total exclusion of water from the eluents used for the 1-BuOH-based RP-UHPLC method is not possible due to buffer solubility issues; the miscibility and solvation properties of 1-BuOH enable the use of high proportions of this solvent. The high proportion of 1-BuOH and low amount of water allows the RP-UHPLC method described here to approach NARP-LC conditions and elute very hydrophobic compounds (e.g. mycolic acids) near the end of the elution gradient (Supplemental Figure 8 and 9) [346,347].

As this method can be run on RPLC configured machines, it is more widely applicable compared to either NPLC or NARPLC. Moreover, in contrast to NARP, this RPLC method is also suitable to analyze less hydrophobic compounds, highlighting

its versatility. Eluent A was designed to allow for separation of less hydrophobic analytes such as fatty acids and lyso-lipids during method development. Combined with the eluotropic strength of 1-BuOH in eluent B, this resulted in a single method capable of separating different analytes ranging from bacterial fatty acids to highly hydrophobic cardiolipins, archaeal tetraethers or methoxy mycolic acid in a single run; making this a useful versatile LC separation method for lipid research. Moreover, the pronounced selectivity towards the acyl chain in bacterial lipids from the C18 column also allows this method to be used for detailed acyl chain composition analysis per lipid class.

Potential improvements

Although the 1-BuOH-based method described here has a larger analytical range compared to conventional mobile phases, the instrument response decreased for the moderately polar lipids in the lipid standard mixture compared to using 2-PrOH-based mobile phases (Supplemental figure 1). Possibly, the ionization efficiency decreased because 1-BuOH is less volatile compared to 2-PrOH. Different ionization conditions or an ionization technique such as heated-ESI (HESI) might be more appropriate for the use of this method for the analysis of moderately hydrophobic lipid mixtures; or possibly atmospheric pressure chemical ionization (APCI) for mixtures containing mostly highly hydrophobic lipids such as tetraether lipid cores or mycolic acids. The use of ammonium formate as buffering agent is suitable while running negative mode ESI-MS to promote the formation of negatively charged adduct ions; especially for lipids with a positively charged headgroup such as PC or neutral glycolipids, which prefer to form formate adduct ions over deprotonation. However, other buffering agents or additives, such as formic acid, acetic acid or ammonium acetate, could be used to increase the ionization efficiency, improve the chromatographic peak shape of particular analytes, or promote positive ion formation in positive mode ESI-MS [105,310]. Furthermore, the gradient used in the 1-BuOH method described in this work can readily be adjusted for particular lipid classes of interest, allowing for shorter analysis times that can at least be reduced by half. The method could potentially be further improved by substituting MeCN with MeOH in the mobile phase, potentially improving peak shapes of specific analytes and GDGT response [105]. Furthermore, if desired, it might be possible to develop a NARP or NARP-capable method based on methanol or MeCN as eluent A, against 1-BuOH as eluent B. With only eluent B or both eluents excluding water. Moreover, formic- or acetic acid instead of ammonium formate or –acetate could possibly be used to avoid buffer precipitation and promote positive ion formation. In theory, other more hydrophobic solvents, equipped with different eluotropic properties, could be used as well. However, special care must be taken to avoid issues with instrument compatibility (e.g.: pump seals, increased back pressure), mobile phase miscibility, buffer solubility, and possibly decreasing ionization efficiency.

Final remarks

1-BuOH has a low long-term health impact compared to several commonly used NPLC solvents and is also less flammable and not known to form dangerous oxidation products over time. This makes 1-BuOH more user friendly compared to many other alternative LC-MS solvents. From a sustainability point of view, 1-BuOH can be obtained from bio-renewable sources in contrast to for example, methylene chloride, tetrahydrofuran, chloroform or hexane, lowering the environmental footprint of this lipid analysis method compared to other common methods.

Data availability statement

All relevant data is available in the manuscript and accompanying supplemental material.

Acknowledgements

The technical discussions regarding method development with T.D. Tiemersma-Wegman (University of Groningen) were highly appreciated.

Funding information

The work was supported by the Building Blocks of Life programme (737.016.006), and the 'BaSyC – Building a Synthetic Cell' Gravitation grant (024.003.019), both programmes are subsidized by NWO. Neither programme influenced the development or submission of this work.

Authors contributions

N.A.W.d.K., A.J.M. and A.J.M.D. developed the concept and designed the study. N.A.W.d.K., M.E. and R.A. performed the study, acquired the data and analyzed the data. R.A. synthetized the lipid molecules that are not commercially available. N.A.W.d.K. drafted the article, and all authors contributed to revising it critically for important intellectual content. All authors approved the final version of the submitted manuscript.

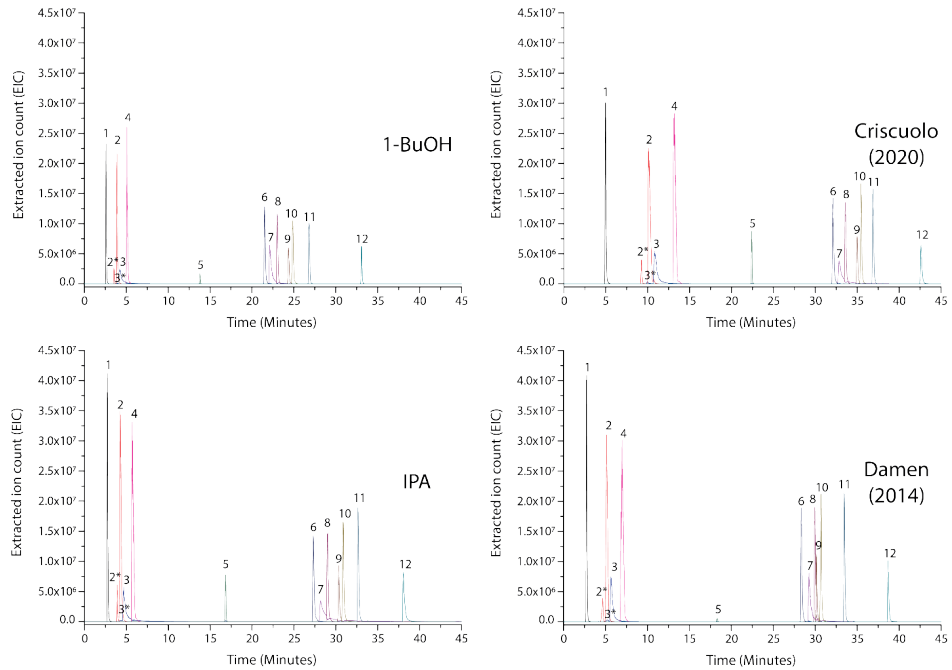
Competing interest

The authors have no competing interest to declare

Supplemental information

Comparing the 1-Butanol with isopropanol as RPLC solvent using the lipid standard

The lipid standard as described in the main text was also separated using a mobile phase containing IPA instead of 1-BuOH and two established IPA-based mobile phases described by Damen [332] and Criscuolo [312] to compare differences in separation (Supplemental figure 1).



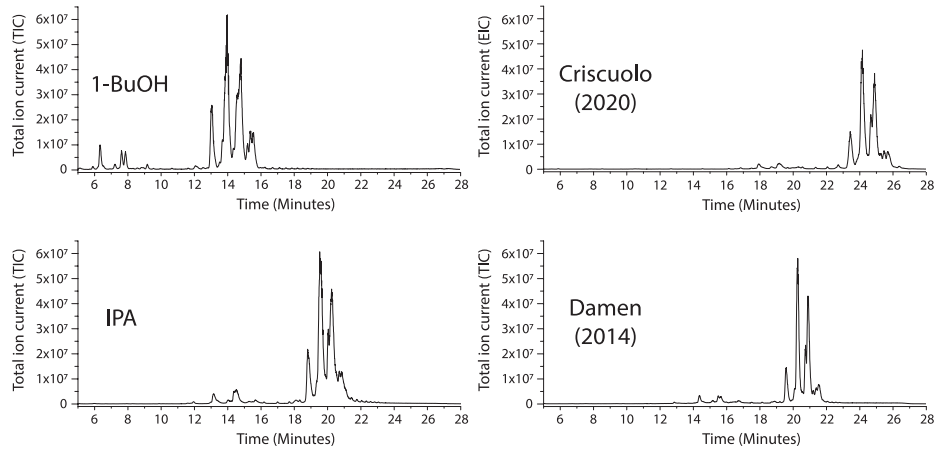
Supplemental figure 1: LC-MS extracted ion chromatogram comparing the separation of a defined lipid standard mix (0.1 mg/ml per lipid in MeOH) containing various selected lipid classes on a well-used (>1500 injections) Acquity UPLC CSH C18 column (2.1 x 150 mm, 1.7 µm; Waters) with various mobile phase compositions using the full gradient (5-90% B). Peak numbers correspond to the compounds listed in table 1. The "1-BuOH" panel is the same as figure 1 from the main text and the mobile phase is described in the materials and methods section. The "IPA" method is the same as the "1-BuOH" mobile phase with 1-BuOH being replaced by 2-propanol (IPA) as a one-to-one comparison. The "Criscuolo" mobile phase is as described by Criscuolo, et al. [332]. The "Damen" mobile phase is as described by Damen, et al.[312].

All mobile phases were able to elute the lipids in this lipid standard. All 2-propanol-based methods elute cardiolipin at the start of, or in the purge phase of the gradient, showing the eluotropic limitations of this eluent; whereas the use of butanol as eluent allowed for the elution of even more hydrophobic compounds using this same gradient (supplemental figure 9). Strikingly, the 1-BuOH based mobile phase elutes all analytes earlier comparing to the other mobile phases. This overall reduction in retention time results in an increase of chromatographic space for more hydrophobic compounds at the end of the gradient. The mobile phase containing 1-BuOH generally maintains or improved on separation quality due to a lower amount of PA (18:1/18:1) (peak 7) peak tailing compared to the "IPA" and "Criscuolo" mobile phases and better peak spacing of peaks 7-10 compared to the "Damen" mobile phase. However, the "Criscuolo" mobile phase is clearly more suitable for the separation of less hydrophobic analytes such as lyso-lipids. Additionally, the "Criscuolo" mobile phase was unable to elute CL (18:1/18:1/18:1/18:1) during the gradient and resulted in eluting during the purge phase, limiting its analytical value for cardiolipins and more hydrophobic lipids such as archaeal cardiolipins and tetraether lipids.

All three 2-propanol-based mobile phases show a higher ionization efficiency for the less hydrophobic analytes (Peaks 1-11) compared to the 1-BuOH mobile phase which could be caused by less suitable ionization conditions or parameters. The use of HESI instead of ESI could yield significant sensitivity improvements due to the lower vapor pressure and higher boiling point of 1-butanol compared to 2-propanol. The "Damen" mobile phase seems to exhibit a higher ionization efficiency for all analytes, but this was not reproducible in subsequent runs with this mobile phase and is likely an injection inaccuracy or detection sensitivity anomaly. The subsequent reproduction runs exhibited extreme tailing specific for Lyso-PA and PA while the ionization efficiency of other analytes was comparable to the other 2-propanol-based mobile phases. The PA peak shape as shown in supplemental figure 1 for the 2-propanol based mobile phases, to our surprise, was much better compared to the PA peak shape observed in our laboratory in the past using the "Damen" mobile phase and a similar, but different column. Notably, the deteriorated PA peak shape was reminiscent of that previously observed poor peak shape (Phosphatidic acid tailing is further addressed in supplemental information section 1.4).

Next, an assessment of mobile phase performance was made using a semi-defined sample containing hydrophobic lipids an *E. coli* CL extract (Supplemental figure 2). Similarly, to the elution of the lipid standards, the 1-BuOH mobile phase elutes the lipid mixture earlier compared to the other tested mobile phases. Every separation

shows the presence of the 4 main cardiolipin peak clusters consisting of several different cardiolipin species. Overall, the separation of CL species in the 1-BuOH method is superior to the other methods. The separation time between the first and the fourth eluting major cardiolipin cluster peak is approximately 2.6 minutes (t_R 1-BuOH 13.2 and 15.8 respectively) using the 1-BuOH mobile phase compared to approximately 2, 2.1 and 2.3 minutes for the "Damen", "IPA" and "Criscuolo" mobile phases respectively. The second and third eluting major cardiolipin clusters (t_R 1-BuOH 14.0 and 14.8 minutes respectively) show to consist of 3 peaks whereas the other methods only reveal the presence of two peaks. Later eluting peaks, mainly the fourth eluting major cardiolipin cluster (t_R 1-BuOH 15.5 minutes), representing more hydrophobic CL species, show to be separated somewhat better. Additionally, the peak shape and separation for lyso-cardiolipins (t_R 1-BuOH 6.3, 7.6 and 7.9 minutes respectively) seems improved over the IPA containing methods.



Supplemental figure 2: LC-MS TIC chromatograms showing the elution of an *E. coli* cardiolipin extract from a well-used (>1500 injections) Acquity UPLC CSH C18 column (2.1 x 150 mm, 1.7 µm; Waters) using the narrow gradient window method with 4 different mobile phase compositions. The various mobile phases used are explained in the legend of Supplemental figure 1.

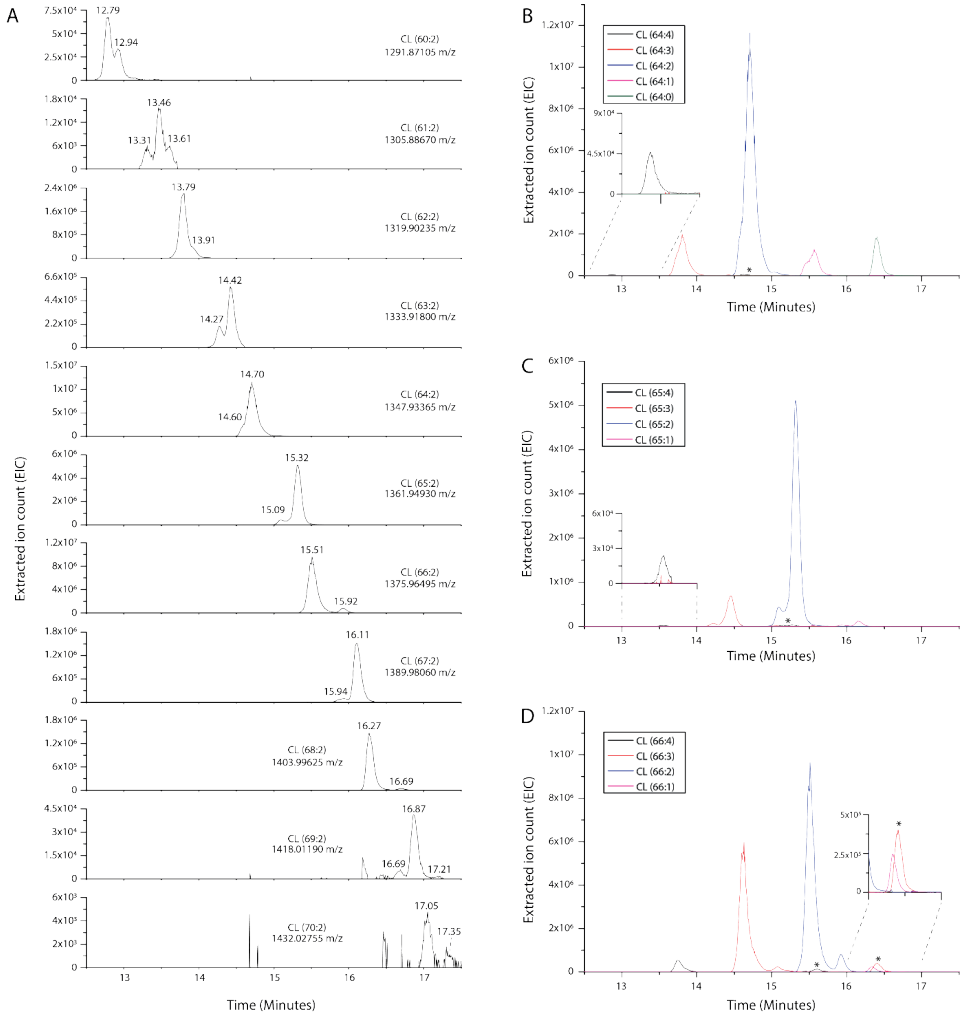
Analysis of the CL elution pattern using the 1-BuOH mobile phase

The increase in the separation of CL species using the 1-butanol mobile phase compared to the 1-propanol containing mobile phases (Supplemental figure 2) allows us to characterize the elution patterns of CL molecule series in more detail (Supplemental figure 3). Notably, the increasing number of carbon atoms does not cause a simple proportional increase of retention time but shows a step-wise change between uneven and even numbers of carbon atoms (A). Further analysis revealed that the retention time increases by approximately 0.62 minutes by adding a CH_2 moiety to an even number of carbon atoms to form an odd number of carbon atoms. However, the increase of adding a CH_2 moiety from an odd to even number of carbon atoms varies between 0.16 and 0.33 minutes with a total average of approximately 0.43 minutes increased retention per CH_2 moiety. This could be caused by the fact that uneven fatty acid chains generally contain cyclopropyl moieties instead of double bonds which have slightly different retention characteristics and can cause extra peaks or peak shifts. Remarkably, EIC traces of CL molecules with the same amount of carbon atoms, but differing numbers of unsaturated bonds show similar elution patterns regarding multiple peaks and their positioning relative to one another (B-D). For example, CLs with 64 fatty acid carbons do not have much spacing between isomer peaks and this spacing is similar regardless of the amount of double bonds present (B). On the contrary, CLs with 66 fatty acid carbons seem to show considerably more distance between the isomeric peaks and for these too the distance remains fairly similar regardless of the amount of double bonds (D). Marked peaks (*) are isobaric $[M\text{-}2H+Na^+]^-$ adducts of other CL species ions; e.g., CL 64:0 (panel B, dark green trace) has a $[M\text{-}2H+Na^+]^-$ adduct at *m/z* 1373.94689, which is within the 5 ppm mass error window of CL 66:3 $[M\text{-}H]^-$ at *m/z* 1373.94930 (Panel D, red trace) and hence these show as extra unrelated peaks in some chromatograms.

Differentiation of lipid isomers using MS/MS

Several past works have described the different fragmentation results of different positional isomers of glycerophospholipids [311,330,331]. This phenomenon was assessed and verified using our LCMS instrumentation (Supplemental figure 4).

Both diesters, PG 18:1/16:0 and PG 16:0-18:1, produce two first-order lyso fragments $[M\text{-}H\text{-}R_1CHCO_{sn1/2}]^-$ (B and D) at *m/z* 483 and *m/z* 509. These fragments correspond to the loss of a ketene group at the *sn1* or *sn2* position. As *sn1* lyso-lipid fragments are more stable, this fragment should be present in a higher abundance compared to the *sn2* lyso-lipid fragment. Masses resulting from subsequent fragmentation are annotated based on whether they resulted from a *sn1* or *sn2* ketene loss in the first fragmentation step, by using the mass of the first-order fragmentation product ions. Lyso-lipid fragments can further fragment to lose glycerol and/or water. For example: *m/z* 509 can lose glycerol to yield *sn1*-lyso-PA (*m/z* 435) and/or water to yield an ion of *m/z* 417 or *m/z* 491 respectively.



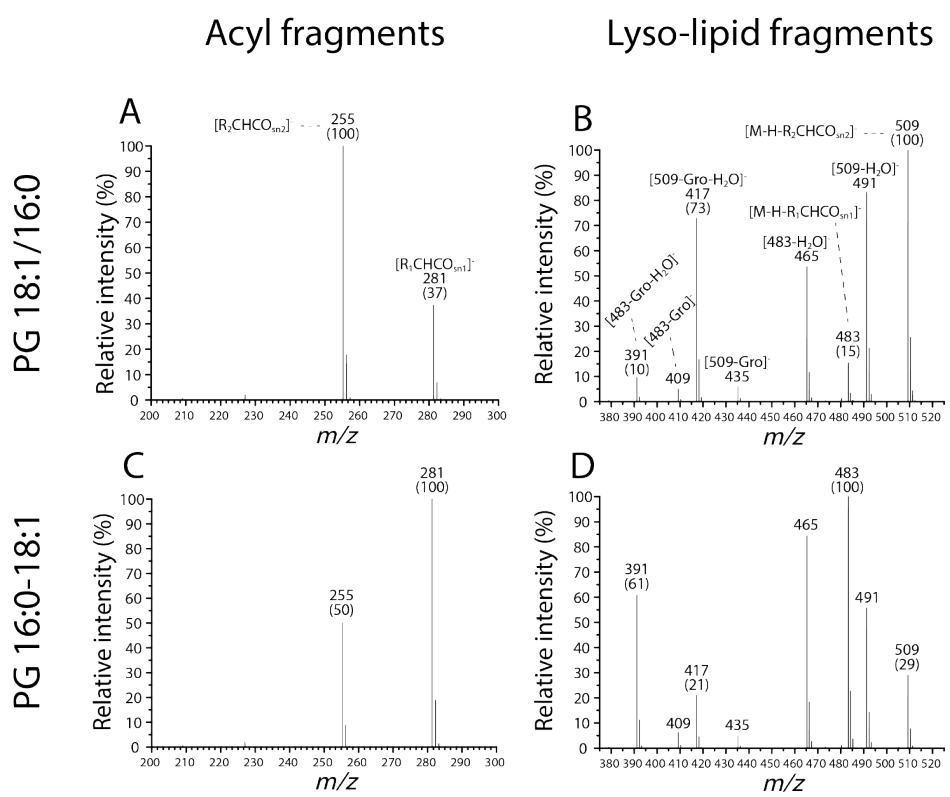
Supplemental figure 3: LC-MS extracted ion chromatograms of an *E. coli* cardiolipin extract showing the retention behavior of different cardiolipin molecules using the narrow gradient window method. The panels highlight the increased retention of cardiolipin molecules with: **(A)** an increasing number of carbon atoms in the fatty acid tails or **(B-D)** a decreasing number of unsaturated bonds with 64, 65 or 66 carbon atoms in the fatty acid tails.

The second-order lyso-PA fragments (*m/z* 409 and *m/z* 435) are not suitable to determine the identity of the acyl chain positional isomer of the parent ion as the abundance of these peaks is relatively low and similar to one another. Likewise, the second-order dehydro lyso-PG fragments (*m/z* 465 and *m/z* 491) are also relatively similar in abundance compared to the difference between first-order lyso-PG fragments (*m/z* 483 and *m/z* 509), and third-order dehydro lyso-PA fragments (*m/z* 391 and *m/z* 417). Therefore, the first-order lyso-PG fragments (*m/z* 483 and *m/z* 509) and third-order dehydro lyso-PA fragments (*m/z* 391 and *m/z* 417) can be used to determine the acyl tail position. Based on these fragments, the parent ion of the MS/MS spectra on the top row must have predominantly had the 18:1 acyl chain on the *sn1* position, as per the compound label (figure 4 panel A and B). The parent ion of the bottom row must have predominantly had 16:0 on the *sn1* position (figure 4 panel C and D). The lower difference of signal intensity between the *sn1* and *sn2* lyso-lipid fragments can be attributed to the lower isomeric purity of the 16:0-18:1 lipid (compare figure 4 panel A versus C and B versus D).

The acyl tail fragment abundance (Figure 4 panel A and C) is also in agreement to the higher stability of *sn1* lyso-lipids; i.e., the acyl moiety attached to the *sn2* position is the more abundant fragment. The acyl fragments are not as suitable to determine the isomeric composition as this is more susceptible to interference of co-eluting lipids compared to first-order lyso lipid fragments and does not discriminate as well as the first-order lyso-lipid or third-order dehydro-lyso-PA fragments.

To determine the isomeric composition of CL species (Supplemental figure 5) only the third-order lyso-lipid fragments were used. The glycerol headgroup of the first-order lyso-lipid fragments is shared in CL molecules and could potentially skew their abundance and further complicate matters. Furthermore, the discriminative power of both the first order and third order fragments is similar (Signal intensity 100% versus 15% and 73% versus 10% for the first- and third-order *sn1* and *sn2* lyso-lipid fragments respectively) and should not negatively impact the quality of analysis.

The next step was to apply the knowledge of phospholipid fragmentation behavior described above to the analysis of cardiolipins. Two CL masses that showed multiple peaks in the extracted ion chromatogram, found in the *E. coli* cardiolipin extract, were selected. Three fragmentation spectrum ranges, for acyl fragments, lyso-lipid fragments and diester lipid fragments were used for identification purposes (Supplemental figure 5).



Supplemental figure 4: MS/MS fragmentation spectra of two PG 34:1 lipid standard compounds. Two mass ranges are shown for each lipid; the acyl mass fragment range (**A, C)** and the lyso-lipid range **(B, D)**. The fragment masses are annotated above the *m/z* value in mass spectra in the panels on the top row and the signal intensity for some fragments is listed in parentheses. The supplier (Avanti Polar Lipids) indicated that PG 16:0-18:1 typically has an 85% isomeric purity, whereas the "isopure" PG 18:1/16:0 should have >97% isomeric purity.

The acyl fragment spectrum range was used for identifying acyl chains that the parent ion might contain. The lyso lipid fragment spectrum range provides information about the *sn1* vs *sn2* positioning of the acyl tails on each side of the cardiolipin molecule as *sn1* fragments are more stable and therefore more abundant (for a pure diester lipid, dehydro-lyso-phospholipid fragments show approximately a 100% vs 14% relative intensity ratio for *sn1* vs *sn2* fragments respectively (Supplemental figure 4, panel B). The diester lipid spectrum range was used to provide constraints, limiting acyl combinations and aiding in final

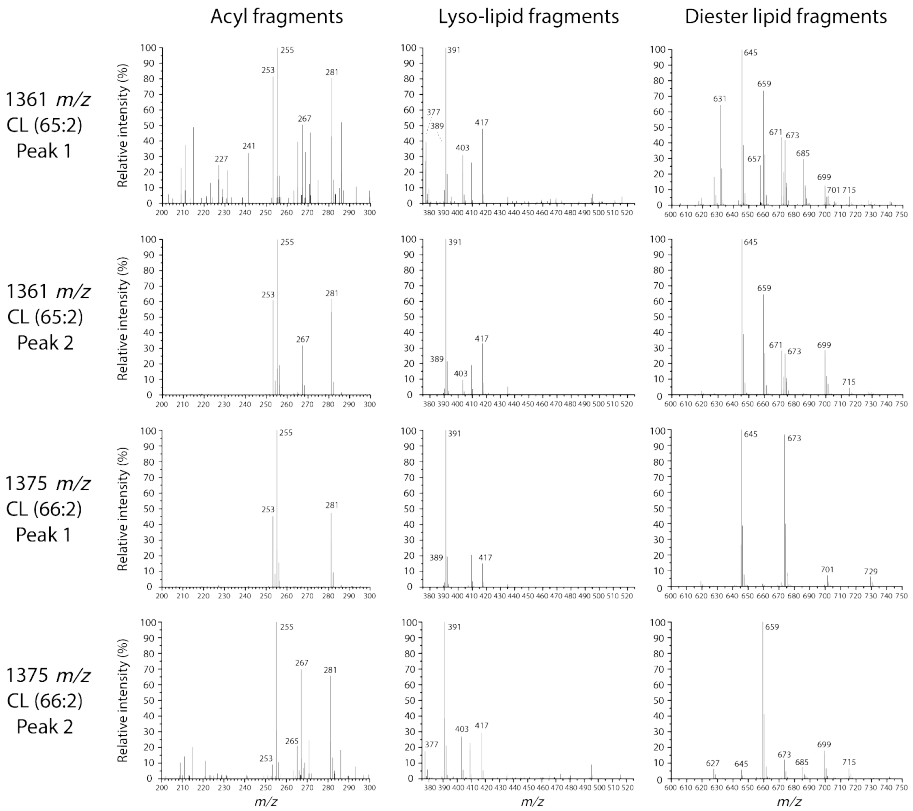
identification. The identification process of the acyl composition of the selected cardiolipin lipid species is described below in increasing order of difficulty and decreased certainty.

The identification of CL 66:2, peak 1: The acyl fragment signals indicate this lipid is relatively pure or much more abundant compared to other lipids during its elution and/or the time period chromatographic peak MS/MS data was extracted. The acyl fragments present are: 16:1, 16:0 and 18:1. The only possible combination of those acyl fragments to match the required amount of carbons and unsaturations (66:2) is 2x 16:0, 1x 16:1 and 1x18:1, as exemplified by the abundance ratios of the acyl fragment peaks. The lyso lipid fragment corresponding to Lyso-PA (16:0) at *m/z* 391 indicates that 16:0 is positioned virtually exclusively on the *sn1* positions resulting in CL 16:0/16:1 ++ 18:1/16:0, which is in agreement with the diester fragments of 32:1 and 34:1 (*m/z* 645 and 673 respectively) and the total CL 66:2 mass.

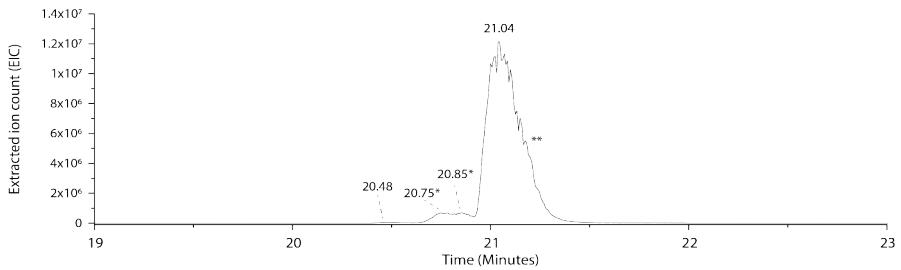
The identification of CL 65:2, peak 2: The acyl fragments show 4 possible acyl chains 16:0, 16:1, Cy17:0 and 18:1. The lyso fragment abundance suggests 16:0 is present predominantly on the *sn1* position. As such the CL is expected to have two 16:0 acyl chains as there are two *sn1* positions. The diester fragments each show 32:1, 33:1, 34:1, 34:2 and 36:2. Where diesters 34:1, 34:2 and 36:2 are not possible fragments due to the complementary diester mass not being present to reach the total of 65 carbon atoms and 2 unsaturated bonds/propyl groups. This leaves 32:1 and 33:1 of which 33:1 must contain a Cy17:0 acyl tail due to the uneven number of carbon atoms; paired with 16:0 on the *sn1* position. Fragment 32:1 thus consists of 16:1 and 16:0. Leading to the final predominant structure CL 16:0/16:1 ++ Cy17:0/16:0 consistent with a CL 65:2 mass.

The identification of CL 66:2, peak 2: The acyl fragment spectrum indicates a larger variety of possible acyl tails with the most abundant tails being 16:0, Cy17:0 and 18:1. The large variety of acyl tail fragments indicate that multiple lipids are eluting and are possibly of a similar abundance to the lipid being analyzed (Also visible in the full-MS spectrum. Not shown.). The lyso fragment spectrum indicates that 16:0 is predominantly present on the *sn1* positions of lipids eluting during the time selected for MS/MS data extraction. The diester fragment corresponding to 33:1 is the most abundant. The 33:1 fragment likely has a 16:0 acyl chain on the *sn1* position with Cy17:0 on the *sn2* position. This would lead to a final structure of CL 16:0/Cy17:0 ++ Cy17:0/16:0 consistent with a CL 66:2 mass. Additionally, this is in agreement with the observation that lipids containing cyclopropyl moieties elute slightly later compared to isomeric lipids containing unsaturated bonds instead (CL 66:2 peak 1).

The identification of CL 65:2, peak 1: The acyl fragment spectrum displays a large variety of possibly acyl tails present in the molecule corresponding to peak 1 of CL 65:2. This indicates the abundance of our lipid of interest is relatively low compared to interfering lipids eluting during the same period (as indicated by the full-MS spectrum, not shown). The most abundant acyl tails eluting during the time selected for MS/MS data extraction are: 16:0, 16:1, Cy17:0 and 18:1. The lyso fragment spectrum still shows 16:0 to be the predominant acyl tail to be present on the *sn1* position, however due to interfering lipid fragments this cannot be established with certainty. Only the following diester fragments specifically co-elute with the parent ion of our lipid of interest: 31:1, 32:1, 32:2, 34:1, 34:2 and 35:2. Fragments containing 2 unsaturations or cyclopropyl moieties are excluded as no diester fragments without unsaturation or cyclopropyl moieties were found to specifically co-elute with the parent ion (Diesters without any unsaturations are required if a single diester fragment of the parent ion would already contain 2 unsaturations to match a total of 2 unsaturations). This leaves diester fragments 31:1, 32:1 and 34:1. Only the combination of 31:1 and 34:1 would be in agreement with the total mass of CL 65:2. Since only one of the fragments has an odd number of carbons in the acyl chains and each have a single unsaturation or double bond; it is likely that 31:1 contains Cy17:0 and thus by extension also 14:0. Even though 14:0 acyl tails are unusual. 16:1 with 15:0 would also be possible. Acyl tails with an uneven amount of carbon atoms in their tail but without a cyclopropyl moiety, such as 15:0, are shown to be rare in *E. coli*. However, the retention time difference between peak 1 and peak 2 of CL 65:2 indicates the substitution of a cyclopropyl moiety for an unsaturated bond leading to a most likely composition of 16:1-15:0 for diester 31:1. The positional isomer of this fragment cannot be established due to similar abundances of the 16:1 and 15:0 *sn1* lyso-lipid fragments. Although, keeping in mind the 15:0 acyl chain is unusual, combined with the relatively high *sn1* lyso-lipid fragment signal corresponding to a 15:0 acyl tail, likely indicates a configuration of 15:0/16:1 for a 31:1 diester fragment. Then, with the abundance of 16:0 and 18:1 in the fatty acid fragments, and with the prevalence of 16:0 being present on the *sn1* position, 16:0/18:1 is the most likely composition of the 34:1 diester fragment. The putative final proposed structure, which is in accordance with the mass of CL 65:2 and eluting earlier than a lipid containing a cyclopropyl moiety, is CL 15:0-16:1 ++ 18:1/16:0. However, ultimately the MS/MS analysis for this structure is not entirely conclusive for this lipid.



Supplemental figure 5: MS/MS fragmentation spectra from two *E. coli* cardiolipin extract masses (CL 65:2 and CL 66:2) that both elute in two peaks. The left column shows the mass range corresponding to acyl fragments. The middle column shows the mass range corresponding to dehydro-lyso-PA fragments. The right column shows the mass range corresponding to diester (PA) fragments of the cardiolipin parent ion. Each row of mass spectrum ranges corresponds to the spectrum collected during elution of the lipid listed on the left.

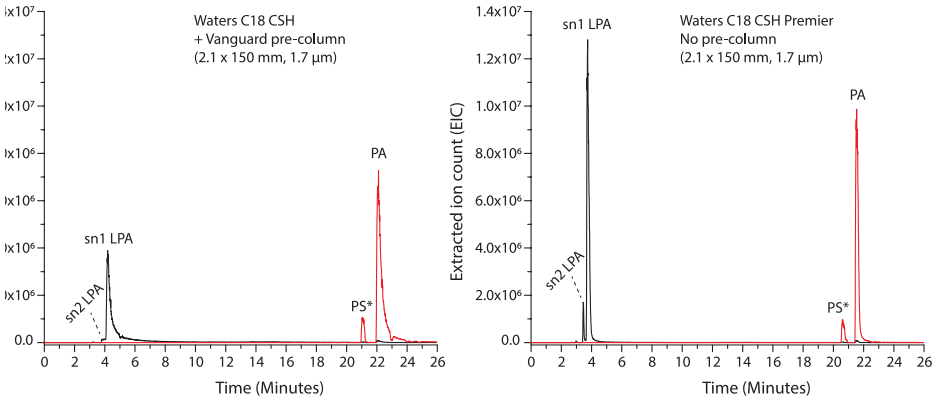


Supplemental figure 6: Extracted ion chromatograms (EIC) of PG (33:1) from the *E. coli* total lipid extract. The two main peaks are shown t_R 20.8 and 21.04. * The decrease in EIC around 20.8 could be caused by ion suppression of eluting PE (28:0) which has its retention time at that time. Therefore, the peaks at t_R 20.75 and 20.85 could be a single peak. ** The shoulder as noted in the main text at approximately 21.20 minutes.

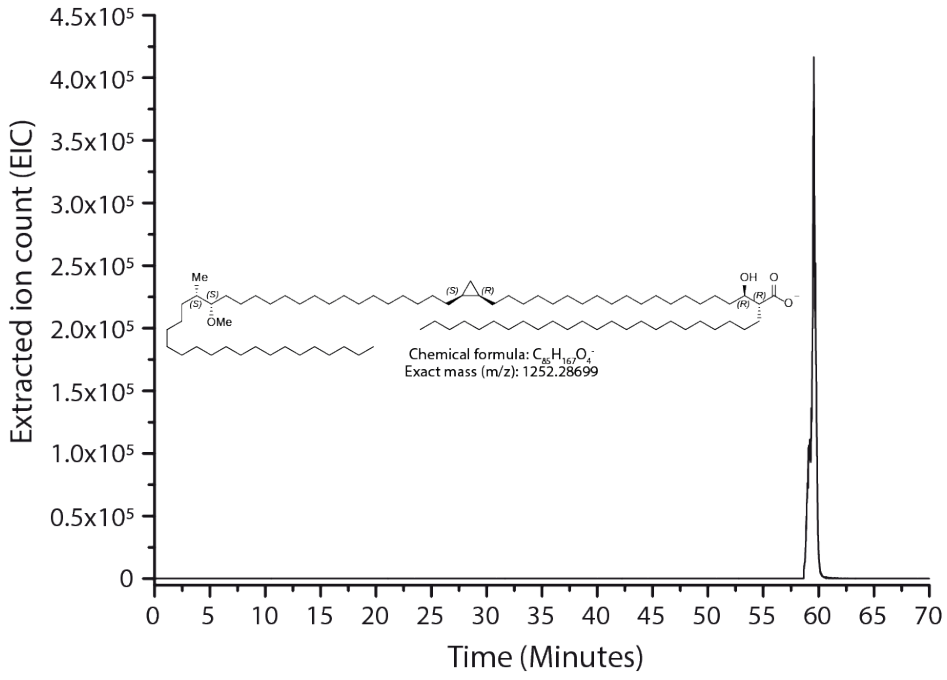
Passivation of column inner metal surface results in improved phosphatidic acid peak shape

Comparing various IPA-based mobile phases to our 1-BuOH containing mobile phase revealed a reproducibility issue. After a number of runs with IPA-based mobile phases it was noticed that the DOPA peak shape in the lipid standard had strongly deteriorated up to the point of giving a semi-continuous DOPA background signal. A new Waters C18 CSH column was tested with the same mobile phase and the DOPA peak was present but showing severe tailing, indicating this issue was related to the column. When running a lipid standard on the previously unused column with a freshly prepared 1-BuOH mobile phase peak shape recovered somewhat but the following run with the IPA-based mobile phase still showed a poor peak shape for DOPA.

As free phosphates are notorious for interacting with metals, we hypothesized that the metal inner column wall of our column could have been passivated over time, preventing DOPA interacting with the column wall metal. This would explain why a new column did not completely solve the issue, it would lack passivation. The improved running behavior of DOPA with 1-BuOH could be attributed to the stronger hydrophobic interactions between DOPA and the mobile phase compared to the column wall interactions attenuating this effect. The disappearance of DOPA on a column that previously showed a decent DOPA peak shape after running many IPA-based mobile phases hints at that the cause could be found in the mobile phase. Perhaps the extensive use of 1-BuOH mobile phases on this column passivated the metal inner surface of the column wall over time. To test whether DOPA peak shape could be improved with a passivated inner surface of the column wall; lipid standard mixtures were run on a Waters C18 CSH Premier column (Supplemental figure 7, cat#: 186009462). This column is virtually identical to the Waters C18 CSH columns described above with the exception that the metal inner surface of the column wall has been passivated as part of the manufacturing process. Using the 1-BuOH mobile phase, LPA and DOPA peak shape in the lipid standard mixture is much improved using the premier column compared to the non-premier Waters C18 CSH column. No clear differences were observed in the running behavior of the other lipids in the lipid standard mixture compared to the non-premier column (not shown). All lipids showed a small retention time shift, which can be attributed to the absence of a guard column on the premier column. This experiment does not explain the previously observed DOPA peak shape degradation during testing of different mobile phases. However, it does show that the inner metal surface of the column can have a large impact on the running behavior of phosphatidic acids while the use of a column with a passivated inner metal surface can significantly improve the peak shape of phosphatidic acids.



Supplemental figure 7: Extracted ion chromatograms (EIC) of LPA (black trace), and PA (red trace) from a lipid mixture containing various lipid species (0.1 mg/ml in MeOH) using the full gradient method and two different columns. * The lipid mix also contained PS, which tends to undergo some in-source fragmentation yielding PA, manifesting as an additional peak in the PA EIC.

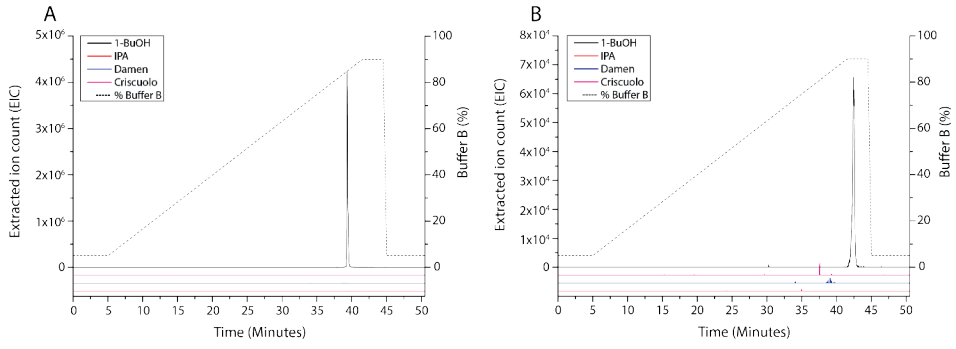


Supplemental figure 8: An extracted ion chromatogram showing the elution profile of synthetic methoxy mycolic acid using two in-line coupled CSH columns (See supplemental materials and methods).

Comparison of different mobile phase compositions for the elution of particularly hydrophobic lipid standards

To emphasize the advantage of a more eluotropic mobile phase solvent such as 1-BuOH; two particularly hydrophobic lipids were selected to test elution behavior with three IPA-based mobile phase compositions using the full 5-90% B gradient method as described in the main text.

The 1-BuOH mobile phase was the only mobile phase capable of eluting synthetic isocaldarchaeol and methoxy mycolic acid using the full, 5-90% B, gradient tested. Taking the system-dwell-volume into account, Isocaldarchaeol elutes during the gradient and methoxy mycolic acid starts to elute at the just after onset of the purge phase. For a more comprehensive comparison, the gradient was extended beyond the 5-90% eluent B gradient described in the manuscript to 5-98% B, while keeping a similar gradient slope (Not shown in figure, method described in supplemental materials and methods). The "IPA" mobile phase was able to elute isocaldarchaeol at the start of the purge phase and methoxy mycolic acid during the purge phase (Taking the ~2 minute system dwell-volume into account). The "Damen" mobile phase was only able to elute both of the analytes during purge phase of the method whereas the "Criscuolo" mobile phase was not able to elute either of the analytes at all (not shown).



Supplemental figure 9: Extracted ion chromatograms of a lipid standard mixture containing (A) isocaldarchaeol and (B) methoxy mycolic acid. The lipids present in the standard sample have a concentration of 0.1 mg/ml in MeOH and were separated using the full gradient method (5-90% B) with different mobile phases on a well-used (>1500 injections) Acquity UPLC CSH C18 column (2.1 x 150 mm, 1.7 µm; Waters). The different methods are described in the legend of Supplemental figure 1.

Supplemental tables

Supplemental table 1: A resolution comparison of the BuOH-based mobile phase versus a mobile phase with BuOH substituted for PrOH using the full gradient method. The data in the top section of the table is derived from chromatograms of the lipid standard and the data in the bottom section is derived from the *E. coli* CL extract. The number in parentheses refers to the eluting order of isomeric peaks. The major peak was selected for analysis unless marked with * in which case a minor peak was selected.

	Mobile phase method resolution		
Compared peaks	**2-PrOH**	**1-BuOH**	**Separation type**
sn2-LPG vs *sn1*-LPG	2.09	1.73	Lyso-lipid isomer
DOPC vs DOPE	2.60	1.65	Diester headgroup
Lyso-CL 50:2 (1) vs Lyso-CL 48:1 (2)	0.49	1.30	Lyso-CL
CL 64:2-OH (2) vs CL 64:2 (1)	2.92	4.45	CL with polar group on radyl tail
CL 65:2 (2) vs CL 66:2 (1)	0.67	0.77	CL fatty acid carbon
CL 66:2 (1) vs CL 66:2 (2)*	1.95	3.39	CL double bond
CL 66:2 (1) vs CL 66:1 (1)	0.85	1.69	CL isomer

Supplemental table 2: A column efficiency comparison of the BuOH-based mobile phase versus a mobile phase with BuOH substituted for PrOH using the full gradient method. The data in the top section of the table is derived from chromatograms of the lipid standard and the data in the bottom section is derived from the *E. coli* CL extract. The number in parentheses refers to the eluting order of isomeric peaks. The major peak was selected for analysis unless marked with * in which case a minor peak was selected.

	Mobile phase	
Chromatographic peaks	**2-PrOH**	**1-BuOH**
sn2-LPG	10807	7643
sn1-LPG	9378	6884
DOPC	362812	149264
DOPE	232818	118495
Lyso-CL 50:2 (1)	242302	347914
Lyso-CL 48:1 (2)	280284	354049
CL 64:2-OH (2)	581893	507615
CL 64:2 (1)	139445	211334
CL 65:2 (2)	386952	389781
CL 66:2 (1)	173843	221877
CL 66:2 (2)*	240195	404365
CL 66:1 (1)	152410	238608

Supplemental table 3: A comparison of limit of detection (LoD) and limit of quantification (LoQ) of the BuOH-based mobile phase versus a mobile phase with BuOH substituted for PrOH using the full gradient method.

	LoD (pmol)		LoQ (pmol)		R^2	
Analyte	**PrOH**	**BuOH**	**PrOH**	**BuOH**	**PrOH**	**BuOH**
DDM	0.021	0.025	0.063	0.076	0.999	0.996
PG (10:0/10:0)	0.016	0.030	0.049	0.091	0.999	0.994
DOPG	0.005	0.023	0.015	0.070	1.000	0.993
DOPA	0.234	0.036	0.709	0.110	0.851	0.987
DOPC	0.036	0.034	0.109	0.103	0.995	0.984
DOPE	0.020	0.028	0.060	0.084	0.999	0.991
CL (4x18:1)	0.008	0.007	0.024	0.020	0.999	0.998

Supplemental table 4: A list of the masses found in the commercial *E. coli* cardiolipin extract. Displayed are the retention time (Narrow gradient method), total carbons in the acyl chain, number of double bonds and cyclopropyl moieties in the acyl tails. In addition, the cardiolipin acyl configuration of some lipids is proposed based on fragmentation data. Most masses eluted in multiple peaks. Two peaks of each mass were analyzed separately; bold retention times indicate the chromatographic peak of that mass that was analyzed for the data in the other columns on the same row of this table. The table breaks indicate the separation between cardiolipin elution clusters.

t_R	Observed *m/z* $[M-H]^-$	Theoretical *m/z* $[M-H]^-$	ppm error	Acyl carbons	Unsaturation / cyclopropyl	Proposed acyl configuration (*sn1+sn2 ++ sn2+sn1*)
12.72/12.86	1265.85663	1265.85540	0.972	58	1	
12.79/12.94	1291.87211	1291.87105	0.821	60	2	14:0+16:1 ++ 16:1+14:0
12.81/12.95	1317.88754	1317.88670	0.637	62	3	14:0+16:1 ++ 16:1+16:1
12.86	1343.90299	1343.90235	0.476	64	4	16:1+16:1 ++ 16:1+16:1
12.72/**12.86**	1265.85673	1265.85540	1.051	58	1	
12.79/**12.94**	1291.87199	1291.87105	0.728	60	2	14:0+14:1 ++ 16:1+16:0
12.81/**12.95**	1317.88755	1317.88670	0.645	62	3	
13.27/13.67	1389.94529	1389.94476	0.385	66	3	66:3(OH)
13.31/13.46	1305.88743	1305.88670	0.559	61	2	
13.33/13.51	1331.90304	1331.90235	0.518	63	3	14:0+16:1 ++ Cy17:0+16:1
13.34/13.74	1363.92958	1363.92856	0.748	64	2	16:0+16:0(OH) ++ 16:1+16:1
13.31/**13.46**	1305.88746	1305.88670	0.582	61	2	
13.33/**13.51**	1331.90317	1331.90235	0.616	63	3	14:0+16:1 ++ Cy17:0+16:1

13.56	1357.91886	1357.91800	0.633	65	4	16:1+16:1 ++ Cy17:0+16:1
13.27/**13.67**	1389.94338	1389.94476	-0.989	66	3	66:3(OH)
13.70	1267.87242	1267.87105	1.081	58	0	
13.34/**13.74**	1363.92846	1363.92856	-0.073	64	2	64:2(OH)
13.74	1371.93222	1371.93365	-1.042	66	4	18:1+16:1 ++ 16:1+16:1
13.74	1293.88691	1293.88670	0.162	60	1	
13.79/13.91	1319.90253	1319.90235	0.136	62	2	14:0+16:1 ++ 16:1+16:0
13.82	1345.91715	1345.91800	-0.632	64	3	16:0+16:1 ++ 16:1+16:1
13.79/**13.91**	1319.90304	1319.90235	0.523	62	2	
14.21/14.46	1359.93445	1359.93365	0.588	65	3	
14.22/14.39	1385.95008	1385.94930	0.563	67	4	
14.24/14.39	1307.90314	1307.90235	0.604	61	1	
14.27/14.42	1333.91873	1333.91800	0.547	63	2	
14.22/**14.39**	1385.94928	1385.94930	-0.014	67	4	16:1+16:1 ++ Cy17:0+18:1
14.24/**14.39**	1307.90327	1307.90235	0.703	61	1	
14.27/**14.42**	1333.91831	1333.91800	0.232	63	2	14:0+16:1 ++ Cy17:0+16:0
14.21/**14.46**	1359.93312	1359.93365	-0.390	65	3	16:0+16:1 ++ Cy17:0+16:1
14.56/15.01	1399.96288	1399.96495	-1.479	68	4	16:1+18:1 ++ 16:1+18:1
14.60/14.71	1347.93310	1347.93365	-0.408	64	2	
14.61	1295.90060	1295.90235	-1.350	60	0	
14.63/15.08	1373.94756	1373.94930	-1.266	66	3	16:0+16:1 ++ 16:1+18:1
14.64	1321.91821	1321.91800	0.159	62	1	16:0+16:1 ++ 14:0+16:0
14.60/**14.71**	1347.93266	1347.93365	-0.734	64	2	16:0+16:1 ++ 16:1+16:0
14.56/**15.01**	1399.96416	1399.96495	-0.564	68	4	
15.02/15.19	1413.98018	1413.98060	-0.297	69	4	
15.05/15.25	1387.96509	1387.96495	0.101	67	3	
14.63/**15.08**	1373.94951	1373.94930	0.153	66	3	16:0+Cy17:0 ++ Cy17:0+16:1
15.09/15.32	1361.94988	1361.94930	0.426	65	2	15:0+16:1 ++ 18:1+16:0
15.10	1309.91773	1309.91800	-0.206	61	0	
15.16/15.28	1335.93445	1335.93365	0.599	63	1	
15.02/**15.19**	1413.98029	1413.98060	-0.219	69	4	16:1+Cy17:0 ++ 18:1+18:1
15.05/**15.25**	1387.96367	1387.96495	-0.922	67	3	16:0+16:1 ++ Cy17:0+18:1
15.16/**15.28**	1335.93455	1335.93365	0.674	63	1	14:0+16:1 ++ Cy17:0+16:0
15.09/**15.32**	1361.94898	1361.94930	-0.235	65	2	16:0+16:1 ++ Cy17:0+16:0

15.36/15.79	1427.99457	1427.99625	-1.176	70	4	
15.42/15.86	1401.97879	1401.98060	-1.291	68	3	16:0+16:1 ++ 18:1+18:1
15.49/15.57	1349.94947	1349.94930	0.126	64	1	
15.51/15.92	1375.96419	1375.96495	-0.552	66	2	16:0+16:1 ++ 18:1+16:0
15.53	1323.93530	1323.93365	1.246	62	0	14:0+16:1++16:0+16:0
15.49/15.57	1349.94946	1349.94930	0.119	64	1	16:0+16:1 ++ 16:0+16:0
15.36/**15.79**	1427.99682	1427.99625	0.399	70	4	
15.86/16.04	1415.99592	1415.99625	-0.233	69	3	
15.42/**15.86**	1401.98092	1401.98060	0.228	68	3	
15.51/**15.92**	1375.96564	1375.96495	0.501	66	2	16:0+Cy17:0 ++ Cy17:0+16:0
15.93/16.15	1363.96529	1363.96495	0.249	65	1	16:0+16:0 ++ 16:0+Cy17:0
15.94/16.11	1389.98082	1389.98060	0.158	67	2	
15.75/**15.97**	1442.01324	1442.01351	-0.184	71	4	
16.01	1337.94898	1337.94930	-0.239	63	0	
15.86/**16.04**	1415.99673	1415.99625	0.339	69	3	
15.94/**16.11**	1389.98086	1389.98060	0.187	67	2	16:0+18:1 ++ Cy17:0+16:0
16.15/16.38	1456.02773	1456.02755	0.124	72	4	
15.93/**16.15**	1363.96564	1363.96495	0.506	65	1	16:0+16:0 ++ Cy17:0+16:0
16.20/16.62	1430.01214	1430.01190	0.168	70	3	
16.27/16.69	1403.99592	1403.99625	-0.235	68	2	16:0+18:1 ++ 18:1+16:0
16.33	1377.98022	1377.98060	-0.276	66	1	16:0+16:0 ++ 18:1+16:0
16.15/16.38	1456.02780	1456.02755	0.172	72	4	
16.41	1351.96506	1351.96495	0.081	64	0	16:0+16:0 ++ 16:0+16:0
16.20/16.62	1430.01347	1430.01190	1.098	70	3	
16.69/16.87	1418.01342	1418.01190	1.072	69	2	
16.27/16.69	1403.99662	1403.99625	0.264	68	2	
16.80	1444.02895	1444.02958	-0.433	71	3	
16.69/16.87	1418.01356	1418.01190	1.171	69	2	16:0+18:1 ++ Cy19:0+16:0
17.05/17.30	1432.02900	1432.02755	1.013	70	2	
17.05/17.30	1432.03003	1432.02755	1.732	70	2	

Supplemental table 5: A list of masses found in the *E. coli* DH5α lipidome. With the other columns showing the retention time (full gradient method), corresponding head group and lipid core. The highest peak is marked with the retention time in bold. Only the isobaric lipids previously identified in *E. coli* are explicitly listed.

t_R	Head-group	Lipid core	Observed *m/z* $[M-H]^-$	Theoretical *m/z* $[M-H]^-$	ppm error	reference
12.31	PG	24:0	609.37578	609.37731	-2.511	
13.44/13.96	PE	24:1	576.36552	576.36708	-2.707	
13.47/**13.87**	PG	25:0	623.39126	623.39296	-2.727	
14.70/**15.00**	PG	27:1	649.40710	649.40861	-2.325	
15.04/**15.39**	PE	25:1	590.38117	590.38273	-2.642	
15.2	PG	26:0	637.40695	637.40861	-2.604	
15.60/15.86/**16.25**/16.97	PG	28:1	663.42282	663.42426	-2.171	
15.82/16.11	PE	24:0	578.38122	578.38273	-2.611	
16.01/**16.47**	PG	27:0	651.42287	651.42426	-2.134	
16.21/16.89/17.33/17.63	PG	30:2	689.43816	689.43991	-2.538	
16.32/**16.51**/16.73/						
16.99/17.15/17.87	PE	26:1	604.39697	604.39838	-2.333	
16.80/**17.07**/17.39	PG	29:1	677.43818	677.43991	-2.554	
17.15/**17.60**	PG	28:0	665.43843	665.43991	-2.224	
17.15/18.04	PE	28:2	630.41196	630.41403	-3.284	
17.26/17.51/**17.66**/18.00/18.67	PG	31:2	703.45441	703.45556	-1.635	
17.27	PE	25:0	592.39694	592.39838	-2.431	
17.60/17.83/17.99/17.10/						
18.30/18.88	PE	27:1	618.41234	618.41403	-2.733	
17.93/**18.16**/19.11	PG	14:0/16:1	691.45361	691.45556	-2.820	D. Oursel, 2007
18.20/18.68/18.58/19.01/19.42	PG	32:2	717.46935	717.47121	-2.592	
18.26/**18.68**	PG	29:0	679.45382	679.45556	-2.561	
18.55/**18.76**/19.30	PE	29:2	644.42802	644.42968	-2.576	
18.57	PE	26:0	606.41268	606.41403	-2.226	
18.79/18.98/**19.19**/20.08	PG	14:0/Cy17:0	705.46943	705.47121	-2.523	D. Oursel, 2007
		15:0/16:1	isobar			D. Oursel, 2007
18.84/**19.07**/19.50/20.02	PE	28:1	632.42823	632.42968	-2.293	
19.14/**19.46**/19.83/20.00/20.60	PG	33:2	731.48503	731.48686	-2.502	
19.30/**19.66**	PG	16:0/14:0	693.46946	693.47121	-2.524	D. Oursel, 2007
19.30/**19.70**	PE	27:0	620.42817	620.42968	-2.434	
19.35/19.57/20.44/20.68	PE	30:2	658.44367	658.44533	-2.521	
19.94/20.18	PG	14:0/18:1	719.48526	719.48686	-2.224	D. Oursel, 2007
		15:0/Cy17:0	isobar			D. Oursel, 2007
		16:0/16:1	isobar			D. Oursel, 2007
20.04/20.33/**20.66**/21.16	PG	18:1/16:1	745.50045	745.50251	-2.763	D. Oursel, 2007
20.24/**20.62**	PG	31:0	707.48510	707.48686	-2.488	
20.27	PE	29:1	646.44392	646.44533	-2.181	
20.39/**20.78**	PE	14:0/14:0	634.44383	634.44533	-2.364	D. Oursel, 2007
		16:0/12:0	isobar			D. Oursel, 2007
20.42/**20.75**	PE	31:2	672.45940	672.46098	-2.350	
20.48/20.75/20.85/**21.04**	PG	16:0/Cy17:0	733.50065	733.50251	-2.536	D. Oursel, 2007
20.85/**21.18**/21.53/22.33	PG	35:2	759.51597	759.51816	-2.883	
20.92/21.7/**21.32**/22.21	PE	14:0/16:1	660.45974	660.46098	-1.877	D. Oursel, 2007
21.29/21.76/22.10/22.47	PE	32:2	686.47474	686.47663	-2.753	
21.53	PG	16:0/16:0	721.50105	721.50251	-2.024	D. Oursel, 2007
21.42/**21.83**	PE	29:0	648.45948	648.46098	-2.313	
21.63/21.99	PG	16:0/18:1	747.51645	747.51816	-2.288	D. Oursel, 2007
21.71/22.04/**22.39**	PG	18:1/18:1	773.53164	773.53381	-2.805	D. Oursel, 2007
22.01/**22.42**	PG	33:0	735.51634	735.51816	-2.474	
22.07/**22.33**	PE	14:0/Cy17:0	674.47496	674.47663	-2.476	D. Oursel, 2007
		15:0/16:1	isobar			D. Oursel, 2007
22.18/**22.53**/22.67/22.91/23.64	PE	33:2	700.49058	700.49228	-2.427	
22.45/**22.81**	PE	30:0	662.47512	662.47663	-2.279	
22.50/**22.77**	PG	16:0/Cy19:0	761.53167	761.53381	-2.810	D. Oursel, 2007
22.86/23.19	PG	37:2	787.54712	787.54946	-2.971	
23.05/23.33	PE	15:0/Cy17:0	688.49060	688.49228	-2.440	D. Oursel, 2007
		16:0/C16:1	isobar			D. Oursel, 2007
23.10/23.38/**23.75**	PE	18:1/16:1	714.50601	714.50793	-2.687	D. Oursel, 2007
		Cy17:0/				
		Cy17:0	isobar			D. Oursel, 2007
23.27	PG	34:0	749.53155	749.53381	-3.015	
23.33/**23.64**	PG	36:1	775.54756	775.54946	-2.450	
23.37/**23.75**	PE	31:0	676.49073	676.49228	-2.291	
23.74/**24.13**	PG	35:0	763.54771	763.54946	-2.292	
23.92/**24.43**/24.69/25.53	PE	35:2	728.52145	728.52358	-2.924	
23.94/**24.36**	PE	16:0/Cy17:0	702.50629	702.50793	-2.334	D. Oursel, 2007
24	PG	38:2	801.56324	801.56511	-2.333	

24.47	PG	37:1	789.56300	789.56511	-2.672	
24.73	PE	16:0/16:0	690.50652	690.50793	-2.042	D. Oursel, 2007
24.82/25.20	PE	16:0/18:1	716.52193	716.52358	-2.303	D. Oursel, 2007
		Cy19:0/				
24.82/25.20/**25.61**	PE	Cy17:0	742.53758	742.53923	-2.222	D. Oursel, 2007
		18:1/18:1	isobar			D. Oursel, 2007
24.96	PG	36:0	777.56333	777.56511	-2.289	
25.24/**25.88**	PE	33:0	704.52225	704.52358	-1.888	
25.60/**26.10**/26.47	PE	37:2	756.55235	756.55488	-3.344	
25.70/**26.10**	PE	16:0/Cy19:0	730.53765	730.53923	-2.163	D. Oursel, 2007
		18:0/Cy17:0	isobar			D. Oursel, 2007
		17:0/18:1	isobar			D. Oursel, 2007
26.57	PE	34:0	718.53775	718.53923	-2.060	
26.58/**26.91**	PE	36:1	744.55324	744.55488	-2.203	
26.96/**27.39**	PE	35:0	732.55390	732.55488	-1.338	
27.25	PE	38:2	770.56875	770.57053	-2.310	
27.25/**27.60/**28.02	PE	39:2	784.58441	784.58618	-2.256	
27.27/**28.10**/28.44	PE	38:0	772.58468	772.58618	-1.942	
27.72	PE	37:1	758.56866	758.57053	-2.465	
28.17	PE	36:0	746.56906	746.57053	-1.969	
30.04	CL	58:1	1265.85313	1265.85540	-1.793	
30.34/**30.54**	CL	59:1	1279.86745	1279.87105	-2.813	
30.55/**30.69**/30.86	CL	61:2	1305.88299	1305.88670	-2.841	
30.79	CL	58:0	1267.86742	1267.87105	-2.863	
30.83/**31.00**	CL	60:1	1293.88310	1293.88670	-2.782	
30.88/**31.00**/31.18	CL	62:2	1319.89915	1319.90235	-2.424	
30.89/31.38	CL	64:3	1345.91486	1345.91800	-2.333	
31.22	CL	59:0	1281.88214	1281.88670	-3.557	
31.30/**31.46**	CL	61:1	1307.89917	1307.90235	-2.431	
31.35/31.49/**31.62**	CL	63:2	1333.91420	1333.91800	-2.849	
31.29/**31.52**	CL	65:3	1359.92963	1359.93365	-2.956	
31.71	CL	60:0	1295.89838	1295.90235	-3.064	
31.72/32.05	CL	66:3	1373.94539	1373.94930	-2.846	
31.75/**31.90**	CL	62:1	1321.91456	1321.91800	-2.602	
31.80/32.08	CL	64:2	1347.92962	1347.93365	-2.986	
32.15/32.37/**32.51**	CL	65:2	1361.94527	1361.94930	-2.959	
32.18/**32.34**	CL	63:1	1335.92966	1335.93365	-2.987	
32.41	CL	70:4	1427.99428	1427.99625	-1.380	
32.55/32.71/**32.93**	CL	66:2	1375.96033	1375.96495	-3.358	
32.59/**32.74**	CL	64:1	1349.94532	1349.94930	-2.948	
32.49/32.61/**32.88**/33.10	CL	68:3	1401.97603	1401.98060	-3.260	
32.92/**33.16**	CL	65:1	1363.96060	1363.96495	-3.189	
32.93/33.10/**33.31**/33.86	CL	67:2	1389.97603	1389.98060	-3.288	
33.32/**33.55**/34.09	CL	66:1	1377.97626	1377.98060	-3.150	
33.93	CL	67:1	1391.98978	1391.99625	-4.648	
32.30/32.45/**32.68**	CL	67:3	1387.96080	1387.96495	-2.990	
33.04/33.23/**33.46**	CL	69:3	1415.99211	1415.99625	-2.924	
33.22/**33.61**/33.81	CL	70:3	1430.00824	1430.01190	-2.559	
33.28/**33.69**	CL	68:2	1403.99195	1403.99625	-3.063	
33.87/**34.05**/34.54	CL	69:2	1418.00881	1418.01190	-2.179	
34.18	CL	71:3	1444.02458	1444.02755	-2.057	
34.32	CL	72:3	1458.04091	1458.04320	-1.571	
34.38	CL	70:2	1432.02459	1432.02755	-2.067	
34.76	CL	68:1	1406.01037	1406.01190	-1.088	
34.73/35.18	CL	71:2	1446.03967	1446.04320	-2.441	
34.83	CL	73:3	1472.05655	1472.05885	-1.562	

Supplemental table 6: A list of the lipid masses found in the *S. acidocaldarius* MW001 lipidome. Listed are the corresponding retention time (Narrow gradient method), ion type, ether lipid core, headgroup and ion abundance. * These lipids have their compound name listed instead of lipid headgroup.

t_R	Headgroup	Lipid core	Observed *m/z*	Theoretical *m/z*	ppm error	Ion type	Ion abundance (10 ppm)
22.46	none	GDGT-0	1346.31030	1346.31359	-2.447	$[M+CHO_2]^-$	6.95E+04
22.49	none	GDGT-1	1344.29290	1344.29794	-3.752	$[M+CHO_2]^-$	1.48E+05
22.46	none	GDGT-2	1342.27893	1342.28229	-2.506	$[M+CHO_2]^-$	4.13E+05
22.48	none	GDGT-3	1340.25978	1340.26664	-5.121	$[M+CHO_2]^-$	1.08E+06
22.46	none	GDGT-4	1338.24616	1338.25099	-3.612	$[M+CHO_2]^-$	2.83E+06
22.39	none	GDGT-5	1336.23335	1336.23534	-1.492	$[M+CHO_2]^-$	4.65E+06
22.28	none	GDGT-6	1334.21978	1334.21969	0.065	$[M+CHO_2]^-$	2.43E+06
22.16	none	GDGT-7	1332.20469	1332.20404	0.485	$[M+CHO_2]^-$	5.39E+03
21	MH	GDGT-0	1508.35894	1508.36642	-4.957	$[M+CHO_2]^-$	3.09E+03
21.03	MH	GDGT-1	1506.34480	1506.35077	-3.962	$[M+CHO_2]^-$	1.26E+04
21.03	MH	GDGT-2	1504.33164	1504.33512	-2.312	$[M+CHO_2]^-$	2.42E+04
21	MH	GDGT-3	1502.31349	1502.31947	-3.979	$[M+CHO_2]^-$	8.27E+04
21	MH	GDGT-4	1500.30421	1500.30382	0.262	$[M+CHO_2]^-$	2.00E+05
20.91	MH	GDGT-5	1498.28856	1498.28817	0.262	$[M+CHO_2]^-$	1.93E+05
20.79	MH	GDGT-6	1496.27372	1496.27252	0.804	$[M+CHO_2]^-$	7.03E+04
20.62	DH	GDGT-0	1670.41020	1670.41924	-5.412	$[M+CHO_2]^-$	9.15E+02
20.64	DH	GDGT-1	1668.39787	1668.40359	-3.429	$[M+CHO_2]^-$	2.97E+05
20.65	DH	GDGT-2	1666.38142	1666.38794	-3.913	$[M+CHO_2]^-$	7.59E+05
20.64	DH	GDGT-3	1664.36588	1664.37229	-3.852	$[M+CHO_2]^-$	3.02E+06
20.61	DH	GDGT-4	1662.35505	1662.35664	-0.957	$[M+CHO_2]^-$	6.94E+06
20.52	DH	GDGT-5	1660.34072	1660.34099	-0.163	$[M+CHO_2]^-$	6.97E+06
20.38	DH	GDGT-6	1658.32602	1658.32534	0.410	$[M+CHO_2]^-$	2.25E+06
20.38	TH	GDGT-1	1830.44226	1830.45641	-7.733	$[M+CHO_2]^-$	3.46E+02
20.35	TH	GDGT-2	1828.42570	1828.44076	-8.239	$[M+CHO_2]^-$	7.29E+02
20.3	TH	GDGT-3	1826.41717	1826.42511	-4.350	$[M+CHO_2]^-$	2.02E+04
20.23	TH	GDGT-4	1824.40476	1824.40946	-2.578	$[M+CHO_2]^-$	5.13E+04
20.21	TH	GDGT-5	1822.39400	1822.39381	0.102	$[M+CHO_2]^-$	4.78E+04
20.05	TH	GDGT-6	1820.37970	1820.37816	0.844	$[M+CHO_2]^-$	7.61E+03
18.57	IP	GDGT-0	1542.32152	1542.32727	-3.729	$[M-H]^-$	2.15E+06
18.57	IP	GDGT-1	1540.30642	1540.31162	-3.377	$[M-H]^-$	4.27E+06
18.52	IP	GDGT-2	1538.29132	1538.29597	-3.023	$[M-H]^-$	6.61E+06
18.49	IP	GDGT-3	1536.27622	1536.28032	-2.669	$[M-H]^-$	1.06E+07
18.47	IP	GDGT-4	1534.26439	1534.26467	-0.183	$[M-H]^-$	1.39E+07
18.33	IP	GDGT-5	1532.24907	1532.24902	0.032	$[M-H]^-$	6.98E+06
18.16	IP	GDGT-6	1530.23405	1530.23337	0.444	$[M-H]^-$	4.67E+05
17.39	MH - PH/IP	GDGT-0	1704.37225	1704.38009	-4.602	$[M-H]^-$	2.57E+03
17.35	MH - PH/IP	GDGT-1	1702.36075	1702.36444	-2.170	$[M-H]^-$	8.95E+03
17.35	MH - PH/IP	GDGT-2	1700.34529	1700.34879	-2.061	$[M-H]^-$	1.70E+04
17.26	MH - PH/IP	GDGT-3	1698.33025	1698.33314	-1.704	$[M-H]^-$	2.91E+04
17.25	MH - PH/IP	GDGT-4	1696.31864	1696.31749	0.676	$[M-H]^-$	2.86E+04
17.1	MH - PH/IP	GDGT-5	1694.30327	1694.30184	0.842	$[M-H]^-$	5.90E+03
17.39	DH - PH/IP	GDGT-0	1866.42165	1866.43292	-6.037	$[M-H]^-$	4.36E+05
17.35	DH - PH/IP	GDGT-1	1864.40887	1864.41727	-4.504	$[M-H]^-$	1.71E+06
17.32	DH - PH/IP	GDGT-2	1862.39481	1862.40162	-3.655	$[M-H]^-$	3.71E+06
17.27	DH - PH/IP	GDGT-3	1860.38003	1860.38597	-3.191	$[M-H]^-$	5.54E+06
17.24	DH - PH/IP	GDGT-4	1858.36884	1858.37032	-0.795	$[M-H]^-$	4.78E+06
17.09	DH - PH/IP	GDGT-5	1856.35470	1856.35467	0.018	$[M-H]^-$	1.03E+06
16.97	DH - PH/IP	GDGT-6	1854.33930	1854.33902	0.153	$[M-H]^-$	3.47E+04
10.74	AA*	DGD (20:0)	731.63391	731.63240	2.064	$[M-H]^-$	9.16E+06
9.29	IP	DGD (20:0)	893.68671	893.68522	1.667	$[M-H]^-$	7.65E+06
7.11	CDP	DGD (20:0)	1036.67517	1036.67369	1.428	$[M-H]^-$	7.50E+04
4.23	DGGGP*	DGD (20:8)	715.50826	715.50720	1.481	$[M-H]^-$	4.31E+05

Supplemental methods

1-BuOH solvent distillation

A 2 L round bottom flask was charged with roughly 1 L of 1-Butanol (at first cat#: 34867, Fisher Scientific and later cat#: 901351, Sigma-Aldrich; as this was found to be more suitable.). The half-filled flask was subsequently connected to a cleaned and dried rotary evaporator, the water bath was set to 55°C and the butanol was evaporated at 10 mbar. When all the butanol was evaporated more was added until the round bottom flask was half full, typically a destination was performed in 2.5 L batches. After the distillation was complete the film of residue in the round bottom flask was discarded.

Eluting methoxy mycolic acid using two in-line C18 CSH columns

Methoxy mycolic acid was eluting using two in-line C18 CSH columns as shown in supplemental figure 8 using the following gradient adapted from the "narrow gradient method". With a flow rate of 150 µl/min the following method was applied: An isocratic focusing step of 45% B for 5 minutes, a gradient from 45% to 100% B over 48 minutes (1.146% B / minute), a purge phase of 100% B for 5 minutes, followed by a transition of 1 minute to re-equilibration conditions of 45% B for 16.5 minutes.

Comparing different mobile phases for elution of methoxy mycolic acid and isocaldarchaeol

The elution of methoxy mycolic acid and isocaldarchaeol using different mobile phase compositions as shown in supplemental figure 9 was performed using the "full gradient method" as described in the main text and a method utilizing a higher maximum percentage of mobile phase B (98%) to facilitate elution with the mobile phase compositions that have less eluotropic power. This adapted method employs a flow rate of 300 µl/min and is described as follows: An isocratic focusing step of 5% B for 2.5 minutes, a gradient from 5% to 98% B over 40 minutes (2.45% B / minute), a purge phase of 98% B for 2.5 minutes, followed by a transition of 2 minutes to re-equilibration conditions of 5% B for 8 minutes.

Determination of chromatographic parameters

The determination of important chromatographic performance parameters was carried out by injecting lipid standards and the *E. coli* CL extract on the premier column. For the determination of the resolution and column efficiency per mobile phase, a lipid standard (10 µl of 1 µg/ml of each lipid in 1:1 MeOH:MQ) and *E. coli* CL extract (5 µl of 0.25 mg/ml in MeOH) were injected on the Waters C18 CSH Premier column and analyzed using the full-gradient method as described in the materials and methods of the main text. For the estimation of the limit of detection (LoD) and limit of quantification (LoQ), serial dilutions of the lipid standard (0.2 to 200 pg of each lipid in 1:1 MeOH:MQ) were injected instead to construct a standard curve.

The resolution and column efficiency was calculated using the retention time and FWHM of the peaks: $R = 1.18((t_{R2}-t_{R1})/(W_{0.5h2}+ W_{0.5h1}))$ and $N = 5.54((t_R/W_{0.5h})^2)$ respectively. The LoD and LoQ were estimated using linear regression based on the slope of the standard curve and the standard deviation of the response, following the ICH definition, using a factor of x3.3 and x10 for the LoD and LoQ respectively [348].

Chapter 3

Structural and Functional Insights into an Archaeal Lipid Synthase

Adapted from Ren et al.:
Cell reports 2020
volume 33, issue 3, 108294
https://doi.org/10.1016/j.celrep.2020.108294

Niels A.W. de Kok [1,6], Sixue Ren [2,6], Yijun Gu [3,6], Weizhu Yan [2], Qiu Sun [2], Yunying Chen [2], Jun He [2], Lejin Tian [2], Ruben L.H. Andringa [4], Xiaofeng Zhu [2], Mei Tang [2], Shiqian Qi [2], Heng Xu [2], Haiyan Ren [2], Xianghui Fu [2], Adriaan J. Minnaard [4], Shengyong Yang [2], Wanjiang Zhang [5], Weimin Li [2], Yuquan Wei [2], Arnold J.M. Driessen [1,*], and Wei Cheng [2,7,*]

[1]Department of Molecular Microbiology, Groningen Biomolecular Sciences and Biotechnology Institute, and The Zernike Institute for Advanced Materials, University of Groningen, Nijenborgh 7, 9747 AG Groningen, the Netherlands

[2]Division of Respiratory and Critical Care Medicine, Respiratory Infection and Intervention Laboratory of Frontiers Science Center for Disease-Related Molecular Network, State Key Laboratory of Biotherapy, West China Hospital of Sichuan University, Chengdu 610041, China

[3]National Facility for Protein Science Shanghai, Shanghai Synchrotron Radiation Facility, Shanghai Advanced Research Institute (Zhangjiang Lab), Zhangheng Road 239, Shanghai 201203, China

[4]Stratingh Institute for Chemistry, University of Groningen, Nijenborgh 7, 9747 AG Groningen, the Netherlands

[5]Department of Pathophysiology, Shihezi University School of Medicine, the Key Laboratory of Xinjiang Endemic and Ethnic Diseases, Shihezi, Xinjiang 832002, China

[6]These authors contributed equally

[7]Lead Contact

*Corresponding author e-mail: *a.j.m.driessen@rug.nl*, chengwei669@scu.edu.cn (W.C.)

Abstract

The UbiA superfamily of intramembrane prenyltransferases catalyzes an isoprenyl transfer reaction in the biosynthesis of lipophilic compounds involved in cellular physiological processes. Di-geranylgeranylglycerol phosphate (DGGGP) synthase (DGGGPS) generates unique membrane core lipids for the formation of the ether bond between the glycerol moiety and the alkyl chains in Archaea and has been confirmed to be a member of the UbiA superfamily. Here, the crystal structure is reported to exhibit nine transmembrane helices along with a large lateral opening covered by a cytosolic cap domain and a unique substrate-binding central cavity. Notably, the lipid-bound states of this enzyme demonstrate that the putative substrate-binding pocket is occupied by the lipidic molecules used for crystallization, indicating the binding mode of hydrophobic substrates. Collectively, these structural and functional studies provide an understanding of lipid biosynthesis by substrate-specific lipid-modifying enzymes.

Introduction

Lipids are highly chemically diverse and are essential for all organisms; they are involved in cell membrane formation [349], energy storage, energy conversion [350], signal transduction [351], protein recruitment [352], and posttranslational modifications [353]. The chemical diversity of lipids allows them to perform these specific functions [349,354,355]. Notably, the compositional and structural diversity of lipids in membranes influences the properties of the cell membrane [349,356]. Membrane lipids mainly consist of glycerophospholipids with chemically diversified structures. For instance, Archaea can be distinguished from Bacteria and Eukarya because in Archaea, a glycerol-1-phosphate (G1P) head group is linked with isoprenoid side chains via an ether bond. In contrast, a glycerol-3-phosphate (G3P) headgroup is linked with fatty acids via an ester bond in Bacteria and Eukarya [354]. Understanding lipid biosynthesis can facilitate the elucidation of the evolution of these molecules and the mechanisms of the evolution of diverse lipid structures [357–360] and the requirement for substrate-specific lipid recognition by these unique enzymes.

Although major insights into lipid biosynthesis have been obtained [101,273,359,361–363], several aspects of the glycerophospholipid biosynthesis pathway in Archaea remain elusive. One reason for this may be that several key steps are mediated by membrane-embedded enzymes [362]. For example, the biosynthesis pathways for 2,3-di-O-geranylgeranylglyceryl phosphate (DGGGP) and 1-CDP-2,3-di-O-geranylgeranylglyceryl phosphate (CDP-archaeol) are catalyzed by DGGGP synthase (DGGGPS) and CDP-archaeol synthase (CarS), respectively. The structure and function of *Aeropyrum pernix* CarS (ApCarS), the downstream enzyme of DGGGPS in this biosynthesis pathway, have recently been elucidated [103]. However, the identity of DGGGPS is still elusive.

The archaeal DGGGPS catalyzes the transfer of a geranylgeranyl group from geranylgeranyl diphosphate (GGPP) to (S)-3-O-geranylgeranylglycerol phosphate (GGGP) (Figure 1A) in the biosynthesis of archaeal membrane lipids [362]. DGGGPS family members contain two characteristic conserved Asp-rich motifs with the consensus sequences DXXDXXXD and DXXXD (D, Asp; X, any amino acid) as well as a conserved YXXXK motif (Y, Tyr; K, Lys; X, any amino acid). Similar features are found in the archetypal UbiA of the UbiA superfamily (Figure 1B), including representative members (e.g., COQ2, UbiA, UBIAD1, and MenA) that catalyze isoprene transfer reactions to synthesize lipophilic compounds associated with diverse functions in organism physiology. Due to sequence similarity and similarity in the reactions catalyzed, DGGGPS is considered to belong to the UbiA superfamily

[94,100] (Figures 1C and S1A). DGGGPS proteins, which can be distinguished from MenA, UbiA, and other subfamilies, are a deeply rooted family from the phylum Archaea (Figure S1). However, the molecular mechanisms of these enzymes remain to be elucidated.

Understanding the function of intramembrane enzymes that recognize hydrophobic substrates is hindered by technical difficulties associated with structural studies. Compared to other transmembrane (TM) enzymes, members of the DGGGPS family are unique in that they are polytopic TM enzymes that catalyze reactions related to two lipid substrates. Thus, the catalytic mechanisms of this unique family of enzymes are still elusive. Here, we functionally characterized the DGGGPS enzyme from *Methanocaldococcus jannaschii* (MjDGGGPS), which is essential for geranylgeranyl group transfer during the biosynthesis of the DGGGP membrane lipid, and we report the crystal structures of MjDGGGPS in apo- and lipid-bound states at resolutions of 3.3 Å and 2.3 Å, respectively. These structures revealed that MjDGGGPS contains nine helices that form a unique cavity containing two lipidic substrate-binding pockets and a wide lateral opening. Remarkably, the substrate-binding pockets of MjDGGGPS in the crystal structure are likely occupied by the lipophilic molecules monoolein and LDAO. Variants based on structural models coupled with functional assays of eukaryotic homologs further verified the importance of the conserved residues. These studies not only provided structural insights into the mechanisms underlying lipid substrate recognition and catalysis by DGGGPS, but also furthered our understanding of the general mechanisms of the UbiA superfamily. Importantly, the identification of substrate-specific lipid modifying enzymes is a key problem in understanding lipid biosynthesis. Notably, this study shed light on the catalytic mechanism of the involvement of the two lipophilic substrates of DGGGPS at the membrane-water interface.

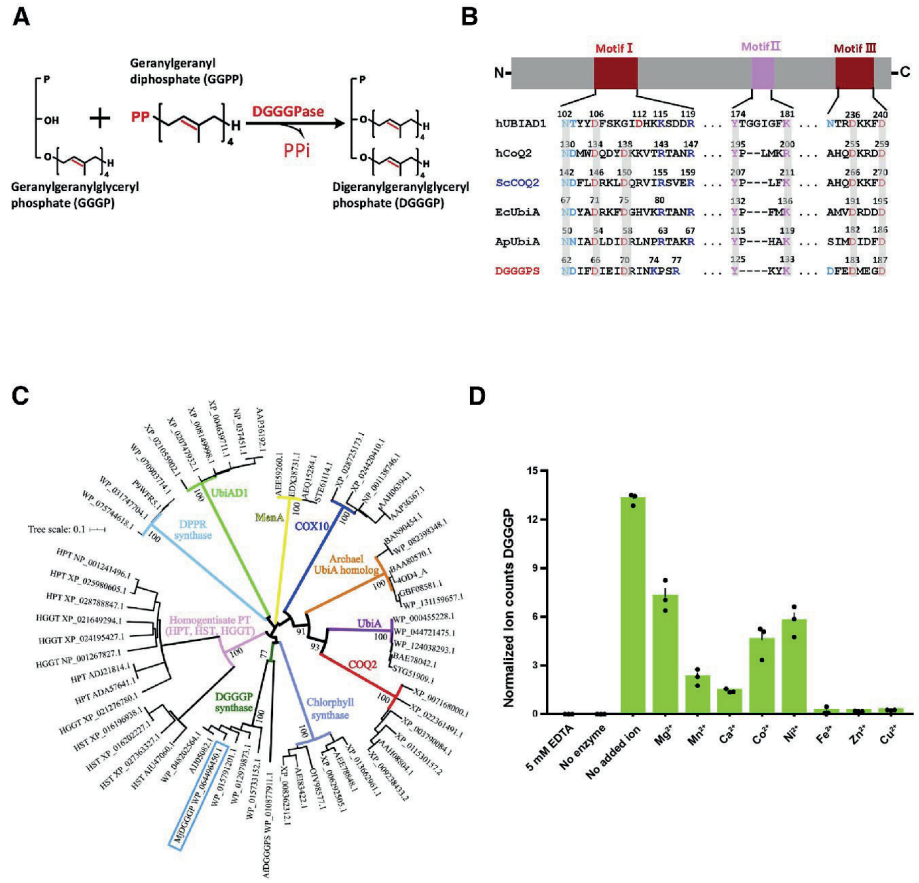


Figure 1: Bioinformatics and functional characterization of MjDGGGPS. (A) Scheme of DGGGPS catalysis. **(B**) Alignment of three conserved sequence motifs. The key residues of two Asp-rich motifs of the DGGGP homologs are highlighted in red, the Y-K motif is indicated in pink, and other residues are colored in blue. **(C)** Maximum likelihood tree based on the protein sequences of the UbiA superfamily prenyltransferases. MjDGGGPS is highlighted in the cyan square. First, representative protein sequences from different subfamilies of the UbiA superfamily were aligned using multiple sequence comparison by log-expectation (MUSCLE). Then, the maximum likelihood phylogenetic tree was computed by MEGA X. The phylogenetic test was calculated with the bootstrap method, and the number of replications was set as 1000. As shown in the figure, high bootstrap values of every subfamily clade mean they have a very high confidence level. Different colors indicate different UbiA subfamilies. Finally, the tree was edited in iTOL (Figure S1C). **(D)** Graphic depiction of DGGGPS activity. The chemically synthesized GGPP and GGGP (STAR Methods) were used as substrates for purified MjDGGGPS activity assays in presence of different divalent metal ions. The production of DGGGP was verified by UHPLC-MS analysis (Figures S2A and S2B; STAR Methods).

Results

Bioinformatics analysis and function characterization of DGGGP Synthase

Using BLAST, we identified a candidate DGGGP synthase in *Methanocaldococcus jannaschii* (MjDGGGPS) that mediates DGGGP biosynthesis. A genetic screen in the crenarchaeote *Sulfolobus islandicus* indicated that the homologous M164_1546 gene is essential [364]. Phylogenetic analysis of representative members of the UbiA superfamily showed considerable sequence similarity (Figure S1) and features highly conserved motifs (Figure 1B). Crystal structures have been reported for members of the UbiA superfamily, i.e., *Aeropyrum pernix* UbiA (ApUbiA, PDB: 4OD4) [100] and *Archaeoglobus fulgidus* UbiA (AfUbiA, PDB: 4TQ5) [96]. However, no enzymatic activity has been demonstrated for enzymes of which a crystal structure was reported. This led us to investigate UbiA homologs with an emphasis on thermophiles as the structures of these enzymes could likely be determined.

MjDGGGPS exhibited activity with two hydrophobic substrates, namely GGPP and GGGP, forming DGGGP. To verify that the thermophilic enzyme was active, a biochemical assay was performed. Purified MjDGGGPS was incubated with synthetically prepared GGPP and GGGP [365] at 80°C for 30 min, and the production of DGGGP was monitored by ultra-high-performance liquid chromatography-mass spectrometry (UHPLC-MS) (Figures 1D, 4B, S2A, and S2B) [366]. As Mg^{2+} is preferred by UbiA members for optimal enzymatic activity [367–369], we examined the enzymatic activity of MjDGGGPS in the presence and absence of Mg^{2+}. We also tested whether EDTA could impair the activity. As anticipated, the enzymatic activity of MjDGGGPS was completely suppressed by EDTA (Figure 1D, S2C), which was consistent with observations of other members of the UbiA superfamily [362,367]. However, the activity of MjDGGGPS was altered to different extents by the addition of Mg^{2+}, Mn^{2+}, Ca^{2+} or Co^{2+} but not Fe^{2+}, Zn^{2+}, or Cu^{2+}. Mg^{2+} was the most favorable metal for enzymatic activity (Figure 1D) [93]. Interestingly, the divalent metal-dependent activity of MjDGGGPS appears broader than that of LePGT1, another UbiA member [367]. Suggesting that divergences occurred during the evolution of this species. Overall, our data indicated that DGGGPS is a unique active UbiA homolog.

Structure of DGGGPS

The purified DGGGPS protein was crystallized using the lipidic cubic phase (LCP), and the crystals suitable for X-ray diffraction were obtained in the presence of the n-dodecyl-β-D-maltoside (DDM) detergent. As shown in Figure 2A, the determined apo-structure of MjDGGGPS (Figure S3; Table S1) contains nine TM helices that form a cupped transmembrane structure with a large central cavity (Figure 2A; Table S1). The cavity capped by a cytoplasmic domain comprises three loop/helix regions, namely CHL23, CHL45, and CHL67, linking TM2 and TM3, TM4 and TM5, and TM6 and TM7, respectively (Figures 2A and 2B). CHL23, CHL45, and CHL67 each contain an Asp-rich motif, $D_{66}XXXD_{70}$,Y_{125}–K_{128}, and $D_{183}XXXD_{187}XXD_{190}$ [356,359,367,370]. All these residues are highly conserved and located in the cytoplasmic domain (Figures 1B and 2A), suggesting their important role in catalysis of the enzyme. Two hydrophobic tunnels adjacent to the central cavity (one formed by TM2, TM4, and TM5 and the other formed by TM6 and TM9) are buried in the membrane (Figure 2B and 3B). The tunnels have the potential to accommodate the geranylgeranyl chains of GGPP and GGGP substrates given their hydrophobicity (Figure 2D). Notably, a laterally V-shaped opening toward the membrane bilayer of the central cavity formed by TM1 and TM9 is almost completely buried in the membrane (Figure 2C). Polar regions adjacent to the lateral opening suggest that they likely participate in the coordination of the phosphate head group of the substrate or of the product (Figure 2C). Notably, this lateral opening feature is also typically observed in soluble prenyltransferases [95,371,372], which catalyze a similar lipid transfer reaction involving isoprenyl chains (Figure S3A). This suggests that the observed lateral opening could be an inlet for substrate binding and an outlet for substrate release.

Active site of DGGGPS

Strikingly, the active site of the MjDGGGPS structure possesses two long hydrophobic tunnels (named tunnel 1 and tunnel 2; Figure 2B, 2D and 3B) with their openings facing the center of the central cavity (Figures 2C and 2D). Compared to the lateral opening observed in the structures of ApUbiA [100] and AfUbiA [96] (Figure S4), the lateral opening of MjDGGGPS is longer, extending from approximately the center of the lipid bilayer toward the extracellular side of the membrane plane. Interestingly, within tunnel 1 and tunnel 2, which open toward the center of the central cavity, two unoccupied elongated regions of electron density were observed (Figure 3A) in the lipid bound structure. Polar residues surrounding the two regions of electron density include Asn62, Asp66, Asp70, and Lys74 from the opening of tunnel 1 as well as Lys175, Glu176, Asp180, and Asp183 from the opening of tunnel 2 (Figure 3B). In addition, hydrophobic residues (Leu122 from TM4 and Leu143 and Val147 from TM5) surround the density in tunnel 1, and other hydrophobic residues (Leu164 and Ile172 from TM6 and Val272 and Leu273 from TM5) surround the density in tunnel 2 (Figures 3B and S5).

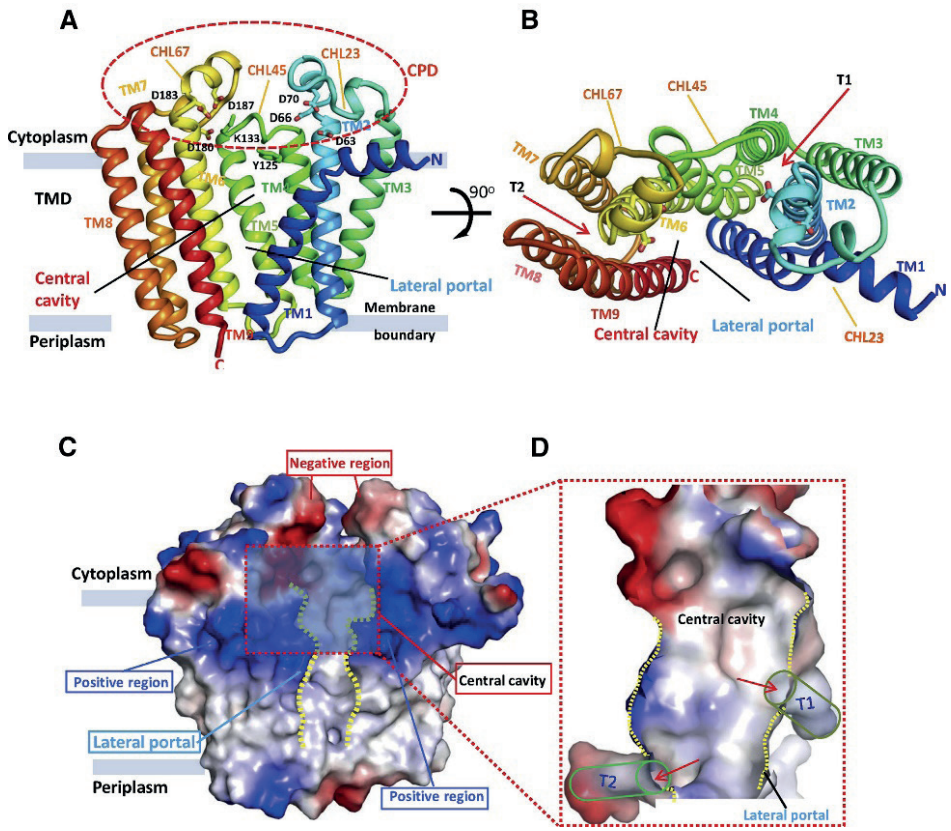


Figure 2: Overall apo-structure of MjDGGGPS. (A) Cartoon of the apo-MjDGGGPS structure viewed from the membrane plane. The cytoplasmic domain is highlighted by a red circle, and the TMs are indicated in different blue colors. Conserved residues in the central cavity are shown as sticks. **(B)** A top view from the cytoplasmic side of the same structure. Tunnels formed by transmembrane helices are indicated by red arrows. **(C)** Electrostatic surface representation of MjDGGGPS. The large central cavity (highlighted in light blue and by a red square) contains two hydrophobic pockets (gray, **D**) for binding to the lipophilic tails of substrates and a unique lateral opening for production release. **(D)** A zoomed view of the central cavity with cavities and pockets culled. TMD refers to the transmembrane domain, while CPD refers to the cytoplasmic domain. T1 and T2, indicated by red arrows refer to tunnel 1 and tunnel 2, respectively. The lateral portal is highlighted in a yellow dashed line.

These residues form two highly hydrophobic tunnels within the membranes and contribute to the polar regions in the cavity that can accommodate the lipophilic tails and phosphate groups of substrates (Figure 3B), respectively. The molecules of n-dodecyl-N,N-dimethylamine-N-oxide (LDAO; used for protein purification) and monoolein (used for crystallization) can fit well into the two regions of electron density (monoolein molecule 1, M1; LDAO molecule, L1) (Figures 3A and 3B). However, we cannot rule out the possibility these elongated regions of electron density were generated by other hydrophobic molecules, such as copurified endogenous lipids. However, the hydrophobic environment and the structural features of these pockets clearly indicate that they function as lipid-binding sites. We speculate that the central cavity and pockets function as continuous ligand-binding pockets and that the GGPP and GGGP chains of substrates are accommodated within these tunnels. Intriguingly, clear regions of electron density at the lateral opening are also identified, which are surrounded by positive residues (Arg16, Lys18, Arg175, and Lys266). A modeling study showed that LDAO molecules fit well with this density (Figures 3C and 3D). The presence of these positive residues suggests that substrate(s) may enter or that the DGGGP product might be released from this lateral opening.

To better understand the substrate binding of DGGGPS, we performed a molecular dynamic simulation using GGPP and GGGP as ligands. In the modeled structure, GGPP and GGGP are coordinated in the central cavity by the polar and hydrophobic regions as observed in homologs [96,100] (Figures 4A and S6), whereas the long hydrophobic tails of GGPP and GGGP are buried in the pockets. To verify the functional role of the key residues that are potentially involved in ligand binding and catalysis, we next performed mutagenesis analyses using an enzymatic assay and a lipid-binding assay. When Asn62, Asp66, and Asp70 in the CHL23 region and Asp180, Asp183, and Asp187 in the CHL67 region of MjDGGGPS were individually mutated to alanine, most of the resulting mutants showed no activity, but the mutant of N62A, of which the activity was ~50% of the wild-type enzyme (Figure 4B). Combined with the observations in the structural model and lipid binding assay, these data suggested that these conserved residues involve in the recognition of the phosphate head groups of the ligands via polar interactions and are particularly crucial for subsequent enzyme catalysis, which agree with the findings of other enzymes with similar binding modes [96,100]. Moreover, we generated several variants by mutating residues from the geranylgeranyl chain-binding site (S146A, F148A, and S171A) and the lateral opening site (I29A and I29F). As anticipated, these variants showed altered catalytic activities compared to that of wild-type MjDGGGPS (Figure 4B).

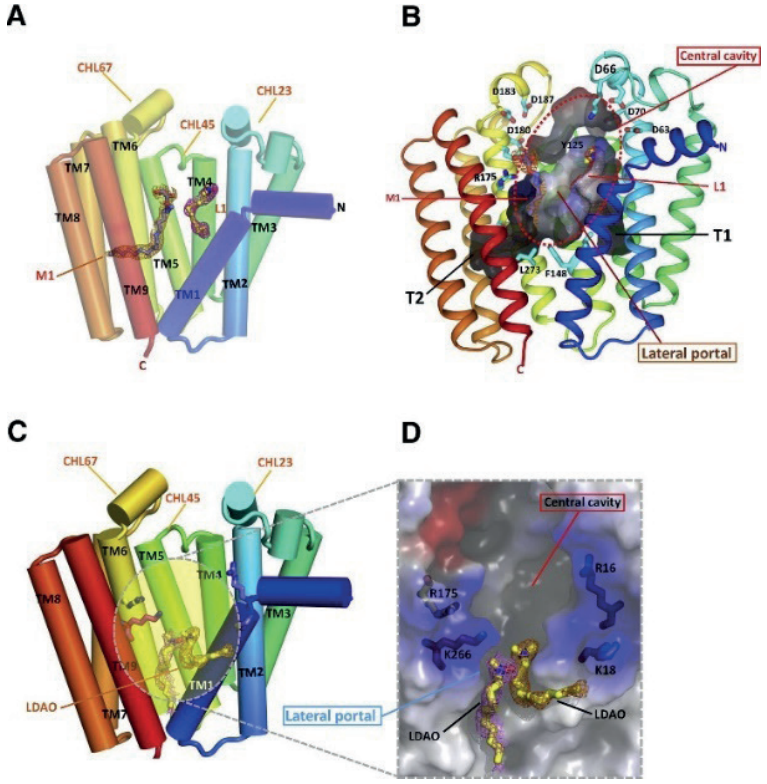


Figure 3: Structure of lipid-like bound MjDGGGPS. (A) The MjDGGGPS structure in complex with LDAO and monoolein. The 2Fo-Fc omit maps of LDAO (L1 in purple) and monoolein (M1 in orange) are contoured at 0.75 s. LDAO and monoolein are indicated in sticks and spheres. **(B)** Electrostatic representation depicting the active site. Tunnel 1 and tunnel 2 (T1 and T2) are shown as two long pockets for lipid tail accommodation, and key residues adjacent to the central cavity are represented as sticks and labeled. **(C)** Lipophilic molecules located at the lateral portal. The 2Fo-Fc omit maps of LDAO colored in pink are contoured at 0.75 s. The lateral portal and LDAO are highlighted in a gray dashed circle. **(D)** A zoomed view of the lateral portal. Electrostatic representation depicting the lateral portal with some positive residues located at the opening represented as sticks and labeled.

Dramatically, the mutant of S171A, which might lead to stronger hydrophobic interactions with the lipophilic tail of GGPP and inhibit the releasing of substrate, showed no activity (Figures 4B and 4D). F148A, crippling the correct positioning of the substrates and product, showed ~50% activity as that of the wild-type enzyme. Ile29 is located at the lower edge of the lateral opening, and mutation of this residue to alanine had no apparent impact on the enzymatic activity. However, when Ile29 was mutated into phenylalanine, a considerable reduction in activity was observed, suggesting that the lateral opening is an important structural feature for activity. The significance of these critical residues involved in ligand binding was further verified using a lipid assay.

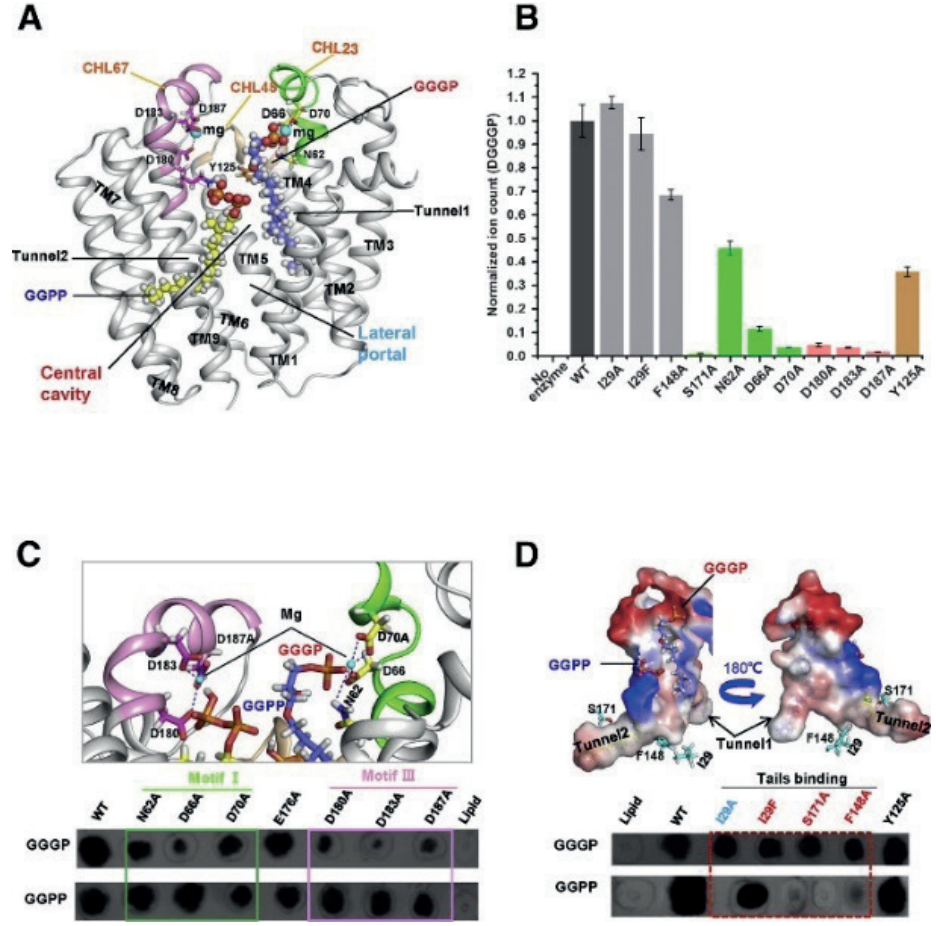


Figure 4: Identification of residues involved in activity and ligand binding. (A) Cartoon depicting the MD model of GGPP and GGGP binding to MjDGGGPS. Conserved residues adjacent to the active site **(C)** are probably for coordination of the GGPP and GGGP phosphate headgroups, which was verified by enzymatic activity assay analyses **(B)**. GGPP and GGGP are represented as sticks and spheres in distinct colors, and Mg^{2+} ions are highlighted as cyan spheres. **(B)** Isoprenyl-transferase activities of MjDGGGPS variants. Residues around the active site were modified to alanines. **(C)** Conserved motifs are involved in binding the phosphate headgroups of GGGP and GGPP. Top: Mg^{2+} is proposed to mediate the interactions between these conserved residues and the polar groups of GGGP and GGPP, the interactions were indicated by blue dashed lines. Bottom: in vitro lipid-binding assays. WT, as a positive control, is known to bind GGGP or GGPP. Negative controls were lipids alone. MjDGGGPS variants are indicated. The reactions were detected by incubating with an anti-6xHis antibody. **(D)** Identified key residues involved in lipid tail binding. Top: electrostatic representation depicting the active site with highlighted tunnels, which potentially accommodate GGGP and GGPP isoprenyl tails. Bottom: in vitro lipid-binding assays. Residues within hydrophobic pockets are verified by lipid binding assays for DGGGPS.

Supporting the results above, the results from the lipid assay showed that mutations of Asp66, Asp70, Asp180, Asp183, and Asp187 reduced the binding affinity with GGGP but had little effect on GGPP binding (Figure 4C). These results implied that the head group of GGGP is probably recognized by Motif I and Motif III. Moreover, the lipid assay also showed that mutations of I29F, F148A, and S171A considerably decreased the GGPP binding but little effect on GGGP binding, these findings further indicated that the lipophilic tail of GGPP is probably buried in the hydrophobic tunnel 2, whereas the tail of GGGP is likely located in tunnel 1 (Figure 4C and 4D). Interestingly, Y125A had little influence on the binding affinity, although the mutation significantly impaired the enzyme activity. This finding is consistent with a previous report showing that Tyr125 could serve as a catalytic base in the enzyme-catalyzed reaction [100].

Structural interpretation for hot-spot mutations associated with diseases

A Dali search identified ApUbiA (PDB: 4OD4 and 4OD5) [100] as being structurally most similar to MjDGGGPS. Together with the conserved motifs and similar substrates, this finding further verified that DGGGPS is a member of the UbiA superfamily. However, DGGGPS is unique in that it prefers linear rather than cyclic substrates, which are the major substrates of the UbiA superfamily. To study the general catalytic mechanism and further understand the evolution of UbiA superfamily enzymes, particularly, to explore mechanisms that diseases associated with quinone-like molecules deficiency. Therefore, we selected COQ2 from yeast and UBIAD1 from human as representatives for ubiquinone and menaquinone biosynthesis, respectively.

The high similarity of the protein sequences of MjDGGGPS, UBIAD1, COQ2, and ApUbiA, particularly involving the functionally conserved motifs $D_{66}XXXD_{70}$,Y_{125}–K_{128}, and $D_{183}XXXD_{187}XXD_{190}$ (Figures 1B, 1C, and S1), allowed us to conduct structural modeling of COQ2 and UBIAD1 with PHYRE2 [203] (Figures 5A and 5B). In the modeled structures, the functionally critical and conserved amino acids in MjDGGGPS were well aligned with those in COQ2 and UBIAD1 (Figure 1B).

To define residues that are essential for catalysis of COQ2 and UBIAD1, we mutated some of the residues, from the DXXXD, Y_{125}–K_{128}, and DXXXDXXD motifs of COQ2 and UBIAD1, and tested the enzymatic activity of the resultant mutants. Supporting a vital role of these residues in a COQ and UBIAD1-catalyzed reaction, mutations, N142A, D146A, D150A, R159A, Y207A, D259A, D266A, and D270A, dramatically decreased the enzymatic activity of COQ2 as well as mutations N102A, D106A, R119A, Y174A, D236A, and D240A of UBIAD1 (Figures 5C and 5D). These results are

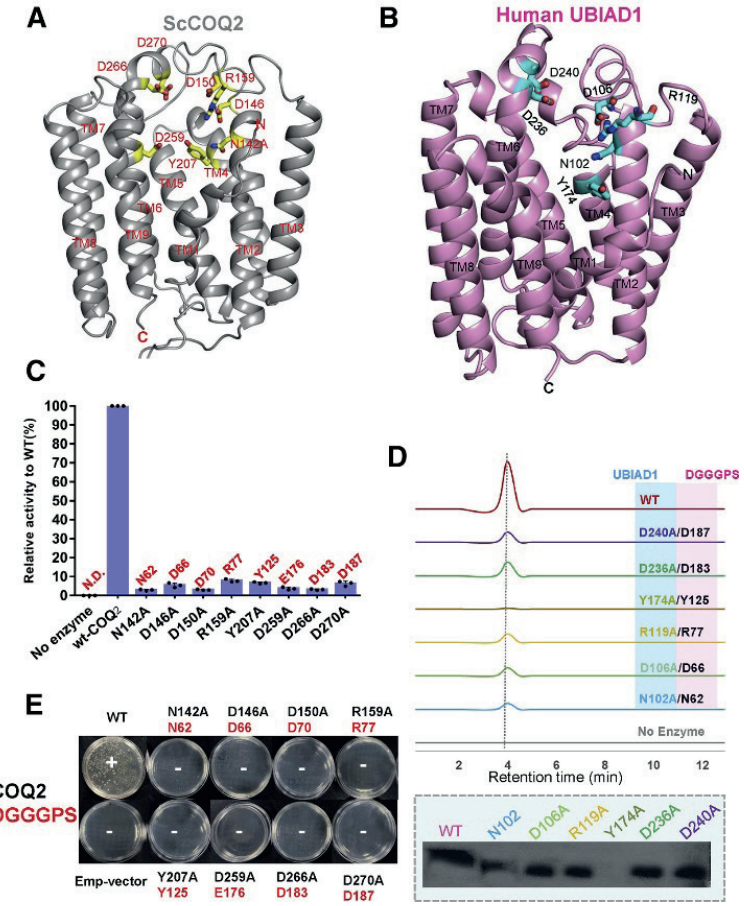


Figure 5: Structural interpretation for hot-spot mutations associated with diseases. The ApUbiA structure (PDB: 4OD5) was used as a template to generate homology models by PHYRE2. **(A)** The functionally conserved residues are indicated in the active site of structural model of *Saccharomyces cerevisiae* COQ2. **(B)** Structure-guided explanations of disease associated residues of human UBIAD1. Hot-spot residues are highlighted in the central cavity. **(C)** Isoprenyl-transferase activities of COQ2 variants. Residues adjacent to the active site are mutated to alanines, and corresponding residues in the MjDGGGPS structure (Figure 2A) are indicated in red (N.D., not detected; WT, wild-type). **(D)** Enzymatic activities of UBIAD1 variants, the corresponding residues of MjDGGGPS are highlighted with light pink; expression of UBIAD1 variants in insect cell was detected by western blot, all mutants expressed well except for Y174A, which could not be detected. The antibodies used in the western blot are shown in the key resources table in the supplementary information. **(E)** Complementation of CoQ2 mutants. Plasmids containing wild-type and mutant CoQ2 were transformed into the *E. coli* BL21 (DE3) strain with a double ubiA$^-$/menA$^-$ deficiency. The active CoQ2 constructs could rescue the growth of this quinone-deficient strain by transformation. Corresponding residues in the MjDGGGPS structure are indicated in red.

in accordance with previous studies [98,367,373,374]. In support of the critical role of those residues in catalysis of COQ2, most of the mutants significantly impaired the growth of the menaquinone-deficient *E. coli* strain, compared with wild-type COQ2 (Figure 5E). Unfortunately, we were unable to rescue this defect strain by expressing UBIAD1 due to poor expression of this protein in *E. coli*. Importantly, these data provide an explanation for the disease-linked mutations in eukaryotic UBIAD1 and COQ2, which further support that DGGGPS is a member of the UbiA superfamily (Figures 5A–5E and S1).

The reported Schnyder corneal dystrophy (SCD) mutations of UBIAD1 are adjacent to the central cavity as shown in Figure 5B, which are probably involved in substrate binding or catalysis. For instance, the Asn102 residue of UBIAD1 corresponding to Asn142 of COQ2 and Asn62 of MjDGGGPS (Figure 1B), respectively, is associated with SCD when mutated to alanine [98,373,374], This association was further evidenced by the enzymatic assays of N142A of COQ2 and N62A of MjDGGGPS, which both almost fully lost activity. In addition, some reported mutations of COQ2 in human caused infantile multisystem diseases and nephropathy, which are highly conserved in yeast COQ2 [374]. The pathogenic mechanism of these mutations also could be explained based on this structure model. Collectively, the structural models of DGGGPS, COQ2, and UBIAD1 could provide insights into the diseases associated with the hot-spot mutations.

Discussion

Here, we report the structural and biochemical characterization of the intramembrane enzyme MjDGGGPS, which plays a critical role in the biosynthesis of archaeal membrane lipids [362,364]. Bioinformatics and biochemical assays demonstrated that MjDGGGPS is a membrane-embedded prenyltransferase and an archaeal member of the UbiA superfamily (Figure 1C and S1). Functional, structural, and modeling studies elucidated a general mechanism by which members of the UbiA superfamily recognize distinct substrates (Figures 4 and 5). The data from these studies collectively indicate that the conserved structural motifs are involved in substrate binding, supporting conserved catalytic mechanisms among members of the UbiA superfamily (Figures 1B, 5A, 5B, and S1).

To date, crystal structures of two UbiA homologs, one from *Aeropyrum pernix* (ApUbiA) [100] and one from *Archaeoglobus fulgidus* (AfUbiA) [96], have been determined. However, the enzymatic activities of these archaeal homologs could not be detected probably due to unknown native substrates.

Although the overall structures of ApUbiA, AfUbiA, and MjDGGGPS are comparable, striking differences exist between MjDGGGPS and the other two enzymes. The large opening and specific tunnels in MjDGGGPS are markedly different from those of ApUbiA and AfUbiA (Figures 2 and S3). The structure of geranyl thiolodiphosphate (GSPP) complexed with ApUbiA showed that the non-cleavable GSPP is coordinated in the central cavity by the conserved residues of the Asp-rich motifs, and these interactions are mediated by Mg^{2+} ions. The large cavity for substrate accommodation and reaction catalysis presumably consists of a hydrophobic bottom wall for isoprenyl chain binding and a small basic pocket for interaction with an aromatic substrate. In contrast, the prenyl-acceptor substrate of MjDGGGPS contains two long tunnels that can specifically bind substrates with two linear lipophilic tails (Figure 2). Consistent with this idea, lipids bind the tunnels, mimicking GGPP and GGGP (Figures 3A and 3B). Overall, our findings reveal that members of the UbiA superfamily have distinct mechanisms of substrate recognition despite their similar structural motifs.

Interestingly, the structure of DGGGPS resembles that of a soluble prenyltransferase (Figure S7), farnesyl diphosphate synthase (FPPS) [95]. This finding implies that these enzymes may have evolved from a common ancestral enzyme, and thus use a common catalytic mechanism. Therefore, the catalytic reaction of DGGGP is likely similar to that of the UbiA enzyme and the soluble prenyltransferase, FPPS. This reaction presumably includes three steps, namely acceptor ionization, condensation, and elimination [96,100]. In addition, the catalytic and binding assays indicated that the $D_{66}XXXD_{70}$

and $D_{183}XXXD_{187}XXD_{190}$ motifs potentially participate only in GGGP binding. First, we postulated that the $D_{66}XXXD_{70}$ motif and tunnel 1 primarily coordinated GGGP with the assistance of the $D_{183}XXXD_{187}XXD_{190}$ motif, and the residue Tyr125 of the Y125–K128 motif likely catalyzed the reaction by nucleophilic attack subsequently followed the condensation. This would suggest that the catalytic reaction of DGGGPS is dependent on conformational changes induced by substrates and intermediate binding. Accordingly, we propose that these steps underlie the structural mechanism of DGGGPS (Figure 6). In summary, we characterized MjDGGGPS, an archaeal lipid synthase, which is a member of the UbiA superfamily, and we reported the structure of an active protein in this superfamily (Figures 1, 2, and S1). Our data provide insights into the functions of members of this superfamily and facilitate the development of different therapeutic strategies for combating UbiA-associated diseases. Moreover, these studies also provide new insights into the evolution of the distinct kingdoms of life, supporting the view that UbiA, from Bacteria and Archaea to Eukaryotes, evolved from a common ancestral enzyme to possess a similar catalytic mechanism but perform different functions.

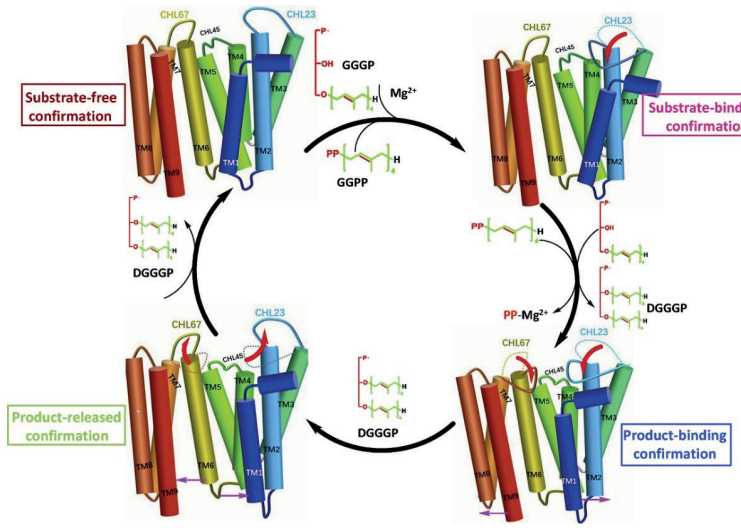


Figure 6: Potential structural mechanism of DGGGPS. The structural conformation mechanism of DGGGPS is distinguished from typical UbiA and FPPS. Based on structural and biochemical assays, we proposed that this can trigger the movement of the flexible fragment CHL23 by substrate GGGP binding; however, the flexible region CHL67 is likely static during this stage even though the other substrate, GGPP, binds. Subsequently, CHL67 seems to move inward to the active site when the catalytic reaction occurs and CHL23 and CHL67 probably coordinate the head group of the intermediate. Then, the helices of TM1 and TM9 synergistically follow the movement of the CHLs to form a larger lateral opening, which is ready for product release affiliating with hydrophobic residues adjacent to this opening.

Materials and methods

Protein expression and purification for crystallization and lipid binding assays

The gene encoding di-geranylgeranylglycerol phosphate synthase from *Methanocaldococcus jannaschii* (MjDGGGPS) was subcloned into the pET-21B vector (Novagen, Madison, WI, USA). The gene fragment of MjDGGGPS was obtained by the PCR with two primers (MjDGGGPS forward primer and reverse primer, primer sequences are presented in the key resources table in the supplementary information). To overexpress MjDGGGPS, freshly transformed *E. coli* C43 (DE3) or *E. coli* Lemo21(DE3) cells were grown in LB media at 37°C to an $OD600_{nm}$ of 1.0, then 0.3 mM isopropyl-β-D-thiogalactopyranoside (IPTG) was added to induce MjDGGGPS expression for 4 h. The cells were resuspended in lysis buffer (containing 15mM Tris-HCl (pH 7.5), 150 mM NaCl, 10% glycerol and 1 mM phenylmethylsulfonyl fluoride (PMSF)) and then lysed by four passes in a French press. After centrifugation (100,000 x g, 55 min, 4°C), the cell membranes were harvested, solubilized in 1% n-dodecyl-β-D-maltoside (DDM), and subsequently incubated with Ni-NTA resin at 4°C for 1-2 h. The resin was washed in three steps with 10, 30, and 50 mM imidazole in buffer (containing 0.03% DDM). The protein was eluted and concentrated to 2 mL at 4°C using a 30 kDa cutoff Amicon Ultra-15 centrifugal filter unit (Merck-Millipore) and then further purified by size-exclusion chromatography using a Superdex 200 column (GE Healthcare, Little Chalfont, UK) in 10 mM Tris-HCl (pH 7.5) buffer containing 0.02% DDM and 100 mM NaCl. Peaks corresponding to MjDGGGPS were pooled and concentrated to 28 mg/ml in a new Amicon Ultra-15 filter unit. To screen for crystal formation, MjDGGGPS was purified with different detergents (DDM, LDAO, n-octyl-β-D-glucoside (OG), n-nonyl-β-D-glucoside (NG) and n-decyl-β-D-maltoside (DM)).

Crystallization

Each purified protein was screened for crystals by the LCP method (Caffrey, 2009) at 20°C; the protein in complex with DDM was in a crystallization buffer containing 100 mM $MgCl_2$, 30% (w/v) PEG600 and 100 mM Tris-HCl (pH 8.0), and the protein in complex with LDAO was in a crystallization buffer containing 100 mM NaCl, 30% (w/v) PEG400, and 100 mM Tris-HCl (pH 8.0). Both of these protein conditions could grow crystals. The crystals were similar and (quadratic plates) grew to a size of approximately 0.03 x 0.3 x 0.1 mm^3 within 4-6 weeks.

Structure determination

The data were collected at the BL19U1 and BL18U1 beamlines at the Shanghai Synchrotron Radiation Facility. All diffraction frames were indexed, integrated, and scaled using the XDS data processing package [375]. The LDAO and DDM crystals diffracted to 2.3 Å and 3.3 Å, respectively and were both in the $P2_12_12$ space group. The statistics of data collection and refinement are shown in Table S1. The preliminary phasing of LDAO crystals was determined by the single-wavelength anomalous dispersion (SAD) method using the SeMet-labeled protein, the SeMet sites were identified by ShelxC/D/E [376], and the density modification was further performed using Resolve [377]. An initial model was built using Resolve in phenix.autobuild, then refined with REFMAC5, and the final model was manually built in COOT [378] with REFMAC5. The apo structures were solved by molecular replacement with the corresponding LDAO structure, built in COOT, and refined by Phenix [379].

Protein expression and purification for MjDGGGPS activity assays

E. coli strains were cultured under aerobic conditions at 37°C with 200 RPM shaking in LB medium supplied with antibiotics where appropriate. For protein expression, *E. coli* Lemo21 previously transformed with the appropriate *Methanocaldococcus jannaschii* DGGGPS expression plasmid was grown in a 1 liter volume of LB medium supplemented with chloramphenicol (25 µg/ml), and kanamycin (50 µg/ml) or ampicillin (100 µg/ml). Protein expression was induced at OD_{600nm} 0.8 by adding IPTG to a final concentration of 0.5 mM for 2 hours.

Induced *E. coli* cultures were harvested by centrifugation at 7,500 x g for 15 minutes at 4°C and washed with chilled basic buffer (50 mM Tris-HCl pH8, 100 mM KCl, 15% glycerol). Washed cells were pelleted (4,000 x g, 15 minutes at 4°C) and resuspended in basic buffer supplemented with an EDTA-free protease inhibitor tablet (Roche) and 0.1 mg/ml DNAse. Cells were then lysed by using a cell press (One-shot cell disruptor, Constant Systems) at a pressure of 12,500 psi. The lysate was centrifuged for 10 minutes at 20,000 x g at 4°C to remove non-lysed cells and cell debris. The resulting supernatant was then centrifuged for 90 minutes at 180,000 x g at 4°C to pellet the total membrane fraction, which was then homogenized in 1 ml basic buffer per liter culture used.

500 ul of membrane fraction was solubilized by the addition of an equal volume of 2x solubilization solution (100 mM Tris-HCl pH 8, 200 mM KCl, 30% glycerol, 4% DDM) followed by mixing and 30 minutes of incubation at 4°C on a rocking roller bench. Solubilized membranes were centrifuged for 10 minutes at 16,000 x g, and the supernatant was added to 0.5 ml column volume of equilibrated (basic buffer + 0.05% DDM) Ni-NTA agarose beads (Qiagen) and incubated for 1 hour at 4°C on a rocking roller bench. Proteins bound to the column were washed with 10 column volumes of wash buffer (basic buffer + 0.05% DDM, 20 mM imidazole). Purified DGGGPS was eluted using 3 column volumes of elution buffer (basic buffer + 0.05% DDM, 250 mM imidazole). Protein purity was assessed by 12% SDS-PAGE stained with Coomassie Brilliant Blue. The protein concentration was then determined by measuring the absorbance at 280 nm using a spectrophotometer (UV-3100PC, VWR).

In vitro MjDGGGPS activity assay and ultra-performance liquid chromatography-ESI-mass spectrometry analysis

DGGGP synthesis reactions were performed in 50 ml of 50 mM MES buffer at pH 7 supplemented with 100 mM KCl, 10 mM $MgCl_2$, 2 mM DDM, 200 µM GGPP, 200 µM GGGP and 1 µM MjDGGGPS enzyme. Reactions were incubated for 30 min at 80°C. In experiments testing the metal dependence, $MgCl_2$ was excluded from the buffer and replaced by 10 mM of other chloride salts of the indicated metal ions.

Reactions were extracted by the addition of 150 µl water-saturated 1-butanol followed by vigorous vortexing. Phase separation was achieved by centrifugation for 2 min at 12,000 x g. An aliquot of 130 µl of the upper (1-butanol) phase was collected in a 1.8 mL glass vial with a PTFE-lined cap. Fresh water-saturated 1-butanol (130 µl) was added, and the extraction was repeated as described above. The collective 1-butanol fractions were evaporated to dryness under a N_2 stream on a heat block at 37°C. The dried residue was dissolved in 50 µl of methanol and analyzed using UHPLC-MS (details in Supplementary experimental procedures).

Butanol extracts of *in vitro* reactions were analyzed by UHPLC-MS using an Accela1250 ultra-performance liquid chromatography system (Thermo Fisher Scientific) coupled to a Thermo Exactive orbitrap mass spectrometer (Thermo Fisher Scientific) equipped with an ESI ion source. Compounds were separated using an Acquity UPLC CSH C18 column (2.1 x 150 mm, 1.7 µm; Waters) maintained at 55°C as described in more detail elsewhere [366]. The injection volume was 5 µl. Mobile phase A consisted of MilliQ:ACN (40:60) containing 5 mM ammonium formate, and mobile phase B consisted of MilliQ:ACN:1-BuOH (0.5:10:90) also containing 5 mM ammonium formate. Compounds were eluted using a gradient elution from 5% to 52.5% buffer B over a period of 22.5 min at a flow rate of 300 µl/minute.

MD simulations

GGGP and GGPP (ligands) were docked into the structure of MjDGGGPS using the GOLD 5.0 program [380] with default settings. The modeled structure was inserted into a previously equilibrated 1-palmitoyl-2-oleoyl-phosphatidylcholine (POPC) membrane, while the orientation and relative position of the protein was determined by online server of Orientations of Proteins in Membranes database (OPM). A 100 x 100 $Å^2$ membrane containing 258 POPC molecules with embedded protein was derived, then 22.5 Å thick water was added on both upper and lower membrane along Z axis.

During the pose refinement stage, the CHARMM36 force field [381] was applied for the enzyme and the CGenFF [382] was applied for GGGP and GGPP. A steepest descent algorithm was performed to optimize the initial docking complex structure. Subsequently, the membrane and GGGP/GGPP was gradually released to free move while protein backbone atoms were restrained. After a total of 2.25 ns equilibrate simulation, the $RMSD_{protein\text{-}heavy}$ between starting structure and last frame protein is 0.993 Å, and corresponding $RMSD_{GGGP/GGPP}$ is 1.501 Å. The last frame was used for illustration. All simulations were run using GROMACS2018 [383].

Lipid binding assay

Membrane lipid strips were utilized as previously described [384], and the detailed experimental procedures were as reported by Ren et al. [103]. Briefly, 0.2 µl (0.005 mM) of synthesized lipids GGPP or GGGP was spotted onto Hybond-C extra membranes (GE Healthcare Life Sciences), and the membrane strips were dried using a hair dryer. The dried membranes were then soaked in blocking buffer containing 5% (w/v) skim milk, 10 mM Tris-HCl (pH 7.5), 100 mM NaCl, 0.02% DDM and 0.01% Tween 20, repeating four times for 8 min each. After this, 0.09 µg of each protein loaded on each strip incubating over-night at 4°C. Subsequently, the membrane strips were incubated with anti-6xHis mouse monoclonal antibodies (Cat# 230001; Zen Bio-Science, Research Triangle Park, NC, USA) for 2-4 h at 4°C, then washed the strips using the blocking buffer 3-5 times for 10 min each and then incubated with HRP-conjugated anti-mouse antibodies (Cat#511103; Zen Bio-Science) for 1 h. After washing several times, they were detected using enhanced chemiluminescence reagents.

In vitro COQ2 enzymatic activity assays

Wild-type and mutant 4-hydroxybenzoate polyprenyltransferase (COQ2) of *Saccharomyces cerevisiae* constructs were cloned in a pET-15b vector. To determine the prenyltransferase activity of COQ2, microsomes containing COQ2 (50 µl) were incubated for 60 min at 37°C in a reaction mixture that included 0.5 mM PHB (Sigma), 1 mM GPP (Echelon Bioscience), 0.4 mM $MgCl_2$ and 50 mM Tris-HCl (pH 7.5) in a total volume of 100 µl [385]. The reaction mixture was extracted by ethyl acetate and dried in air, and the remainder was dissolved in methanol. The geranylated PHB was HPLC-separated, and its identity was confirmed by mass spectrometry. In addition, the prenyltransferase activity of COQ2 was measured by an *in vivo* rescue assay in which the growth-defective ubiA$^-$menA$^-$ *E. coli* strain was transformed with the COQ2 variant constructs.

In vitro UBIAD1 enzymatic activity assays and analysis

The gene coding UbiA prenyltransferase from human was cloned into pFastBac series vector to produce recombinant plasmid. The recombinant plasmid of pFastBac was transformed into *E. coli* of DH10Bac and cultured at 37°C for 48 h on LB agar tablets containing 50 µg/mL of kanamycin, 7 µg/mL of gentamicin, 10 µg/mL of tetracycline, 100 µg/mL of X-gal and 40 µg/mL of IPTG to screen by blue and white spots. Picked white spots and identified of recombinant bacmid by PCR. The extracted recombinant Bacmid was transfected into Sf9 insect (Thermo Scientific) cells and cultured at 28°C for 96 h. The supernatant of the culture medium was collected and centrifuged at 500 x g for 5 min to precipitate cells and cell fragments containing UBIAD1 prenyltransferase. The supernatant was then identified as P1 baculovirus. P1 generation virus can acquire P2 generation virus by reinfecting insect cells. Subsequently, we treated insect cells and performed the activity assay using microsome in the same way as COQ2.

Sequence alignment and homology clustering

Multiple sequence alignments were generated using ClustalW software and were edited with the ESPript 3.0 program [386]. The homology clustering of the UbiA superfamily prenyltransferases was first generated by the PSI-BLAST [387] search of the representative sequences and finally conducted by the CLANS (CLuster ANalysis of Sequences) program [388] on the MPI Bioinformatics Toolkit website (https://toolkit.tuebingen.mpg.de/). In Figure 1C, protein sequences of UbiA superfamily member were aligned using MUSCLE (MUltiple Sequence Comparison by Log-Expectation) [389], then, the maximum likelihood phylogenetic tree was computed by MEGA X [188], the tree was edited in Itol [390].

In Vitro MjDGGGPS / COQ2/ UBIAD1 enzymatic activity assay

Data acquired for the enzymatic activity assay experiments were analyzed by the software GraphPad Prism 7. All of the statistical details of experiments can be found in the figure legends of Figures 1D, 4B, 5C, and S2C.

Acknowledgements

We thank the staff from BL19U1 beamline of the National Center for Protein Sciences Shanghai (NCPSS) at the Shanghai Synchrotron Radiation Facility for assistance during data collection. We thank J. Chai for critical comments. We thank D. Su for providing the equipment. This work was funded by grants to W.C. from the National Key Research and Development Program of China (2018YFC1002802) National Natural Science Foundation of China (31570842 and 31870836), and the 1.3.5 Project for Disciplines Excellence of West China Hospital, Sichuan University (ZYYC20005). W.C. and W. Z were supported by Key Science and Technology Research Projects in Key Areas of the Corps (Grant No. 2018AB019). X.F. was supported by the Ministry of Science and Technology of China (2018ZX09201018-005). N.A.W.d.K and R.A. are supported by a grant of the Dutch Research Council (NWO) within the framework of the Building Block of Life program.

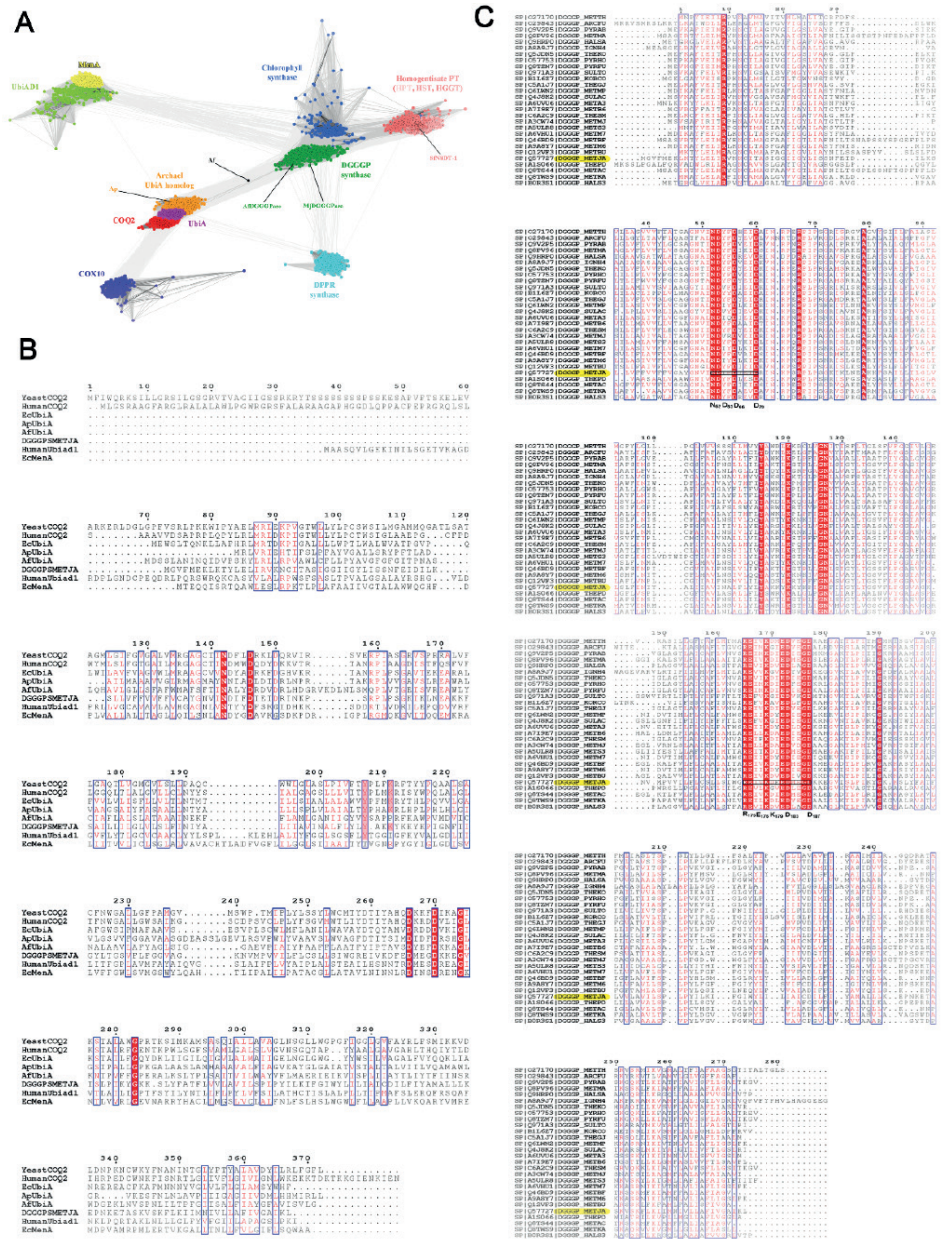
Author contributions

W.C. performed the research design. S.R., M.T., and W.Y. made constructs. S.R., W.Y., and Y.C. purified the proteins and performed in vitro assays and lipid binding assay. S.R. and Y.C. grew and optimized the crystals. Y.G. collected the data. Y.G., S.Q., and X.Z. determined the structures. Q.S. and J.H. created the molecular docking and performed the MD simulations. L.T. did the bioinformatics assay. N.A.W.d.K. made expression constructs, purified the enzymes, and examined their activity by mass spectrometry. R.L.H.A. and A.J.M. performed the lipid synthesis. W.C., H.X., X.F., H.R., A.J.M., W.Z., W.L., Y.W., and A.J.M.D. analyzed the data. W.C. wrote the manuscript. Other authors participated in writing the paper.

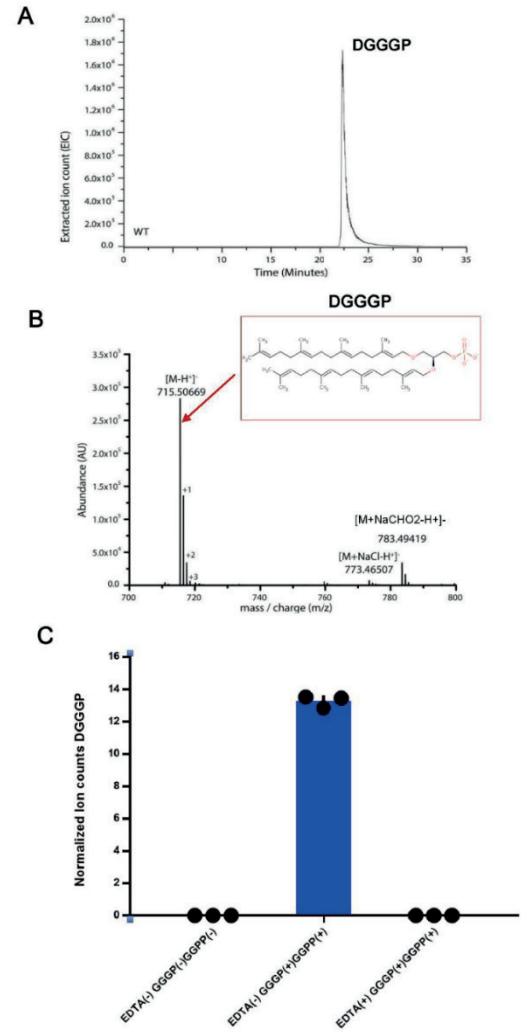
Declaration of Interests

The authors declare no competing interests.

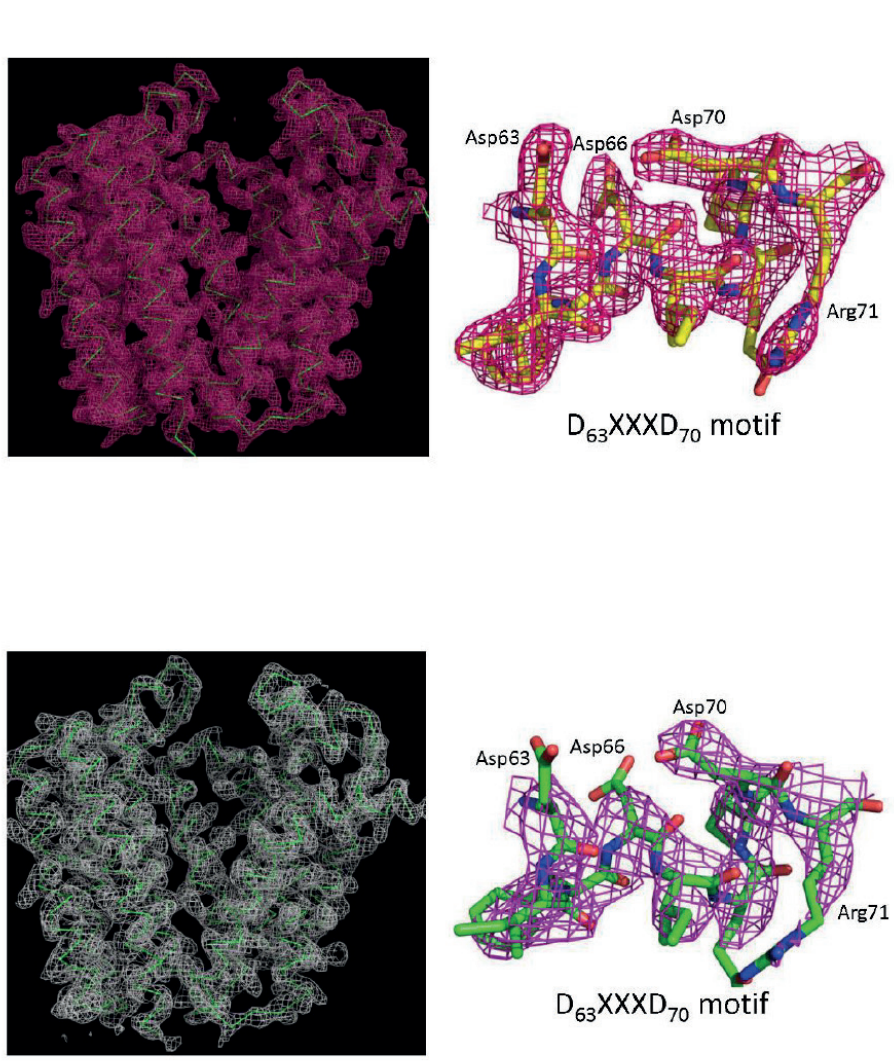
Supplementary information



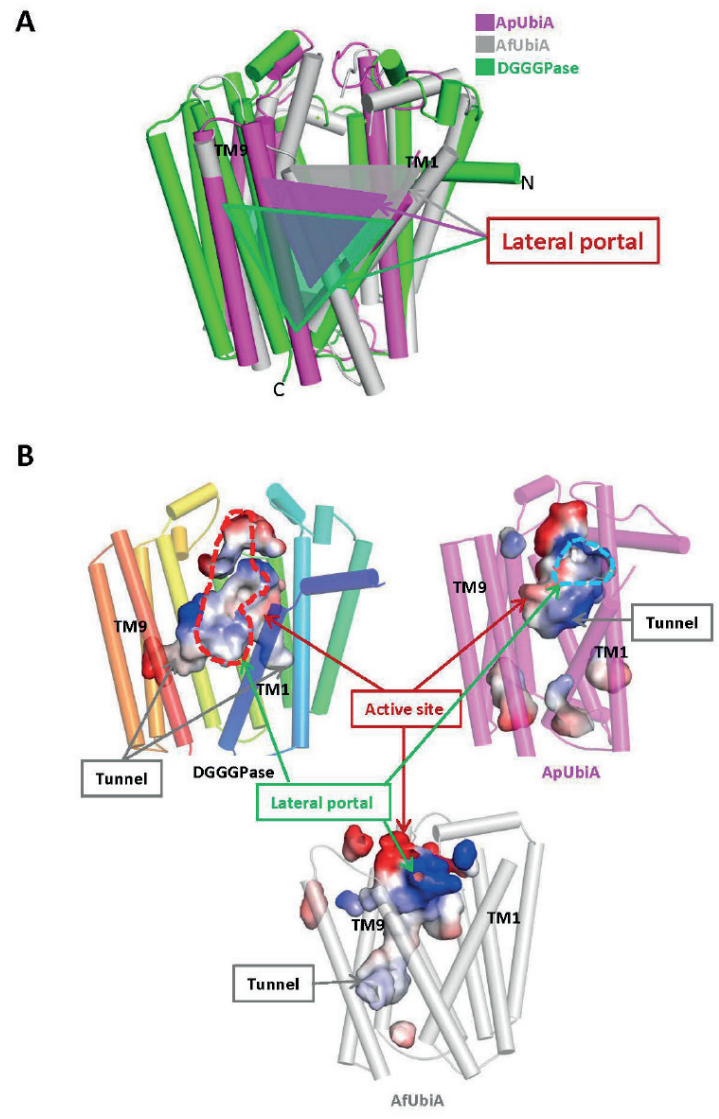
Supplemental figure 1 (previous page): Homology clustering and alignment of the UbiA superfamily prenyltransferases. Related to Figure 1 and Figure 5. **(A)** Dots of the same color mean they are of the same subfamily, and each line shows the similarity of two sequences, with darker lines indicating higher similarity (lower E values). The UbiA superfamily includes the following subfamilies: COQ2 (colored in red) and UbiA (magenta); UBIAD1 (yellow green) and MenA (yellow); homologous prenyltransferases (HPT, HST, and HGGT in different pink colors); chlorophyll synthase (blue); COX10 (dark blue); DGGGP synthase (dark green); and DPPR synthase (cyan). The Ap homolog represents an archaeal UbiA homolog group (orange) close to UbiA and COQ2. However, a cluster cannot be generated for the *Archaeoglobus fulgidus* homolog (annotated as bacteriochlorophyll synthase in NCBI) because of its low homology to other proteins. The homolog search started from the following representative protein sequences: UbiA (NCBI GI: 85676792) and MenA (332345926) from *Escherichia coli*; the ApUbiA homolog (14601492) from *Aeropyrum pernix*; COQ2 (14250676), UBIAD1 (7019551), and COX10 (13623563) from *Homo sapiens;* chlorophyll synthases (332645327) and HPT (281193026) from *Arabidopsis thaliana;* HGGT (526117663) from *Vitis vinifera*; HST (699979766) from *Medicago sativa*; DPPR synthase (618730398) from *Mycobacterium tuberculosis;* and DGGGPS (2495885) from *Methanocaldococcus jannaschii.* The searching database is NCBI protein database of non-redundant proteins filtered for 70% maximum sequence identity (nr70). **(B)** Representative members of the UbiA superfamily were aligned by the CLUSTAL W2 algorithm. ScCOQ2 from *Saccharomyces cerevisiae* (SsCOQ2), hCOQ2 from Homo sapiens (hCOQ2), EcUbiA from *Escherichia coli* (EcUbiA), EcMenA from *Escherichia coli* (EcMenA), hUBIAD1 from *Homo sapiens* (hUBIAD1), ApUbiA from *Aeropyrum pernix* (ApUbiA), AfUbiA from *Archaeoglobus fulgidus* (AfUbiA), MjDGGGPS from *Methanocaldococcus jannaschii* (MjDGGGPS). **(C)** Pairwise sequence alignment of representative DGGGPS sequences, which are all the members of archaeal DGGGPS family. The DGGGPS of *Methanocaldococcus jannaschii,* which was analyzed in the current study, is highlighted in yellow. The conserved residues probably for GGGP and GGPP binding are indicated in black box, the protein is indicated as following: DGGGPS_METJA, di-geranylgeranylglycerol phosphate synthase from *Methanocaldococcus jannaschii*. The alignment was generated using the MultAlign and ENDscript programs.



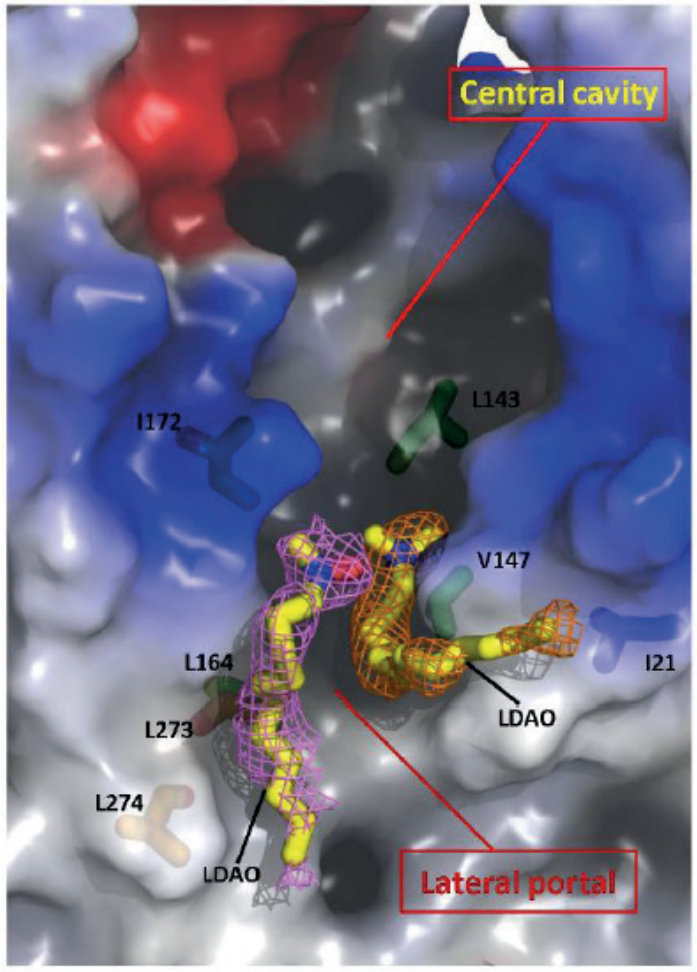
Supplemental figure 2: Detection of products catalyzed by DGGGPS HPLC-MS. Related to Figure 1D and STAR Methods. **(A)** Products from the DGGGPS activity assay using GGGP (2,3-bis-O-geranylgeranyl *sn*-glycerol-phosphate) and GGPP as substrates, as depicted in Figure 1d, were separated by HPLC. **(B)** The total Ion Current Chromatogram filtered for the di-geranylgeranylglycerol phosphate (DGGGP) *m/z* value to detect the product by MS. The DGGGP peak, with a *m/z* of 715.50669 [M-H], is pointed by a red arrow. **(C)** Detection of DGGGPS activity added with EDTA. The activity of purified DGGGPS was analyzed with synthesized GGGP and GGPP as substrates and in the presence or absence of EDTA. The production of DGGGP [*m/z* 1,020.55 (M-H)] was verified by high performance liquid chromatography-mass spectrometry (HPLC-MS) analysis similar to Figure 1d.



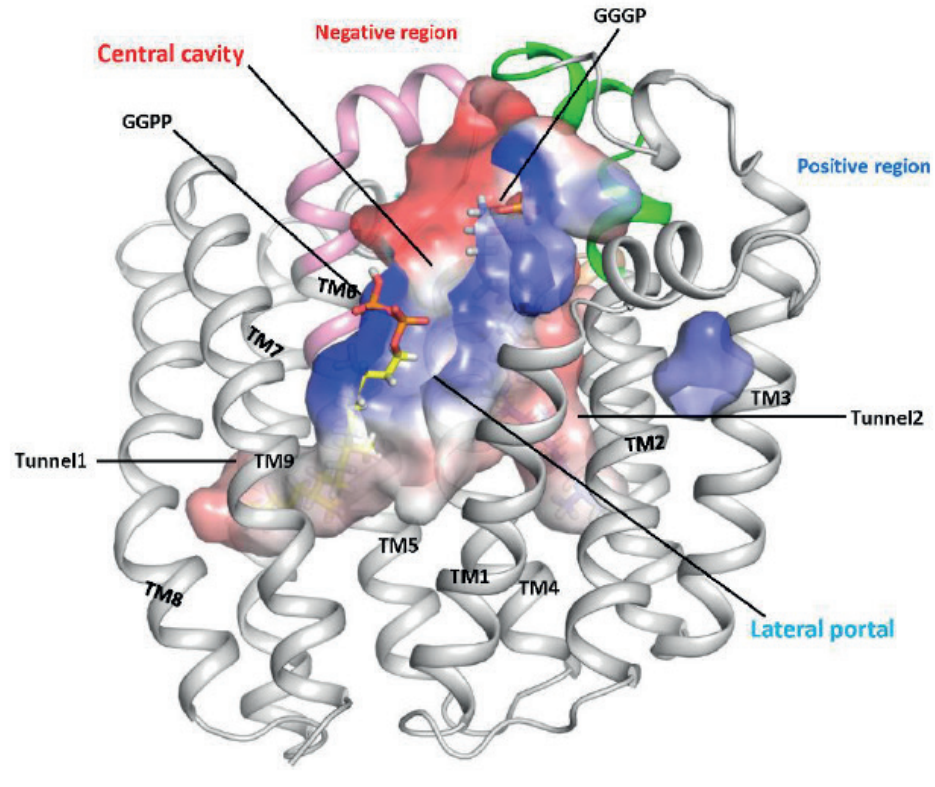
Supplemental figure 3: Representative electron density map of DGGGPS. Related to Figure 2. (A) The 2Fo-Fc omit map of apo-DGGGPS colored in red is contoured at 1.5 σ. With the green backbone ribbon in left panel, the representative D63-D70 motif is highlighted in right panel. **(B)** The 2Fo-Fc omit map of lipid-like bound DGGGPS structure colored in gray is contoured at 1.5 σ with the green backbone in ribbon in left panel, the representative D63-D70 motif is highlighted in light panel.



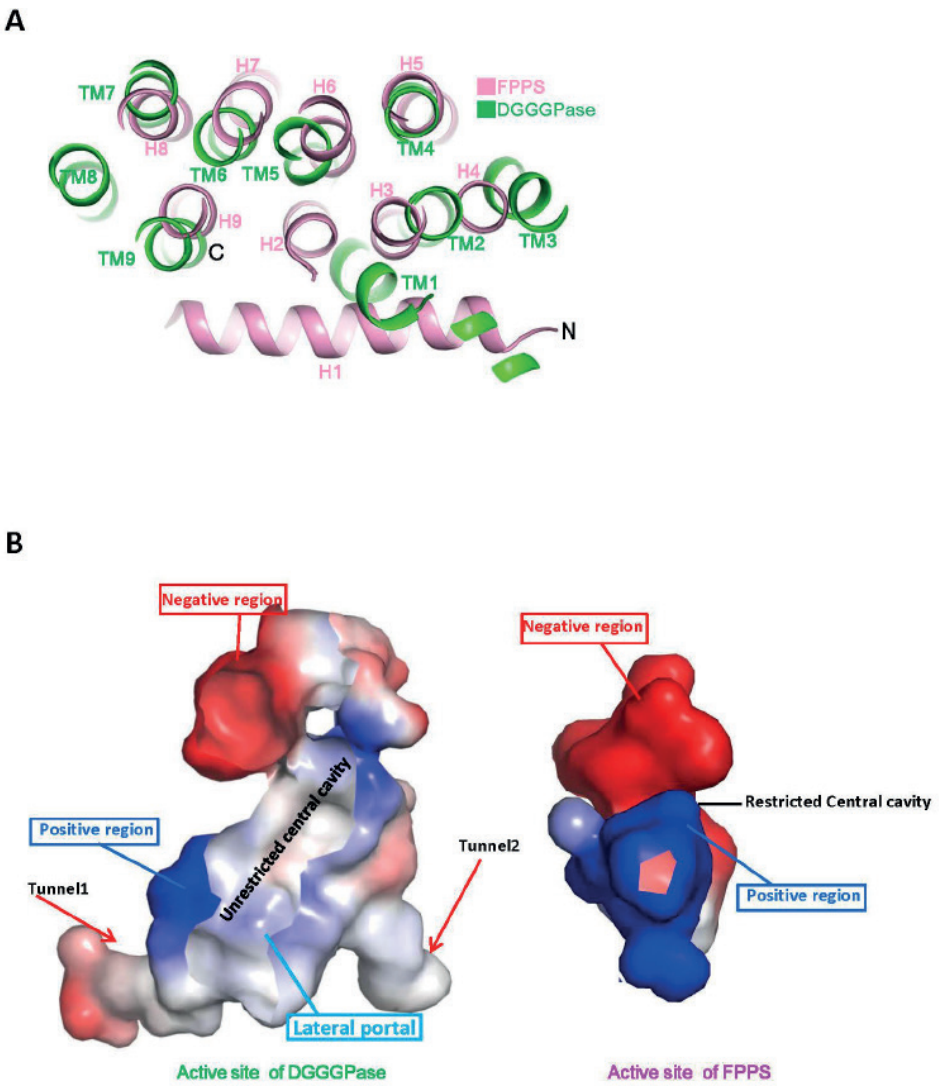
Supplemental figure 4: Structural comparisons among DGGGPS, ApUbiA and AfUbiA of UbiA superfamily. Related to Figure 3. **(A)** The superimposed structures of DGGGPS (green), ApUbiA (red) and AfUbiA (gray). The lateral openings are indicated in different colored triangles. **(B)** Electrostatic representation of substrates binding pockets of DGGGPS, ApUbiA and AfUbiA, respectively. The obviously observations that the larger central cavity consisted of a larger V-shape lateral opening and additional longer tunnels of DGGGPS structure, but in ApUbiA and AfUbiA structures, the lateral openings and tunnels are smaller and shorter, respectively.



Supplemental figure 5: Key residues surrounded the lipid tail at lateral opening. Related to Figure 4A. Hydrophobic residues lining the lateral opening probably are involved in substrate binding and/or product release.



Supplemental figure 6: Electrostatic representation depicting the ligands binding to DGGGPS of sMD structure. Related to Figure 4A.



Supplemental figure 7: Structural comparison between DGGGPS and a soluble prenyltransferase FPPS. Related to Figure 3. The FPPS (1RQI (20)) structure is top by DALI search using the DGGGP structure of soluble prenyltransferase family. **(A)** The superimposed structures of DGGGPS (green) and FPPS (pink). **(B)** Electrostatic representation of central cavity of DGGGP (left) and FPPS (right). The central cavity of DGGGPS, comprising of a larger V-shape lateral opening towards to membrane lipid and two additional long tunnels embedded in membrane lipid, is without restriction, in FPPS structure, the central cavity of the substrate binding is a small pocket that restricts the chain length of substrates.

Supplemental table 1: Data collect and refinement statistics. Related to the STAR Methods and Figure 2, 3.

	SeMet-DGGGPase Lipid-like binding	Apo DGGGPase
PDB code	**6M31**	**7BPU**
Data collect		
Space group	P 21 21 2	P 21 21 2
Cell dimensions		
a, b, c (Å)	70.00, 133.49, 79.50	70.36 134.06 79.80
α, β, γ (°)	90.00, 90.00, 90.00	90.00, 90.00, 90.00
Resolution (Å)	48.89 - 2.3 (2.32 - 2.29)	48.53°– 3.32 (3.52 – 3.32)
Rmerge (%)	9.2(83.4)	21.8(81.5)
CC (1/2)	0.999(0.930)	0.997(0.945)
I/σ (I)	13.40 (2.63)	10.41 (2.90)
Completeness (%)	99.80 (99.5)	99.7 (98.2)
Redundancy	7.02	10.20
Refinement		
No. reflections	33780	11662
R_{work}/R_{free} **(%)**	20.10/23.00	24.60/27.70
No. atoms	5336	4436
Protein	4448	4426
Ligand/ion	2	10
Average B-factors	54.51	62.67
R.m.s deviations		
Bond length (Å)	0.010	0.015
Bond angles (°)	1.06	1.13
Ramachandran Plot		
Favoured Regions	538(96.76%)	530(96.0%)
Allowed Regions	14(2.52%)	22(4.0%)
Outliers	4(0.72%)	2(0.36%)
Protein Sidechains		
Favoured Rotamers	484(99%)	475(97.0%)
Outliers	6(1%)	13(3.0%)
Clashscore/Close contact	6/71	2/22

Key resources table

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies Mouse monoclonal anti-His	Zen Bio-Science,	Cat#: 230001
Goat anti-Mouse IgG(H+L)	Zen Bio-Science,	Cat#: 511103
Bacterial and Virus Strains		
E. coli DH10Bac	Shanghai Weidi	Cat#: DL1071
E. coli BL21 (DE3)	NEB	Cat#: C2527H
C43(DE3) pLysS	Shanghai Weidi	Cat#: EC1041
Biological Samples		
Annatto seeds	Amazon.de	https://www.amazon.com
Chemicals, Peptides, and Recombinant Proteins		
Geranylgeranylglycerol pyrophosphate (GGGP)	In house synthesized [365]	N/A
Geranylgeranyl pyrophosphate (GGPP)	In house synthesized [365]	CAS: 6699-20-3
Geranyl pyrophosphate (GPP)	Echelon Bioscience	CAS: 763-10-0
p-hydroxybenzoate (PHB)	Sigma-Aldrich	CAS: 123-08-0
Dodecyl-b-D-maltoside (DDM)	Anatrace	CAS: 69227-93-6
n-Dodecyl-N,N-Dimethylamine-N-Oxide (LDAO)	Anatrace	CAS: 1643-20-5
Selenomethionine (SeMet)	Sigma-Aldrich	CAS: 1464-42-2
X-Gal	Sigma-Aldrich	CAS: 7240-90-6
isopropyl-b-d-thiogalactopyranoside (IPTG)	Sigma-Aldrich	CAS: 367-93-1
Monoolein	Sigma-Aldrich	CAS: 111-03-5
Nickel-NTA resin	GE Healthcare	LOT: 10236606
Superdex 200 Increase 10/300 GL	GE Healthcare	LOT: 10263259
Hybond-C extra membranes	GE Healthcare	Cat#: RPN303B
30 kDa Amicon Ultra-15	Merck-Millipore	Cat#: PLTK04310
HP5 column	Agilent technologies	Cat#: 19091J-101
Critical Commercial Assays		
2 x High-Fidelity PCR Mix	Beijing TSINGKE	LOT: TP001
Deposited Data		
DGGGPS (apo) structure	This manuscript	PDB: 7BPU; Table S1
SeMet-DGGGPS lipids binding	This manuscript	PDB: 6M31; Table S1
Experimental Models: Cell lines		
Sf9 Cells	Thermo Fisher Scientific	Cat#: 10359016
Oligonucleotides		
MjDGGGPS forward primer: 5'-ATGATATTTTTGATATTGAGATAGCTAGAATAAACAAGCCATCCCGTC-3'	This manuscript	N/A
MjDGGGPS reverse primer: 5'-GACGGGATGGCTTGTTTATTCTAGCTATCTCAATATCAAAAATATCAT-3'	This manuscript	N/A
MjDGGGPS forward primer for pRSF from pET15B: 5'-CGCGCCATGGGGGTTTTTATGGAGAAG-3'	This manuscript	N/A
MjDGGGPS forward primer for pRSF from pET15B: 5'-GCGCCTCGAGCTAGTGATGGTGATGGTGGTGATGATGTAGTTTTATGGCTCCAACAATAAATG-3'	This manuscript	N/A

Recombinant DNA		
pET-21B vector	Novagen	Cat#: 69741
pRSF-Duet vector	Novagen	Cat#: 71341-3
pET-15B vector	Novagen	Cat#: 69661-3
pFastBac vector	Thermo Fisher Scientific	Cat#: 10360014
Plasmid: pET21B-MjDGGGPS	This manuscript	N/A
Plasmid: pRSF-MjDGGGPS	This manuscript	N/A
Plasmid: pET15b-COQ2	This manuscript	N/A
Plasmid: pFastBac-UbiA	This manuscript	N/A
Software and Algorithms		
XDS	[375]	http://xds.mpimf-heidelberg.mpg.de/ html_doc/ xds_prepare.html
ShelX	[376]	http://shelx.uni-goettingen.de/
Phenix	[379]	http://www.phenix-online.org/
COOT	[378]	https://www2.mrc-lmb.cam.ac.uk/personal/pemsley/coot/
MEGAX	[188]	https://www.megasoftware.net/
ESPript 3.0	[386]	http://espript.ibcp.fr/ESPript/ESPript/
ClustalW	[386]	https://www.genome.jp/tools-bin/clustalw
Itol	[390]	https://itol.embl.de/
PSI-BLAST	[387]	https://www.ebi.ac.uk/seqdb/confluence/ display/THD/PSI-BLAST
CLANS	[388]	https://toolkit.tuebingen.mpg.de/
MUSCLE	[389]	http://www.drive5.com/muscle/index.htm
GOLD 5.0 program	[380]	https://www.goldensoftware.com/

Chapter 4

A promiscuous archaeal cardiolipin synthase enables construction of diverse natural and unnatural phospholipids

Journal of Biological Chemistry (2021)
volume 296, 100691
https://doi.org/10.1016/j.jbc.2021.100691

Marten Exterkate a, Niels A. W. de Kok a, Ruben L. H. Andringa b, Niels H. J. Wolbert a, Adriaan J. Minnaard b, Arnold J. M. Driessen a*

a Department of Molecular Microbiology, Groningen Biomolecular Sciences and Biotechnology Institute and Zernike Institute for Advanced Materials, University of Groningen, Nijenborgh 7, 9747 AG, Groningen, The Netherlands

b Department of Chemical Biology, Stratingh Institute for Chemistry, University of Groningen, Nijenborgh 7, 9747 AG Groningen, The Netherlands

* Corresponding author: a.j.m.driessen@rug.nl; Tel. + 31-50-3632164; Fax. +31-50-3632154.

Abstract

Cardiolipins (CL) are a class of lipids involved in the structural organization of membranes, enzyme functioning, and osmoregulation. Biosynthesis of CLs has been studied in Eukaryota and Bacteria, but has barely been explored in Archaea. Unlike the common fatty acyl chain-based ester phospholipids, Archaeal membranes are made up of the structurally different isoprenoid-based ether phospholipids, possibly involving a different cardiolipin biosynthesis mechanism. Here, we identified a phospholipase D motif-containing cardiolipin synthase (MhCls) from the methanogen *Methanospirillum hungatei*. The enzyme was overexpressed in *Escherichia coli,* purified, and its activity was characterized by LC-MS analysis of substrates/products. MhCls utilizes two archaetidylglycerol (AG) molecules in a transesterification reaction to synthesize glycerol-di-archaetidyl-cardiolipin (Gro-DACL) and glycerol. The enzyme is non-selective to the stereochemistry of the glycerol-backbone and the nature of the lipid tail, as it also accepts phosphatidylglycerol (PG) to generate glycerol-di-phosphatidyl-cardiolipin (Gro-DPCL). Remarkably, in the presence of AG and PG, MhCls formed glycerol-archaetidyl-phosphatidyl-cardiolipin (Gro-APCL), an archaeal-bacterial hybrid cardiolipin species that so far has not been observed in nature. Due to the reversibility of the transesterification, in the presence of glycerol, Gro-DPCL can be converted back into two PG molecules. In the presence of other compounds that contain primary hydroxyl groups (e.g., alcohols, water, sugars), various natural and unique unnatural phospholipid species could be synthesized, including multiple di-phosphatidyl-cardiolipin species. Moreover, MhCls can utilize a glycolipid in the presence of phosphatidylglycerol to form a glycosyl-mono-phosphatidyl-cardiolipin species, emphasizing the promiscuity of this cardiolipin synthase, that could be of interest for bio-catalytic purposes.

Introduction

Cardiolipin (CL) is present in lipid membranes throughout all three domains in life. This lipid class comprises lipid species that contain two pairs of lipid tails that are each linked to a glycerol(phosphate) backbone, with a bridging polar headgroup. 1,3-bis(*sn*-3'-phosphatidyl)-*sn*-glycerol, or simply glycerol-di-phosphatidyl-cardiolipin (Gro-DPCL) is the most prominent and studied cardiolipin species. Gro-DPCL is usually a relatively minor component (< 10 mol%) of the total membrane lipid composition, and its primary role appears to be supporting the function of various membrane proteins [162,391]. Unlike most naturally occurring glycerophospholipids, Gro-DPCL consists of two 1,2-diacyl-glycerolphosphate moieties esterified to the 1- and 3-hydroxyl groups of a glycerol molecule. Because CL can contain up to two phosphates it can carry up to at least two negative charges [392]. Furthermore, the polar head group is relatively small compared to the four bulky hydrophobic tails, giving Gro-DPCL its characteristic inverted conical shape in the presence of divalent cations [393]. Due to this structural feature, Gro-DPCL may induce membrane curvature. Indeed, Gro-DPCL is believed to be located in lipid-domains at the cell poles and division site in bacteria [171,394–396], and in eukaryotes, it is an important constituent of the curvy mitochondrial membrane [397–399]. Besides a tight association with cytochrome *c* oxidase, a part of the respiratory chain complex [400–402], other specific Gro-DPCL - protein interactions seem not to be conserved among the three domains of life. In bacteria, Gro-DPCL is often not an essential membrane constituent, and in several organisms its production is upscaled under certain specific circumstances, such as during stationary phase, or when cells are exposed to certain stressors, e.g. osmotic shock [161,186,403,404].

Cardiolipins have also been described in archaea, most notably in Euryarchaeota. Specifically, in halophiles the production of glycerol-di-archaetidyl-cardiolipin (Gro-DACL) is influenced by changes in the ionic composition of the environment [181]. For instance, in the halophilic organism *Halorubrum sp.*, hypotonic stress resulted in increased production of Gro-DACL, as well as the mono-archaetidyl-cardiolipin species S-di-glycosyl-mono-archaetidyl-cardiolipin (S-2glyco-MACL), which only contains one phosphate-group. This cardiolipin species (originally referred to as: S-DGD-5-PA), consists of an archaetidic acid (AA) molecule attached to a sulfated diglycosyl diphytanylglycerol diether (S-2Glyco-DGD) [160]. Further, a variety of other glycosyl-mono-archaetidyl-cardiolipin (glyco-MACL) species have been identified in archaea [181]. For example, *Halobacterium salinarum* produces a S-tri-glycosyl-diether glycolipid fused to AA (S-3Glyco-MACL, originally referred to as: S-TGD-1-PA) [168], while *Haloferax volcanii* also contains the glycosyl

cardiolipin analogue S-2Glyco-MACL (originally referred to as: S-GL-2) [177]. The polar head group of these glyco-MACL species is structurally very different from the classical di-phosphatidyl/archaetidyl cardiolipin species, which raises questions on the specific function of these lipids, as well as the enzymatic mechanism of their synthesis.

Whereas cardiolipin synthesis in Eukaryotes and Bacteria has been studied extensively, the mechanism of cardiolipin biosynthesis in Archaea has largely remained unexplored. Currently, two phylogenetically distinct type of enzyme families are known in Bacteria and Eukaryotes to synthesize di-phosphatidyl-cardiolipin. One family is characterized by the presence of a CDP-alcohol phosphatidyltransferase domain (CAPT), which catalyzes Gro-DPCL synthesis by utilizing the substrates cytidine diphosphate diacylglycerol (CDP-DAG) and phosphatidylglycerol (PG) [405]. In this reaction, the first phosphate group connected to DAG is coupled to the polar glycerol head of PG, and a cytidine monophosphate (CMP) is released. CAPT Cls enzymes are predominantly found in eukaryotes, but have more recently also been identified in bacteria [182]. The second family features two phospholipase D (PLD) domains, which enable Gro-DPCL synthesis by transferring the phosphatidyl-group, originating from a PG molecule, to another PG, whereby a glycerol molecule is released. This type of Gro-DPCL biosynthesis is predominantly present in Bacteria, but has been found in eukaryotes as well [406,407]. This latter reaction is catalyzed by the ClsA and ClsB-type cardiolipin synthases (Cls) [193,408]. However, an alternative synthesis method has been identified in *Escherichia coli* in which phosphatidylethanolamine (PE) is used together with PG, to form Gro-DPCL and ethanolamine, a reaction catalyzed by the PLD-containing ClsC-type enzymes [186,409]. Homology searches in genomes of archaea revealed Cls-like members from both PLD-type and CAPT-type cardiolipin synthase families. However, until now no archaeal enzyme has been experimentally associated with any cardiolipin biosynthesis. Here, we report on the identification and characterization of a PLD-type cardiolipin synthase of the Euryarchaeote *Methanospirillum hungatei*. The enzyme showed a remarkable promiscuity with respect to accepted and produced lipid species, ranging from a wide variety of phospholipids and di-phosphatidyl/archaetidyl-cardiolipins to include even glycolipid and glycosyl-mono-phosphatidyl-cardiolipin species.

Results

Bioinformatic identification of cardiolipin synthases in Archaea

Cardiolipins have only been identified in some archaea, most notably in halophiles; a group of organisms belonging to the phylum of the Euryarchaeota [181]. To identify possible cardiolipin synthesizing enzymes in Archaea, we performed a BLAST homology search with both CAPT-type Cls and PLD-type Cls query sequences to the domain of Archaea. For the identification of archaeal CAPT-type Cls candidates, the cardiolipin synthase from *Streptomyces coelicolor A3* (Sco1389) was used, for which mostly hypothetical proteins were found that could only be appointed to a specific (sub-)phylum (e.g., candidatus Woesearchaeota archaeon) (data not shown) [120]. Although these hypothetic proteins contain the CAPT-motif, no individually cultured archaeal species were identified, and therefore they were not further used in this study. Likewise, to identify putative PLD-containing archaeal cardiolipin synthases, the three Cls enzymes (ClsA, ClsB, and ClsC) of the bacterium *E. coli* were used as query sequence in the BLAST search. This resulted in similar hits, with the best sequence coverage and identity with ClsA (see supporting information). After comparing the BLAST results obtained for all three PLD-type Cls query sequences, the sequence coverage and identity with *E. coli* ClsA (EcClsA) was largest, so it was decided to continue with these putative archaeal candidates. Although the crystal structure of EcClsA has not been determined, a membrane topology prediction based on the average hydropathy-profile of the amino acid sequences of a family of bacterial ClsA proteins reveals the presence of two hydrophobic regions at the N-terminus of this membrane protein (corresponding to amino acids 7-29 and 39-61 in ClsA) (Figure 1A). These could represent transmembrane segments that are linked to a C-terminal globular domain (Figure 1B, left panel) that contains two HKD motifs ($HxKx_4D$), which are universally present in cardiolipin synthases that belong to the phospholipase D superfamily (PLD-type Cls) [408].

The BLAST homology search revealed 2 main clusters of archaeal homologs that also contain the two HKD motifs (Figure 1C), which are discussed in more detail in the supporting information (Figure S1-2). One cluster consists of halophilic archaea and the other concerns methanogenic archaea, both groups belonging to the Euryarchaeota. The enzymes from the cluster of methanogenic archaea show about 20-25% sequence identity with the *E. coli* ClsA, including the predicted two hydrophobic regions at the N-terminus (Figure S3). Specifically, the second predicted N-terminal hydrophobic domain contains the conserved motif: $Wx_7Px_2Gx_3Yx_3G$ ("x" represents a hydrophobic amino acid), present also in bacterial ClsA-type proteins (Figure 1D), but not found in enzymes belonging to

the halophilic cluster (Figure S2). The high incidence of hydrophobic amino acids suggests that this region is likely embedded in the membrane. The conserved proline and glycine (PxxG) residues are located in the middle of the hydrophobic stretch. These two amino acids are often found in turns and loops and thus may introduce flexibility in this predicted helix-region [410,411]. Furthermore, this hydrophobic region is flanked on both sides by multiple positive charges, which are known to inhibit translocation across the membrane according to the positive-inside rule, and thus may affect the membrane topology of this region. As a consequence, this predicted helix-domain may not span across the membrane, but represent a re-entrance loop that enters and leaves the membrane at the same leaflet side (Figure 1B, right panel).

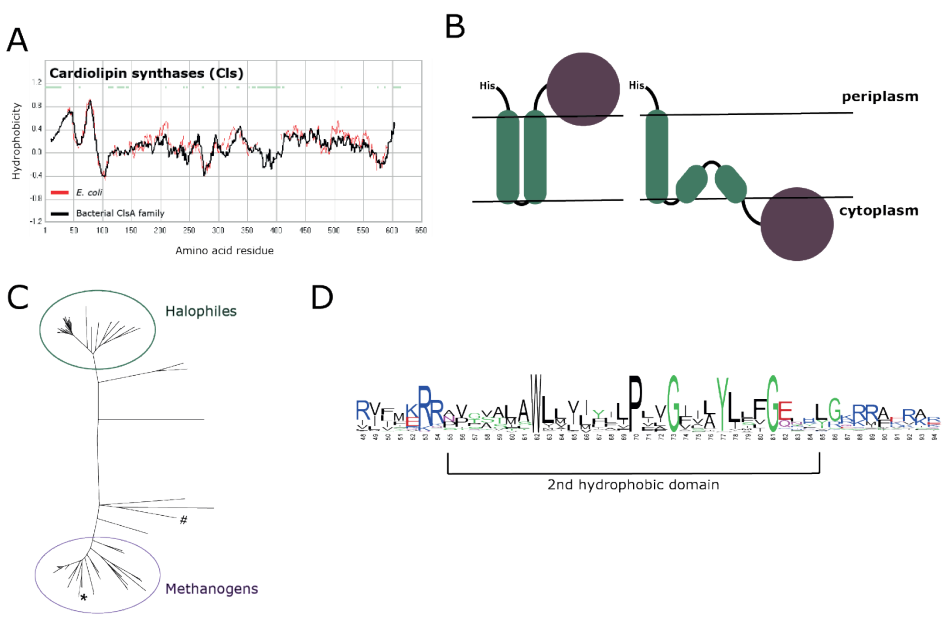


Figure 1: Bioinformatic identification of an archaeal cardiolipin synthase (Cls). (A) Hydropathy profile alignment of *E. coli* ClsA (red line), with the averaged hydropathy profile of its bacterial protein family (black line). **(B)** Schematic representation of potential membrane topologies of cardiolipin synthase with two predicted transmembrane anchors and a globular active domain. **(C)** Unrooted tree of putative archaeal cardiolipin synthases with two main clusters. # *E. coli* ClsA; * M. hungatei Cls. A detailed tree can be found in the supporting information Figure S1. **(D)** Consensus sequence logo of the second hydrophobic region of cardiolipin synthase type A enzymes from bacteria and the group of methanogenic archaea.

Taken together, this group of methanogenic ClsA-like proteins all appear to be promising candidates for potential archaeal cardiolipin synthases. The putative *Methanospirillum hungatei* Cls (WP_011448254) was selected for further characterization as it is the only candidate present in this organism. Furthermore, it grows at mesophilic temperatures and at a pH of around 7, with an osmotic pressure requirement that is comparable to *E. coli*, which was used for protein overexpression.

Glycerol-di-archaetidyl-cardiolipin (Gro-DACL) synthesis from archaetidylglycerol (AG)

The putative *M. hungatei* ClsA homologue (MhCls) was ordered as an *E. coli* codon-optimized synthetic gene, cloned into a His-tag containing overexpression vector and expressed in *E. coli*. Overexpressed MhCls was recovered from the membrane fraction, solubilized with the detergent n-dodecyl-β-D-maltoside (DDM; 2%, w/v), and purified by Ni-NTA agarose affinity chromatography (Figure 2A). To confirm that MhCls is indeed an archaeal cardiolipin synthase, the activity of purified enzyme towards the substrate archaetidylglycerol (AG) was tested *in vitro*. Since this archaeal equivalent of phosphatidylglycerol (PG) is not commercially available, both the *sn1-sn1* **(14)** and the *sn1-sn3* **(14')** diastereomers of AG were chemically synthesized (Figure 2B) (see the supporting information; Figure S4-6). The bis-phytanyl glycerol core was readily synthesized from commercially available phytol and glycidol. Both enantiomers of the glycerol-phosphate headgroup were also produced by asymmetric synthesis and were individually coupled to the aforementioned chiral lipid core via a phosphor-amidite coupling.

After the successful synthesis of *sn1-sn1* and *sn1-sn3* AG, first a mixture of both diastereomers was incubated overnight at 37°C together with purified MhCls. LC-MS analysis revealed that the majority of AG was consumed which coincided with the production of an ion of *m/z* 1520.30 $[M-H]^-$, corresponding to glycerol-di-archaetidyl-cardiolipin (Gro-DACL) (Figure 2C; for fragmentation data see Figure S7). In the absence of MhCls no such conversion was observed. Subsequently the activity of MhCls towards the individual AG diastereomers (*sn1* and *sn3*) was tested and compared, but both acted as substrates in a similar manner (data not shown). Aside from producing Gro-DACL, MhCls additionally synthesized the lipid species archaetidic acid (AA) (Figure 2C). The latter is probably the result of an unsuccessful transfer of the archaetidyl-group, whereby the enzyme hydrolyzes the terminal phosphodiester bond in a phospholipase D-like manner.

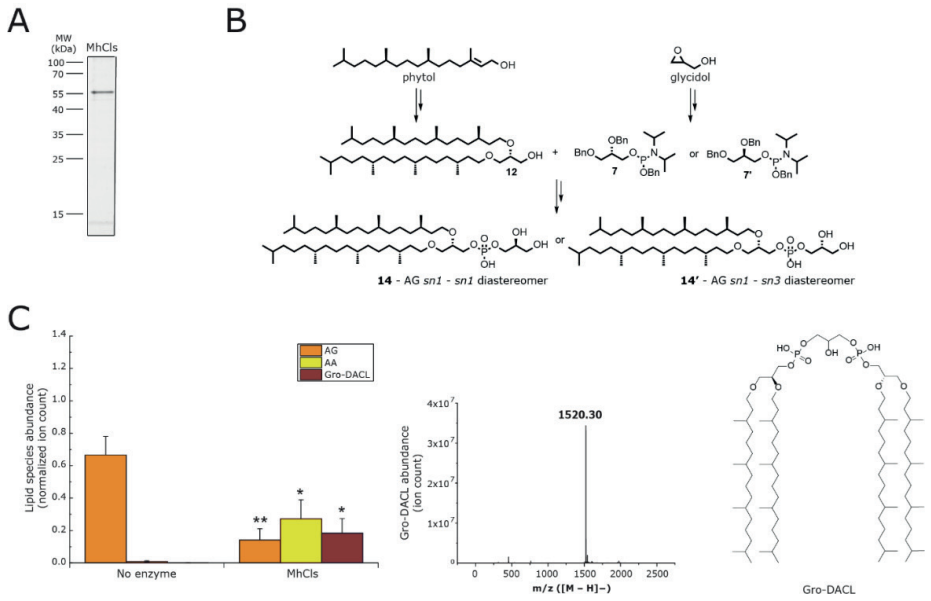


Figure 2: Purification and activity of the cardiolipin synthase from *M. hungatei* (MhCls). **(A)** Coomassie stained SDS-PAGE gel of the MhCls purified by Ni-NTA chromatography. **(B)** Schematic representation of chemical archaetidylglycerol (AG) *sn1-sn1* **(14)**, and *sn1-sn3* **(14')** synthesis. **(C)** *In vitro* activity of purified MhCls. AG was incubated overnight (18h) in the absence and presence of the enzyme MhCls. Lipid species were analyzed by LC-MS. Levels of AG, archaetidic acid (AA) and glycerol-di-archaetidyl-cardiolipin (Gro-DACL) were normalized for the internal standard DDM, and plotted on the y-axis. Data are mean ± SD (n = 3). Statistical significance is shown for the enzymatic reaction (MhCls) compared to the control (no enzyme), for each individual lipid species, by using the Student's t-test analysis; *p ≤ 0.05; **p ≤ 0.01. Mass spectrum showing the presence of Gro-DACL as *m/z* 1520.30 [M-H]$^-$), and its structure.

To verify that MhCls functions as a cardiolipin synthase *in vivo*, the gene was expressed in an *E. coli clsABC* null strain [186]. Subsequent analysis of the lipidome revealed the presence of multiple Gro-di-phosphatidyl-cardiolipin (Gro-DPCL) species (differing in acyl-chain composition), indicating that this enzyme can also utilize phosphatidyl-containing lipids (Figure 3B). As, overexpression of an empty vector resulted in the complete absence of any CL-species (Figure 3A), it can be concluded that MhCls is a cardiolipin synthase.

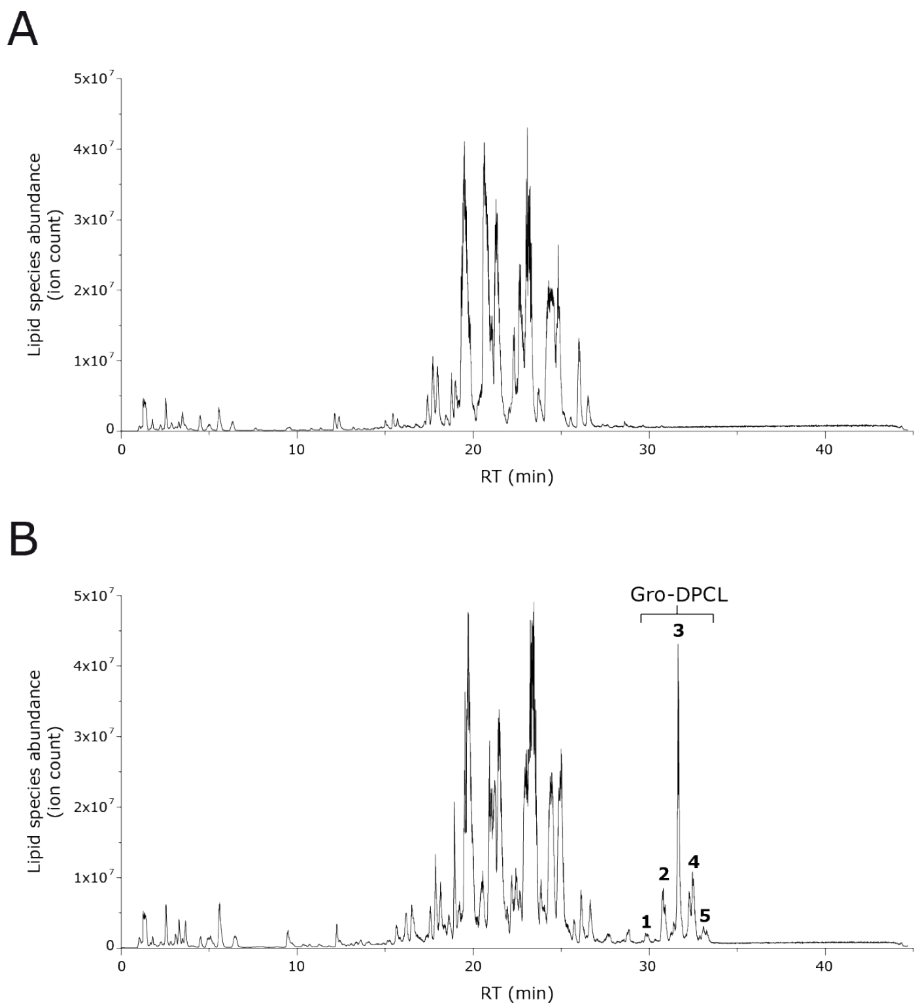


Figure 3: Gro-DPCL biosynthesis by MhCls expressed in the *E. coli* clsABC null strain. LC-MS chromatograms of **(A)** lipidome of the *E. coli* clsABC null strain overexpressing an empty vector, showing a wide variety of lyso-phospholipids (RT 0-8 min.), and di-acyl phospholipids (RT 17-23 min.) **(B)** Lipidome of the *E. coli* clsABC null strain overexpressing MhCls, showing the additionally formed glycerol-di-phosphatidyl-cardiolipin (Gro-DPCL) species (RT 29-34 min.), that elute in 5 main peaks. Predominantly cardiolipin species present in peak 1: Gro-DPCL 60:2 (*m/z* 1291.87 [M-H]$^-$); peak 2: Gro-DPCL 62:2 (*m/z* 1319.90 [M-H]$^-$); peak 3: Gro-DPCL 64:2 (*m/z* 1347.93 [M-H]$^-$); peak 4: Gro-DPCL 66:2 (*m/z* 1375.96 [M-H]$^-$); peak 5: Gro-DPCL 67:2 (*m/z* 1389.98 [M-H]$^-$).

A glycerol-dependent dynamic equilibrium between PG and Gro-DPCL

Next, the activity of MhCls towards the bacterial equivalents of the archaeal lipids was further tested *in vitro*. To perform the measurements under optimal conditions, adequate membrane reconstitution of MhCls is required for which a bacterial lipid/detergent molar ratio-profile was recorded for the substrate PG (Figure S8). Subsequently, MhCls activity was monitored in the presence of PG, which resulted in the formation of Gro-DPCL, showing again that the enzyme accepts both bacterial and archaeal phospholipids as a substrate. Similar to AG, utilization of PG not only resulted in the production of Gro-DPCL, but also phosphatidic acid (PA). Initially most of the PG is converted into Gro-DPCL, although some PA is produced as well (Figure 4A, purple lines). Eventually the PG levels reach a plateau, and Gro-DPCL levels start to drop, concomitantly with the continuous production of PA. In the presence of a high concentration (100 mM) of glycerol (Figure 4A, green lines), the initial conversion of PG into Gro-DPCL is similar, but production of PA is significantly reduced, while PG reaches a plateau level at a higher concentration. This indicates that glycerol most likely stimulates the formation of PG in a reverse transesterification reaction, and thereby competes with the hydrolysis of Gro-DPCL (Figure 4D). To confirm this hypothesis, the same reaction was performed with Gro-DPCL as substrate. In the presence of 100 mM glycerol (Figure 4B, green lines), Gro-DPCL is initially predominantly converted into PG, while only low levels of PA are noted. In contrast, in the absence of glycerol, PA formation is stimulated, while lower levels of PG are observed (Figure 4B, purple lines). These data demonstrate that PA is formed together with PG by hydrolysis of Gro-DPCL (Figure 4D).

Hydrolysis of a Gro-DPCL should yield stoichiometric amounts of PA and PG, but released PG can be directly re-utilized for production of this cardiolipin species. To examine the reaction in more detail and to further address the transesterification step, PG and Gro-DPCL, both with a different acyl-chain composition, were introduced. This enables the detection of a single lipid type as a substrate or a product, based on the detected mass. Incubation of MhCls with an equimolar ratio of palmitoyl-oleoyl phosphatidylglycerol (POPG 16:0/18:1) and di-oleoyl-di-oleoyl-glycerol-di-phosphatidyl-cardiolipin (Gro-DPCL 18:1/18:1/18:1/18:1), resulted in the expected production of palmitoyl-oleoyl-palmitoyl-oleoyl-glycerol-di-phosphatidyl-cardiolipin (Gro-DPCL 16:0/18:1/18:1/16:0), di-oleoyl phosphatidylglycerol (DOPG 18:1/18:1) and di-oleoyl phosphatidic acid (DOPA 18:1/18:1), but also in the production of palmitoyl-oleoyl phosphatidic acid (POPA 16:0/18:1), demonstrating that the enzymatic reaction occurs in both directions (Figure 4C). Moreover, a Gro-DPCL species with the mixed 16:0/18:1 and 18:1/18:1 acyl-chain configuration could be identified as well, which is the synthesis product of POPG combined with DOPG. The latter originated from the reverse-transesterification reaction of Gro-DPCL 18:1/18:1/18:1/18:1, thereby showing the dynamic character of the PG - Gro-DPCL equilibrium (Figure 4D). On the other hand, the accumulation of PA suggests that this reaction eventually becomes unidirectional as PA cannot be further used. Indeed, when POPA 16:0/18:1 was added to a reaction containing DOPG 18:1/18:1, the MhCls-mediated reaction resulted in the formation of only Gro-DPCL 18:1/18:1/18:1/18:1 and DOPA, whereas no POPG or Gro-DPCL 16:0/18:1/16:0/18:1 was detected, showing that PA cannot be reutilized by MhCls (Figure S9).

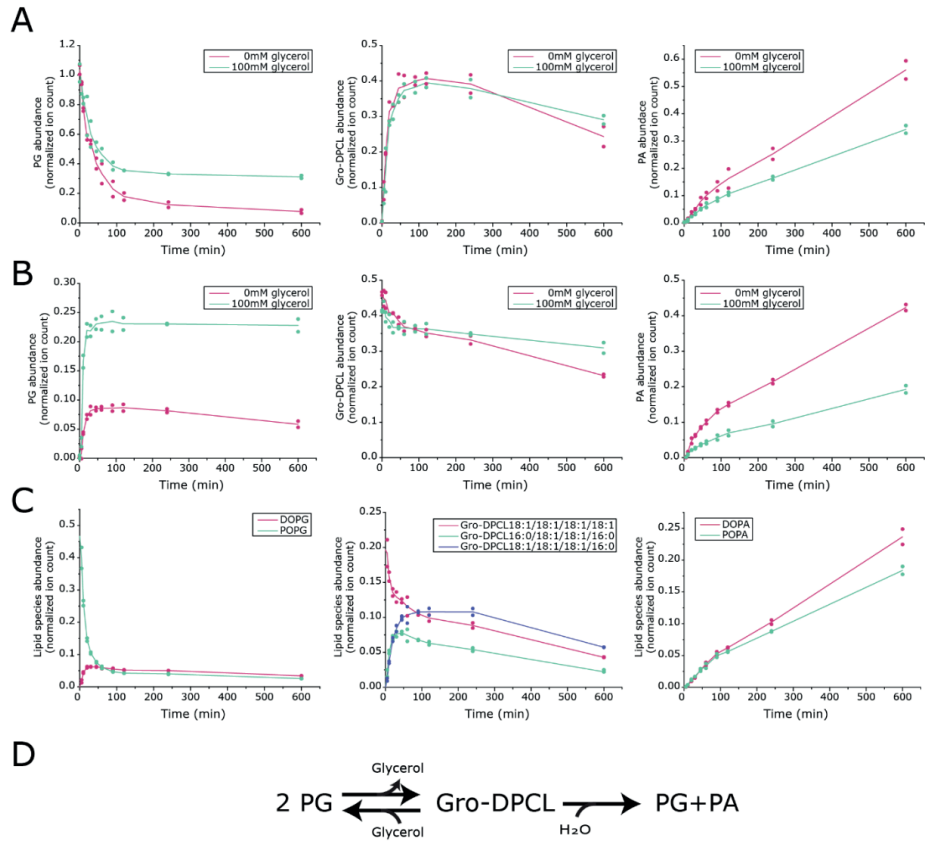


Figure 4: MhCls activity in the presence or absence of glycerol starting with various substrate(s). (A) di-oleoyl-phosphatidylglycerol (DOPG), **(B)** glycerol-di-phosphatidyl-cardiolipin (Gro-DPCL) 18:1/18:1/18:1/18:1 and **(C)** palmitoyl-oleoyl-phosphatidylglycerol (POPG) together with Gro-DPCL 18:1/18:1/18:1/18:1. The formed lipid species PG, Gro-DPCL and phosphatidic acid (PA) were analyzed by LC-MS, normalized for the internal standard DDM and ion counts are plotted on the y-axis. Lines represent the mean of the data points (n = 2). **(D)** Schematic representation of the MhCls mediated glycerol-dependent dynamic equilibrium in Gro-DPCL formation.

Formation of a bacterial-archaeal hybrid cardiolipin species

Like most archaeal lipids, AG consists of two isoprenoid chains that are coupled to a glycerol-1-phosphate (G1P) backbone via an ether-bond, whereas PG, present in Bacteria and Eukaryotes, exists of fatty acid tails that are ester-linked to a glycerol-3-phosphate (G3P). This makes that the glycerol backbone of the archaeal AG has the opposite chirality compared to bacterial/eukaryotic PG. To further examine the lipid specificity of MhCls, its activity was examined in the presence of a mixture containing both PG and AG (Figure 5A). Remarkably, not only Gro-DPCL and Gro-DACL were produced, but an additional cardiolipin species was detected: *m/z* 1488.16 $[M-H]^-$, that contains one set of isoprenoid lipid-tails and one set of fatty acid lipid-tails (Figure 5B-C; for fragmentation data see Figure S7). This lipid species represents a unique archaeal-bacterial hybrid glycerol-archaetidyl-phosphatidyl-cardiolipin (Gro-APCL). Moreover, Gro-APCL seems to be the major produced CL-species, while about the same amount of ions are detected for Gro-DACL and Gro-DPCL, suggesting that there is no clear preference for any of the lipid substrates PG or AG.

Diverse polar headgroup incorporation

In the MhCls mediated conversion of Gro-DPCL, either glycerol or H_2O is utilized. This raises the question which other molecules can be used by this enzyme. Therefore, molecules structurally related to glycerol were tested in the Gro-DPCL transesterification reaction (Figure 6A). In the presence of 1-propanol, Gro-DPCL consumption resulted in the production of substantial amounts of phosphatidyl-1-propanol (P-1-PrOH). On the other hand, the addition of 2-propanol resulted only in low levels (~1%) of phosphatidyl-2-propanol (P-2-PrOH), illustrating the importance of a primary hydroxyl (–OH) group for the transesterification. Similar to 1-propanol, 1,2-propanediol and 1,3-propanediol functioned together with Gro-DPCL as substrates for MhCls, resulting in the formation of phosphatidyl-1,2-propanol (P-1,2-PrOH) and phosphatidyl-1,3-propanol (P-1,3-PrOH), respectively. However, in the case of 1,3-PrOH, a clear additional ion of *m/z* 1440.20 $[M-H]^-$ could be identified, which corresponds to the molecule 1,3-propanediol-di-phosphatidyl-cardiolipin (1,3-PrOH-DPCL), a Gro-DPCL analogue containing a propanediol head group instead of a glycerol. This remarkable cardiolipin species could have only been formed if P-1,3-PrOH functioned as a phosphatidyl-acceptor instead of PG in the cardiolipin forming reaction.

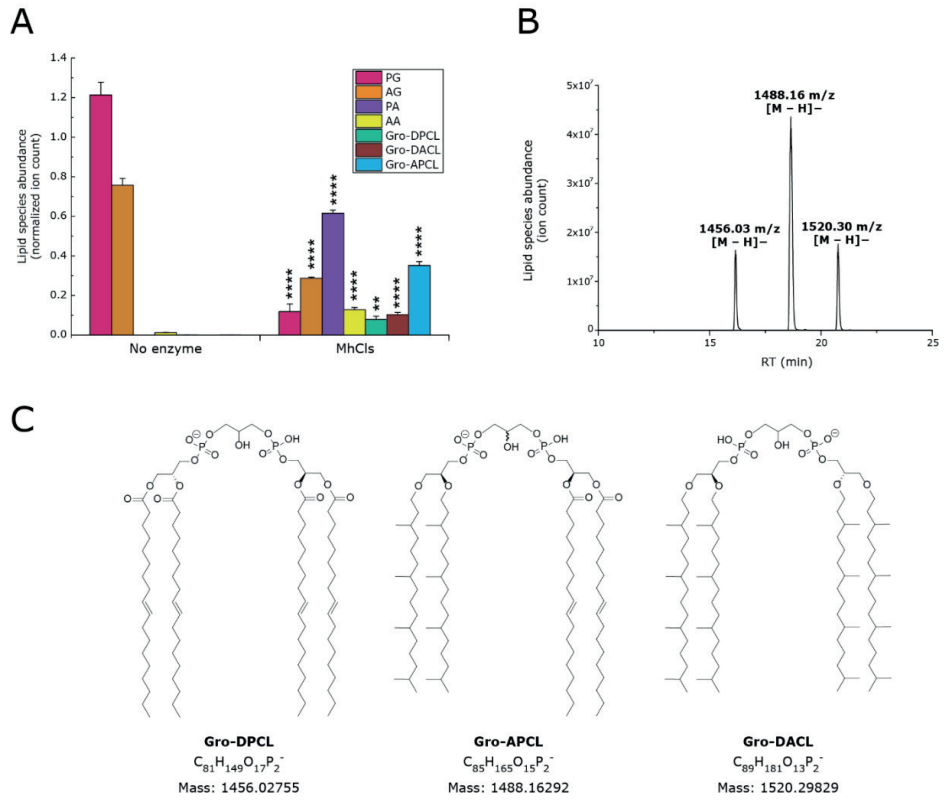


Figure 5: Synthesis of a bacterial-archaeal hybrid cardiolipin species. (A) Activity of the archaeal MhCls in the presence of both archaetidylglycerol (AG) and phosphatidylglycerol (PG). Lipid species PG, AG, phosphatidic acid (PA), archaetidic acid (AA), glycerol-di-phosphatidyl-cardiolipin (Gro-DPCL), glycerol-archaetidyl-phosphatidyl-cardiolipin (Gro-APCL), and glycerol-di-archaetidyl-cardiolipin (Gro-DACL). were analyzed by LC-MS, normalized for the internal standard, and plotted. Data are mean ± SD (n = 3). Statistical significance is shown for the enzymatic reaction (MhCls) compared to the control (no enzyme), for each individual lipid species, by using the Student's t-test analysis; **p ≤ 0.01; ****p ≤ 0.0001. **(B)** LC-MS chromatogram showing the separation of the produced bacterial Gro-DPCL, hybrid Gro-APCL, and archaeal Gro-DACL. **(C)** Structures of the three cardiolipin species. Note that in Gro-APCL, the presence of the archaetidyl-group and the phosphatidyl-group makes the central carbon atom of the glycerol head group a chiral center.

As in the presence of 1,2-propanol only trace amounts of 1,2-PrOH-DPCL could be detected, it seems that a second primary hydroxyl group is essential for the formation of cardiolipins. This was further tested with the substrates 1,2-ethanediol, 1,3-butanediol, and 1,4-butanediol. All three molecules functioned as a substrate for MhCls in the Gro-DPCL consuming reaction, which resulted in the production of the respective diester phospholipids, P-1,2-ethanediol (P-1,2-EtOH), P-1,3-butanediol (P-1,3-BuOH, and P-1,4-butanediol (P-1,4-BuOH). Moreover, for all these reaction conditions a cardiolipin equivalent could be detected as well. However, in the presence of 1,3-butanediol, instead of 1,4-butanediol, only 15% of butanediol-di-phosphatidyl-cardiolipin (BuOH-DPCL) could be produced, indicating that a primary hydroxyl group is preferred over a secondary hydroxyl group as a phosphatidyl-acceptor. This was further confirmed by the incorporation of mannitol (a six-carbon polyol), which resulted in the production of both phosphatidyl-mannitol (P-mannitol) and mannitol-di-phosphatidyl-cardiolipin (mannitol-DPCL), respectively. These experiments do not only show that the enzyme can utilize a wide variety of primary alcohols, but also exemplify its versatility. By starting with an isomerically pure, symmetric, DPCL-species and a substrate containing two primary hydroxyl groups, up to four additional lipid species can be synthesized (Figure 6B). The number of species can be further increased by using an isomerically pure, asymmetric DPCL-species instead, in which also variations in the acyl chain configuration contribute.

Next, we introduced molecules similar to glycerol, but with different bulky side-chains at the C-2 atom (Figure 7A). The presence of two methyl groups or a phenyl group at the C-2 position of the propanediol did not prevent these compounds from being used as a substrate, which resulted in the production of phosphatidyl-2,2-dimethyl-1,3-propanediol (P-2,2-Me-1,3-PrOH and phosphatidyl-2-phenyl-1,3-propanediol (P-2-Phe-1,3-PrOH). However, only trace amounts of the respective cardiolipin equivalents could be formed, which indicates that these phosphatidyl-alcohol lipids do not function as suitable phosphatidyl-acceptors, possibly because of the steric hindrance of the substituents on the lipid headgroup. Furthermore, a series of primary alcohols of biological relevance were tested. Herein, choline, L-serine, and ethanolamine were included in the reaction. In the case of choline and L-serine, the phospholipid species phosphatidylcholine (PC) and phosphatidylserine (PS) were formed together with PG and PA (Figure 7B). However, in the presence of ethanolamine, no phosphatidylethanolamine (PE) was formed. Notably, the production of PA and PG was largely abolished as well, suggesting that ethanolamine is an inhibitor of MhCls. An inhibiting effect was also observed in the presence of 3-amino-propanol (Figure S10A). Likewise, MhCls is unable to synthesize Gro-DPCL or PA from a reaction mixture of PG in the presence of either 3-amino-propanol or ethanolamine, suggesting that primary amines inhibit the enzyme (Figure S10B). Finally, the sugar inositol was also tested even though this molecule does not contain any primary hydroxyl groups, and as expected no phosphatidylinositol (PI) was found.

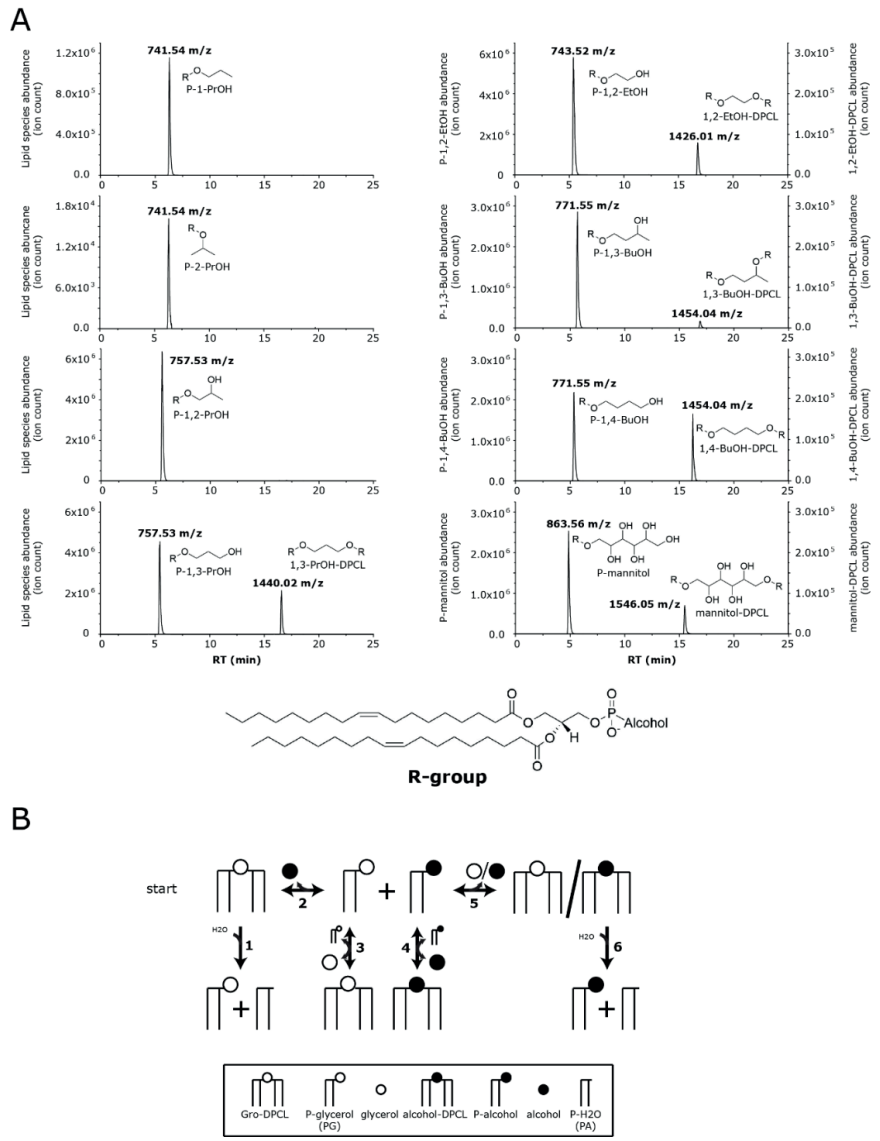


Figure 6: MhCls activity in the presence of various alcohols I. (A) Glycerol-like alcohols: 1-propanol (1-PrOH), 2-propanol (2-PrOH), 1,2-propanediol (1,2-PrOH), 1,3-propanediol (1,3-PrOH), 1,2-ethanediol (1,2-EtOH), 1,3-butanediol 1,3-BuOH), 1,4-butanediol (1,4-BuOH) and mannitol. Lipid species are displayed as LC-MS chromatograms. **(B)** Schematic representation of all possible enzymatic reactions performed and products formed by MhCls, starting with glycerol-di-phosphatidyl-cardiolipin (Gro-DPCL) and a free alcohol containing 2 primary hydroxyl groups as substrates (left top). Reaction 1 and 6: hydrolysis of Gro-DPCL and alcohol-DPCL, respectively. Reactions 2,3,4 and 5: reversible transesterification of Gro-DPCL/alcohol-DPCL in the presence of glycerol/alcohol into P-alcohol/P-glycerol.

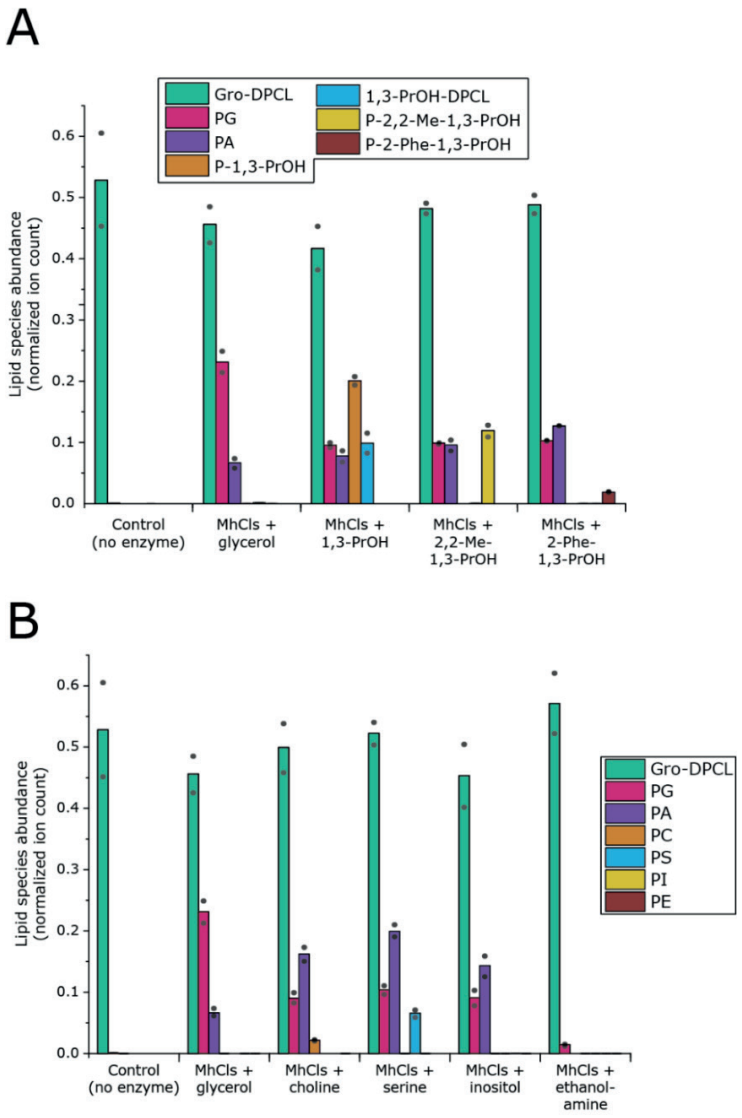


Figure 7: MhCls activity in the presence of various alcohols II. (A) 1,3-propanediol (1,3-PrOH) and its derivatives: glycerol (1,2,3-propanetriol), 2,2-dimethyl-1,3-propanediol (2,2-Me-1,3-ProH) and 2-phenyl-1,3-propanediol (2-Phe-1,3-PrOH). **(B)** Common polar lipid headgroup alcohols: choline, serine, inositol, ethanolamine. The lipid species glycerol-di-phosphatidyl-cardiolipin (Gro-DPCL), phosphatidylglycerol (PG), phosphatidic acid (PA), and all other phosphatidyl-alcohol (P-alcohol) lipid species were analyzed by LC-MS, normalized for the internal standard DDM, and plotted. Bars represent the mean of the data points (n = 2).

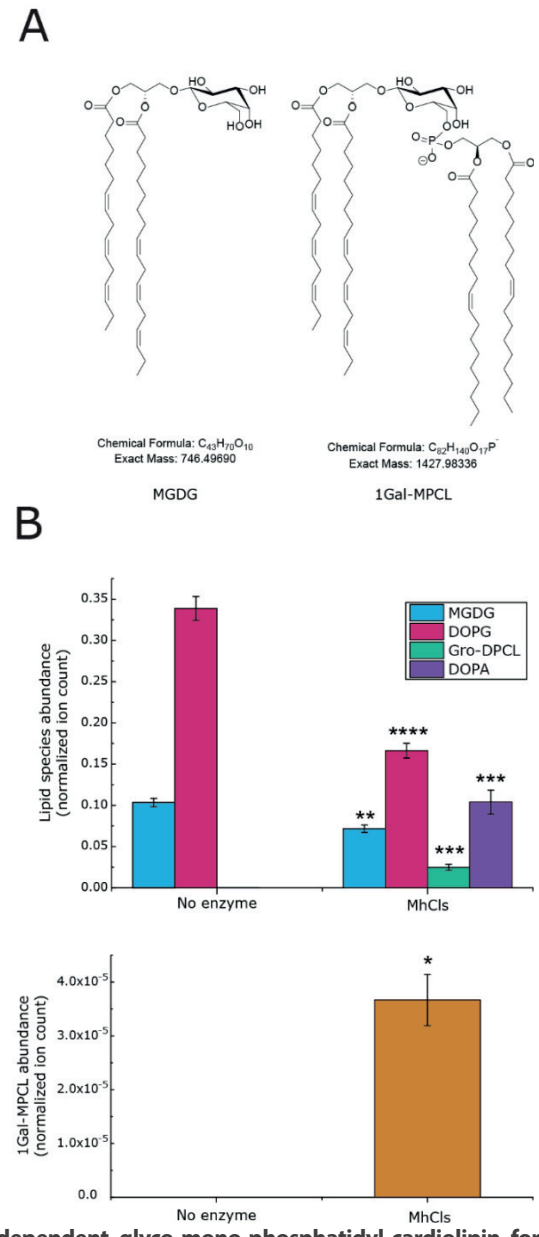


Figure 8 MhCls-dependent glyco-mono-phosphatidyl-cardiolipin formation. (A) Structures of glycolipid mono-galactosyl-diacylglycerol (MGDG) and 1-galactosyl-mono-phosphatidyl-cardiolipin (1Gal-MPCL). **(B)** MhCls mediated formation of glycerol-di-phosphatidyl-cardiolipin (Gro-DPCL), di-oleoyl-phosphatidic acid (DOPA) and 1Gal-MPCL from the substrates MGDG and di-oleoyl-phosphatidylglycerol (PG). Lipid species were analyzed by LC-MS, normalized for the internal standard DDM, and plotted. Data are mean ± SD (n = 3). Statistical significance is shown for the enzymatic reaction (MhCls) compared to the control (no enzyme), for each individual lipid species, by using the Student's t-test analysis; *p ≤ 0.05; **p ≤ 0.01; ***p ≤ 0.001; ****p ≤ 0.0001.

Glycocardiolipin formation

Since archaea also contain glycosyl-mono-archaetidyl-cardiolipin (glyco-MACL) species, the question arises if MhCls could also catalyze their synthesis [181]. These molecules basically consist of a glycolipid ester-bonded to a phospholipid, thus containing only one instead of two phosphate moieties and a sugar headgroup instead of a glycerol (Figure 8A). To examine the ability of MhCls to make a glyco-MACL/MPCL, monogalactosyldiacylglycerol (MGDG), a glycolipid species present in the thylakoid membrane of higher plant chloroplasts, was tested as a possible substrate together with PG. Noteworthy, the MGDG used is a mixture of natural lipids with different fatty acid compositions. For simplicity we focused on the most abundant MGDG species (65-70%) with the acyl-chain configuration: 16:3-18:3 (*m/z* 745.49 $[M-H]^-$). To promote glyco-MPCL formation over Gro-DPCL, MGDG was added in a two-fold excess compared to PG. However, since MGDG is a non-bilayer forming lipid, an excess of PC was also added to ensure bilayer formation for enzyme reconstitution, which resulted in a molar PG:MGDG:PC ratio of 1:2:17. In the presence of MhCls, the expected products, Gro-DPCL and PA were formed in substantial amounts with concomitant utilization of PG (Figure. 8B). Furthermore, small amounts of another compound (*m/z* 1427.98 $[M-H]^-$, mass error: 0.91 ppm) could be detected, corresponding to the glyco-MPCL species mono-galactosyl-mono-phosphatidyl-cardiolipin (1Gal-MPCL) (Figure 8B). Although the signal is low, it is clearly detectable and elutes during the expected retention time range. MGDG elutes earlier from the column compared to PG, and therefore it is expected that the retention time of 1Gal-MPCL is also shorter than that of Gro-DPCL (Figure S11). Moreover, this mass is not present in the control condition without enzyme (Figure 8B), or without MGDG (data not shown), confirming that this molecule can only be formed by MhCls in the presence of MGDG.

Discussion

Cardiolipins are lipid species found in the membranes of all domains of life. In contrast to Eukaryotes and Bacteria, the enzymes responsible for cardiolipin biosynthesis have not been studied in Archaea. Here, we identified the cardiolipin synthase from *Methanospirillum hungatei* (MhCls), which is a member of the phospholipase D superfamily [195], and characterized its function. Archaeal phospholipids consist of isoprenoid tails that are ether-linked to a glycerol-1-phosphate backbone. As a result, their chemical composition and chirality differ from the bacterial and eukaryotic lipids made of fatty acid-tails that are coupled to a glycerol-3-phosphate backbone via an ester-bond. Nevertheless, MhCls indiscriminately uses AG or PG to generate Gro-DACL and Gro-DPCL, respectively. This promiscuous feature of the enzyme enables it to simultaneously utilize PG and AG, resulting in the production of a novel archaeal-bacterial hybrid-cardiolipin species, Gro-APCL, that contains one archaetidyl- and one phosphatidyl-moiety bridged by a glycerol headgroup. So far, such a molecule has not been observed in natural membranes, which can be attributed to an evolutionary event in which our last universal common ancestor (LUCA) evolved into the domains of Archaea and Bacteria. This event is also known as the "lipid divide" and distinguishes both domains with respect to the chemical composition and chirality of their membrane phospholipids [412,413]. The ability of MhCls to synthesize Gro-DACL, Gro-DPCL and Gro-APCL implies that critical substrate recognition only involves the polar headgroup of AG and PG [71,273]. Our data further shows that MhCls can substitute for the cardiolipin synthases in cardiolipin synthesis when expressed in the *E. coli clsABC* null strain. This demonstrates that the enzyme also has the predicted activity *in vivo*. Preferentially, the role of MhCls should be tested in *M. hungatei*, but there are no genetic tools that enable gene inactivation for this strictly anaerobic organism.

During the *in vitro* formation of these cardiolipin species, substantial levels of archaetidic acid (AA) and phosphatidic acid (PA) were noted, that emerged from the hydrolytic degradation of Gro-DACL and Gro-DPCL, respectively. As a consequence, cardiolipin synthases may act in lipid remodeling. For instance, the PA produced through the hydrolysis of Gro-DPCL, can be re-utilized by the enzyme CDP-diacylglycerol synthase (CdsA), for the formation of other phospholipid species [414,415]. The same accounts for the produced (fully-saturated) AA, which might be recycled back into the lipid biosynthesis route by CDP-archaeol synthase CarS [416]. However, *in vitro* we show that in the presence of only MhCls, PA cannot be re-utilized and accumulates in time, eventually depleting the PG - Gro-DPCL pool.

Our data show that MhCls catalyzes a glycerol-dependent dynamic equilibrium between its product Gro-DPCL, and the substrate PG. However, the enzyme exhibits a remarkable substrate promiscuity towards the lipid head group. Besides glycerol (and H_2O), MhCls can incorporate various other substrates in the cardiolipin-utilizing reaction, yielding a wide variety of phosphatidyl-containing lipid species. Various primary alcohols can be attached to a phosphatidyl-group in the Gro-DPCL utilizing reaction, which results in the formation of PG and the specific phosphatidyl-alcohol. This is illustrated by the ability of MhCls to incorporate ethanediol, a two-carbon-diol, as well as the six-carbon-polyol mannitol, showing the enzyme its flexibility towards the length of the carbon-chain. Moreover, 1,3-propanediol derivatives with varying bulky side-groups at the second carbon can serve as phospholipid head group as well. Besides that, compounds with more biological relevance were tested, in which PC and PS could be formed during the conversion of Gro-DPCL in the presence of choline and serine, respectively. However, substrates that contain a primary amine (e.g., ethanolamine), act as inhibitors of the enzyme both in the cardiolipin biosynthetic and hydrolytic reaction. In addition, some phosphatidyl-alcohol species that are synthesized from a substrate that contain a second primary hydroxyl-group, can be further converted into a Gro-DPCL analogue with the alcohol as bridging head group, thereby forming atypical and non-natural cardiolipin species. The versatility of MhCls is further exemplified by the production of the glycosyl-mono-phosphatidyl-cardiolipin 1Gal-MPCL from the glycolipid substrate MGDG and the phospholipid substrate PG. This molecule was only detected in small amounts, indicating that MGDG (originating from plants) is a poor substrate, but the data support the notion that this enzyme can produce a glycocardiolipin and that their synthesis might not involve a separate class of enzymes. Thus in halophiles, the identified glycosyl-mono-archaetidyl-cardiolipin species likely arise from a reaction that involves archaetidylglycerol and the respective glycolipid precursor [160,168,177].

Our *in vitro* assays show that MhCls can perform a wide variety of catalytic reactions. They are all based on the reversible transfer of a primary alcohol to a phosphatidyl-group. The implications of this promiscuity for the biological function of MhCls is as yet unclear. So far, no cardiolipin species have been reported in the *M. hungatei* lipidome [417,418], but those lipidomics studies were performed under growth conditions where the *MhCls* gene is barely expressed [419]. This is not an uncommon finding, as in many bacteria, and insofar studied, archaea; the cardiolipin synthases are predominantly expressed during specific conditions (e.g. osmotic shock, specific growth phase, etc.), yielding different levels of cardiolipins [161,186,403,404]. In this respect, the observed reversibility of the phosphatidyl-transfer makes the abundance of cardiolipins flexible, which could aid in the environmental response. Moreover, the adaptive ability of the membrane might be diversified with the promiscuous behavior of MhCls, illustrated by the ability of the enzyme to accept a wide variety of primary alcohols and lipids, which could be a general feature for PLD-type cardiolipin synthases [183,192,193,200].

Finally, the promiscuity of MhCls could be utilized for bioengineering purposes. As an example, the ability of this enzyme to incorporate a wide variety of non-natural polar head groups into phospholipid species may lead to new bio-catalytic applications for the synthesis of unique phospholipid species. Furthermore, MhCls potentially could be used for the bottom-up construction of a synthetic cellular membrane, in which this enzyme could diversify the phospholipid head group composition of an expanding phospholipid bilayer [414].

Materials and methods

Bioinformatic identification of MhCls

Using *E. coli* K12; MG1655 ClsA (EcClsA: NP_415765.1), ClsB (EcClsB: WP_187790083), or ClsC (EcClsC: WP_188006884.1) as query sequences, BLAST homology searches to the domain of Archaea were performed with the following result: EcClsA: query coverage 75-98%, and sequence identity 23-31%; EcClsB: query coverage 76-88%, and sequence identity 25-33%; EcClsC: query coverage 69-91%, and sequence identity 21-28%. The BLAST results were further analyzed using MEGA X and filtered for sequences that contain at least two HKD domains. Next, sequences were aligned using the MUSCLE algorithm (default settings) and a phylogenetic tree was estimated using the LG+G model [420]. Subsequently, a putative archaeal cardiolipin synthase (MhCls: WP_011448254) from *Methanospirillum hungatei* JF-1 was selected. The consensus sequence logo was created with the program Weblogo *(https://weblogo.berkeley.edu/logo.cgi),* for which the group of methanogens was selected together with a group of bacterial EcClsA homologs (see supplemental information). This resulted in a group of 90 species (57 bacteria and 33 archaea), from which one archaeal sequence, containing many additional amino acids within the second hydrophobic domain, was removed.

Bacterial strains and cloning procedures

An *E. coli* codon-optimized synthetic gene of *M. hungatei* Cls (MhCls) was ordered (GeneArt, Thermo Fisher scientific) and used as a template for the amplification of MhCls, during which a N-terminal 6xHis-tag was added. The 6His-MhCls fragment was cloned into a pRSF-Duet expression vector, using the NcoI and SacI restriction

enzymes and T4 DNA ligase, resulting in pRSF-6His-MhCls (pNDK001). *E. coli* DH5α (Invitrogen) was used as a host for Cloning procedures. *E. coli* Lemo21 (DE3) was used as the overexpression strain for pNDK001. The 6His-MhCls fragment was also cloned into a pBAD expression vector, using the EcoRI and HindIII restriction enzymes and T4 DNA ligase, resulting in pBAD-6his-MhCls (pME006). *E. coli* DH5α (Invitrogen) was used as a host for Cloning procedures. *E. coli clsABC* null strain (K12 Δ*clsA*, Δ*clsB*, Δ*clsC*::*kanR* (BKT12)) [186] was used as the overexpression strain for pME006. All primers and plasmids used in the present study are listed in Supplemental Table 1 and supplemental Table 2. All strains were grown under aerobic conditions at 37°C in LB medium supplemented with the required antibiotics: kanamycin (50 µg/ml), chloramphenicol (34 µg/ml), Zeocin™ (25 µg/ml).

Expression and purification of MhCls

MhCls was overexpressed in *E. coli* Lemo21 (DE3) strain in the presence of 250 µM rhamnose and induced with 0.5 mM isopropyl β-D-1-thiogalactopyranoside (IPTG). After 2.5 hours of induction, cytoplasmic and membrane fractions were separated as described [101]. The total membranes were resuspended in buffer A (50 mM Tris/HCl, pH 8.0, 100 mM KCl and 15% glycerol) after which they could be stored at -80°C. For further purification, 0.5 mg/ml of membranes were solubilized in 2% n-dodecyl-β-D-maltoside (DDM) detergent for 1 hr. at 4°C. The material was subjected to a centrifugation (17,000 x g) step for 15 min at 4°C to remove insolubilized material and the supernatant was incubated with Ni-NTA agarose beads (Qiagen, cat: 30230) for 2 hrs. at 4°C. The Ni-NTA beads were washed 10 times with 6 column volumes (CV) of buffer B (50 mM Tris/HCl, pH 8.0, 100 mM KCl, 15% glycerol and 0.05% DDM) supplemented with 10 mM imidazole, and the protein was eluted three times with 0.5 CV of buffer B supplemented with 300 mM imidazole. To remove the imidazole and glycerol, the purified protein was passed over a Zebra™ Spin Desalting column 40K; 0.5 ml (Thermo scientific), and eluted in buffer C (50 mM MES pH 7.0, 100 mM KCl and 0.05% DDM). Purity of the eluted protein was assessed on 15% SDS/PAGE stained with Coomassie Brilliant Blue and the protein concentration was determined by measuring the absorbance at 280 nm and calculating the molar concentration using the calculated extinction coefficient. Extinction coefficients were obtained from the ProtParam tool from the ExPASy website (https://web.expasy.org/protparam/).

Total lipid extraction

Lipids were extracted from 10 mg freeze-dried cells from *E. coli clsABC* null strain pellets using an adapted Bligh and Dyer method employing 5% trichloroacetic acid as described elsewhere [329]. The crude chloroform fraction was dried and the lipid film was re-extracted with 400 µl chloroform - methanol (1:2), dried again and finally re-extracted with 50 µl methanol.

Preparation of liposomes

Chloroform stocks of the lipid species DOPG, POPG, DOPA, POPA, DOPE, DOPC, Gro-DPCL 18:1/18:1/18:1/18:1 and MGDG were purchased from Avanti (Avanti Polar Lipids, Birmingham, AL). The chemical synthesis of AG was performed in house and is described in detail in the supporting information. For liposomes with a heterogeneous lipid mixture, the required amount of lipid chloroform stocks were mixed together in the stated molar ratio. Next the lipid solution was dried under a nitrogen gas stream for multiple hours, after which the dry lipid film was resuspended in a 50 mM 2-(N-morpholino)ethanesulfonic acid (MES) buffer, pH 7.0 yielding a translucent suspension. For formation of liposomes, a probe sonicator was employed (30 seconds cycle time with a 50% duty cycle for 10-20 cycles) until the suspension became transparent.

In vitro assays for phospholipid synthesis

All *in vitro* reactions were performed in 100 µl of buffer D containing a final concentration of 50 mM MES pH 7.0 and 100 mM KCl in the presence of 1 µM MhCls. The activity of MhCls with archaeal substrate was assayed in the presence of 250 µM AG (*sn*1-*sn*1 and *sn*1-*sn*3; ratio 1:1) and 0.4 mM DDM. The glycerol-dependent dynamic equilibrium of MhCls was assayed in the presence of 1.8 mM DDM, and either 1 mM DOPG, 0.5 mM Gro-DPCL, or 0.5 mM POPG together with 0.25 mM DPCL. The activity of MhCls with AG and PG (molar ratio 1:1) was assayed with 250 µM of each lipid substrate in the presence of 0.8 mM DDM. The promiscuity toward primary-hydroxyl-containing compounds was assayed with 0.5 mM Gro-DPCL or 1 mM DOPG, 100 mM primary-hydroxyl-containing substrate and 1.8 mM DDM. Formation of 1Gal-MPCL was performed in the presence of 1 µM MhCls, 0.5 mM lipid (PG:MGDG:PC, molar ratio 1:2:17) and 0.8 mM DDM. All reactions were incubated overnight at 37°C unless stated differently. Lipids were extracted from the reaction mixtures two times with 0.3 ml of 1-butanol, and evaporated under a stream of nitrogen gas and resuspended in 50 µl of methanol for LC-MS analysis.

LC–MS analysis of lipids

Samples from the *in vitro* reactions were analyzed using an Accela1250 high-performance liquid chromatography (HPLC) system coupled with a Heated electrospray ionization mass spectrometry (HESI–MS) Orbitrap Exactive (Thermo Fisher Scientific). A sample of 5 µl was injected into an ACQUITY UPLC® CSH™ C18 1.7 µm Column, 2.1x150 mm (Waters Chromatography Ireland Ltd) operating at 55°C with a flow rate of 300 µl/min [366]. Separation of the compounds was achieved by a changing gradient of eluent A (5 mM ammonium formate in water/acetonitrile 40:60, v/v), and eluent B (5 mM ammonium formate in 1-butanol/acetonitrile/water, 90:10:1, v/v). The following linear gradient was applied: 45% eluent B for 2.5 minutes; a gradient from 45% to 90% eluent B over 19.5 minutes; holding for 3 minutes; returning to 45% eluent B in 0.5 minutes; and holding for 8 minutes. For the *E. coli clsABC* null strain total lipid extracts, 10 µl was injected and the following linear gradient was applied: 5% eluent B for 2.5 minutes; a gradient from 5% to 90% eluent B over 36.5 minutes; holding for 3 minutes; returning to 5% eluent B in 0.5 minutes; and holding for 8 minutes. The column effluent was injected directly into the Exactive ESI-MS Orbitrap operating in negative ion mode. Voltage parameters of 3 kV (spray), −75 V (capillary), −190 V (tube lens) and -46 V (Skimmer voltage) were used. Capillary temperature of 300°C, sheath gas flow 60, and auxiliary gas flow of 5 was maintained during the analysis.

Spectral data constituting total ion counts were analyzed using the Thermo Scientific XCalibur processing software by applying the Genesis algorithm based automated peak area detection and integration. The total ion counts of the extracted lipid products (Supplemental Table 3) were normalized for DDM (*m/z* 509.3 $[M-H]^-$) and plotted on the y-axis as normalized ion count in a bar graph.

Acknowledgments

We would like to thank Dr. M. Bogdanov for kindly providing us with the *E. coli K12 clsABC* null strain.

Funding information

This work was supported and funded by the 'BaSyC – Building a Synthetic Cell' Gravitation grant (024.003.019) of the Netherlands Ministry of Education, Culture and Science (OCW) and the Netherlands Organization for Scientific Research (NWO), and by the Dutch NWO Building Blocks of Life programme (737.016.006).

Author contributions

Conceptualization, M.E., N.A.W.d.K., and A.J.M.D.; Investigation: M.E., N.A.W.d.K., R.L.H.A., N.H.J.W.; Resources: A.J.M., A.J.M.D.; Writing –Original Draft: M.E., R.L.H.A.; Writing –Review & Editing: M.E., N.A.W.d.K., A.J.M., A.J.M.D.; Visualization: M.E., R.L.H.A., N.A.W.d.K.; Supervision, M.E., A.J.M., A.J.M.D.; Project Administration, M.E.; Funding Acquisition, A.J.M.D., A.J.M.

Conflicts of interest

The authors declare that they have no conflicts of interest with the contents of this article.

List of abbreviations

AA	archaetidic acid
AG	archaetidylglycerol
CDP-DAG	cytidine diphosphate diacylglycerol
CL	cardiolipin (lipid class)
Cls	cardiolipin synthase
ClsA	cardiolipin synthase A
ClsB	cardiolipin synthase B
ClsC	cardiolipin synthase C
DDM	n-dodecyl-β-d-maltoside
DOPA	di-oleoyl phosphatidic acid
DOPC	di-oleoyl phosphatidylcholine
DOPE	di-oleoyl phosphatidylethanolamine
DOPG	di-oleoyl phosphatidylglycerol
DOPI	di-oleoyl phosphatidylinositol
DOPS	di-oleoyl phosphatidylserine
glyco-MACL	glycosyl-mono-archaetidyl-cardiolipin
glyco-MPCL	glycosyl-mono-phosphatidyl-cardiolipin
Gro-APCL	glycerol-archaetidyl-phosphatidyl-cardiolipin
Gro-DACL	glycerol-di-archaetidyl-cardiolipin
Gro-DPCL	glycerol-di-phosphatidyl-cardiolipin
LUCA	last universal common ancestor
mannitol-DPCL	mannitol-di-phosphatidyl-cardiolipin
MGDG	monogalactosyldiacylglycerol
MhCls	*Methanospirillum hungatei* cardiolipin synthase
Ni-NTA	Nickel-nitrilotriacetic Acid
PA	phosphatidic acid
PG	phosphatidylglycerol
POPA	palmitoyl-oleoyl phosphatidic acid
POPG	palmitoyl-oleoyl phosphatidylglycerol
P-mannitol	phosphatidyl-mannitol
P-NH2-PrOH	phosphatidyl-aminopropanol
P-1-PrOH	phosphatidyl-1-propanol
P-1,2-EtOH	phosphatidyl-1,2-ethanol
P-1,2-PrOH	phosphatidyl-1,2-propandiol
P-1,3-BuOH	phosphatidyl-1,3-butanediol
P-1,3-PrOH	phosphatidyl-1,3-propandiol
P-1,4-BuOH	phosphatidyl-1,4-butanediol
P-2-Phe-1,3-PrOH	phosphatidyl-2-phenyl-1,3-propanediol
P-2-PrOH	phosphatidyl-2-propanol
P-2,2-Me-1,3-PrOH	phosphatidyl-2,2-dimethyl-1,3-propanediol
S-DGD-5-PA	S-diphytanylglycerol diether-5-phosphatidic acid
S-2Glyco-aMPCL	S-di-glycosyl-archaeal mono-phosphate cardiolipin
S-2Glyco-DGD	S-di-glycosyl diphytanylglycerol diether
S-3Glyco-aMPCL	S-tri-glycosyl-archaeal mono-phosphate cardiolipin
S-GL-2	S-glycosylcardiolipin-2
S-TGD-1-PA	S-tri-glycosyl-diether-1-phosphatidic acid
1Gal-MPCL	monogalactosyl-mono-phosphatidyl-cardiolipin
1,2-EtOH-DPCL	1,2-ethanol-di-phosphatidyl-cardiolipin
1,2-PrOH-DPCL	1,2-propanediol-di-phosphatidyl-cardiolipin
1,3-BuOH-DPCL	1,3-butanediol-di-phosphatidyl-cardiolipin
1,3-PrOH-DPCL	1,3-propanediol-di-phosphatidyl-cardiolipin
1,4-BuOH-DPCL	1,4-butanediol-di-phosphatidyl-cardiolipin

Supplemental information

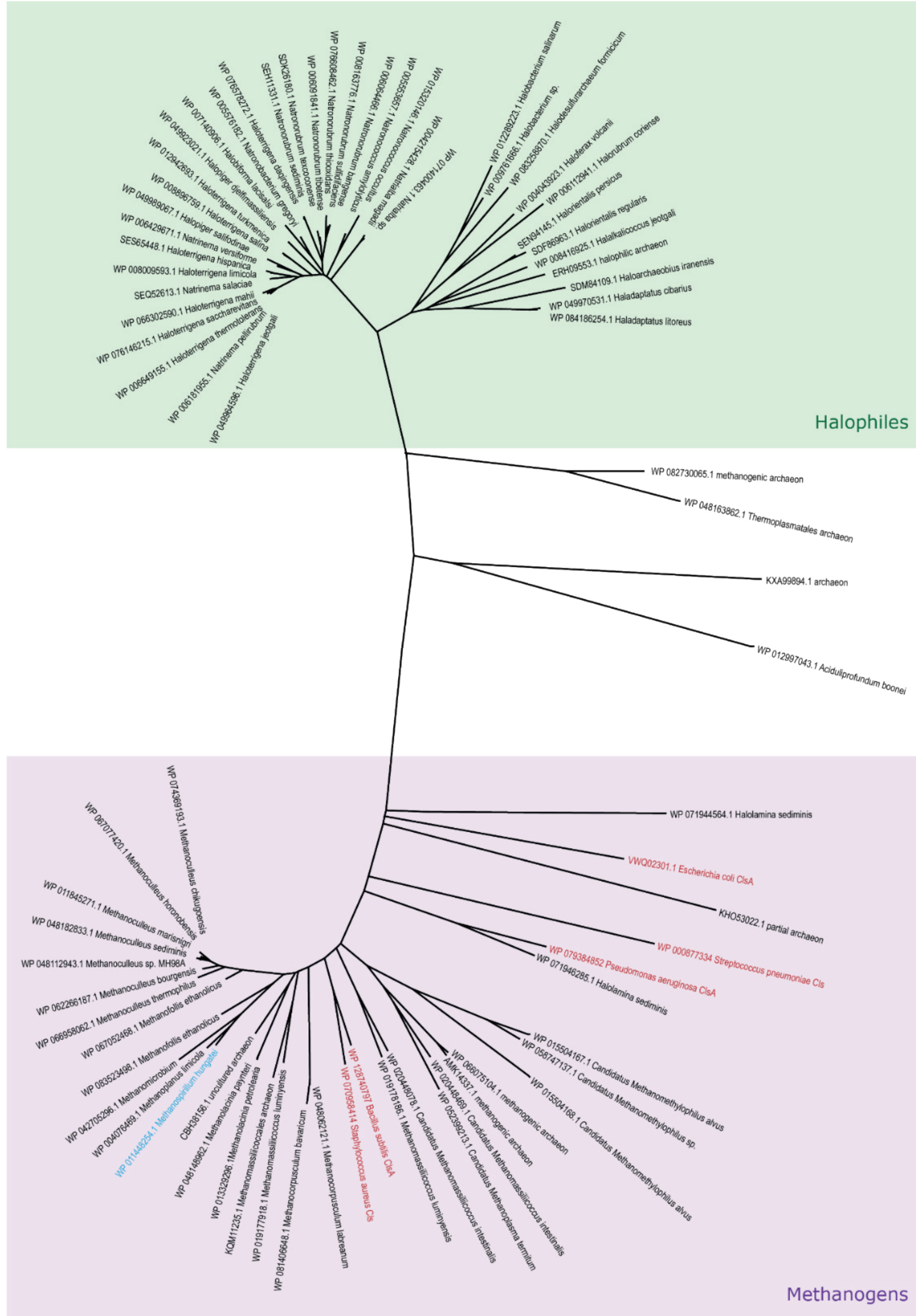
Additional information on the identification of putative archaeal cardiolipin synthases

PLD-type Cls homologs

The three Cls enzymes (ClsA, ClsB, and ClsC) present in the bacterium *Escherichia coli* were used as template in individual BLAST searches to identify putative archaeal cardiolipin synthases. For these enzymes hits were selected for EcClsA: query coverage 75-98%, and sequence identity 23-31%; EcClsB: query coverage 76-88%, and sequence identity 25-33%; EcClsC: query coverage 69-91%, and sequence identity 21-28%. As the hits obtained with EcClsA showed the best query coverage and sequence identity, it was decided to only continue with the results from EcClsA. Next, the results were filtered for sequences containing at least two HKD motifs, which is a common feature among PLD-type Cls enzymes. Additionally, some cardiolipin species from bacteria were added. Subsequently the sequences were aligned and a phylogenetic tree was estimated using the LG+G model (Figure S1) [420].

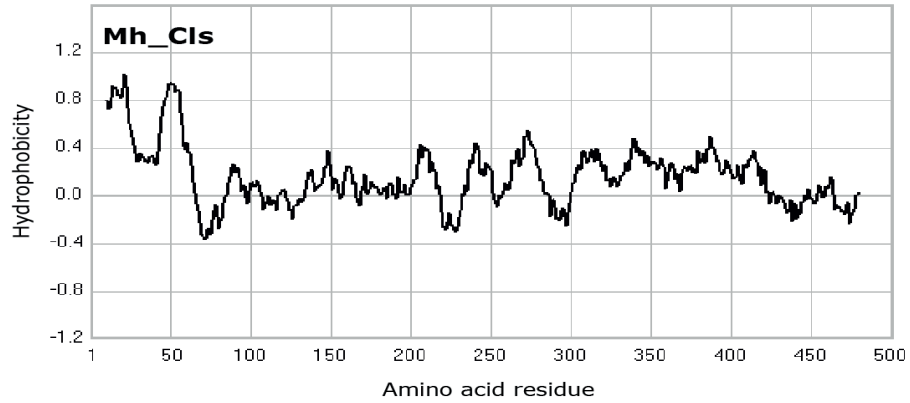
The putative archaeal cardiolipin synthases are divided into two main clusters. One cluster mostly exists of methanogens, whereas the other cluster predominantly contains halophiles that belong to the class of Haloarchaea (Halobacteria). The latter cluster is further subdivided into two groups. An alignment of *Haloferax volcanii* with *Haloterrigena salina* (each representing one of the two clusters), shows there is a query coverage of 84% in which there is a sequence identity of 44%.

To investigate the identity between the methanogenic and halophilic clusters, an alignment of the putative Cls from *Methanospirillum hungatei* and *H. volcanii* was made. The alignment has a query coverage of only 32% in which there is a sequence identity of 26%, clearly indicating two distinct groups. Noteworthy, the bacterial cardiolipin synthases are all located in the methanogenic cluster, in which *E. coli* ClsA seems to be most distinct. An alignment of *E. coli* ClsA with *M. hungatei* Cls shows a query coverage of 95% in which there is a sequence identity of 27%, whereas an alignment of *E. coli* ClsA with *H volcanii* Cls only shows 11% query coverage containing a sequence identity of 27%. These alignments indicate that the methanogenic cluster of putative archaeal cardiolipin synthases are relatively closely related to the bacterial ClsA type, in contrast to the halophilic cluster.



Supplemental figure 1: Unrooted tree of putative archaeal cardiolipin synthases that contain two HKD motifs (PLD-type Cls). Representative bacterial species are highlighted in red and locate in the methanogenic cluster. The putative Cls from Methanospirillum hungatei (MhCls) that is characterized in this work is highlighted in blue.

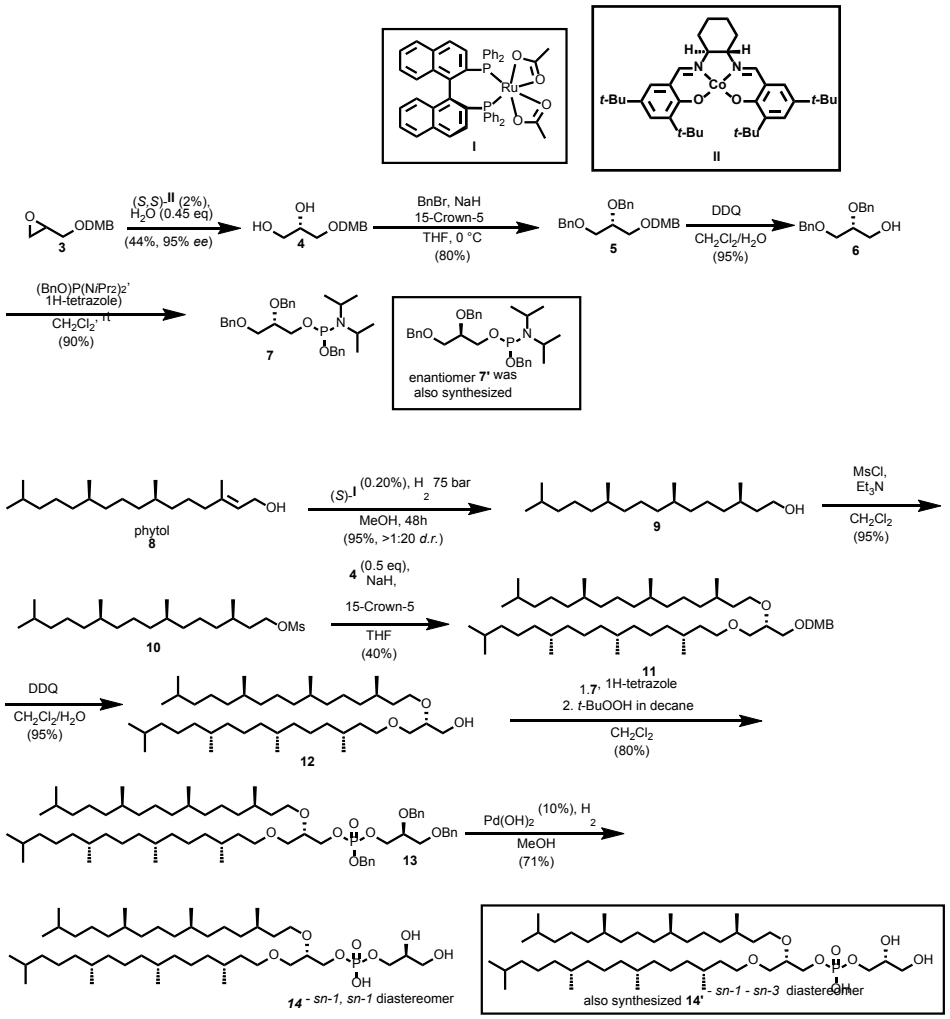
Supplemental figure 2: Alignment of *Escherichia coli* ClsA with a selection of bacterial and archaeal (putative) cardiolipin synthases. ClsA from *Escherichia coli* (NP_415765.1) was aligned together with ClsA from *Bacillus subtilis* (WP_128740797), a cardiolipin synthase from *Staphylococcus aureus*, and four putative archaeal cardiolipin synthases from *Methanospirillum hungatei* (WP_011448254), *Methanoculleus bourgensis* (WP_014868452), *Haloferax volcanii* (WP_004043923), *Halobacterium salinarum* (WP_012289223), which represent the methanogenic and halophilic clusters presented in figure 1C. The light grey boxes outline the two conserved HKD motifs. The second hydrophobic domain (Figure 1D) is conserved among the bacterial and methanogenic cardiolipin synthases, but not in the halophiles. * Indicates strictly conserved amino acid positions.



Supplemental figure 3: Hydropathy profile of MhCls.

AG synthesis

We synthesized both the *sn*1-*sn*1 (**14**) and the *sn*1-*sn*3 (**14'**) diastereomers of AG (Archaetidyl glycerol), as shown in Scheme 1. Both enantiomers of the glycerol headgroup were synthesized from dimethoxybenzyl-protected glycidol via the Jacobsen hydrolytic kinetic resolution, obtaining high yields and *ee*. Benzylation of the two free hydroxy groups followed by deprotection of the DMB-group gave **6** and **6'**, which were converted into their respective benzyloxy-phosphoramidites **7** and **7'**. It should be noted that several attempts were made to prepare **7** with DCI (dicyanoimidazole), a more recently introduced phosphoramidite coupling catalyst, however we observed superior yields with solid 1H-tetrazole



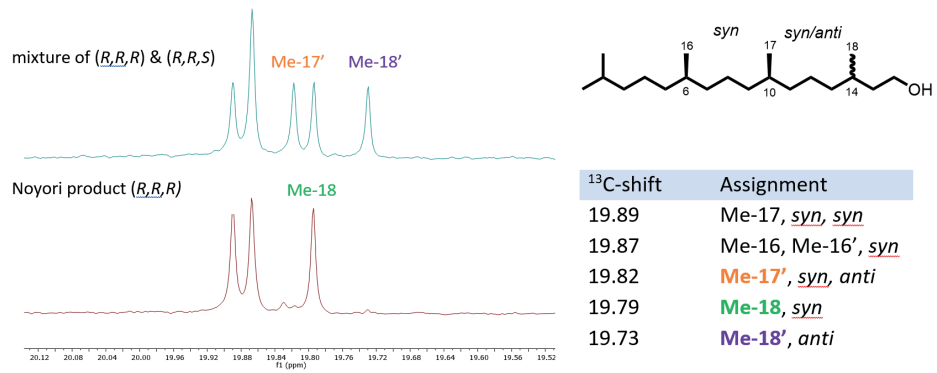
Scheme 1: Total synthesis of the *sn1-sn1* and *sn1-sn3* diastereomers of AG (DDQ: Dicyanodichlorobenzoquinone, 15-C-5: 15-crown-5-ether, DMB: dimethoxybenzyl).

The phytanyl unit could be readily prepared from commercially available *trans*-phytol, as described by Sita *et al.* [421]. It was envisioned that the archaeal diether lipid precursor **12** would be accessible via dialkylation of DMB (dimethoxybenzyl) protected glycerol **(4)**. Dialkylation of glycerol derivatives is very challenging due to steric hindrance of the secondary alkoxide, leading to competing elimination of the electrophile. The highest yields were obtained in our hands utilizing the corresponding mesylate and adding crown ether. Smith and co-workers speculate that α-chelating groups make very stable chelates with the alkoxide salt [422]. This

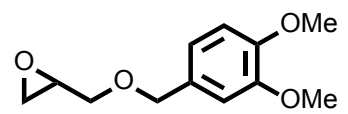
could explain the increased yield due to disassociation of the chelate when crown ether is added. The DMB group was readily cleaved under oxidative conditions giving the free alcohol in high yield.

The headgroup and lipid fragments were coupled via a similar phosphoramidite coupling, and subsequently oxidized to the phosphate with *t*-BuOOH in one pot. Reductive cleavage of the benzyl groups yielded the desired AGs in good yield.

We were able to determine the selectivity of the ruthenium catalyzed hydrogenation by ^{13}C-NMR-spectroscopy by comparing a diastereomeric mixture of phytanol (*R,R,R + R,R,S)* with the product of the Noyori asymmetric hydrogenation (*R,R,R*) which is shown in Figure S4. The characteristic signal of methyl **18** for the *R,R,R* diastereomer is found at 19.79 ppm and for the *R,R,S* diastereomer at 19.73 ppm. The signal pattern observed was comparable to that observed in previous studies on 1-5 *syn*/*anti* methyl systems performed by Curran and co-workers [423]. Integration of the signals showed a >20:1 selectivity for the all *syn*-methyl-phytanol. Sita observed a similar selectivity by HPLC analysis of chiral carbamate derivatives.



Supplemental figure 4: d.r. determination of phytanol via the comparison of the diastereomeric mixture to the enriched sample employing the Curran method [423].



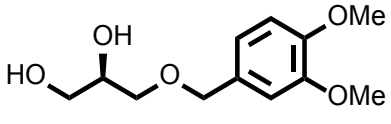
2-(((3,4-dimethoxybenzyl)oxy)methyl)oxirane (3). To a 100 mL 3-necked flask equipped with magnetic stirrer bar was added 25 mL of a 50% NaOH solution, epichlorohydrin (18.5 g, 15.6 mL, 0.2 mol) and Bu_4NHSO_4 (1.5 mmol, 525 mg, 4 mol%). The resulting solution was cooled to 0°C (ice/water-bath) after 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 to rt 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%, 7.9 g, 35.2 mmol).

^{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.

The spectral data correspond to that previously reported [424].



***(S)*-3-((3,4-dimethoxybenzyl)oxy)propane-1,2-diol (4).** [425] A 25 mL flask equipped with a magnetic stirrer bar was charged with (*R,R*)-II (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, 9.1 mmol).[4] The remaining epoxide (50%) was recovered and could be enriched to 95% *ee* (*R*-enantiomer) with (*S,S*)-II as catalyst.

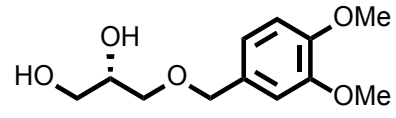
Chiral HPLC analysis on a Lux® 5 µm Cellulose-3 column, *n*-heptane : *i*-PrOH = 90 : 10, 40°C, flow = 1 mL/min, UV detection at 274 nm, t_R(major): 25.29 min, t_R(minor): 29.06 min, 97% ee

^{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).

4 The diol was isolated together with small amounts of the catalyst co-eluting on the column. For optical rotation determination the diol was separated from the catalyst by distillation, yielding a colourless oil. For further reactions the initial column purification was found to be sufficient.

^{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 in $CHCl_3$).



***(R)*-3-((3,4-dimethoxybenzyl)oxy)propane-1,2-diol (4').** 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 (*S,S*)-**II** as catalyst (45% yield).

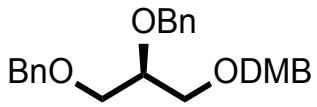
Chiral HPLC analysis on a Lux® 5 µm Cellulose-3 column, *n*-heptane: *i*-PrOH = 90 : 10, 40°C, flow = 1 mL/min, UV detection at 274 nm, t_R(minor): 26.01 min, t_R(major): 29.21 min, 95% ee.

^{1}H NMR (400 MHz, Chloroform-*d*): Same as reported for compound **4**

^{13}C NMR (101 MHz, Chloroform-*d*): Same as reported for compound **4**

$[\alpha]_D^{20}$ = -2.4 (c = 0.1 in $CHCl_3$).

The spectral data correspond to previously reported [426].

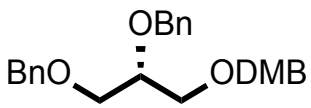


***(R)*-3-((3,4-dimethoxybenzyl)oxy)propane-1,2-di-benzylether (5).**
Di-methoxybenzyl protected glycerol **4** (500 mg, 2.06 mmol), benzyl bromide (0.736 mL, 6.19 mmol, 3 equiv.) and 15-crown-5 (1.4 mL, 7.22 mmol, 3 equiv.) were dissolved in THF (4.1 mL) and cooled to 0°C. NaH (60% in mineral oil, 250 mg, 6.19 mmol, 3 equiv.) was added to the mixture in 3 portions over 10 min. After 5 h the reaction was quenched by the careful addition of NH_4Cl_{aq} (sat), the layers were separated and the aqueous layer washed with ether (3x). The organic layers were combined, dried over $MgSO_4$, filtered and concentrated *in vacuo*. The crude oil was further purified by column chromatography (10% ether in pentane), which yielded the desired product as a thick clear oil (80%, 690 mg, 1.62 mmol).

^{1}H NMR (400 MHz, Chloroform-*d*) δ 7.39 – 7.22 (m, 10H), 6.91 – 6.78 (m, 3H), 4.71 (s, 2H), 4.54 (s, 2H), 4.48 (s, 2H), 3.88 (s, 3H), 3.83 (s, 3H), 3.85 – 3.79 (m, 1H), 3.74 – 3.47 (m, 4H).

^{13}C NMR (101 MHz, Chloroform-*d*) δ 149.1, 148.7, 138.8, 138.4, 131.0, 128.5, 128.4, 127.8, 127.7, 127.7, 127.6, 120.33, 111.1, 111.0, 77.4, 73.5, 73.4, 72.4, 70.5, 70.2, 56.1, 55.9.

$[\alpha]_D^{20}$ = -1.0 (c = 0.1 in $CHCl_3$).

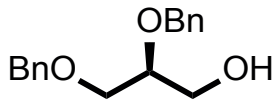


***(S)*-3-((3,4-dimethoxybenzyl)oxy)propane-1,2-di-benzylether (5').**
This compound was prepared with the same synthetic procedure that was used for **5**

^{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**

$[\alpha]_D^{20}$ = +1.1 (c = 0.1 in $CHCl_3$).



***(S)*-2,3-bis(benzyloxy)propan-1-ol (6).** Dimethoxy-benzyl-di-benzyl glycerol (**5**) (220 mg, 0.52 mmol) was dissolved in CH_2Cl_2/H_2O (9:1, 6 mL) and cooled to 0°C (ice/water bath). DDQ (142 mg, 0.625 mmol, 1.2 equiv.) was added in 3 portions over 5 min. The resulting orange biphasic solution was allowed to stir for 3 h before filtration over celite. The filtrate was concentrated and further purified by column chromatography (20% EtOAc in pentane) to yield the desired alcohol as a clear thick oil (85%, 135 mg, 0.5 mmol).

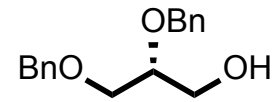
^{1}H NMR (400 MHz, Chloroform-*d*) δ 7.45 – 7.28 (m, 10H), 4.73 (H_a, d, *J* = 11.7 Hz, 1H), 4.64 (H_b, d, *J* = 11.8 Hz, 1H), 4.61 – 4.52 (m, 2H), 3.86 – 3.53 (m, 5H), 2.30 (s, 1H).

^{13}C NMR (101 MHz, Chloroform-*d*) δ 138.4, 138.1, 128.5, 128.5, 127.9, 127.9, 127.8, 127.7, 78.2, 73.6, 72.2, 70.3, 62.9.

$[\alpha]_D^{20}$ = -17.8 (c = 0.1 in $CHCl_3$).

HRMS-ESI+ (*m/z*): $[M + Na]^+$ calculated for $C_{17}H_{20}O_3Na$, 292.131; found, 292.130.

The spectral data correspond to previously reported [427].



***(S)*-2,3-bis(benzyloxy)propan-1-ol (6').** This compound was prepared with the same synthetic procedure that was used for **6**

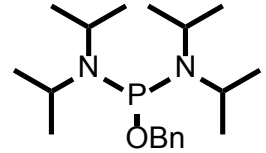
^{1}H NMR (400 MHz, Chloroform-*d*): Same as reported for compound **6**

^{13}C NMR (101 MHz, Chloroform-*d*): Same as reported for compound **6**

$[\alpha]_D^{20}$ = +18.1 (c = 0.1 in $CHCl_3$).

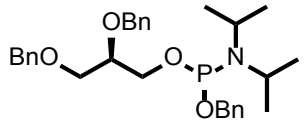
HRMS-ESI+ (*m/z*): $[M + Na]^+$ calculated for $C_{17}H_{20}O_3Na$, 292.131; found, 292.130.

The spectral data correspond to previously reported [428].



1-(benzyloxy)-N,N,N',N'-tetraisopropylphosphanediamine. To a stirring suspension of 1-chloro-N,N,N',N'-tetraisopropylphosphanediamine (1.33 g, 5 mmol) in dry Et_2O (20 mL) at 0°C, was added a solution of benzyl alcohol (0.57 mL, 5.5 mmol, 1.1 equiv.) and Et_3N (0.84 mL, 6 mmol, 1.2 equiv.) in dry Et_2O (5 mL) over 1/2 h (syringe pump). The resulting milky mixture was stirred for 3 h. The ammonium salts were filtered, the filtrate concentrated *in vacuo* and further purified by column chromatography (5% Et_3N in hexane) which yielded the desired phosphoramidite as a colorless oil (95%, 1.52 g, 4.49 mmol)

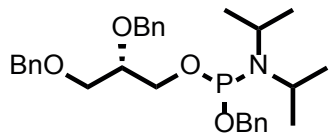
Spectral data were identical to previously reported [429].



benzyl *((R)*-2,3-bis(benzyloxy)propyl) diisopropylphosphoramidite (7). A flame dried Schlenk flask was charged with tetrazole (24 mg, 0.33 mmol, 0.95 eq), and to this was added dibenzylglycerol **6** (95 mg, 0.35 mmol) in dry toluene. The toluene was evaporated under Schlenk line vacuum while stirring vigorously, the co-evaporation with toluene was repeated three times. CH_2Cl_2 (1.2 mL) was added and the resulting solution was cooled to 0°C, after which the benzyloxy-phosphoramidite (237 mg, 0.7 mmol, 2 equiv.) in CH_2Cl_2 (0.6 mL) was added in a dropwise fashion. After 16 h, all volatiles were evaporated *in vacuo* and the crude was further purified by column chromatography (10% ether, 5% Et_3N in pentane) which yielded the desired phosphoramidite as a thick oil. (95%, 170 mg, 0.33 mmol)

^{1}H NMR (400 MHz, Chloroform-*d*) δ 7.57 – 7.18 (m, 15H), 4.85 – 4.66 (m, 4H), 4.58 (t, *J* = 2.8 Hz, 2H), 3.92 – 3.81 (m, 2H), 3.82 – 3.74 (m, 1H), 3.75 – 3.55 (m, 4H), 1.29 – 1.15 (m, 12H).

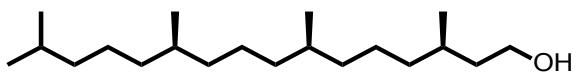
^{13}C NMR (101 MHz, cdcl$_3$) δ 138.88, 138.51, 128.41, 128.36, 128.33, 127.84, 127.79, 127.70, 127.69, 127.59, 127.55, 127.32, 127.30, 127.06, 127.04, 78.17, 78.13, 78.09, 78.06, 73.46, 72.33, 70.56, 65.48, 65.30, 63.31, 63.24, 63.15, 63.08, 43.19, 43.06, 24.80, 24.76, 24.73, 24.68.



benzyl *((S)*-2,3-bis(benzyloxy)propyl) diisopropylphosphoramidite (7'). This compound was prepared with the same synthetic procedure that was used for **7**

^{1}H NMR (400 MHz, Chloroform-*d*): Same as reported for compound **7**

^{13}C NMR (101 MHz, Chloroform-*d*): Same as reported for compound **7**



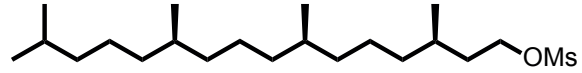
(3*R*,7*R*,11*R*)-3,7,11,15-tetramethylhexadecan-1-ol (9). To an oven dried 20 mL vial equipped with magnetic stirrer bar was added trans-phytol (1 g, 3.4 mmol) in 3 mL of dry methanol under nitrogen atmosphere. The resulting mixture was degassed by the freeze-pump-thaw technique after which (*S*)-BINAP-ruthenium(II) dicarboxylate (**II**) (6 mg, 0.07 mmol, 0.2 mol%) was added. The vial containing the reaction mixture was placed inside a 0.5 L Parr bomb and subjected to 75 bar of H_2 pressure for 72 h. The resulting brown solution was concentrated *in vacuo* and passed over a small silica column eluted with pentane. Evaporation of the solvent yielded diastereomerically pure (*R,R,R*)-phytanol as a yellow oil (96% yield, 0.97 g, d.r. >1:20, determined by NMR-spectroscopy).

^{1}H NMR (400 MHz, Chloroform-*d*) δ 3.76 – 3.61 (m, 2H), 1.99 (s, 1H), 1.73 – 1.45 (m, 3H), 1.45 – 0.96 (m, 21H), 0.95 – 0.76 (m, 15H).

^{13}C NMR (151 MHz, Chloroform-*d*) δ 61.2, 40.1, 39.5, 37.7, 37.6, 37.5, 37.5, 37.4, 32.9, 32.9, 29.7, 28.1, 24.9, 24.6, 24.5, 22.8, 22.7, 19.9, 19.9, 19.8.

$[\alpha]_D^{20}$ = +1.90 (c = 0.1 in $CHCl_3$).

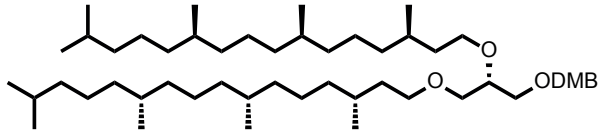
The spectral data correspond to previously reported [421].



***(3R,7R,11R)*-3,7,11,15-tetramethylhexadecyl methanesulfonate (10).** To a stirring solution of phytanol (2.99 g, 10 mmol) in CH_2Cl_2 (20 mL) at 0°C (ice/water bath) was added Et_3N (4.2 mL, 30 mmol, 3 equiv.), DMAP (611 mg, 5 mmol, 0.5 equiv.) and MsCl (0.93 mL, 12 mmol, 1.2 equiv.). After 2 h the mixture was quenched by the addition of $NaHCO_3$ (sat) and the aqueous layer was extracted with ether. All the organic layers were combined, dried over $MgSO_4$, filtered and concentrated *in vacuo*, the crude oil was further purified by column chromatography (15% ether in pentane) to give the desired mesylate as a pale oil (90%, 3.4 g, 8.95 mmol).

^{1}H NMR (400 MHz, Chloroform-*d*) δ 4.34 – 4.13 (m, 2H), 2.99 (s, 3H), 1.84 – 1.70 (m, 1H), 1.70 – 1.44 (m, 3H), 1.44 – 0.97 (m, 20H), 0.91 (d, *J* = 6.4 Hz, 3H), 0.88 – 0.80 (m, 12H).

^{13}C NMR (101 MHz, Chloroform-*d*) δ 68.7, 39.5, 37.5, 37.5, 37.5, 37.5, 37.4, 37.3, 37.2, 36.1, 32.9, 32.9, 29.5, 28.1, 24.9, 24.6, 24.3, 22.8, 22.7, 19.7, 19.4.



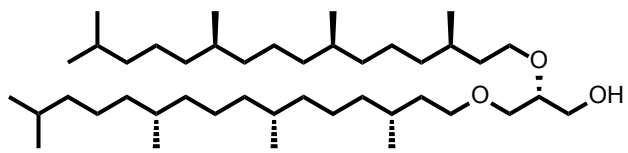
2,3-bisphytanol-*sn*-dimethoxybenzyl-glycerol (11). To a stirring solution of 15-crown-5 (0.55 mL, 2.8 mmol, 2.2 equiv.) and *sn*-1-DMB-glycerol (200 mg, 0.83 mmol) in dry THF (1.2 mL) at 0°C (ice/water bath) was added NaH (60% in mineral oil, 73mg, 1.81 mmol, 2.2 equiv.) in 4 portions over 5 min. The mixture was allowed to warm up to rt over 1/2 h, after which mesylate **10** (685 mg, 1.82 mmol, 2.2 equiv.) in dry THF (0.5 mL) was added in a dropwise fashion. After 16 h the reaction was quenched by the careful addition of NH_4Cl_{aq} (sat). The aqueous layer was extracted with Et_2O (3x), the organic layers combined, dried over $MgSO_4$, filtered and concentrated *in vacuo*. The crude oil was further purified by column chromatography (10% ether in pentane) which yielded the desired product as a pale-yellow oil (40%, 267 mg, 3.33 mmol)

^{1}H NMR (400 MHz, Chloroform-*d*) δ 6.92 – 6.79 (m, 3H), 4.48 (s, 2H), 3.87 (s, 3H), 3.86 (s, 3H), 3.69 – 3.38 (m, 9H), 1.73 – 1.46 (m, 6H), 1.44 – 0.96 (m, 42H), 0.85 (m, 30H).

^{13}C NMR (101 MHz, Chloroform-*d*) δ 149.1, 148.6, 131.1, 120.3, 111.1, 111.0, 78.1, 73.4, 71.0, 70.2, 70.1, 69.0, 56.0, 55.9, 39.5, 37.7, 37.7, 37.6, 37.5, 37.5, 37.4, 37.3, 36.8, 33.0, 32.9, 30.1, 30.0, 28.1, 24.9, 24.6, 24.5, 22.9, 22.8, 19.9, 20.0, 19.9, 19.8.

HRMS-ESI+ (*m/z*): $[M + Na]^+$ calculated for $C_{52}H_{98}O_5Na$, 825.7307; found, 825.730.

$[\alpha]_D^{20}$ = +0.70 (c = 0.1 in $CHCl_3$).



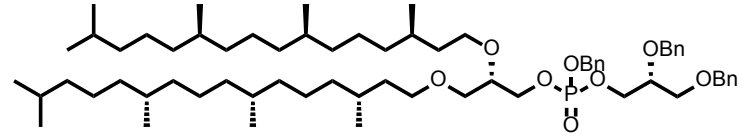
2,3-bisphytanol-*sn*-glycerol (11). 1-(3,4-dimethoxybenzyl)-2,3-bisphytanol-*sn*-glycerol (**12**) (400 mg, 0.6 mmol) was dissolved in CH_2Cl_2 (10 mL) to which water (1 mL) was added. The biphasic mixture was cooled to 0°C (ice/water bath) and DDQ (136 mg, 0.6 mmol, 1.2 equiv.) was added in three portions. The resulting green biphasic mixture was stirred for 2 h at 0°C during which it turned light brown. The mixture was quenched by the addition of $NaHCO_3$ (sat), the layers were separated and the aqueous layer was washed with ether (3x). The organic layers were combined, dried over $MgSO_4$, filtered and concentrated in vacuo. The crude oil was further 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 (90%, 290 mg, 0.44 mmol).

^{1}H NMR (400 MHz, Chloroform-*d*) δ 3.78 – 3.40 (m, 9H), 2.21 (br s, 1H), 1.56 (m, 6H), 1.43 – 0.96 (m, 42H), 0.92 – 0.75 (m, 30H).

^{13}C NMR (101 MHz Chloroform-*d*) δ 78.5, 71.1, 70.3, 68.8, 63.2, 39.5, 37.6, 37.6, 37.5, 37.5, 37.5, 37.4, 37.2, 36.7, 32.9, 30.0, 30.0, 28.1, 24.9, 24.6, 24.5, 22.9, 22.8, 19.9, 19.9, 19.9, 19.8.

HRMS-ESI+ (*m/z*): $[M + H]^+$ calculated for $C_{43}H_{89}O_3H$, 653.6812; found, 653.679.

$[\alpha]_D^{20}$ = +4.1 (c = 0.1 in $CHCl_3$).



Benzyl (2,3-bisphytanol-*sn*-glycerol)-*((R)*-2,3-bis(benzyloxy)propyl) phosphate (13). 2,3-bisphytanol-*sn*-glycerol (20 mg, 30 µmol) and tetrazole (7 mg, 92 µmol, 3 equiv.) were dissolved in dry toluene. The toluene was evaporated under Schlenk-line vacuum while stirring vigorously, this process was repeated three times. 200 µL of CH_2Cl_2 was added and the solution was cooled to 0°C (ice/water bath). Phosphoramidite **7** (39 mg, 77 µmol, 2.5 equiv.) in 100 µL CH_2Cl_2 was added in a drop-wise fashion and the mixture was stirred for 16 h during which it was allowed to warm up to rt. To the resulting mixture a 5M solution of *t*-BuOOH (60 µL, 300 µmol, 10 eq) in decane was added in a drop-wise fashion and the mixture was allowed to stir for 30 minutes. The crude mixture was poured into phosphate buffer (1 M, pH = 7) and was extracted with Et_2O (3x). The organic extracts were combined, dried over $MgSO_4$, filtered and concentrated. The crude was further purified by column chromatography (20% ether in pentane) to give the pure protected glycerol-phosphate **13** as a yellow oil (75%, 25 mg, 23 µmol)

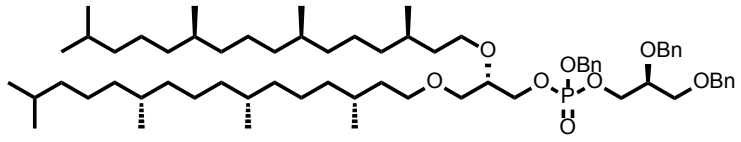
^{1}H NMR (400 MHz, Chloroform-*d*) δ 7.67 – 7.01 (m, 15H), 5.13 – 4.99 (m, 2H), 4.75 – 4.59 (m, 2H), 4.52 (d, 2H), 4.30 – 3.91 (m, 5H), 3.89 – 3.73 (m, 1H), 3.66 – 3.52 (m, 4H), 3.49 – 3.33 (m, 4H), 1.77 – 1.42 (m, 6H), 1.41 – 0.97 (m, 42H), 0.93 – 0.77 (m, 30H).

^{13}C NMR (101 MHz, Chloroform-*d*) δ 138.1, 137.98, 135.96, 135.9, 128.54, 128.53, 128.50, 128.47, 128.45, 128.43, 128.38, 128.3, 127.94, 127.85, 127.82, 127.80, 127.75, 127.7, 127.7, 127.6, 78.1, 77.3, 77.2, 76.6, 76.56, 76.52, 76.49, 73.4, 72.2, 70.1, 69.9, 69.8, 69.3, 69.23, 69.22, 69.21, 69.2, 69.0, 67.1, 67.1, 67.0, 66.9, 39.4, 37.6, 37.5, 37.5, 37.4, 37.3, 37.0, 36.6, 32.8, 30.3, 29.9, 29.8, 29.7, 28.0, 24.8, 24.5, 24.4, 22.7, 22.6, 19.8.

^{31}P NMR (162 MHz, Chloroform-*d*) δ -0.89,

HRMS-ESI+ (*m/z*): $[M + Na]^+$ calculated for $C_{67}H_{113}O_{8P}$, 1077.825; found, 1077.826.

$[\alpha]_D^{20}$ = -3.7 (c = 0.1 in $CHCl_3$).



Benzyl (2,3-bisphytanol-*sn*-glycerol)-*((S)*-2,3-bis(benzyloxy)propyl) phosphate (13'). This compound was prepared with the same synthetic procedure that was used for **13**

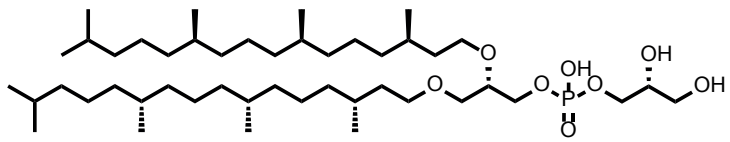
^{1}H NMR (400 MHz, Chloroform-*d*) δ 7.67 – 7.01 (m, 15H), 5.13 – 4.99 (m, 2H), 4.75 – 4.59 (m, 2H), 4.52 (d, 2H), 4.30 – 3.91 (m, 5H), 3.89 – 3.73 (m, 1H), 3.66 – 3.52 (m, 4H), 3.49 – 3.33 (m, 4H), 1.77 – 1.42 (m, 6H), 1.41 – 0.97 (m, 42H), 0.93 – 0.77 (m, 30H).

^{13}C NMR (101 MHz, Chloroform-*d*) δ 138.1, 137.98, 135.96, 135.9, 128.54, 128.53, 128.50, 128.47, 128.45, 128.43, 128.38, 128.3, 127.94, 127.85, 127.82, 127.80, 127.75, 127.7, 127.7, 127.6, 78.1, 77.3, 77.2, 76.6, 76.56, 76.52, 76.49, 73.4, 72.2, 70.1, 69.9, 69.8, 69.3, 69.23, 69.22, 69.21, 69.2, 69.0, 67.1, 67.1, 67.0, 66.9, 39.4, 37.6, 37.5, 37.5, 37.4, 37.3, 37.0, 36.6, 32.8, 30.3, 29.9, 29.8, 29.7, 28.0, 24.8, 24.5, 24.4, 22.7, 22.6, 19.8.

^{31}P NMR (162 MHz, Chloroform-*d*) δ -0.89,

HRMS-ESI+ (*m/z*): $[M + Na]^+$ calculated for $C_{67}H_{113}O_{8P}$, 1077.825; found, 1077.827.

$[\alpha]_D^{20}$ = +3.60 (c = 0.1 in $CHCl_3$).



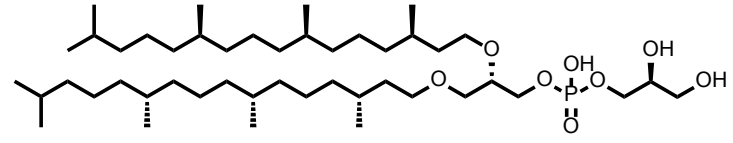
2,3-bisphytanol-*sn*-glycerol-*((R)*-2,3-dihydroxypropyl) hydrogen phosphate (14) (Figure S5). 13 was dissolved in a mixture of ethanol/THF (1:1) and was degassed via the freeze-pump-thaw technique. $Pd(OH)_2$/C (20%, 8.5 mg, 12 µmol, 1 eq) was added to the stirring solution under N_2-atmosphere. The resulting suspension was placed under H_2 atmosphere and was allowed to stir for 16 h at rt. The mixture was filtered over celite and further purified by column chromatography (100% $CHCl_3$ → 30% MeOH in $CHCl_3$, Davisil Grade 635 silica) which yielded the desired phospholipid **14** as a white film (71%, 7 mg, 9 µmol).

^{1}H NMR (400 MHz, Methanol-d_4) δ 4.05 – 3.82 (m, 2H), 3.75 (p, *J* = 6.8 Hz, 4H), 3.67 – 3.48 (m, 5H), 3.49 – 3.38 (m, 2H), 3.38 – 3.27 (m, 1H), 2.68 (s, 1H), 2.39 (s, 1H), 1.60 – 1.42 (m, 5H), 1.41 – 1.12 (m, 24H), 1.13 – 0.95 (m, 32H), 0.89 – 0.75 (m, 18H).

^{13}C NMR (151 MHz, MeOD) δ 71.5, 71.0, 70.7, 69.5, 67.0, 65.6, 62.8, 58.1, 39.9, 37.9, 37.8, 37.5, 37.2, 33.3, 32.4, 30.5, 30.4, 30.2, 29.8, 28.5, 26.0, 25.3, 25.0, 24.9, 23.1, 23.0, 20.15, 20.06.

^{31}P NMR (243 MHz, MeOD) δ -0.41.

HRMS-ESI- (*m/z*): $[M - H]^-$ calculated for $C_{46}H_{94}O_8P$, 805.669; found, 805.669.



2,3-bisphytanol-*sn*-glycerol-*((S)*-2,3-dihydroxypropyl) hydrogen phosphate (14') (Figure S6). This compound was prepared with the same synthetic procedure that was used for **14**.

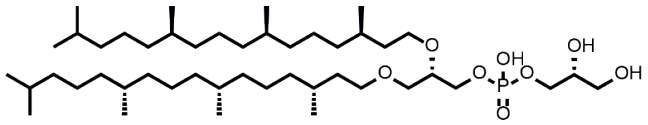
^{1}H NMR (400 MHz, Methanol-d_4) δ 4.05 – 3.82 (m, 2H), 3.75 (p, *J* = 6.8 Hz, 4H), 3.67 – 3.48 (m, 5H), 3.49 – 3.38 (m, 2H), 3.38 – 3.27 (m, 1H), 2.68 (s, 1H), 2.39 (s, 1H), 1.60 – 1.42 (m, 5H), 1.41 – 1.12 (m, 24H), 1.13 – 0.95 (m, 32H), 0.89 – 0.75 (m, 18H).

^{13}C NMR (151 MHz, MeOD) δ 71.5, 71.0, 70.7, 69.5, 67.0, 65.6, 62.8, 58.1, 39.9, 37.9, 37.8, 37.5, 37.2, 33.3, 32.4, 30.5, 30.4, 30.2, 29.8, 28.5, 26.0, 25.3, 25.0, 24.9, 23.1, 23.0, 20.15, 20.06.

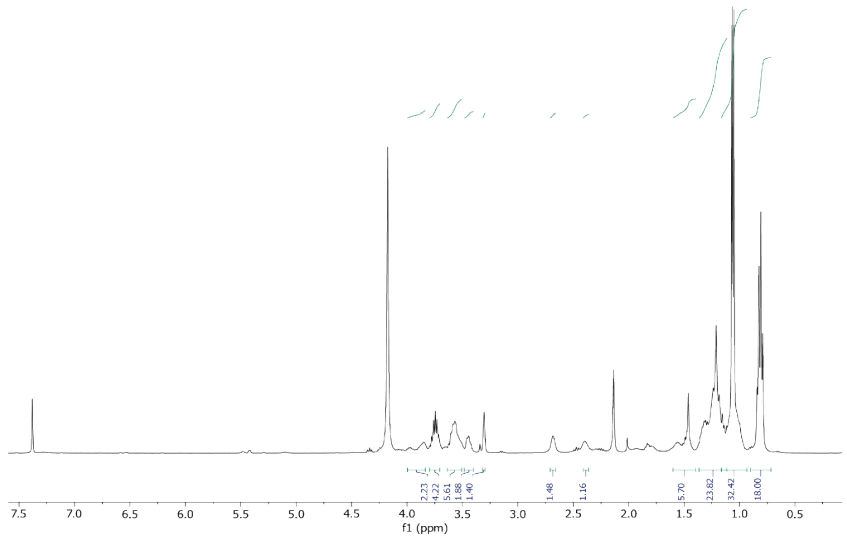
^{31}P NMR (243 MHz, MeOD) δ -0.41.

HRMS-ESI- (*m/z*): $[M - H]^-$ calculated for $C_{46}H_{94}O_8P$, 805.669; found, 805.669.

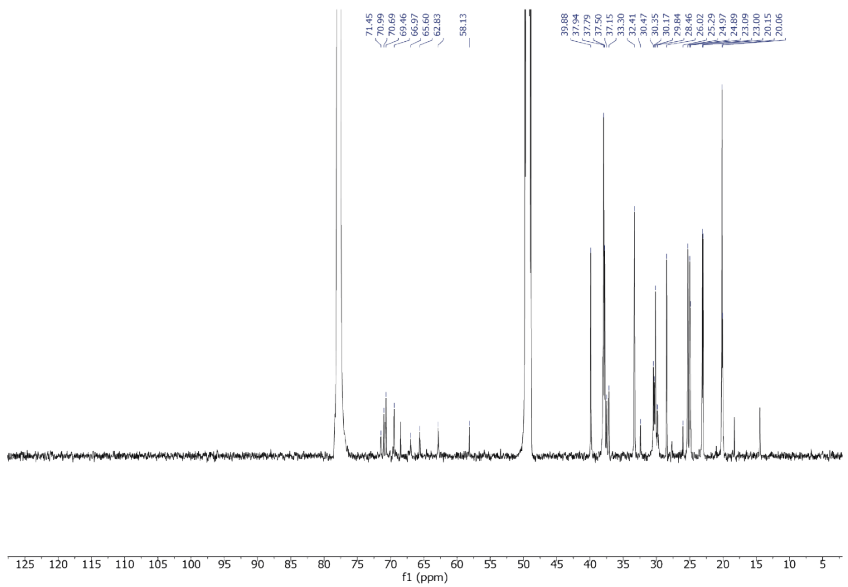
A



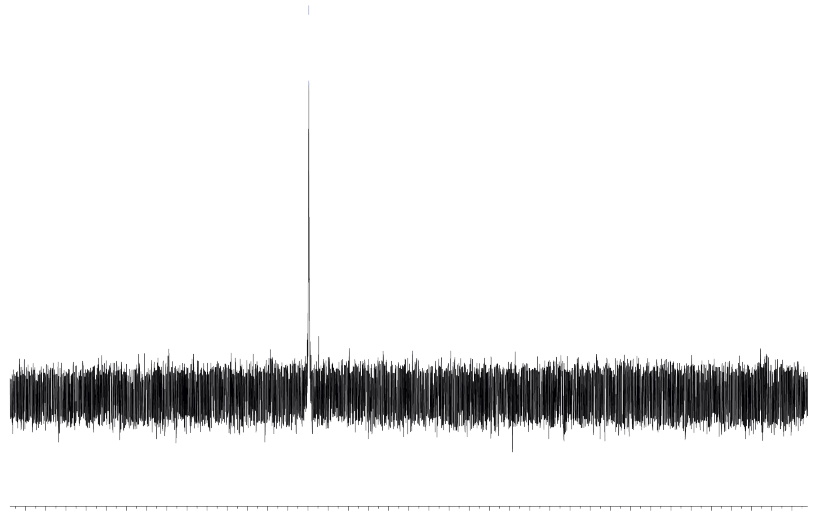
B



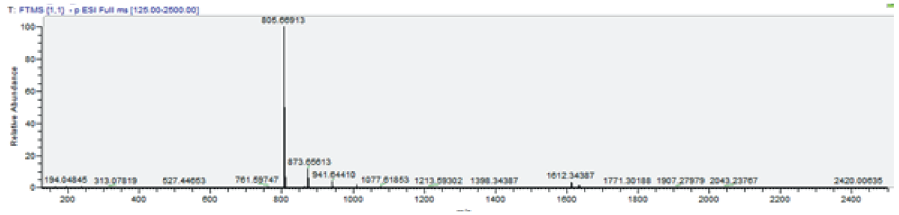
C



D

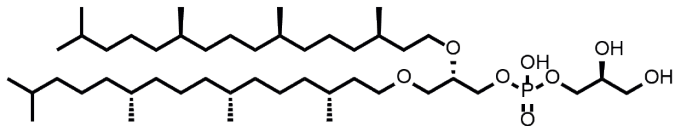


E

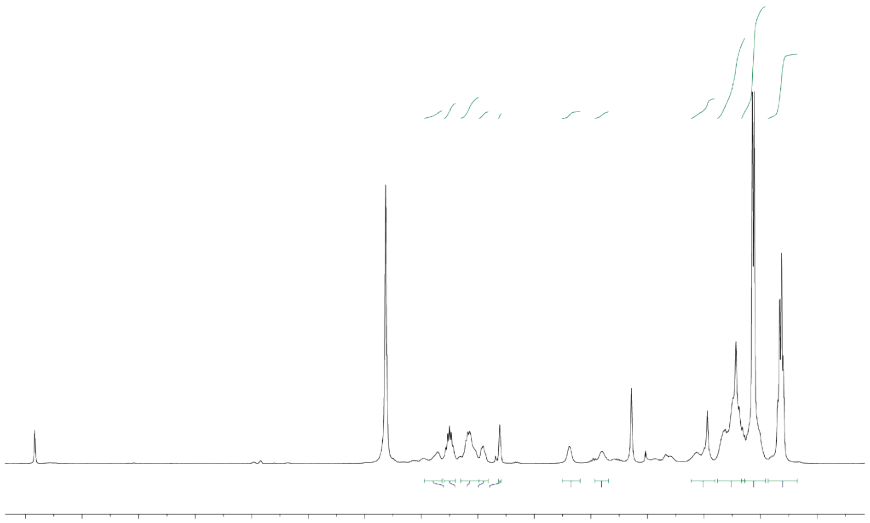


Supplemental figure 5: 2,3-bisphytanol-*sn*-glycerol-(*(R)*-2,3-dihydroxypropyl) hydrogen phosphate (14) **(A)** Structure of 14. **(B)** 1H-NMR of 14. **(C)** 13C-NMR of 14. **(D)** 31P-NMR of 14. **(E)** HRMS of 14.

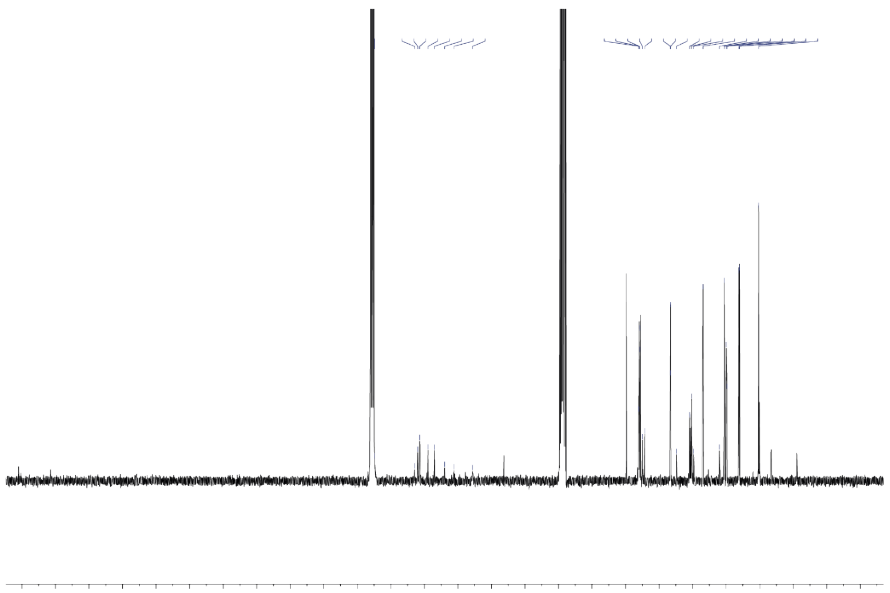
A



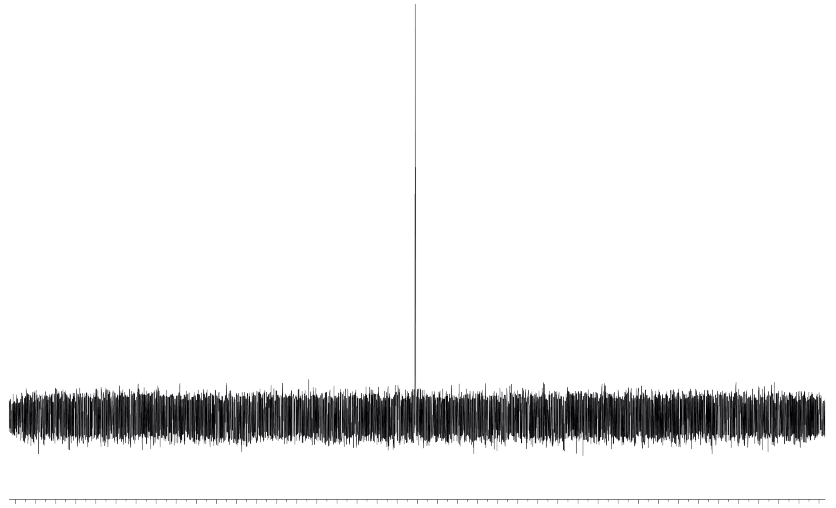
B



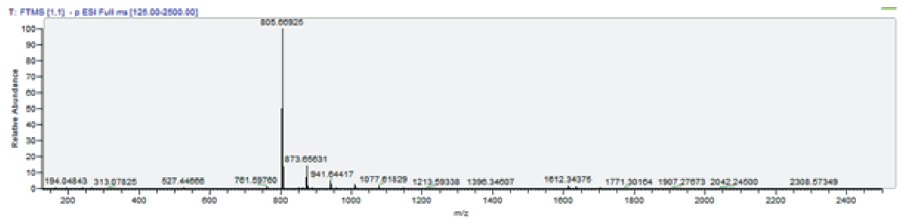
C



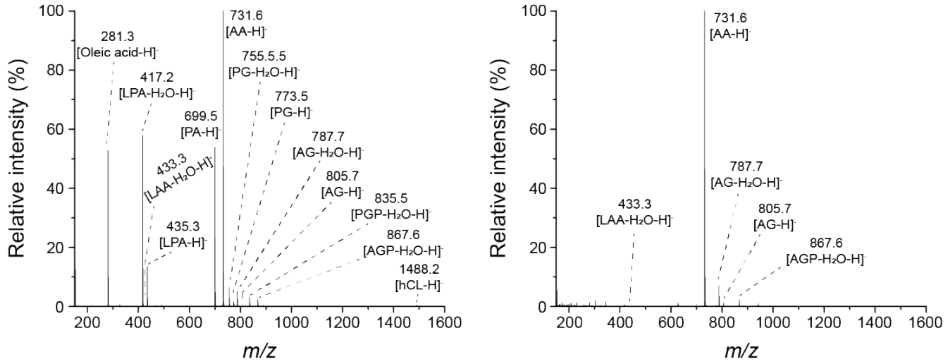
D



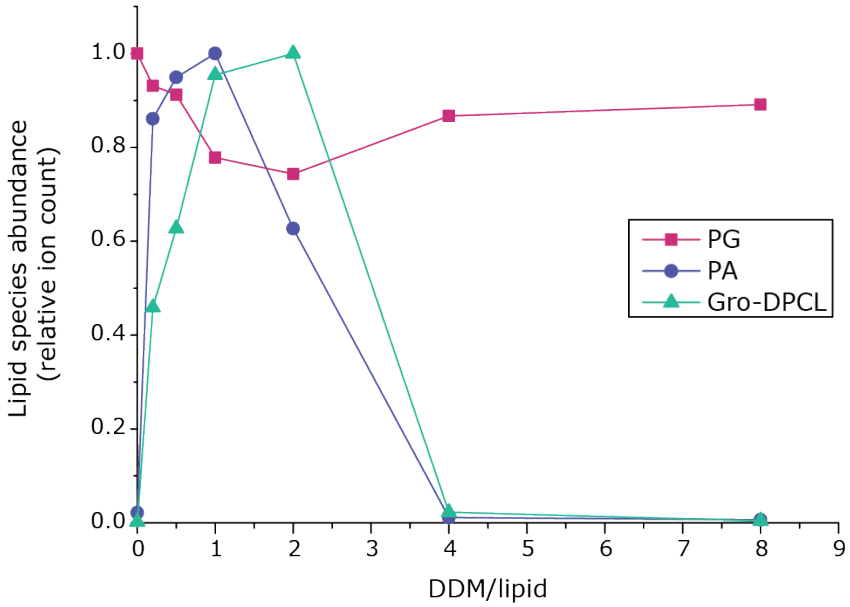
E



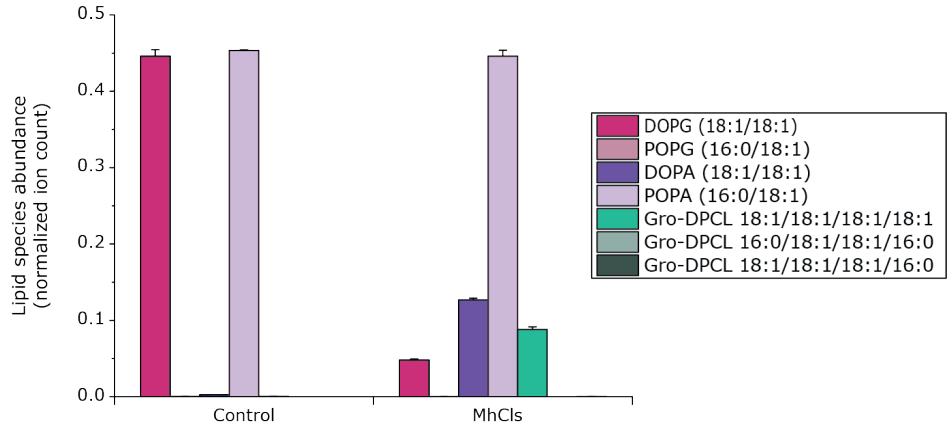
Supplemental figure 6: 2,3-bisphytanol-*sn*-glycerol-((*S*)-2,3-dihydroxypropyl) hydrogen phosphate (14′) **(A)** Structure of 14′. **(B)** 1H-NMR of 14′. **(C)** 13C-NMR of 14′. **(D)** 31P-NMR of 14′. **(E)** HRMS of 14′.



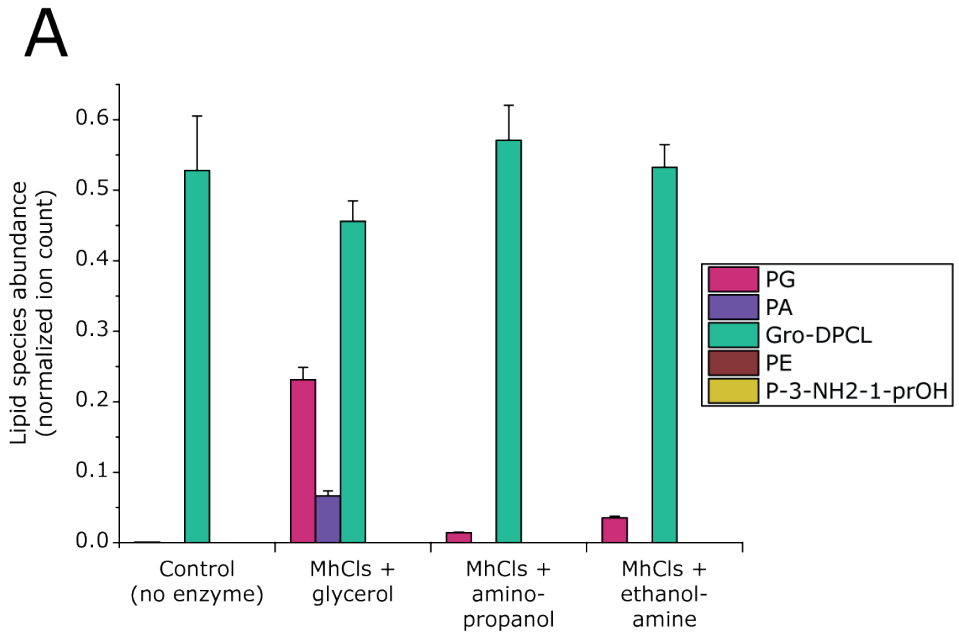
Supplemental figure 7: LC-MS fragmentation spectra of hybrid archaetidyl-phosphatidyl-cardiolipin (APCL; left panel) and di-archaetidyl-cardiolipin (DACL; right panel). Terms: lyso-phosphatidic acid (LPA), lyso-archaetidic acid (LAA), phosphatidic acid (PA), archaetidic acid (AA), phosphatidylglycerol (PG), archaetidylglycerol (AG), phosphatidylglycerol-phosphate (PGP), archaetidylglycerol-phosphate (AGP), hybrid cardiolipin (hCL; also referred to as APCL).



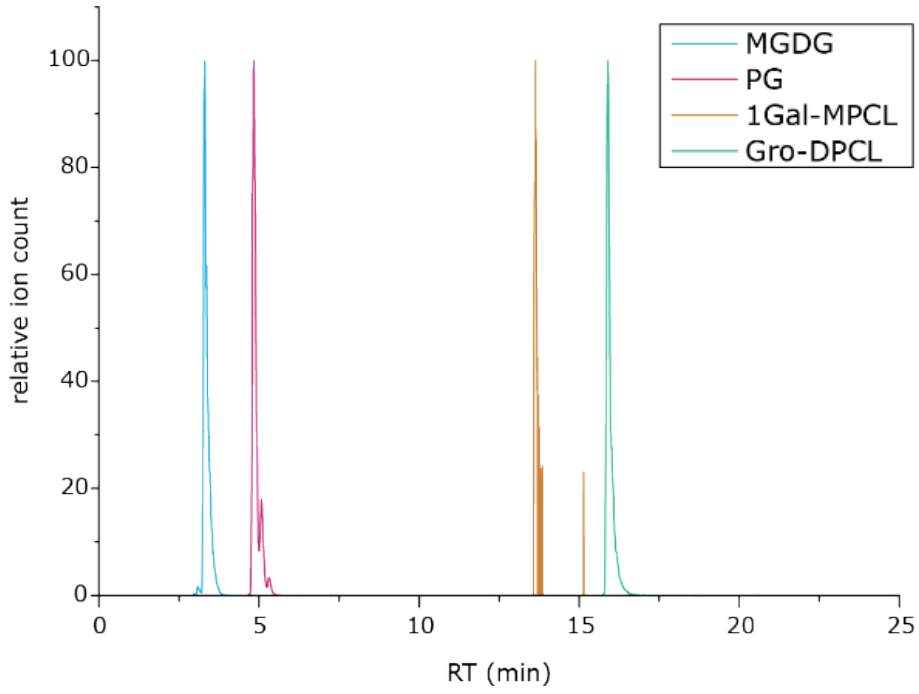
Supplemental figure 8: n-Dodecyl β-D-maltoside (DDM)-dependence of the activity of MhCls. The activity of MhCls was tested with different DDM-lipid molar ratios. Individual lipid species phosphatidylglycerol (PG), phosphatidic acid (PA) and glycerol-di-phosphatidyl-cardiolipin (Gro-DPCL) were analyzed by LC-MS, and plotted on the Y-axis relative to the highest detected amount of that particular lipid species.



Supplemental figure 9: MhCls activity in the presence of DOPG and POPA. Lipid species di-oleoyl-phosphatidylglycerol (DOPG), palmitoyl-oleoyl-phosphatidylglycerol (POPG), di-oleoyl-phosphatidic acid (DOPA), palmitoyl-oleoyl-phosphatidic acid (POPA) and the glycerol-diphosphatidyl-cardiolipin species (Gro-DPCL), were analyzed by LC-MS, normalized for the internal standard DDM and plotted.



Supplemental figure 10: MhCls activity in the presence of glycerol and the amines 3-amino-propanol and ethanolamine. Lipid species phosphatidylglycerol (PG), phosphatidic acid (PA), glycerol-di-phosphatidyl-cardiolipin (Gro-DPCL), phosphatidylethanolamine (PE) and phosphatidyl-3-amino-propanol (P-3-NH2-1-PrOH) were analyzed by LC-MS, normalized for the internal standard DDM, and plotted.



Supplemental figure 11: Retention times of the lipid species mono-galactosyl-diacylglycerol (MGDG), phosphatidylglycerol (PG), 1-galactosyl-mono-phosphatidyl-cardiolipin (1Gal-MPCL) and glycerol-di-phosphatidyl-cardiolipin (Gro-DPCL), during LC-MS analysis. The extracted ion chromatogram peaks are all normalized to 100%.

Supplemental tables

Supplemental table 1: Cloning and expression vectors used in this study.

Plasmid	Description	Reference
pRSF-Duet-1	Expression vector (Kan^R), T7 promoter	Novagen
pNDK001	*MhCls* gene with N-terminus 6x His-tag from *M. hungatei JF-1* cloned into pRSF-Duet vector using the primers NDKo027 and NDKo028.	This study
pBAD	Expression vector, pBR322 ori; araC; pBAD, Zeo^R	This study
pME006	*MhCls* gene with N-terminus 6x His-tag from *M. hungatei JF-1* cloned into pBAD-vector using the primers prME010 and prME011 to amplify the *MhCls* gene, and primers prME012 and prME013 to amplify the pBAD backbone.	This study

Supplemental table 2: Oligonucleotide primers used in this study. Introduced restriction sites are underlined.

Primers	Primer sequence 5′ -> 3′	Restriction site
NDKo027	ACAGTTCCATGGCCCATCACCATCATCACCACATCCATGATCTGATTCTG GTGATCCACAATTTTC	NcoI
NDKo028	ACTTACGAGCTCTTATTATTACAGCAGCGGACTAAACAGACG	SacI
prME010	TATCGAATTCATGCATCACCATCATCACCACATCCATGATCTGA	EcoRI
prME011	CTATAAGCTTTTATTATTACAGCAGCGGACTAAACAGACGGCTAAT	HindIII
prME012	ACTCGAATTCTTCCTCCTAGCCTGCTTTTTTGT	EcoRI
prME013	ACTGAAGCTTGATATCGTTTAAACGGTCTCCAGCTTGG	HindIII

Supplemental table 3: lipid species detected with LC-MS.

Lipid species	*m/z* $[M-H]^-$
archaetidylglycerol (AG)	805.66
archaetidic acid (AA)	731.63
glycerol-di-archaetidyl-cardiolipin (Gro-DACL)	1520.30
di-oleoyl phosphatidic acid (DOPA), phosphatidic acid (PA)	699.49
di-oleoyl phosphatidylglycerol (DOPG), phosphatidylglycerol (PG)	773.53
palmitoyl-oleoyl phosphatidic acid (POPA)	673.48
palmitoyl-oleoyl phosphatidylglycerol (POPG)	747.52
di-phosphatidyl-cardiolipin (Gro-DPCL), Gro-DPCL18:1/18:1/18:1/18:1	1456.03
Gro-DPCL16:0/18:1/16:0/18:1	1403.99
Gro-DPCL18:1/18:1/16:0/18:1	1430.01
Gro-archaetidyl-phosphatidyl-cardiolipin (APCL)	1488.16
phosphatidyl-1-propanol (P-1-PrOH), and phosphatidyl-2-propanol (P-2-PrOH)	741.54
phosphatidyl-1,2-propandiol (P-1,2-PrOH), and phosphatidyl-1,3-propandiol (P-1,3-PrOH)	757.53
1,2-propanediol-di-phosphatidyl-cardiolipin (1,2-PrOH-DPCL), and 1,3-propanediol-di-phosphatidyl-cardiolipin (1,3-PrOH-DPCL)	1440.02
phosphatidyl-1,2-ethanol (P-1,2-EtOH)	743.52
1,2-ethanol-di-phosphatidyl-cardiolipin (1,2-EtOH-DPCL)	1426.01
phosphatidyl-1,3-butanediol (P-1,3-BuOH), and phosphatidyl-1,4-butanediol (P-1,4-BuOH)	771.55
1,3-butanediol-di-phosphatidyl-cardiolipin (1,3-BuOH-DPCL), and 1,4-butanediol-di-phosphatidyl-cardiolipin (1,4-BuOH-DPCL)	1454.04
phosphatidyl-mannitol (P-mannitol)	863.56
mannitol-di-phosphatidyl-cardiolipin (mannitol-DPCL)	1546.05
phosphatidyl-2,2-dimethyl-1,3-propanediol (P-2,2-Me-1,3-PrOH)	786.56
phosphatidyl-2-phenyl-1,3-propanediol (P-2-Phe-1,3-PrOH)	833.56
di-oleoyl phosphatidylethanolamine (DOPE)	742.54
phosphatidyl-aminopropanol (P-NH2-PrOH)	756.55
di-oleoyl phosphatidylcholine (DOPC)	830.59
di-oleoyl phosphatidylserine (DOPS)	786.53
di-oleoyl phosphatidylinositol (DOPI)	861.55
monogalactosyldiacylglycerol (MGDG)	746.49
monogalactosyl-mono-phosphatidyl-cardiolipin (1Gal-MPCL)	1427.98

Supplemental methods

Bioinformatic identification of bacterial homologs of *E. coli* ClsA.

Using *E. coli* K12; MG1655 ClsA (EcClsA: NP_415765.1) as the query sequence, BLAST homology searches to the domain of bacteria were performed, and sequences were selected from the classes Gammaproteobacteria, Deltaproteobacteria and Betaproteobacteria (all belonging to the phylum proteobacteria), to get a versatile group of bacterial ClsA. The BLAST results were further analyzed using MEGA X, where the sequences were combined with the group of archaeal methanogens (Figure 2C) and aligned using the MUSCLE algorithm (default settings).

Assay conditions

The detergent-dependent activity of MhCls was assayed in 300 mM MES pH 7.0, 100mM KCl, in the presence of 1 µM MhCls and 1 mM DOPG, for 6 minutes at 37°C. The glycerol-dependent dynamic equilibrium of MhCls in the presence of POPA was assayed in presence of, 1.8 mM DDM, and 0.5 mM DOPG together with 0.5 mM POPA. The activity of MhCls in the presence of glycerol and the amines 3-amino-1-propanol and ethanolamine was tested using either 100 mM Glycerol, 3-amino-1-propanol, or ethanolamine and 0.5 mM Gro-DPCL or 1 mM DOPG, with addition of 1.8 mM DDM.

Chapter 5

Summary

Samenvatting

Summary

The Archaea comprise the third domain of life next to Bacteria and Eukarya [1]. Archaea can be found in a diverse range of environments but are often associated with extreme conditions, such as high-temperatures, high- or low-pH, high-salt, high-pressure or a combination of these [169,275,284,327,430]. These organisms are often classified as (hyper-)thermophiles, alkali- or acidophiles, halophiles, barophiles or for example, thermoacidophiles respectively. However, members of the Archaea are also found in mesophilic environments and thus are not extremophiles *per se*. One of the most salient features that distinguishes Archaea from Bacteria and Eukaryotes is the difference in membrane phospholipid structure. Archaea utilize isoprenoid radyl groups ether-bound to glycerol-1-phosphate (G1P) lipid backbones; whereas Bacteria and Eukarya predominantly synthesize membrane phospholipids based on fatty-acid based radyl groups, ester-bound to glycerol-3-phosphate lipid backbones. These fundamental differences in the structure of membrane phospholipids between the different domains of life has been termed the "lipid divide". It is one of the aspects that may have enabled Archaea to survive and thrive in more extreme environments compared to Bacteria and Eukarya [431].

The archaeal and bacterial lipids are each characterized by their own phospholipid biosynthesis pathways. While the two pathways exhibit an overall similar organization, the different nature of the ether- and ester-bound radyl tails requires fundamentally different enzymatic reactions. Further, there is a remarkable difference in chirality of the phospholipid backbone between those of Archaea and the other domains. Therefore, the formation of the lipid core involves unique enzymes that not only differ in their catalytic mechanism but also their stereoselectivity, thereby upholding the lipid divide. Comparison of the structures of the biosynthetic enzymes of the archaeal and bacterial phospholipid pathways thus have led to new insights on the lipid divide and the evolution of early life (**Chapter 1**).

Briefly, the archaeal lipid core is synthesized as follows: The building blocks for the isoprenyl radyl groups originate from the mevalonate (MVA) pathway which uses acetyl-CoA to form isopentenyl pyrophosphate (IPP) and dimethylallyl pyrophosphate (DMAPP) [12]. The isoprenyl radyl moieties are formed by iterative addition of three IPP molecules to one DMAPP molecule resulting in geranylgeranyl pyrophosphate (GGPP) by geranylgeranyl pyrophosphate synthase (GGPPS) [43]. The first committing step is the formation of the first radyl-ether bond. Geranylgeranylglycerol phosphate synthase (GGGPS) catalyzes the formation of geranylgeranylglycerol phosphate (GGGP) through the addition of GGPP to the C-3 hydroxyl of G1P [57]. This is followed by the second critical step, the addition of a second GGPP moiety to C-2 of the G1P backbone of GGGP by di-geranylgeranyl phosphate synthase (DGGGPS), forming di-geranylgeranylglycerol phosphate (DGGGP). This reaction completes the archaeal dialkyl glycerol diether (DGD) lipid core [94], which regardless of the saturation state of the radyl chains, is commonly also referred to as archaeol. The activity of DGGGPS has recently been characterized and coupled to a crystal structure, providing novel insights in the catalytic mechanism of this enzyme with significance for the entire UbiA superfamily to which, of which this enzyme is a member (**Chapter 3**) [82]. Similar to other UbiA prenyltransferases, DGGGPS is Mg^{2+} dependent and shares a similar overall structure. DGGGPS was shown to be an integral membrane protein with nine helices forming a transmembrane domain and a cytosolic domain. The transmembrane domain contains two hydrophobic tunnels which serve to bind the hydrophobic isoprenyl moieties of both substrates, i.e., GGPP and GGGP. The cytosolic domain contains two conserved aspartate rich motifs that are involved in the binding of Mg^{2+} ions that play an essential role in enzymatic activity, coordinating the phosphate headgroups of the two substrates. Substrate binding assays with various DGGGPS mutants allowed for the assignment of a specific substrate to each of the hydrophobic substrate binding tunnels. Furthermore, through mutagenesis and activity assays, other residues important for activity and substrate binding were identified. The inclusion of mutants corresponding to the mutations in the human UBIAD1 protein that have been linked to disease provided an explanation of the molecular mechanism behind human UbiA-associated diseases and will facilitate the development of new therapeutic strategies.

After the formation of the core phospholipid structure, the phospholipid phosphate headgroup of DGGGP is activated by CDP-archaeol synthase (CarS) using CTP to form CDP-archaeol. This is followed by polar headgroup attachment by CDP-alcohol phosphatidyl transferases yielding to produce a variety of headgroups. This group of enzymes appears highly promiscuous with respect to the phospholipid core structure, and includes both archaeal and bacterial proteins. In archaea, it is not known at which stage of lipid synthesis the saturation of the radyl groups occur, but it is thought to occur after polar headgroup activation by CarS (**Chapter 1**).

In Archaea, there is not as much structural diversity in the diether phospholipid core compared to bacterial lipids, as most archaea produce DGDs exclusively with C_{20} isoprenyl radyl groups which are typically fully saturated in mature lipids [106,275,432,433]. In contrast, the radyl chains of bacterial lipids can vary in carbon-number, degree of saturation and other modifications such as

hydroxylation and cyclopropane ring formation. Archaea have been shown to change the phospholipid polar headgroup composition in the membrane in response to different cultivation temperatures [236]. Many thermophilic archaea [231,434], some mesophilic archaea [106,110,240,417], and even some bacteria [230,234], produce macrocyclic membrane-spanning tetraether lipids (or glycerol dialkyl glycerol tetraether lipids, GDGTs). Moreover, an increased proportion of GDGTs in the cell membrane [435], and pentacyclic rings in GDGTs [249,250], increases membrane stability and decreases (proton) permeability. GDGTs can be modified to contain up to eight pentacyclic rings, and specifically lipids from Thaumarchaeota can contain a hexacyclic ring. These modifications are thought to modulate lipid packing and stability [231,240]. To some extent akin to bacterial homeoviscous adaptation, Archaea appear to increase the proportion of GDGTs in the membrane with increasing temperature to improve lipid packing, membrane stability, and decrease proton permeability at elevated temperatures [241,242] or increase the amount of pentacyclic rings in GDGTs at elevated temperatures or sub-optimal pH values [236,243,247]. The enzymes responsible for the formation of the pentacyclic rings in GDGTs have been recently identified as members of the radical S-adenosyl methionine (rSAM) family [251]. While the biosynthetic pathway of GDGT synthesis has so far not been resolved. However, it is thought that the macrocyclic membrane-spanning ring of GDGT is formed through a tail-to-tail addition of two DGD lipids [259,260,263,265,436] by the action of a recently identified radical SAM enzyme (**Chapter 1**) [252,256].

Cardiolipins (CL) are a class of lipids involved in osmoregulation, membrane organization and is associated with proteins participating in bioenergetic processes [160–173]. The structure of cardiolipin is rather unusual, as it resembles a phospholipid linked head-to-head to another phospholipid through a shared polyol headgroup or a phospholipid directly linked to the polar headgroup of a glycolipid. While present in all three domains of life, prototypical glycerol-cardiolipins can be synthesized in two different ways [113,168,169,174–177]. In archaea and bacteria, glycerol-CL is typically synthesized from two phosphatidylglycerol (PG) molecules by a phospholipase D type cardiolipin synthase (PLD-type Cls). Recently, we reported on the first characterization of an archaeal PLD-type Cls from the methanogen *Methanospirillum hungatei* (MhCls) (**Chapter 4**) [184]. MhCls did not show appreciable selectivity towards the core lipid structure and accepted both archaetidylglycerol (AG, the archaeal analogue of bacterial phosphatidylglycerol) and PG. Remarkably, a novel hybrid-cardiolipin species was synthesized in the presence of both AG and PG, containing both the archaeal and bacterial glycerol backbone stereoisomers and radyl configurations. It was further shown that Cls not only has a transphosphatidylase and hydrolase activity, but in presence of compatible alcoholic substrates, it can catalyze a reverse "alcoholysis" reaction on CL, splitting the molecule into two diesters with each their own headgroup. In the presence of water, this results in the formation of PG (AG) and PA (AA), but in the presence of a suitable alcoholic substrate novel phospholipid species are generated. MhCls was found to be a promiscuous enzyme, accepting a broad range of lipids and alcoholic substrates resulting in the synthesis of several more non-natural cardiolipin species and diester lipids. These properties make Cls an interesting enzyme for bioengineering, such as for the synthesis of novel non-natural lipids with customized properties.

Generally, archaeal phospholipids with an equivalent headgroup are more hydrophobic compared to bacterial phospholipids due to the presence of ether bonds as opposed to ester bonds, and because of the bulkier saturated isoprenoid radyl groups. At the time of writing, reversed-phase (RP) high-performance liquid chromatography (HPLC) is the most common analysis technique. As this is also the most common HPLC system configuration, this presents a new analytic challenge for laboratories having only this technique available and wanting to perform lipidomic analyses. As many archaeal phospholipids are more hydrophobic compared to bacterial phospholipids with equivalent headgroups, also taking into account the increased size of phospholipid species such as cardiolipin and GDGTs, the elution of such lipids from typically used C18-modified silica columns can be challenging. Moreover, phospholipids with free phosphate groups, such as GGPP, GGGP, DGGGP (but also the bacterial lyso-phosphatidic acid and phosphatidic acid) typically exhibit significant chromatographic tailing with classical chromatographic methods [335]. To address these issues, a new RP-HPLC method was developed based on 1-butanol as mobile phase solvent (**Chapter 2**) [366]. The higher hydrophobicity of 1-butanol compared to traditional acetonitrile-, methanol- or 2-propanol-based mobile phases gives it a higher eluotropic strength and improves elution characteristics. Using the same chromatographic gradient and ionization conditions, the use of the 1-butanol mobile phase composition allowed for the elution and analysis of a larger range of lipids, improved chromatographic resolution and overall peak shape. However, decreased ionization efficiency leads to reduced sensitivity and higher limits of detection and quantification for most less-hydrophobic phospholipids. It was shown that peak tailing of free phosphate containing analytes is in large part caused by phosphate-metal interactions of the analyte with the inner column wall. These interactions can be significantly reduced by using columns with passivated inner-metal surfaces, which significantly improves peak shape for these analytes.

In summary, the works described in this thesis contribute to our knowledge on the catalytic and structural mechanism of archaeal lipid biosynthesis. However, several pressing questions remain unanswered. First, the biosynthetic pathway of GDGTs and other closely related lipids such as, cyclic DGDs, bridged H-GDGTs and macrocyclic diabolic acid based tetraesters and tetraethers in bacteria, remain to be elucidated. Due to the similar nature of their probable DGD precursors, and the C-C bond formed, macrocyclic tetraether and tetraester lipids are likely synthesized through the same, or a similar mechanism, by related radical SAM enzymes. Second, other unresolved questions relate to specific subclades in the archaea. Lokiarchaea not only contain the genes encoding the archaeal phospholipid biosynthesis pathway enzymes but also appear contain genes that encode for enzymes typically found in bacterial phospholipid biosynthesis. This unique group of archaea — that is most closely related to Eukaryota — may bridge the lipid divide either by producing phospholipids with G1P and G3P stereochemistries and have been suggested to produce chimeric lipids with mixed ester/ether bound acyl/alkyl and isoprenyl radyl groups [3]. Since Lokiarchaeota are difficult to culture, compelling evidence for such an unusual lipid composition is still missing. The study of their lipid biosynthesis enzymes *in vitro* may potentially offer an alternative opportunity to learn more about how the lipid divide occurred. Third, taking the sofar reported *in vitro* substrate specificities of archaeal phospholipid biosynthesis into account, unregulated catalysis would result in dead-end products such as unsaturated GDGTs or saturated isoprenyl pyrophosphates. However, the lack of detectable levels of these compounds in lipid extracts indicates that there is a mechanism preventing their formation. Particularly, the substrate specificities of CarS, GGR and GDGT synthase are of interest, as it is not known at which point during phospholipid biosynthesis saturation occurs. It has been suggested that CarS might not accept saturated DGGGP (AA) and that GGR would not accept unsaturated GDGTs; lending support to the notion that DGD saturation happens after headgroup activation and before GDGT formation. Additionally, there are some indications that lipid biosynthesis enzymes co-localize in cells [131,133,148]. The spatial organization of phospholipid biosynthesis (i.e., the formation of synthesis complexes) could provide some order to the biosynthetic pathway and thus potentially prevent counterproductive side reactions.

Samenvatting

De Archaea omvatten een derde domain van leven naast de domeinen van de Bacteria en Eukarya [1]. Archaea kunnen worden gevonden in diverse omgevingen en worden vaak geassocieerd met extreme omstandigheden, zoals hoge temperaturen, hoge- of lage-pH, hoge zoutconcentraties, hoge druk, of een combinatie van deze zaken [169,275,284,327,430]. Deze organismen zijn hierom vaak geclassificeerd als (hyper-)thermofielen, alkali- of acidofielen, halofielen, barofielen of bijvoorbeeld thermoacidofielen, respectievelijk. Echter, er zijn ook leden van de Archaea die in mesofiele omgevingen gevonden worden en dus niet *per se* extremofiel zijn. Een van de meest in het oog springende kenmerken die Archaea onderscheiden van Bacteria en Eukarya is het verschil in de chemische structuur van fosfolipiden in het membraan van de cel. Archaea gebruiken isoprenoïde radylgroepen die ether-gebonden worden glycerol-1-phosphaat (G1P) lipide ruggengraten; waar Bacteria en Eukarya voornamelijk membraan fosfolipiden synthetiseren die gebaseerd zijn op vetzuur radylgroepen, die ester-gebonden zijn met glycerol-3-phosphate lipide ruggengraten. Deze fundamentele verschillen in de chemische structuur van membraan fosfolipiden tussen de verschillende domeinen van leven wordt beschreven met de term "lipid divide" ("*lipidekloof*"). De lipidekloof is een van de aspecten die Archaea mogelijk in staat hebben gesteld om te overleven en te gedijen onder extremere omstandigheden dan Bacteria en Eukarya [431].

Hoewel de twee routes in het algemeen een vergelijkbare organisatie vertonen, vereist de verschillende aard van de ether- en ester-gebonden radylgroepen fundamenteel verschillende enzymatische reacties. Verder is er een opmerkelijk verschil in chiraliteit van de fosfolipide-ruggengraat tussen die van Archaea en de andere domeinen. Daarom zijn bij de vorming van de lipidekern unieke enzymen betrokken die niet alleen verschillen in hun katalytische mechanisme, maar ook in hun stereoselectiviteit, waardoor de lipidekloof tussen de domeinen van leven in stand wordt gehouden. De vergelijking van de structuren van biosynthetische enzymen in de archaeale en bacteriële fosfolipide syntheseroutes hebben geleid tot nieuwe inzichten over de lipidenverdeling en de evolutie van het vroege leven (**Hoofdstuk 1**).

In het kort wordt de archaeale lipidekern als volgt gesynthetiseerd: De bouwstenen voor de isoprenyl radylgroepen zijn afkomstig van de mevalonaat (MVA)-route die acetyl-CoA gebruikt om isopentenylpyrofosfaat (IPP) en dimethylallylpyrofosfaat (DMAPP) te vormen [12]. De isoprenyl radylgroepen worden gevormd door iteratieve toevoeging van drie IPP-moleculen aan één DMAPP-molecuul, wat

resulteert in geranylgeranylpyrofosfaat (GGPP) door geranylgeranylpyrofosfaat synthase (GGPPS) [43]. De eerste committerende stap is de vorming van de eerste radyl-etherbinding. Geranylgeranylglycerolfosfaat synthase (GGGPS) katalyseert de vorming van geranylgeranylglycerolfosfaat (GGGP) door de koppeling van GGPP aan de C-3-hydroxyl van G1P [57]. Dit wordt gevolgd door de tweede kritische stap; de koppeling van GGPP aan de C-2-hydroxyl van de G1P-ruggengraat van GGGP door di-geranylgeranylfosfaat synthase (DGGGPS), waarbij di-geranylgeranylglycerolfosfaat (DGGGP) wordt gevormd. Deze reactie voltooit de formatie van de archaeale dialkylglyceroldiether (DGD) lipidekern [94], die ongeacht de verzadigingstoestand van de radylketens, gewoonlijk ook wel archaeol wordt genoemd. De activiteit van DGGGPS is recentelijk gekarakteriseerd en gekoppeld aan een kristalstructuur, wat nieuwe inzichten verschafte in het katalytisch mechanisme van dit enzym met betekenis voor de hele UbiA-superfamilie waarvan dit enzym lid is (**Hoofdstuk 3**) [82]. Net als andere UbiA-prenyltransferasen, is DGGGPS Mg^{2+}-afhankelijk en heeft het een vergelijkbare algemene structuur. DGGGPS bleek een integraal membraaneiwit te zijn met negen helices die een transmembraandomein en een cytosolisch domein vormen. Het transmembraandomein bevat twee hydrofobe tunnels die dienen om de hydrofobe isoprenyl-groepen van beide substraten, d.w.z. GGPP en GGGP, te binden. Het cytosolische domein bevat twee geconserveerde aspartaat-rijke motieven die betrokken zijn bij de binding van Mg^{2+}-ionen die de fosfaatkopgroepen van de twee substraten coördineren en een essentiële rol spelen bij katalyse. Substraatbinding-analyses met verschillende DGGGPS-mutanten maakten het mogelijk om een specifiek substraat aan elk van de hydrofobe substraatbindingstunnels toe te wijzen. Verder werden door middel van mutagenese en activiteit-analyses andere residuen geïdentificeerd die belangrijk zijn voor activiteit en substraatbinding. Mutanten van DGGGPS, die overeenkomen met mutaties in het menselijke UBIAD1-eiwit die geassocieerd worden met bepaalde ziektebeelden, werden ook meegenomen; dit verschafte een verklaring van het moleculaire mechanisme achter UbiA-geassocieerde ziekten in mensen en zal de ontwikkeling van nieuwe therapeutische strategieën bevorderen.

Na de vorming van de fosfolipidekernstructuur, wordt de fosfolipidefosfaatkopgroep van DGGGP geactiveerd door CDP-archaeolsynthase (CarS) met behulp van CTP om CDP-archaeol te vormen. Dit wordt gevolgd door de koppeling van polaire kopgroepen door CDP-alcoholfosfatidyltransferasen, wat resulteert in het produceren van fosfolipiden een verscheidenheid aan kopgroepen. Deze groep enzymen lijkt zeer promiscue met betrekking tot de fosfolipide-kernstructuur en omvat zowel archaeale als bacteriële eiwitten. In archaea is het niet bekend in welk stadium van lipidesynthese de verzadiging van de radylgroepen optreedt, maar het wordt gedacht dat dit optreedt na activering van de polaire kopgroep door CarS (**Hoofdstuk 1**).

In Archaea is er niet zoveel structurele diversiteit in de diether-fosfolipidekern in vergelijking met bacteriële lipiden, aangezien de meeste archaea uitsluiten DGD's produceren met C_{20}-isoprenyl radylgroepen die doorgaans volledig verzadigd zijn in volledig gevormde lipiden [106,275,432,433]. Daarentegen kunnen de radylketens van bacteriële lipiden variëren in koolstofgetal, verzadigingsgraad en andere modificaties zoals hydroxylering en cyclopropanering. Het is aangetoond dat Archaea de samenstelling van de polaire kopgroep van fosfolipiden in het membraan veranderen als reactie op verschillende kweektemperaturen [236]. Vele thermofiele archaea [231,434], sommige mesofiele archaea [106,110,240,417], en zelfs sommige bacterien [230,234], produceren macrocyclische membraan-overspannende tetraether lipiden (ook wel glyceroldialkylglyceroltetraether lipiden of GDGTs genoemd). Bovendien, een vergroot aandeel GDGT's in het celmembraan [435], en pentacyclische ringen in GDGTs [249,250], verhoogt de stabiliteit en proton permeabiliteit van het membraan. GDGT's kunnen worden gemodificeerd om maximaal acht pentacyclische ringen te bevatten, en specifiek lipiden uit Thaumarchaeota kunnen een hexacyclische ring bevatten. Van deze modificaties wordt gedacht dat ze de lipidepakking en -stabiliteit van het membraan moduleren [231,240]. Tot op zekere hoogte vergelijkbaar met bacteriële homeovisceuze membraan adaptaties, lijken Archaea het aandeel van GDGT's in het membraan te verhogen met toenemende temperatuur om de lipidepakking te bevorderen, of de hoeveelheid pentacyclische ringen in GDGT's verhogen bij verhoogde temperaturen of suboptimale pH-waarden [236,243,247], en daar mee membraanstabiliteit te verbeteren en de protonpermeabiliteit bij verhoogde temperaturen te verminderen [241,242]. De enzymen die verantwoordelijk zijn voor de vorming van de pentacyclische ringen in GDGT's zijn onlangs geïdentificeerd als leden van de radicale S-adenosylmethionine (SAM) familie [251]. De biosynthetische route van GDGT-synthese is tot nu toe niet opgehelderd. De huidige gedachte is dat de macrocyclische membraan-overspannende ring van GDGTs wordt gevormd door een staart-aan-staart koppeling van twee DGD-lipiden [259,260,263,265,436] door de werking van een recentelijk geidentificeerd radicaal SAM enzym (**Hoofdstuk 1**) [252,256].

Cardiolipines (CL) zijn een klasse van lipiden die betrokken zijn bij osmoregulatie, membraanorganisatie en worden geassocieerd met eiwitten die deelnemen aan bio-energetische processen [160–173]. De structuur van cardiolipine is nogal ongebruikelijk, omdat het lijkt op een fosfolipide dat rechtstreeks aan een ander

fosfolipide is gekoppeld via een gedeelde polyol-kopgroep of een fosfolipide dat rechtstreeks is gekoppeld aan de polaire kopgroep van een glycolipide. Hoewel ze in alle drie de levensdomeinen aanwezig zijn, kunnen prototypische glycerol-cardiolipinen op twee verschillende manieren worden gesynthetiseerd [113,168,169,174–177]. In archaea en bacteriën wordt glycerol-CL normaliter gesynthetiseerd uit twee fosfatidylglycerol (PG) moleculen door een fosfolipase D-type cardiolipinesynthase (PLD-type Cls). Onlangs hebben we de eerste karakterisering gerapporteerd van een archaeale PLD-type Cls uit de methanogeen Methanospirillum hungatei (MhCls) (**Hoofdstuk 4**) [184]. MhCls vertoonde geen merkbare selectiviteit naar de kernlipidestructuur en accepteerde zowel archaetidylglycerol (AG, de archaeale analoog van bacteriële fosfatidylglycerol) als PG. Opmerkelijk is dat een nieuwe hybride-cardiolipine-soort werd gesynthetiseerd in de aanwezigheid van zowel AG als PG, die zowel de archaeale als bacteriële glycerol-ruggengraat stereo-isomeren en radylconfiguraties bevat. Verder werd aangetoond dat Cls niet alleen een transfosfatidylase- en hydrolase-activiteit heeft, maar in aanwezigheid van compatibele alcoholische substraten ook een omgekeerde "alcoholyse" -reactie op CL kan katalyseren, waarbij het molecuul in twee diesters wordt gesplitst met elk hun eigen kopgroep. In aanwezigheid van water resulteert dit in de vorming van PG (AG) en PA (AA), maar in aanwezigheid van een geschikt alcoholisch substraat worden nieuwe fosfolipiden gevormd. MhCls bleek een promiscue enzym te zijn, dat een breed scala aan lipiden en alcoholische substraten accepteert, wat resulteerde in de synthese van verschillende niet-natuurlijke cardiolipine-soorten en diester-lipiden. Deze eigenschappen maken Cls een interessant enzym voor bio-engineering, zoals voor de synthese van nieuwe niet-natuurlijke lipiden met aangepaste eigenschappen.

Over het algemeen zijn archaeale fosfolipiden met een equivalente kopgroep meer hydrofoob in vergelijking met bacteriële fosfolipiden vanwege de aanwezigheid van etherbindingen in tegenstelling tot esterbindingen en vanwege de grotere verzadigde isoprenoïde radylgroepen. Op het moment van schrijven is omgekeerde-fase (RP) hoge-prestatie vloeistofchromatografie (HPLC) de meest gebruikte analysetechniek. Aangezien dit ook de meest voorkomende HPLC-systeemconfiguratie is, vormt dit een nieuwe analytische uitdaging voor laboratoria die alleen over deze techniek beschikken en lipidomische analyses willen uitvoeren. Omdat veel archaeale fosfolipiden meer hydrofoob zijn in vergelijking met bacteriële fosfolipiden met equivalente kopgroepen, ook rekening houdend met de toegenomen grootte van fosfolipidensoorten zoals cardiolipine en GDGT's, kan de elutie van dergelijke lipiden uit typisch gebruikte C18-gemodificeerde silicakolommen een uitdaging zijn. Bovendien vertonen fosfolipiden met vrije fosfaatgroepen, zoals GGPP, GGGP, DGGGP (maar ook het bacteriële lyso-fosfatidezuur en fosfatidezuur) typisch significante chromatografische staartvorming met klassieke chromatografische methoden [335]. Om deze problemen aan te pakken, is een nieuwe RP-HPLC-methode ontwikkeld op basis van 1-butanol als oplosmiddel voor de mobiele fase (**Hoofdstuk 2**) [366]. De hogere hydrofobiciteit van 1-butanol in vergelijking met traditionele mobiele fasen op basis van acetonitril, methanol of 2-propanol geeft het een hogere eluotrope sterkte en verbetert de elutie-eigenschappen. Door gebruik te maken van dezelfde chromatografische gradiënt en ionisatiecondities, maakte het gebruik van de 1-butanol mobiele fase samenstelling de elutie en analyse mogelijk van een groter bereik aan lipiden, verbeterde de chromatografische resolutie en algehele piekvorm van analyten. Een verminderde ionisatie-efficiëntie leidt echter tot verminderde gevoeligheid en hogere detectie- en kwantificatielimieten voor de meeste minder hydrofobe fosfolipiden. Er werd aangetoond dat staartvorming van de chromatografische pieken van vrije fosfaatbevattende analyten voor een groot deel wordt veroorzaakt door fosfaat-metaal-interacties van de analyt met de binnenwand van de kolom. Deze interacties kunnen aanzienlijk worden verminderd door kolommen te gebruiken met gepassiveerde metalen binnenoppervlakken, wat de piekvorm voor deze analyten aanzienlijk verbetert.

Samenvattend, de in dit proefschrift beschreven werken dragen bij aan onze kennis over het katalytische en structurele mechanisme van de biosynthese van archaeale lipiden. Er blijven echter een aantal prangende vragen onbeantwoord. Ten eerste moet de biosynthetische route van GDGT's en andere nauw verwante lipiden, zoals cyclische DGD's, gebrugde H-GDGT's en op macrocyclische diabolische zuren gebaseerde tetra-esters en tetraethers in bacteriën, nog worden opgehelderd. Vanwege de vergelijkbare aard van hun waarschijnlijke DGD-precursoren en de gevormde CC-binding, worden macrocyclische tetra-ether- en tetraester-lipiden waarschijnlijk gesynthetiseerd via hetzelfde of een vergelijkbaar mechanisme door verwante enzymen. Ten tweede hebben andere onopgeloste vragen betrekking op specifieke subclades in de archaea. Lokiarchaea bevatten niet alleen de genen die coderen voor de archaea-enzymen voor de biosynthese van fosfolipiden, maar lijken ook genen te bevatten die coderen voor enzymen die typisch worden aangetroffen in de biosynthese van bacteriële fosfolipiden. Deze unieke groep archaea - die het nauwst verwant is aan Eukaryota – zou de lipidekloof kunnen overbruggen door fosfolipiden te produceren met G1P- en G3P-stereochemie en er is gesuggereerd dat ze chimerische lipiden kunnen produceren met gemengde ester-/ether-gebonden acyl/alkyl- en isoprenyl radylgroepen [3]. Omdat Lokiarchaeota moeilijk te kweken zijn, ontbreekt er nog steeds overtuigend bewijs voor een dergelijk ongebruikelijke samenstelling van membraanlipiden. De studie van hun lipide biosynthese-enzymen in vitro kan mogelijk een alternatieve mogelijkheid bieden

om meer te leren over hoe de lipidekloof ontstaan is. Ten derde, rekening houdend met de tot nu toe gerapporteerde in vitro substraatspecificiteiten van archaeale fosfolipidebiosynthese, zou ongereguleerde katalyse resulteren in doodlopende producten zoals onverzadigde GDGT's of verzadigde isoprenylpyrofosfaten. Het ontbreken van detecteerbare niveaus van deze verbindingen in lipide-extracten geeft echter aan dat er een mechanisme is die hun vorming verhindert. In het bijzonder zijn de substraatspecificiteiten van CarS en GGR van belang, aangezien het niet bekend is op welk punt tijdens de fosfolipidebiosynthese verzadiging van de isoprenyl ketens optreedt. Er is gesuggereerd dat CarS mogelijk geen verzadigde DGGGP (AA) accepteert en dat GGR geen onverzadigde GDGT's zou accepteren; dit ondersteunt het idee dat DGD-verzadiging plaatsvindt na activering van de kopgroep en vóór GDGT-vorming. Bovendien zijn er enkele aanwijzingen dat lipide-biosynthese-enzymen co-lokaliseren in cellen. [131,133,148]. De ruimtelijke organisatie van de biosynthese van fosfolipiden (d.w.z. de vorming van synthesecomplexen) zou enige orde kunnen scheppen in de biosyntheseroute en zo mogelijk contraproductieve nevenreacties kunnen voorkomen.

References

[1] C.R. Woese, G.E. Fox, Phylogenetic structure of the prokaryotic domain: The primary kingdoms, Proc. Natl. Acad. Sci. U. S. A. 74 (1977) 5088–5090. https://doi.org/10.1073/pnas.74.11.5088.

[2] Y. Koga, From promiscuity to the lipid divide: On the evolution of distinct membranes in archaea and bacteria, J. Mol. Evol. 78 (2014) 234–242. https://doi.org/10.1007/s00239-014-9613-4.

[3] L. Villanueva, S. Schouten, J.S.S. Damsté, Phylogenomic analysis of lipid biosynthetic genes of Archaea shed light on the 'lipid divide,' Environ. Microbiol. 19 (2017) 54–69. https://doi.org/10.1111/1462-2920.13361.

[4] Y. Koga, Early evolution of membrane lipids: How did the lipid divide occur?, J. Mol. Evol. 72 (2011) 274–282. https://doi.org/10.1007/s00239-011-9428-5.

[5] J. Lombard, P. López-García, D. Moreira, The early evolution of lipid membranes and the three domains of life, Nat. Rev. Microbiol. 10 (2012) 507–515. https://doi.org/10.1038/nrmicro2815.

[6] J.S. Lolkema, D.J. Slotboom, Estimation of structural similarity of membrane proteins by hydropathy profile alignment, Mol. Membr. Biol. 15 (1998) 33–42. https://doi.org/10.3109/09687689809027516.

[7] J.S. Lolkema, D.-J. Slotboom, Hydropathy profile alignment: a tool to search for structural homologues of membrane proteins, FEMS Microbiol. Rev. 22 (1998) 305–322. https://doi.org/10.1111/j.1574-6976.1998.tb00372.x.

[8] J. Lombard, D. Moreira, Origins and early evolution of the mevalonate pathway of isoprenoid biosynthesis in the three domains of life, Mol. Biol. Evol. 28 (2011) 87–99. https://doi.org/10.1093/molbev/msq177.

[9] J.C. VanNice, D.A. Skaff, A. Keightley, J.K. Addo, G.J. Wyckoff, H.M. Miziorko, Identification in haloferax volcanii of phosphomevalonate decarboxylase and isopentenyl phosphate kinase as catalysts of the terminal enzyme reactions in an archaeal alternate mevalonate pathway, J. Bacteriol. 196 (2014) 1055–1063. https://doi.org/10.1128/JB.01230-13.

[10] J.M. Vinokur, T.P. Korman, Z. Cao, J.U. Bowie, Evidence of a novel mevalonate pathway in archaea, Biochemistry. 53 (2014) 4161–4168. https://doi.org/10.1021/bi500566q.

[11] H. Hayakawa, K. Motoyama, F. Sobue, T. Ito, H. Kawaide, T. Yoshimura, H. Hemmi, Modified mevalonate pathway of the archaeon Aeropyrum pernix proceeds via transanhydromevalonate 5-phosphate, Proc. Natl. Acad. Sci. U. S. A. 115 (2018) 10034–10039. https://doi.org/10.1073/pnas.1809154115.

[12] R. Yoshida, T. Yoshimura, H. Hemmi, Reconstruction of the "archaeal" mevalonate pathway from the methanogenic archaeon Methanosarcina mazei in Escherichia coli cells, Appl. Environ. Microbiol. 86 (2020) 1–12. https://doi.org/10.1128/AEM.02889-19.

[13] N. Dellas, S.T. Thomas, G. Manning, J.P. Noel, Discovery of a metabolic alternative to the classical mevalonate pathway, Elife. 2013 (2013). https://doi.org/10.7554/eLife.00672.001.

[14] J.M. Vinokur, M.C. Cummins, T.P. Korman, J.U. Bowie, An Adaptation to Life in Acid Through A Novel Mevalonate Pathway, Sci. Rep. 6 (2016) 1–11. https://doi.org/10.1038/srep39737.

[15] Y. Azami, A. Hattori, H. Nishimura, H. Kawaide, T. Yoshimura, H. Hemmi, (R)-mevalonate 3-phosphate is an intermediate of the mevalonate pathway in Thermoplasma acidophilum, J. Biol. Chem. 289 (2014) 15957–15967. https://doi.org/10.1074/jbc.M114.562686.

[16] S. Watanabe, Y. Murase, Y. Watanabe, Y. Sakurai, K. Tajima, Crystal structures of aconitase X enzymes from bacteria and archaea provide insights into the molecular evolution of the aconitase superfamily, Commun. Biol. 4 (2021) 1–15. https://doi.org/10.1038/s42003-021-02147-5.

[17] H. Hemmi, S. Ikejiri, S. Yamashita, T. Nishino, Novel medium-chain prenyl diphosphate synthase from the thermoacidophilic archaeon Sulfolobus solfataricus, J. Bacteriol. 184 (2002) 615–620. https://doi.org/10.1128/JB.184.3.615-620.2002.

[18] H.Y. Sun, T.P. Ko, C.J. Kuo, R.T. Guo, C.C. Chou, P.H. Liang, A.H.J. Wang, Homodimeric hexaprenyl pyrophosphate synthase from the thermoacidophilic crenarchaeon Sulfolobus solfataricus displays asymmetric subunit structures, J. Bacteriol. 187 (2005) 8137–8148. https://doi.org/10.1128/JB.187.23.8137-8148.2005.

[19] Y.P. Lu, H.G. Liu, P.H. Liang, Different reaction mechanisms for cis- and trans-prenyltransferases, Biochem. Biophys. Res. Commun. 379 (2009) 351–355. https://doi.org/10.1016/j.bbrc.2008.12.061.

[20] S.I. Ohnuma, T. Nakazawa, H. Hemmi, A.M. Hallberg, T. Koyama, K. Ogura, T. Nishino, Conversion from farnesyl diphosphate synthase to geranylgeranyl diphosphate synthase by random chemical mutagenesis, J. Biol. Chem. 271 (1996) 10087–10095. https://doi.org/10.1074/jbc.271.17.10087.

[21] L.C. Tarshis, P.J. Proteau, B.A. Kellogg, J.C. Sacchettini, C.D. Poulter, Regulation of product chain length by isoprenyl diphosphate synthases, Proc. Natl. Acad. Sci. U. S. A. 93 (1996) 15018–15023. https://doi.org/10.1073/pnas.93.26.15018.

[22] C.D. Poulter, J.C. Argyle, E.A. Mash, Farnesyl pyrophosphate synthetase. Mechanistic studies of the 1'-4 coupling reaction with 2-fluorogeranyl pyrophosphate., J. Biol. Chem. 253 (1978) 7227–7233. https://doi.org/10.1016/s0021-9258(17)34489-7.

[23] C.D. Poulter, D.M. Satterwhite, Mechanism of the Prenyl Transfer Reaction. Studies with (E)- and (Z)-3-Trifluoromethyl-2-buten-1-yl Pyrophosphate, Biochemistry. 16 (1977) 5470–5478. https://doi.org/10.1021/bi00644a012.

[24] J.C. Smart, B.L. Pinsky, Prenyltransferase. New Evidence for an Ionization-Condensation-Elimination Mechanism with 2-Fluorogeranyl Pyrophosphate, J. Am. Chem. Soc. 99 (1977) 957–959. https://doi.org/10.1021/ja00445a056.

[25] A. Tachibana, Y. Yano, S. Otani, N. Nomura, Y. Sako, M. Taniguchi, Novel prenyltransferase gene encoding farnesylgeranyl diphosphate synthase from a hyperthermophilic archaeon, Aeropyrum pernix. Molecular evolution with alteration in product specificity, Eur. J. Biochem. 267 (2000) 321–328. https://doi.org/10.1046/j.1432-1327.2000.00967.x.

[26] E. Oldfield, F.Y. Lin, Terpene biosynthesis: Modularity rules, Angew. Chemie - Int. Ed. 51 (2012) 1124–1137. https://doi.org/10.1002/anie.201103110.

[27] P.H. Liang, T.P. Ko, A.H.J. Wang, Structure, mechanism and function of prenyltransferases, Eur. J. Biochem. 269 (2002) 3339–3354. https://doi.org/10.1046/j.1432-1033.2002.03014.x.

[28] K.L. Kavanagh, J.E. Dunford, G. Bunkoczi, R.G.G. Russell, U. Oppermann, The crystal structure of human geranylgeranyl pyrophosphate synthase reveals a novel hexameric arrangement and inhibitory product binding, J. Biol. Chem. 281 (2006) 22004–22012. https://doi.org/10.1074/jbc.M602603200.

[29] T.H. Chang, R.T. Guo, T.P. Ko, A.H.J. Wang, P.H. Liang, Crystal structure of type-III geranylgeranyl pyrophosphate synthase from Saccharomyces cerevisiae and the mechanism of product chain length determination, J. Biol. Chem. 281 (2006) 14991–15000. https://doi.org/10.1074/jbc.M512886200.

[30] A. Tachibana, T. Tanaka, M. Taniguchi, S. Oi, Purification and Characterization of Geranylgeranyl Diphosphate Synthase from Methanobacterium Thermoformicicum SF-4, Biosci. Biotechnol. Biochem. 57 (1993) 1129–1133. https://doi.org/10.1271/bbb.57.1129.

[31] K. Wang, S. ichi Ohnuma, Chain-length determination mechanism of isoprenyl diphosphate synthases and implications for molecular evolution, Trends Biochem. Sci. 24 (1999) 445–451. https://doi.org/10.1016/S0968-0004(99)01464-4.

[32] R.T. Guo, R. Cao, P.H. Liang, T.P. Ko, T.H. Chang, M.P. Hudock, W.Y. Jeng, C.K.M. Chen, Y. Zhang, Y. Song, C.J. Kuo, F. Yin, E. Oldfield, A.H.J. Wang, Bisphosphonates target multiple sites in both cis- and trans- prenyltransferases, Proc. Natl. Acad. Sci. U. S. A. 104 (2007) 10022–10027. https://doi.org/10.1073/pnas.0702254104.

[33] C.M. Lacbay, D.D. Waller, J. Park, M.G. Palou, F. Vincent, X.F. Huang, V. Ta, A.M. Berghuis, M. Sebag, Y.S. Tsantrizos, Unraveling the Prenylation-Cancer Paradox in Multiple Myeloma with Novel Geranylgeranyl Pyrophosphate Synthase (GGPPS) Inhibitors, J. Med. Chem. 61 (2018) 6904–6917. https://doi.org/10.1021/acs.jmedchem.8b00886.

[34] J.D. Artz, A.K. Wernimont, J.E. Dunford, M. Schapira, A. Dong, Y. Zhao, J. Lew, R.G.G. Russell, F.H. Ebetino, U. Oppermann, R. Hui, Molecular characterization of a novel geranylgeranyl pyrophosphate synthase from Plasmodium Parasites, J. Biol. Chem. 286 (2011) 3315–3322. https://doi.org/10.1074/jbc.M109.027235.

[35] F. Zhou, C.Y. Wang, M. Gutensohn, L. Jiang, P. Zhang, D. Zhang, N. Dudareva, S. Lua, A recruiting protein of geranylgeranyl diphosphate synthase controls metabolic flux toward chlorophyll biosynthesis in rice, Proc. Natl. Acad. Sci. U. S. A. 114 (2017) 6866–6871. https://doi.org/10.1073/pnas.1705689114.

[36] C. Wang, Q. Chen, D. Fan, J. Li, G. Wang, P. Zhang, Structural Analyses of Short-Chain Prenyltransferases Identify an Evolutionarily Conserved GFPPS Clade in Brassicaceae Plants, Mol. Plant. 9 (2016) 195–204. https://doi.org/10.1016/j.molp.2015.10.010.

[37] M.N. Ashby, P.A. Edwards, Elucidation of the deficiency in two yeast coenzyme Q mutants. Characterization of the structural gene encoding hexaprenyl pyrophosphate synthetase, J. Biol. Chem. 265 (1990) 13157–13164. https://doi.org/10.1016/s0021-9258(19)38280-8.

[38] A. Chen, C. Dale Poulter, P.A. Kroon, Isoprenyl diphosphate synthases: Protein sequence comparisons, a phylogenetic tree, and predictions of secondary structure, Protein Sci. 3 (1994) 600–607. https://doi.org/10.1002/pro.5560030408.

[39] L. Song, C.D. Poulter, Yeast farnesyl-diphosphate synthase: Site-directed mutagenesis of residues in highly conserved prenyltransferase domains I and II, Proc. Natl. Acad. Sci. U. S. A. 91 (1994) 3044–3048. https://doi.org/10.1073/pnas.91.8.3044.

[40] A. Joly, P.A. Edwards, Effect of site-directed mutagenesis of conserved aspartate and arginine residues upon farnesyl diphosphate synthase activity, J. Biol. Chem. 268 (1993) 26983–26989. https://doi.org/10.1016/s0021-9258(19)74207-0.

[41] K.M. Chang, S.H. Chen, C.J. Kuo, C.K. Chang, R.T. Guo, J.M. Yang, P.H. Liang, Roles of amino acids in the Escherichia coli octaprenyl diphosphate synthase active site probed by structure-guided site-directed mutagenesis, Biochemistry. 51 (2012) 3412–3419. https://doi.org/10.1021/bi300069j.

[42] P.H. Liang, Reaction kinetics, catalytic mechanisms, conformational changes, and inhibitor design for prenyltransferases, Biochemistry. 48 (2009) 6562–6570. https://doi.org/10.1021/bi900371p.

[43] Y. Feng, R.M.L. Morgan, P.D. Fraser, K. Hellgardt, P.J. Nixon, Crystal Structure of Geranylgeranyl Pyrophosphate Synthase (CrtE) Involved in Cyanobacterial Terpenoid Biosynthesis, Front. Plant Sci. 11 (2020) 589. https://doi.org/10.3389/fpls.2020.00589.

[44] S.I. Ohnuma, K. Hirooka, H. Hemmi, C. Ishida, C. Ohto, T. Nishino, 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 (1996) 18831–18837. https://doi.org/10.1074/jbc.271.31.18831.

[45] S.I. Ohnuma, H. Hemmi, T. Koyama, K. Ogura, T. Nishino, Recognition of allylic substrates in Sulfolobus acidocaldarius geranylgeranyl diphosphate synthase: Analysis using mutated enzymes and artificial allylic substrates, J. Biochem. 123 (1998) 1036–1040. https://doi.org/10.1093/oxfordjournals.jbchem.a022040.

[46] S.I. Ohnuma, K. Hirooka, N. Tsuruoka, M. Yano, C. Ohto, H. Nakane, T. Nishino, A pathway where polyprenyl diphosphate elongates in prenyltransferase: Insight into a common mechanism of chain length determination of prenyltransferases, J. Biol. Chem. 273 (1998) 26705–26713. https://doi.org/10.1074/jbc.273.41.26705.

[47] S.I. Ohnuma, K. Hirooka, C. Ohto, T. Nishino, Conversion from archaeal geranylgeranyl diphosphate synthase to farnesyl diphosphate synthase: Two amino acids before the first aspartate-rich motif solely determine eukaryotic farnesyl diphosphate synthase activity, J. Biol. Chem. 272 (1997) 5192–5198. https://doi.org/10.1074/jbc.272.8.5192.

[48] M. Lisnyansky, N. Kapelushnik, A. Ben-Bassat, M. Marom, A. Loewenstein, D. Khananshvili, M. Giladi, Y. Haitin, Reduced activity of geranylgeranyl diphosphate synthase mutant is involved in bisphosphonate-induced atypical fractures, Mol. Pharmacol. 94 (2018) 1391–1400. https://doi.org/10.1124/mol.118.113670.

[49] C.K.M. Chen, M.P. Hudock, Y. Zhang, R.T. Guo, R. Cao, H.N. Joo, P.H. Liang, T.P. Ko, T.H. Chang, S.C. Chang, Y. Song, J. Axelson, A. Kumar, A.H.J. Wang, E. Oldfield, Inhibition of geranylgeranyl diphosphate synthase by bisphosphonates: A crystallographic and computational investigation, J. Med. Chem. 51 (2008) 5594–5607. https://doi.org/10.1021/jm800325y.

[50] H. Hemmi, M. Noike, T. Nakayama, T. Nishino, An alternative mechanism of product chain-length determination in type III geranylgeranyl diphosphate synthase, Eur. J. Biochem. 270 (2003) 2186–2194. https://doi.org/10.1046/j.1432-1033.2003.03583.x.

[51] V.E. Ahn, E.I. Lo, C.K. Engel, L. Chen, P.M. Hwang, L.E. Kay, R.E. Bishop, G.G. Privé, A hydrocarbon ruler measures palmitate in the enzymatic acylation of endotoxin, EMBO J. 23 (2004) 2931-2941. https://doi.org/10.1038/sj.emboj.7600320.

[52] J. Payandeh, M. Fujihashi, W. Gillon, E.F. Pai, The crystal structure of (S)-3-O-geranylgeranylglyceryl phosphate synthase reveals an ancient fold for an ancient enzyme, J. Biol. Chem. 281 (2006) 6070–6078. https://doi.org/10.1074/jbc.M509377200.

[53] C. Kropp, K. Straub, M. Linde, P. Babinger, Hexamerization and thermostability emerged very early during geranylgeranylglyceryl phosphate synthase evolution, Protein Sci. 30 (2021) 583–596. https://doi.org/10.1002/pro.4016.

[54] P.N. Blank, A.A. Barnett, T.A. Ronnebaum, K.E. Alderfer, B.N. Gillott, D.W. Christianson, J.A. Himmelberger, Structural studies of geranylgeranylglyceryl phosphate synthase, a prenyltransferase found in thermophilic Euryarchaeota, Acta Crystallogr. Sect. D Struct. Biol. 76 (2020) 542–557. https://doi.org/10.1107/S2059798320004878.

[55] M. Linde, K. Heyn, R. Merkl, R. Sterner, P. Babinger, Hexamerization of Geranylgeranylglyceryl Phosphate Synthase Ensures Structural Integrity and Catalytic Activity at High Temperatures, Biochemistry. 57 (2018) 2335–2348. https://doi.org/10.1021/acs.biochem.7b01284.

[56] N. Nemoto, K.I. Miyazono, M. Tanokura, A. Yamagishi, Crystal structure of (S)-3-O-geranylgeranylglyceryl phosphate synthase from Thermoplasma acidophilum in complex with the substrate sn-glycerol 1-phosphate, Acta Crystallogr. Sect. F Struct. Biol. Commun. 75 (2019) 470–479. https://doi.org/10.1107/S2053230X19007453.

[57] D. Peterhoff, B. Beer, C. Rajendran, E.P. Kumpula, E. Kapetaniou, H. Guldan, R.K. Wierenga, R. Sterner, P. Babinger, 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 (2014) 885–899. https://doi.org/10.1111/mmi.12596.

[58] D. Peterhoff, H. Zellner, H. Guldan, R. Merkl, R. Sterner, P. Babinger, Dimerization Determines Substrate Specificity of a Bacterial Prenyltransferase, ChemBioChem. 13 (2012) 1297–1303. https://doi.org/10.1002/cbic.201200127.

[59] J. Yao, C.O. Rock, Phosphatidic acid synthesis in bacteria, Biochim. Biophys. Acta - Mol. Cell Biol. Lipids. 1831 (2013) 495–502. https://doi.org/10.1016/j.bbalip.2012.08.018.

[60] J.E. Cronan, R.M. Bell, Mutants of Escherichia coli defective in membrane phospholipid synthesis: Mapping of the structural gene for L glycerol 3 phosphate dehydrogenase, J. Bacteriol. 118 (1974) 598–605. https://doi.org/10.1128/jb.118.2.598-605.1974.

[61] R.M. Bell, Mutants of Escherichia coli defective in membrane phospholipid synthesis. Properties of wild type and Km defective sn glycerol 3 phosphate acyltransferase activities, J. Biol. Chem. 250 (1975) 7147–7152. https://doi.org/10.1016/s0021-9258(19)40921-6.

[62] V.A. Lightner, T.J. Larson, P. Tailleur, G.D. Kantor, C.R. Raetz, R.M. Bell, P. Modrich, Membrane phospholipid synthesis in Escherichia coli. Cloning of a structural gene (plsB) of the sn-glycerol-3-phosphate acyltransferase, J. Biol. Chem. 255 (1980) 9413–9420. https://doi.org/10.1016/s0021-9258(19)70578-x.

[63] J.B. Parsons, C.O. Rock, Bacterial lipids: Metabolism and membrane homeostasis, Prog. Lipid Res. 52 (2013) 249–276. https://doi.org/10.1016/j.plipres.2013.02.002.

[64] T.J. Larson, D.N. Ludtke, R.M. Bell, sn-Glycerol-3-phosphate auxotrophy of plsB strains of Escherichia coli: Evidence that a second mutation, plsX, is required, J. Bacteriol. 160 (1984) 711–717. https://doi.org/10.1128/jb.160.2.711-717.1984.

[65] Y.J. Lu, Y.M. Zhang, K.D. Grimes, J. Qi, R.E. Lee, C.O. Rock, Acyl-Phosphates Initiate Membrane Phospholipid Synthesis in Gram-Positive Pathogens, Mol. Cell. 23 (2006) 765–772. https://doi.org/10.1016/j.molcel.2006.06.030.

[66] Y. Kim, H. Li, T.A. Binkowski, D. Holzle, A. Joachimiak, Crystal structure of fatty acid/phospholipid synthesis protein PlsX from Enterococcus faecalis, J. Struct. Funct. Genomics. 10 (2009) 157–163. https://doi.org/10.1007/s10969-008-9052-9.

[67] L. Paoletti, Y.J. Lu, G.E. Schujman, D. De Mendoza, C.O. Rock, Coupling of fatty acid and phospholipid synthesis in Bacillus subtilis, J. Bacteriol. 189 (2007) 5816–5824. https://doi.org/10.1128/JB.00602-07.

[68] A. Chen, D. Zhang, C.D. Poulter, (S)-geranylgeranylglyceryl phosphate synthase. Purification and characterization of the first pathway-specific enzyme in archaebacterial membrane lipid biosynthesis, J. Biol. Chem. 268 (1993) 21701–21705. https://doi.org/10.1016/s0021-9258(20)80598-5.

[69] N. Nemoto, T. Oshima, A. Yamagishi, Purification and characterization of geranylgeranylglyceryl phosphate synthase from a thermoacidophilic archaeon, Thermoplasma acidophilum, J. Biochem. 133 (2003) 651–657. https://doi.org/10.1093/jb/mvg083.

[70] D. Zhang, C.D. Poulter, Biosynthesis of Archaebacterial Lipids in Halobacterium halobium and Methanobacterium thermoautotrophicum, J. Org. Chem. 58 (1993) 3919–3922. https://doi.org/10.1021/jo00067a025.

[71] A. Caforio, M.F. Siliakus, M. Exterkate, S. Jain, V.R. Jumde, R.L.H. Andringa, S.W.M. Kengen, A.J. Minnaard, A.J.M. Driessen, J. van der Oost, Converting Escherichia coli into an archaebacterium with a hybrid heterochiral membrane, Proc. Natl. Acad. Sci. U. S. A. 115 (2018) 3704–3709. https://doi.org/10.1073/pnas.1721604115.

[72] M.C. Vega, E. Lorentzen, A. Linden, M. Wilmanns, Evolutionary markers in the (β/α)8-barrel fold, Curr. Opin. Chem. Biol. 7 (2003) 694–701. https://doi.org/10.1016/j.cbpa.2003.10.004.

[73] N. Nagano, C.A. Orengo, J.M. Thornton, One fold with many functions: The evolutionary relationships between TIM barrel families based on their sequences, structures and functions, J. Mol. Biol. 321 (2002) 741–765. https://doi.org/10.1016/S0022-2836(02)00649-6.

[74] F. Ren, X. Feng, T.P. Ko, C.H. Huang, Y. Hu, H.C. Chan, Y.L. Liu, K. Wang, C.C. Chen, X. Pang, M. He, Y. Li, E. Oldfield, R.T. Guo, Insights into TIM-Barrel Prenyl Transferase Mechanisms: Crystal Structures of PcrB from Bacillus subtilis and Staphylococcus aureus, ChemBioChem. 14 (2013) 195–199. https://doi.org/10.1002/cbic.201200748.

[75] J. Payandeh, E.F. Pai, Enzyme-driven speciation: Crystallizing Archaea via lipid capture, J. Mol. Evol. 64 (2007) 364–374. https://doi.org/10.1007/s00239-006-0141-8.

[76] F. Ren, T.-P. Ko, X. Feng, C.-H. Huang, H.-C. Chan, Y. Hu, K. Wang, Y. Ma, P.-H. Liang, A.H.-J. Wang, E. Oldfield, R.-T. Guo, Insights into the Mechanism of the Antibiotic-Synthesizing Enzyme MoeO5 from Crystal Structures of Different Complexes, Angew. Chemie. 124 (2012) 4233–4236. https://doi.org/10.1002/ange.201108002.

[77] T.J.O. Wyckoff, S. Lin, R.J. Cotter, G.D. Dotson, C.R.H. Raetz, Hydrocarbon rulers in UDP-N-acetylglucosamine acyltransferases, J. Biol. Chem. 273 (1998) 32569–32372. https://doi.org/10.1074/jbc.273.49.32369.

[78] H. Guldan, F.M. Matysik, M. Bocola, R. Sterner, P. Babinger, Functional assignment of an enzyme that catalyzes the synthesis of an archaea-type ether lipid in bacteria, Angew. Chemie - Int. Ed. 50 (2011) 8188–8191. https://doi.org/10.1002/anie.201101832.

[79] Y. Boucher, M. Kamekura, W.F. Doolittle, Origins and evolution of isoprenoid lipid biosynthesis in archaea, Mol. Microbiol. 52 (2004) 515–527. https://doi.org/10.1111/j.1365-2958.2004.03992.x.

[80] C. Vieille, G.J. Zeikus, Hyperthermophilic Enzymes: Sources, Uses, and Molecular Mechanisms for Thermostability, Microbiol. Mol. Biol. Rev. 65 (2001) 1–43. https://doi.org/10.1128/mmbr.65.1.1-43.2001.

[81] R. Sterner, W. Liebl, Thermophilic adaptation of proteins, Crit. Rev. Biochem. Mol. Biol. 36 (2001) 39–106. https://doi.org/10.1080/20014091074174.

[82] S. Ren, N.A.W. de Kok, Y. Gu, W. Yan, Q. Sun, Y. Chen, J. He, L. Tian, R.L.H. Andringa, X. Zhu, M. Tang, S. Qi, H. Xu, H. Ren, X. Fu, A.J. Minnaard, S. Yang, W. Zhang, W. Li, Y. Wei, A.J.M. Driessen, W. Cheng, Structural and Functional Insights into an Archaeal Lipid Synthase, Cell Rep. 33 (2020) 108294. https://doi.org/10.1016/j.celrep.2020.108294.

[83] R.M. Robertson, J. Yao, S. Gajewski, G. Kumar, E.W. Martin, C.O. Rock, S.W. White, A two-helix motif positions the lysophosphatidic acid acyltransferase active site for catalysis within the membrane bilayer, Nat. Struct. Mol. Biol. 24 (2017) 666–671. https://doi.org/10.1038/nsmb.3436.

[84] J. Coleman, Characterization of Escherichia coli cells deficient in 1-acyl-sn-glycerol-3-phosphate acyltransferase activity, J. Biol. Chem. 265 (1990) 17215–17221. https://doi.org/10.1016/s0021-9258(17)44891-5.

[85] J. Coleman, Characterization of the Escherichia coli gene for 1-acyl-sn-glycerol-3-phosphate acyltransferase (plsC), MGG Mol. Gen. Genet. 232 (1992) 295–303. https://doi.org/10.1007/BF00280009.

[86] K. Saiki, T. Mogi, K. Ogura, Y. Anraku, In vitro heme O synthesis by the cyoE gene product from Escherichia coli, J. Biol. Chem. 268 (1993) 26041–26045. https://doi.org/10.1016/s0021-9258(19)74272-0.

[87] K. Saiki, T. Mogi, Y. Anraku, Heme O biosynthesis in Escherichia coli: The cyoe gene in the cytochrome BO operon encodes a protoheme IX farnesyltransferase, Biochem. Biophys. Res. Commun. 189 (1992) 1491–1497. https://doi.org/10.1016/0006-291X(92)90243-E.

[88] E. Collakova, D. DellaPenna, Isolation and functional analysis of homogentisate phytyltransferase from synechocystis sp. PCC 6803 and arabidopsis, Plant Physiol. 127 (2001) 1113–1124. https://doi.org/10.1104/pp.010421.

[89] B. Savidge, J.D. Weiss, Y.H.H. Wong, M.W. Lassner, T.A. Mitsky, C.K. Shewmaker, D. Post-Beittenmiller, H.E. Valentin, Isolation and characterization of homogentisate phytyltransferase genes from Synechocystis sp. PCC 6803 and Arabidopsis, Plant Physiol. 129 (2002) 321–332. https://doi.org/10.1104/pp.010747.

[90] K. Suvarna, D. Stevenson, R. Meganathan, M.E.S. Hudspeth, Menaquinone (vitamin K2) biosynthesis: Localization and characterization of the menA gene from Escherichia coli, J. Bacteriol. 180 (1998) 2782–2787. https://doi.org/10.1128/jb.180.10.2782-2787.1998.

[91] M. Siebert, A. Bechthold, M. Melzer, U. May, U. Berger, G. Schröder, J. Schröder, K. Severin, L. Heide, Ubiquinone biosynthesis Cloning of the genes coding for chorismate pyruvate-lyase and 4-hydroxybenzoate octaprenyl transferase from Escherichia coli, European Biochemical Societies, 1992. https://doi.org/10.1016/0014-5793(92)80710-X.

[92] U. Oster, W. Rüdiger, The G4 gene of Arabidopsis thaliana encodes a chlorophyll synthase of etiolated plants, Bot. Acta. 110 (1997) 420–423. https://doi.org/10.1111/j.1438-8677.1997.tb00658.x.

[93] N. Roy, N. Nemoto, A. Yamagishi, Molecular cloning and Characterization of a membrane-intrinsic (S)-2,3-di-O-digeranylgeranylglyceryl phosphate synthase involved in the biosynthesis of archaeal ether-linked membrane lipids, JJBS Jordan J. Biol. Sci. 3 (2010) 57–64. http://www.jjbs.hu.edu.jo/files/v3n2/Molecular cloning and Characterization of a membrane-intrinsic.pdf%0Ahttp://platform.almanhal.com/Article/Preview.aspx?ID=6200 (accessed June 27, 2017).

[94] H. Hemmi, K. Shibuya, Y. Takahashi, T. Nakayama, T. Nishino, (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 (2004) 50197–50203. https://doi.org/10.1074/jbc.M409207200.

[95] D.J. Hosfield, Y. Zhang, D.R. Dougan, A. Broun, L.W. Tari, R. V. Swanson, J. Finn, Structural Basis for Bisphosphonate-mediated Inhibition of Isoprenoid Biosynthesis, J. Biol. Chem. 279 (2004) 8526–8529. https://doi.org/10.1074/jbc.C300511200.

[96] H. Huang, E.J. Levin, S. Liu, Y. Bai, S.W. Lockless, M. Zhou, Structure of a Membrane-Embedded Prenyltransferase Homologous to UBIAD1, PLoS Biol. 12 (2014) 1–11. https://doi.org/10.1371/journal.pbio.1001911.

[97] H. Zhang, K. Shibuya, H. Hemmi, T. Nishino, G.D. Prestwich, Total synthesis of geranylgeranylglyceryl phosphate enantiomers: Substrates for characterization of 2,3-O-digeranylgeranylglyceryl phosphate synthase, Org. Lett. 8 (2006) 943–946. https://doi.org/10.1021/ol0530878.

[98] M.L. Nickerson, B.N. Kostiha, W. Brandt, W. Fredericks, K.P. Xu, F.S. Yu, B. Gold, J. Chodosh, M. Goldberg, D.W. Lu, M. Yamada, T.M. Tervo, R. Grutzmacher, C. Croasdale, M. Hoeltzenbein, J. Sutphin, S.B. Malkowicz, L. Wessjohann, H.S. Kruth, M. Dean, J.S. Weiss, UBIAD1 mutation alters a mitochondrial prenyltransferase to cause schnyder corneal dystrophy, PLoS One. 5 (2010) e10760. https://doi.org/10.1371/journal.pone.0010760.

[99] Mutations in COQ2 in Familial and Sporadic Multiple-System Atrophy , N. Engl. J. Med. 369 (2013) 233–244. https://doi.org/10.1056/nejmoa1212115.

[100] W. Cheng, W. Li, Structural insights into ubiquinone biosynthesis in membranes, Science (80-.). 343 (2014) 878–881. https://doi.org/10.1126/science.1246774.

[101] S. Jain, A. Caforio, P. Fodran, J.S. Lolkema, A.J. Minnaard, A.J.M. Driessen, Identification of CDP-archaeol synthase, a missing link of ether lipid biosynthesis in Archaea, Chem. Biol. 21 (2014) 1392–1401. https://doi.org/10.1016/j.chembiol.2014.07.022.

[102] H. Morii, M. Nishihara, Y. Koga, CTP:2,3-di-O-geranylgeranyl-sn-glycero-1-phosphate cytidyltransferase in the methanogenic archaeon Methanothermobacter thermoautotrophicus, J. Biol. Chem. 275 (2000) 36568–36574. https://doi.org/10.1074/jbc.M005925200.

[103] S. Ren, A. Caforio, Q. Yang, B. Sun, F. Yu, X. Zhu, J. Wang, C. Dou, Q. Fu, N. Huang, Q. Sun, C. Nie, S. Qi, X. Gong, J. He, Y. Wei, A.J.M. Driessen, W. Cheng, Structural and mechanistic insights into the biosynthesis of CDP-archaeol in membranes, Cell Res. 27 (2017) 1378–1391. https://doi.org/10.1038/cr.2017.122.

[104] S.M. Jensen, M. Brandl, A.H. Treusch, C.S. Ejsing, Structural characterization of ether lipids from the archaeon Sulfolobus islandicus by high-resolution shotgun lipidomics, J. Mass Spectrom. 50 (2015) 476–487. https://doi.org/10.1002/jms.3553.

[105] L. Wörmer, J.S. Lipp, J.M. Schröder, K.U. Hinrichs, Application of two new LC-ESI-MS methods for improved detection of intact polar lipids (IPLs) in environmental samples, Org. Geochem. 59 (2013) 10–21. https://doi.org/10.1016/j.orggeochem.2013.03.004.

[106] K.W. Becker, F.J. Elling, M.Y. Yoshinaga, A. Söllinger, T. Urich, K.U. Hinrichs, Unusual butane- and pentanetriol-based tetraether lipids in Methanomassiliicoccus Luminyensis, a representative of the seventh order of methanogens, Appl. Environ. Microbiol. 82 (2016) 4505–4516. https://doi.org/10.1128/AEM.00772-16.

[107] K.P. Law, C.L. Zhang, Current progress and future trends in mass spectrometry-based archaeal lipidomics, Org. Geochem. 134 (2019) 45–61. https://doi.org/10.1016/j.orggeochem.2019.04.001.

[108] C. Knappy, D. Barillà, J. Chong, D. Hodgson, H. Morgan, M. Suleman, C. Tan, P. Yao, B. Keely, Mono-, di- and trimethylated homologues of isoprenoid tetraether lipid cores in archaea and environmental samples: Mass spectrometric identification and significance, J. Mass Spectrom. 50 (2015) 1420–1432. https://doi.org/10.1002/jms.3709.

[109] H. Morii, T. Eguchi, M. Nishihara, K. Kakinuma, H. König, Y. Koga, A novel ether core lipid with H-shaped C80-isoprenoid hydrocarbon chain from the hyperthermophilic methanogen Methanothermus fervidus, Biochim. Biophys. Acta - Lipids Lipid Metab. 1390 (1998) 339-345. https://doi.org/10.1016/S0005-2760(97)00183-5.

[110] S. Schouten, E.C. Hopmans, R.D. Pancost, J.S. Sinninghe Damsté, Widespread occurrence of structurally diverse tetraether membrane lipids: Evidence for the ubiquitous presence of low-temperature relatives of hyperthermophiles, Proc. Natl. Acad. Sci. U. S. A. 97 (2000) 14421–14426. https://doi.org/10.1073/pnas.97.26.14421.

[111] M.J.L. Hoefs, S. Schouten, J.W. De Leeuw, L.L. King, S.G. Wakeham, J.S. Sinninghe Damsté, Ether lipids of planktonic archaea in the marine water column?, Appl. Environ. Microbiol. 63 (1997) 3090–3095. https://doi.org/10.1128/aem.63.8.3090-3095.1997.

[112] H.A. Boumann, E.C. Hopmans, I. Van De Leemput, H.J.M. Op Den Camp, J. Van De Vossenberg, M. Strous, M.S.M. Jetten, J.S. Sinninghe Damsté, S. Schouten, Ladderane phospholipids in anammox bacteria comprise phosphocholine and phosphoethanolamine headgroups, FEMS Microbiol. Lett. 258 (2006) 297–304. https://doi.org/10.1111/j.1574-6968.2006.00233.x.

[113] N.J. Bale, D.Y. Sorokin, E.C. Hopmans, M. Koenen, W.C. Irene Rijpstra, L. Villanueva, H. Wienk, J.S. Sinninghe Damsté, New insights into the polar lipid composition of extremely halo(alkali)philic euryarchaea from hypersaline lakes, Front. Microbiol. 10 (2019) 377. https://doi.org/10.3389/fmicb.2019.00377.

[114] J. Lombard, P. López-García, D. Moreira, Phylogenomic investigation of phospholipid synthesis in archaea, Archaea. 2012 (2012) 13. https://doi.org/10.1155/2012/630910.

[115] H. Daiyasu, K.I. Kuma, T. Yokoi, H. Morii, Y. Koga, H. Toh, A study of archaeal enzymes involved in polar lipid synthesis linking amino acid sequence information, genomic contexts and lipid composition, Archaea. 1 (2005) 399–410. https://doi.org/10.1155/2005/452563.

[116] A. Caforio, S. Jain, P. Fodran, M. Siliakus, A.J. Minnaard, J. Van Der Oost, A.J.M. Driessen, Formation of the ether lipids archaetidylglycerol and archaetidylethanolamine in Escherichia coli, Biochem. J. 470 (2015) 343–355. https://doi.org/10.1042/BJ20150626.

[117] N. Hulo, C.J.A. Sigrist, V. Le Saux, P.S. Langendijk-Genevaux, L. Bordoli, A. Gattiker, E. De Castro, P. Bucher, A. Bairoch, Recent improvements to the PROSITE database, Nucleic Acids Res. 32 (2004) D134. https://doi.org/10.1093/nar/gkh044.

[118] J.G. Williams, C.R. McMaster, Scanning alanine mutagenesis of the CDP-alcohol phosphotransferase motif of Saccharomyces cerevisiae cholinephosphotransferase, J. Biol. Chem. 273 (1998) 13482–13487. https://doi.org/10.1074/jbc.273.22.13482.

[119] P. Nogly, I. Gushchin, A. Remeeva, A.M. Esteves, N. Borges, P. Ma, A. Ishchenko, S. Grudinin, E. Round, I. Moraes, V. Borshchevskiy, H. Santos, V. Gordeliy, M. Archer, X-ray structure of a CDP-alcohol phosphatidyltransferase membrane enzyme and insights into its catalytic mechanism, Nat. Commun. 5 (2014) 1–10. https://doi.org/10.1038/ncomms5169.

[120] G. Sciara, O.B. Clarke, D. Tomasek, B. Kloss, S. Tabuso, R. Byfield, R. Cohn, S. Banerjee, K.R. Rajashankar, V. Slavkovic, J.H. Graziano, L. Shapiro, F. Mancia, Structural basis for catalysis in a CDP-alcohol phosphotransferase, Nat. Commun. 5 (2014) 1–10. https://doi.org/10.1038/ncomms5068.

[121] K. Gräve, M.D. Bennett, M. Högbom, Structure of Mycobacterium tuberculosis phosphatidylinositol phosphate synthase reveals mechanism of substrate binding and metal catalysis, Commun. Biol. 2 (2019) 1–11. https://doi.org/10.1038/s42003-019-0427-1.

[122] W.. Cleland, The kinetics of enzyme-catalyzed reactions with two or more substrates or products, Biochim. Biophys. Acta - Spec. Sect. Enzymol. Subj. 67 (1963) 104–137. https://doi.org/10.1016/0926-6569(63)90211-6.

[123] A. Dutt, W. Dowhan, Purification and Characterization of a Membrane-Associated Phosphatidylserine Synthase from Bacillus licheniformis, Biochemistry. 24 (1985) 1073–1079. https://doi.org/10.1021/bi00326a001.

[124] M. Aktas, S. Köster, S. Kizilirmak, J.C. Casanova, H. Betz, C. Fritz, R. Moser, Ö. Yildiz, F. Narberhaus, Enzymatic properties and substrate specificity of a bacterial phosphatidylcholine synthase, FEBS J. 281 (2014) 3523–3541. https://doi.org/10.1111/febs.12877.

[125] T. Hirabayashi, T.J. Larson, W. Dowhan, Membrane-Associated Phosphatidylglycerophosphate Synthetase from Escherichia coli: Purification by Substrate Affinity Chromatography on Cytidine 5′ Diphospho-1,2-diacyl-sn-glycerol Sepharose, Biochemistry. 15 (1976) 5205–5211. https://doi.org/10.1021/bi00669a002.

[126] M.S. Bae-Lee, G.M. Carman, Phosphatidylserine synthesis in Saccharomyces cerevisiae. Purification and characterization of membrane-associated phosphatidylserine synthase, J. Biol. Chem. 259 (1984) 10857–10862. https://doi.org/10.1016/s0021-9258(18)90592-2.

[127] H. Morii, S. Kiyonari, Y. Ishino, Y. Koga, 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 (2009) 30766–30774. https://doi.org/10.1074/jbc.M109.034652.

[128] H. Morii, Y. Koga, CDP-2,3-di-O-geranylgeranyl-sn-glycerol:L-serine O-archaetidyltransferase (archaetidylserine synthase) in the methanogenic archaeon Methanothermobacter thermautotrophicus, J. Bacteriol. 185 (2003) 1181–1189. https://doi.org/10.1128/JB.185.4.1181-1189.2003.

[129] J. Gidden, J. Denson, R. Liyanage, D.M. Ivey, J.O. Lay, Lipid compositions in Escherichia coli and Bacillus subtilis during growth as determined by MALDI-TOF and TOF/TOF mass spectrometry, Int. J. Mass Spectrom. 283 (2009) 178–184. https://doi.org/10.1016/j.ijms.2009.03.005.

[130] M. Okada, H. Matsuzaki, I. Shibuya, K. Matsumoto, Cloning, sequencing, and expression in Escherichia coli of the Bacillus subtilis gene for phosphatidylserine synthase, J. Bacteriol. 176 (1994) 7456–7461. https://doi.org/10.1128/jb.176.24.7456-7461.1994.

[131] A. Nishibori, J. Kusaka, H. Hara, M. Umeda, K. Matsumoto, Phosphatidylethanolamine domains and localization of phospholipid synthases in Bacillus subtilis membranes, J. Bacteriol. 187 (2005) 2163–2174. https://doi.org/10.1128/JB.187.6.2163-2174.2005.

[132] A. Dutt, W. Dowhan, Characterization of a membrane-associated cytidine diphosphate-diacylglycerol-dependent phosphatidylserine synthase in bacilli, J. Bacteriol. 147 (1981) 535–542. https://doi.org/10.1128/jb.147.2.535-542.1981.

[133] M.F. Abdul-Halim, S. Schulze, A. DiLucido, F. Pfeiffer, A.W.B. Filho, M. Pohlschroder, Lipid anchoring of archaeosortase substrates and midcell growth in haloarchaea, MBio. 11 (2020) 1–14. https://doi.org/10.1128/mBio.00349-20.

[134] Y. Watanabe, Y. Watanabe, S. Watanabe, Structural Basis for Phosphatidylethanolamine Biosynthesis by Bacterial Phosphatidylserine Decarboxylase, Structure. 28 (2020) 799-809. e5. https://doi.org/10.1016/j.str.2020.04.006.

[135] G. Cho, E. Lee, J. Kim, Structural insights into phosphatidylethanolamine formation in bacterial membrane biogenesis, Sci. Rep. 11 (2021) 5785. https://doi.org/10.1038/s41598-021-85195-5.

[136] J.Y. Choi, M.T. Duraisingh, M. Marti, C. Ben Mamoun, D.R. Voelker, From protease to decarboxylase: The molecular metamorphosis of phosphatidylserine decarboxylase, J. Biol. Chem. 290 (2015) 10972–10980. https://doi.org/10.1074/jbc.M115.642413.

[137] O.B. Ogunbona, O. Onguka, E. Calzada, S.M. Claypool, Multitiered and Cooperative Surveillance of Mitochondrial Phosphatidylserine Decarboxylase 1, Mol. Cell. Biol. 37 (2017). https://doi.org/10.1128/mcb.00049-17.

[138] M. Satre, E.P. Kennedy, Identification of bound pyruvate essential for the activity of phosphatidylserine decarboxylase of Escherichia coli, J. Biol. Chem. 253 (1978) 479–483. https://doi.org/10.1016/s0021-9258(17)38234-0.

[139] Q.X. Li, W. Dowhan, Structural characterization of Escherichia coli phosphatidylserine decarboxylase, J. Biol. Chem. 263 (1988) 11516–11522. https://doi.org/10.1016/s0021-9258(18)37988-2.

[140] Q.X. Li, W. Dowhan, Studies on the mechanism of formation of the pyruvate prosthetic group of phosphatidylserine decarboxylase from Escherichia coli, J. Biol. Chem. 265 (1990) 4111–4115. https://doi.org/10.1016/s0021-9258(19)39709-1.

[141] J.L. Ekstrom, W.D. Tolbert, H. Xiong, A.E. Pegg, S.E. Ealick, Structure of a human S-adenosylmethionine decarboxylase self-processing ester intermediate and mechanism of putrescine stimulation of processing as revealed by the H243A mutant, Biochemistry. 40 (2001) 9495–9504. https://doi.org/10.1021/bi010736o.

[142] S.J. Singer, G.L. Nicolson, The fluid mosaic model of the structure of cell membranes, Science (80-.). 175 (1972) 720–731. https://doi.org/10.1126/science.175.4023.720.

[143] Y. Jiang, X. Dai, M. Qin, Z. Guo, Identification of an amphipathic peptide sensor of the Bacillus subtilis fluid membrane microdomains, Commun. Biol. 2 (2019) 1–9. https://doi.org/10.1038/s42003-019-0562-8.

[144] D. Lingwood, K. Simons, Lipid rafts as a membrane-organizing principle, Science (80-.). 327 (2010) 46–50. https://doi.org/10.1126/science.1174621.

[145] S. Sonnino, A. Prinetti, Membrane Domains and the "Lipid Raft" Concept, Curr. Med. Chem. 20 (2012) 4–21. https://doi.org/10.2174/09298673130103.

[146] J.C.M. Holthuis, G. Van Meer, K. Huitema, Lipid microdomains, lipid translocation and the organization of intracellular membrane transport, Mol. Membr. Biol. 20 (2003) 231–241. https://doi.org/10.1080/0988768031000100768.

[147] N. Mori, T. Moriyama, M. Toyoshima, N. Sato, Construction of global acyl lipid metabolic map by comparative genomics and subcellular localization analysis in the red alga cyanidioschyzon merolae, Front. Plant Sci. 7 (2016) 958. https://doi.org/10.3389/fpls.2016.00958.

[148] D. Gully, E. Bouveret, A protein network for phospholipid synthesis uncovered by a variant of the tandem affinity purification method in Escherichia coli, Proteomics. 6 (2006) 282–293. https://doi.org/10.1002/pmic.200500115.

[149] W. Dowhan, W.T. Wickner, E.P. Kennedy, Purification and properties of phosphatidylserine decarboxylase from Escherichia coli, J. Biol. Chem. 249 (1974) 3079–3084. https://doi.org/10.1016/s0021-9258(19)42640-9.

[150] H. Morii, M. Ogawa, K. Fukuda, H. Taniguchi, Y. Koga, A revised biosynthetic pathway for phosphatidylinositol in Mycobacteria, J. Biochem. 148 (2010) 593–602. https://doi.org/10.1093/jb/mvq093.

[151] M. Belcher Dufrisne, C.D. Jorge, C.G. Timóteo, V.I. Petrou, K.U. Ashraf, S. Banerjee, O.B. Clarke, H. Santos, F. Mancia, Structural and Functional Characterization of Phosphatidylinositol-Phosphate Biosynthesis in Mycobacteria, J. Mol. Biol. 432 (2020) 5137–5151. https://doi.org/10.1016/j.jmb.2020.04.028.

[152] Y.Y. Chang, E.P. Kennedy, Biosynthesis of phosphatidyl glycerophosphate in Escherichia coli., J. Lipid Res. 8 (1967) 447–455. https://doi.org/10.1016/s0022-2275(20)38901-x.

[153] B. Yang, H. Yao, D. Li, Z. Liu, The phosphatidylglycerol phosphate synthase PgsA utilizes a trifurcated amphipathic cavity for catalysis at the membrane-cytosol interface, Curr. Res. Struct. Biol. 3 (2021) 312–323. https://doi.org/10.1016/j.crstbi.2021.11.005.

[154] O.B. Clarke, D. Tomasek, C.D. Jorge, M.B. Dufrisne, M. Kim, S. Banerjee, K.R. Rajashankar, L. Shapiro, W.A. Hendrickson, H. Santos, F. Mancia, Structural basis for phosphatidylinositol-phosphate biosynthesis, Nat. Commun. 6 (2015) 1–11. https://doi.org/10.1038/ncomms9505.

[155] H. Morii, M. Ogawa, K. Fukuda, H. Taniguchi, Ubiquitous distribution of phosphatidylinositol phosphate synthase and archaetidylinositol phosphate synthase in Bacteria and Archaea, which contain inositol phospholipid, Biochem. Biophys. Res. Commun. 443 (2014) 86–90. https://doi.org/10.1016/j.bbrc.2013.11.054.

[156] T. Icho, C.R.H. Raetz, Multiple genes for membrane-bound phosphatases in Escherichia coli and their action on phospholipid precursors, J. Bacteriol. 153 (1983) 722–730. https://doi.org/10.1128/jb.153.2.722-730.1983.

[157] Y.H. Lu, Z. Guan, J. Zhao, C.R.H. Raetz, Three phosphatidylglycerol-phosphate phosphatases in the inner membrane of Escherichia coli, J. Biol. Chem. 286 (2011) 5506–5518. https://doi.org/10.1074/jbc.M110.199265.

[158] C.R. Funk, L. Zimniak, W. Dowhan, The pgpA and pgpB genes of Escherichia coli are not essential: Evidence for a third phosphatidylglycerophosphate phosphatase, J. Bacteriol. 174 (1992) 205–213. https://doi.org/10.1128/jb.174.1.205-213.1992.

[159] T. Icho, Membrane-bound phosphatases in Escherichia coli: sequence of the pgpB gene and dual subcellular localization of the pgpB product., J. Bacteriol. 170 (1988) 5117–5124. https://doi.org/10.1128/jb.170.11.5117-5124.1988.

[160] P. Lopalco, S. Lobasso, F. Babudri, A. Corcelli, Osmotic shock stimulates de novo synthesis of two cardiolipins in an extreme halophilic archaeon, J. Lipid Res. 45 (2004) 194–201. https://doi.org/10.1194/jlr.M300329-JLR200.

[161] T. Romantsov, Z. Guan, J.M. Wood, Cardiolipin and the osmotic stress responses of bacteria, Biochim. Biophys. Acta - Biomembr. 1788 (2009) 2092–2100. https://doi.org/10.1016/j.bbamem.2009.06.010.

[162] M. Bogdanov, E. Mileykovskaya, W. Dowhan, Lipids in the assembly of membrane proteins and organization of protein supercomplexes: Implications for lipid-linked disorders, Subcell. Biochem. 49 (2008) 197–239. https://doi.org/10.1007/978-1-4020-8831-5_8.

[163] M. Schlame, M. Ren, The role of cardiolipin in the structural organization of mitochondrial membranes, Biochim. Biophys. Acta - Biomembr. 1788 (2009) 2080–2083. https://doi.org/10.1016/j.bbamem.2009.04.019.

[164] A. Mühleip, S.E. McComas, A. Amunts, Structure of a mitochondrial ATP synthase with bound native cardiolipin, Elife. 8 (2019). https://doi.org/10.7554/eLife.51179.

[165] M. Tsai, R.L. Ohniwa, Y. Kato, S.L. Takeshita, T. Ohta, S. Saito, H. Hayashi, K. Morikawa, Staphylococcus aureus requires cardiolipin for survival under conditions of high salinity, BMC Microbiol. 11 (2011) 1–12. https://doi.org/10.1186/1471-2180-11-13.

[166] T.H. Haines, N.A. Dencher, Cardiolipin: A proton trap for oxidative phosphorylation, FEBS Lett. 528 (2002) 35–39. https://doi.org/10.1016/S0014-5793(02)03292-1.

[167] M. Schlame, Cardiolipin synthesis for the assembly of bacterial and mitochondrial membranes, J. Lipid Res. 49 (2008) 1607–1620. https://doi.org/10.1194/jlr.R700018-JLR200.

[168] S. Lobasso, P. Lopalco, V.M.T. Lattanzio, A. Corcelli, Osmotic shock induces the presence of glycocardiolipin in the purple membrane of Halobacterium salinarum, J. Lipid Res. 44 (2003) 2120–2126. https://doi.org/10.1194/jlr.M300212-JLR200.

[169] A. Corcelli, M. Colella, G. Mascolo, F.P. Fanizzi, M. Kates, A novel glycolipid and phospholipid in the purple membrane, Biochemistry. 39 (2000) 3318–3326. https://doi.org/10.1021/bi992462z.

[170] R. Arias-Cartin, S. Grimaldi, P. Arnoux, B. Guigliarelli, A. Magalon, Cardiolipin binding in bacterial respiratory complexes: Structural and functional implications, Biochim. Biophys. Acta - Bioenerg. 1817 (2012) 1937–1949. https://doi.org/10.1016/j.bbabio.2012.04.005.

[171] T. Romantsov, K. Gonzalez, N. Sahtout, D.E. Culham, C. Coumoundouros, J. Garner, C.H. Kerr, L. Chang, R.J. Turner, J.M. Wood, Cardiolipin synthase A colocalizes with cardiolipin and osmosensing transporter ProP at the poles of Escherichia coli cells, Mol. Microbiol. 107 (2018) 623–638. https://doi.org/10.1111/mmi.13904.

[172] M. Klingenberg, Cardiolipin and mitochondrial carriers, Biochim. Biophys. Acta - Biomembr. 1788 (2009) 2048–2058. https://doi.org/10.1016/j.bbamem.2009.06.007.

[173] M. Schlame, M.L. Greenberg, Cardiolipin synthase from yeast, Biochim. Biophys. Acta - Lipids Lipid Metab. 1348 (1997) 201–206. https://doi.org/10.1016/S0005-2760(97)00117-3.

[174] R. Angelini, P. Corral, P. Lopalco, A. Ventosa, A. Corcelli, Novel ether lipid cardiolipins in archaeal membranes of extreme haloalkaliphiles, Biochim. Biophys. Acta - Biomembr. 1818 (2012) 1365–1373. https://doi.org/10.1016/j.bbamem.2012.02.014.

[175] V.M.T. Lattanzio, A. Corcelli, G. Mascolo, A. Oren, Presence of two novel cardiolipins in the halophilic archaeal community in the crystallizer brines from the salterns of Margherita di Savoia (Italy) and Eilat (Israel), Extremophiles. 6 (2002) 437–444. https://doi.org/10.1007/s00792-002-0279-2.

[176] M.Y. Yoshinaga, L. Wörmer, M. Elvert, K.U. Hinrichs, Novel cardiolipins from uncultured methane-metabolizing Archaea, Archaea. 2012 (2012) 832097. https://doi.org/10.1155/2012/832097.

[177] G.D. Sprott, S. Larocque, N. Cadotte, C.J. Dicaire, M. McGee, J.R. Brisson, Novel polar lipids of halophilic eubacterium Planococcus H8 and archaeon Haloferax volcanii, Biochim. Biophys. Acta - Mol. Cell Biol. Lipids. 1633 (2003) 179–188. https://doi.org/10.1016/j.bbalip.2003.08.001.

[178] D. Desmarais, P.E. Jablonski, N.S. Fedarko, M.F. Roberts, 2-Sulfotrehalose, a Novel Osmolyte in Haloalkaliphilic Archaea, J. Bacteriol. 179 (1997) 3146–3153. https://doi.org/10.1128/jb.179.10.3146-3153.1997.

[179] L. Chen, E.T. Spiliotis, M.F. Roberts, Biosynthesis of di-myo-inositol-1,1′-phosphate, a novel osmolyte in hyperthermophilic archaea, J. Bacteriol. 180 (1998) 3785–3792. https://doi.org/10.1128/jb.180.15.3785-3792.1998.

[180] M.F. Roberts, Osmoadaptation and osmoregulation in archaea: update 2004., Front. Biosci. 9 (2004) 1999–2019. https://doi.org/10.2741/1366.

[181] A. Corcelli, The cardiolipin analogues of Archaea, Biochim. Biophys. Acta - Biomembr. 1788 (2009) 2101–2106. https://doi.org/10.1016/j.bbamem.2009.05.010.

[182] M. Sandoval-Calderón, O. Geiger, Z. Guan, F. Barona-Gómez, C. Sohlenkamp, A eukaryote-like cardiolipin synthase is present in Streptomyces coelicolor and in most actinobacteria, J. Biol. Chem. 284 (2009) 1738317390. https://doi.org/10.1074/jbc.M109.006072.

[183] A. Jeucken, J.B. Helms, J.F. Brouwers, Cardiolipin synthases of Escherichia coli have phospholipid class specific phospholipase D activity dependent on endogenous and foreign phospholipids, Biochim. Biophys. Acta - Mol. Cell Biol. Lipids. 1863 (2018) 1345–1353. https://doi.org/10.1016/j.bbalip.2018.06.017.

[184] M. Exterkate, N.A.W. De Kok, R.L.H. Andringa, N.H.J. Wolbert, A.J. Minnaard, A.J.M. Driessen, A promiscuous archaeal cardiolipin synthase enables construction of diverse natural and unnatural phospholipids, J. Biol. Chem. 296 (2021) 100691. https://doi.org/10.1016/j.jbc.2021.100691.

[185] D. Guo, B.E. Tropp, A second Escherichia coli protein with CL synthase activity, Biochim. Biophys. Acta - Mol. Cell Biol. Lipids. 1483 (2000) 263–274. https://doi.org/10.1016/S1388-1981(99)00193-6.

[186] B.K. Tan, M. Bogdanov, J. Zhao, W. Dowhan, C.R.H. Raetz, Z. Guan, Discovery of a cardiolipin synthase utilizing phosphatidylethanolamine and phosphatidylglycerol as substrates, Proc. Natl. Acad. Sci. U. S. A. 109 (2012) 16504–16509. https://doi.org/10.1073/pnas.1212797109.

[187] G.E. Crooks, G. Hon, J.M. Chandonia, S.E. Brenner, WebLogo: A sequence logo generator, Genome Res. 14 (2004) 1188–1190. https://doi.org/10.1101/gr.849004.

[188] S. Kumar, G. Stecher, M. Li, C. Knyaz, K. Tamura, MEGA X: Molecular evolutionary genetics analysis across computing platforms, Mol. Biol. Evol. 35 (2018) 1547–1549. https://doi.org/10.1093/molbev/msy096.

[189] D.T. Jones, W.R. Taylor, J.M. Thornton, The rapid generation of mutation data matrices from protein sequences, Bioinformatics. 8 (1992) 275–282. https://doi.org/10.1093/bioinformatics/8.3.275.

[190] F. Madeira, Y.M. Park, J. Lee, N. Buso, T. Gur, N. Madhusoodanan, P. Basutkar, A.R.N. Tivey, S.C. Potter, R.D. Finn, R. Lopez, The EMBL-EBI search and sequence analysis tools APIs in 2019, Nucleic Acids Res. 47 (2019) W636–W641. https://doi.org/10.1093/nar/gkz268.

[191] R.L. Ohniwa, K. Kitabayashi, K. Morikawa, Alternative cardiolipin synthase Cls1 compensates for stalled Cls2 function in Staphylococcus aureus under conditions of acute acid stress, FEMS Microbiol. Lett. 338 (2013) 141–146. https://doi.org/10.1111/1574-6968.12037.

[192] P. Reinink, J. Buter, V.K. Mishra, E. Ishikawa, T.Y. Cheng, P.T.J. Willemsen, S. Porwollik, P.J. Brennan, E. Heinz, J.A. Mayfield, G. Dougan, C.A. Van Els, V. Cerundolo, G. Napolitani, S. Yamasaki, A.J. Minnaard, M. McClelland, D. Branch Moody, I. Van Rhijn, Discovery of Salmonella trehalose phospholipids reveals functional convergence with mycobacteria, J. Exp. Med. 216 (2019) 757–771. https://doi.org/10.1084/jem.20181812.

[193] C. Li, B.K. Tan, J. Zhao, Z. Guan, In vivo and in vitro synthesis of phosphatidylglycerol by an Escherichia coli cardiolipin synthase, J. Biol. Chem. 291 (2016) 25144–25153. https://doi.org/10.1074/jbc.M116.762070.

[194] E. V. Koonin, A duplicated catalytic motif in a new superfamily of phosphohydrolases and phospholipid synthases that includes poxvirus envelope proteins, Trends Biochem. Sci. 21 (1996) 242–243. https://doi.org/10.1016/S0968-0004(96)30024-8.

[195] C.P. Ponting, I.D. Kerr, A novel family of phospholipase D homologues that includes phospholipid synthases and putative endonucleases: Identification of duplicated repeats and potential active site residues, Protein Sci. 5 (1996) 914–922. https://doi.org/10.1002/pro.5560050513.

[196] J.A. Stuckey, J.E. Dixon, Crystal structure of a phospholipase D family member, Nat. Struct. Biol. 6 (1999) 278–284. https://doi.org/10.1038/6716.

[197] T.C. Sung, R.L. Roper, Y. Zhang, S.A. Rudge, R. Temel, S.M. Hammond, A.J. Morris, B. Moss, J.A. Engebrecht, M.A. Frohman, Mutagenesis of phospholipase D defines a superfamily including a trans-Golgi viral protein required for poxvirus pathogenicity, EMBO J. 16 (1997) 4519–4530. https://doi.org/10.1093/emboj/16.15.4519.

[198] B.R. Quigley, B.E. Tropp, *E. coli* cardiolipin synthase: Function of N-terminal conserved residues, Biochim. Biophys. Acta - Biomembr. 1788 (2009) 2107–2113. https://doi.org/10.1016/j.bbamem.2009.03.016.

[199] L. Ragolia, B.E. Tropp, The effects of phosphoglycerides on Escherichia coli cardiolipin synthase, Biochim. Biophys. Acta (BBA)/Lipids Lipid Metab. 1214 (1994) 323–332. https://doi.org/10.1016/0005-2760(94)90080-9.

[200] I. Shibuya, S. Yamagoe, C. Miyazaki, H. Matsuzaki, A. Ohta, Biosynthesis of novel acidic phospholipid analogs in Escherichia coli, J. Bacteriol. 161 (1985) 473–477. https://doi.org/10.1128/jb.161.2.473-477.1985.

[201] J. Li, F. Yu, H. Guo, R. Xiong, W. Zhang, F. He, M. Zhang, P. Zhang, Crystal structure of plant PLDα1 reveals catalytic and regulatory mechanisms of eukaryotic phospholipase D, Cell Res. 30 (2020) 61–69. https://doi.org/10.1038/s41422-019-0244-6.

[202] I. Leiros, F. Secundo, C. Zambonelli, S. Servi, E. Hough, The first crystal structure of a phospholipase D, Structure. 8 (2000) 655–667. https://doi.org/10.1016/S0969-2126(00)00150-7.

[203] L.A. Kelley, S. Mezulis, C.M. Yates, M.N. Wass, M.J.E. Sternberg, The Phyre2 web portal for protein modeling, prediction and analysis, Nat. Protoc. 10 (2015) 845–858. https://doi.org/10.1038/nprot.2015.053.

[204] I. Leiros, S. McSweeney, E. Hough, The reaction mechanism of phospholipase D from Streptomyces sp. strain PMF. Snapshots along the reaction pathway reveal a pentacoordinate reaction intermediate and an unexpected final product, J. Mol. Biol. 339 (2004) 805–820. https://doi.org/10.1016/j.jmb.2004.04.003.

[205] C. Ogino, H. Daido, Y. Ohmura, N. Takada, Y. Itou, A. Kondo, H. Fukuda, N. Shimizu, Remarkable enhancement in PLD activity from Streptoverticillium cinnamoneum by substituting serine residue into the GG/GS motif, Biochim. Biophys. Acta - Proteins Proteomics. 1774 (2007) 671–678. https://doi.org/10.1016/j.bbapap.2007.04.004.

[206] Y. Nakazawa, M. Uchino, Y. Sagane, H. Sato, K. Takano, Isolation and characterization of actinomycetes strains that produce phospholipase D having high transphosphatidylation activity, Microbiol. Res. 164 (2009) 43–48. https://doi.org/10.1016/j.micres.2006.11.003.

[207] T. Hagishita, M. Nishikawa, T. Hatanaka, Isolation of phospholipase D producing microorganisms with high transphosphatidylation activity, Biotechnol. Lett. 22 (2000) 1587–1590. https://doi.org/10.1023/A:1005644032415.

[208] R. Sato, Y. Itabashi, T. Hatanaka, A. Kuksis, Asymmetric in vitro synthesis of diastereomeric phosphatidylglycerols from phosphatidylcholine and glycerol by bacterial phospholipase D, Lipids. 39 (2004) 1013–1018. https://doi.org/10.1007/s11745-004-1324-1.

[209] Y. Uesugi, K. Mori, J. Arima, M. Iwabuchi, T. Hatanaka, Recognition of phospholipids in Streptomyces phospholipase D, J. Biol. Chem. 280 (2005) 26143–26151. https://doi.org/10.1074/jbc.M414319200.

[210] Y. Uesugi, J. Arima, M. Iwabuchi, T. Hatanaka, C-terminal loop of Streptomyces phospholipase D has multiple functional roles, Protein Sci. 16 (2006) 197–207. https://doi.org/10.1110/ps.062537907.

[211] A. Masayama, T. Takahashi, K. Tsukada, S. Nishikawa, R. Takahashi, M. Adachi, K. Koga, A. Suzuki, T. Yamane, H. Nakano, Y. Iwasaki, Streptomyces phospholipase D mutants with altered substrate specificity capable of phosphatidylinositol synthesis, ChemBioChem. 9 (2008) 974–981. https://doi.org/10.1002/cbic.200700528.

[212] E.B. Gottlin, A.E. Rudolph, Y. Zhao, H.R. Matthews, J.E. Dixon, Catalytic mechanism of the phospholipase D superfamily proceeds via a covalent phosphohistidine intermediate, Proc. Natl. Acad. Sci. U. S. A. 95 (1998) 9202–9207. https://doi.org/10.1073/pnas.95.16.9202.

[213] Y. Nishimura, T. Eguchi, Biosynthesis of archaeal membrane lipids: Digeranylgeranylglycerophospholipid reductase of the thermoacidophilic archaeon Thermoplasma acidophilum, J. Biochem. 139 (2006) 1073–1081. https://doi.org/10.1093/jb/mvj118.

[214] Y. Nishimura, T. Eguchi, Stereochemistry of reduction in digeranylgeranylglycerophospholipid reductase involved in the biosynthesis of archaeal membrane lipids from Thermoplasma acidophilum, Bioorg. Chem. 35 (2007) 276–283. https://doi.org/10.1016/j.bioorg.2006.12.001.

[215] Q. Xu, T. Eguchi, I.I. Mathews, C.L. Rife, H.J. Chiu, C.L. Farr, J. Feuerhelm, L. Jaroszewski, H.E. Klock, M.W. Knuth, M.D. Miller, D. Weekes, M.A. Elsliger, A.M. Deacon, A. Godzik, S.A. Lesley, I.A. Wilson, 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 (2010) 403–417. https://doi.org/10.1016/j.jmb.2010.09.032.

[216] L.E. Cybulski, D. Albanesi, M.C. Mansilla, S. Altabe, P.S. Aguilar, D. De Mendoza, Mechanism of membrane fluidity optimization: Isothermal control of the Bacillus subtilis acyl-lipid desaturase, Mol. Microbiol. 45 (2002) 1379–1388. https://doi.org/10.1046/j.1365-2958.2002.03103.x.

[217] Y. Fujita, H. Matsuoka, K. Hirooka, Regulation of fatty acid metabolism in bacteria, Mol. Microbiol. 66 (2007) 829–839. https://doi.org/10.1111/j.1365-2958.2007.05947.x.

[218] R.J. Heath, C.O. Rock, Roles of the FabA and FabZ β-hydroxyacyl-acyl carrier protein dehydratases in Escherichia coli fatty acid biosynthesis, J. Biol. Chem. 271 (1996) 27795–27801. https://doi.org/10.1074/jbc.271.44.27795.

[219] Y. Feng, J.E. Cronan, Escherichia coli unsaturated fatty acid synthesis. Complex trancription of the fabA gene and in vivo identification of the essential reaction catalyzed by FabB, J. Biol. Chem. 284 (2009) 29526–29535. https://doi.org/10.1074/jbc.M109.023440.

[220] G.J. Dodge, A. Patel, K.L. Jaremko, J.A. McCammon, J.L. Smith, M.D. Burkart, Structural and dynamical rationale for fatty acid unsaturation in Escherichia coli, Proc. Natl. Acad. Sci. U. S. A. 116 (2019) 6775–6783. https://doi.org/10.1073/pnas.1818686116.

[221] D. Sasaki, M. Fujihashi, Y. Iwata, M. Murakami, T. Yoshimura, H. Hemmi, K. Miki, Structure and mutation analysis of archaeal geranylgeranyl reductase, J. Mol. Biol. 409 (2011) 543–557. https://doi.org/10.1016/j.jmb.2011.04.002.

[222] Y. Kung, R.P. McAndrew, X. Xie, C.C. Liu, J.H. Pereira, P.D. Adams, J.D. Keasling, Constructing tailored isoprenoid products by structure-guided modification of geranylgeranyl reductase, Structure. 22 (2014) 1028–1036. https://doi.org/10.1016/j.str.2014.05.007.

[223] L. Holm, C. Sander, Dali: a network tool for protein structure comparison, Trends Biochem. Sci. 20 (1995) 478–480. https://doi.org/10.1016/S0968-0004(00)89105-7.

[224] R. Cervinka, D. Becker, S. Lüdeke, S.V. Albers, T. Netscher, M. Müller, Enzymatic Asymmetric Reduction of Unfunctionalized C=C Bonds with Archaeal Geranylgeranyl Reductases, ChemBioChem. 22 (2021) 2693–2696. https://doi.org/10.1002/cbic.202100290.

[225] M. Murakami, K. Shibuya, T. Nakayama, T. Nishino, T. Yoshimura, H. Hemmi, Geranylgeranyl reductase involved in the biosynthesis of archaeal membrane lipids in the hyperthermophilic archaeon Archaeoglobus fulgidus, FEBS J. 274 (2007) 805–814. https://doi.org/10.1111/j.1742-4658.2006.05625.x.

[226] S. Sato, M. Murakami, T. Yoshimura, H. Hemmi, 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 (2008) 3923–3929. https://doi.org/10.1128/JB.00082-08.

[227] C.W. Meadows, F. Mingardon, B.M. Garabedian, E.E.K. Baidoo, V.T. Benites, A. V. Rodrigues, R. Abourjeily, A. Chanal, T.S. Lee, Discovery of novel geranylgeranyl reductases and characterization of their substrate promiscuity 06 Biological Sciences 0605 Microbiology 06 Biological Sciences 0601 Biochemistry and Cell Biology, Biotechnol. Biofuels. 11 (2018) 1–17. https://doi.org/10.1186/s13068-018-1342-2.

[228] K. Isobe, T. Ogawa, K. Hirose, T. Yokoi, T. Yoshimura, H. Hemmi, 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 (2014) 417–423. https://doi.org/10.1128/JB.00927-13.

[229] J. Oberto, SyntTax: A web server linking synteny to prokaryotic taxonomy, BMC Bioinformatics. 14 (2013). https://doi.org/10.1186/1471-2105-14-4.

[230] J.S.S. Damsté, W.I.C. Rijpstra, E.C. Hopmans, S. Schouten, M. Balk, A.J.M. Stams, Structural characterization of diabolic acid-based tetraester, tetraether and mixed ether/ester, membrane-spanning lipids of bacteria from the order Thermotogales, Arch. Microbiol. 188 (2007) 629–641. https://doi.org/10.1007/s00203-007-0284-z.

[231] S. Schouten, M.T.J. Van Der Meer, E.C. Hopmans, W.I.C. Rijpstra, A.L. Reysenbach, D.M. Ward, J.S.S. Damsté, Archaeal and bacterial glycerol dialkyl glycerol tetraether lipids in hot springs of Yellowstone National Park, Appl. Environ. Microbiol. 73 (2007) 6181-6191. https://doi.org/10.1128/AEM.00630-07.

[232] J.S. Sinninghe Damsté, W.I.C. Rijpstra, E.C. Hopmans, B.U. Foesel, P.K. Wüst, J. Overmann, M. Tank, D.A. Bryant, P.F. Dunfield, K. Houghton, M.B. Stott, Ether- and ester-bound iso-diabolic acid and other lipids in members of Acidobacteria subdivision 4, Appl. Environ. Microbiol. 80 (2014) 5207–5218. https://doi.org/10.1128/AEM.01066-14.

[233] M. de Rosa, S. de Rosa, A. Gambacorta, L. Minale, J.D. Bu'lock, Chemical structure of the ether lipids of thermophilic acidophilic bacteria of the Caldariella group, Phytochemistry. 16 (1977) 1961–1965. https://doi.org/10.1016/0031-9422(77)80105-2.

[234] B.D.A. Naafs, D. McCormick, G.N. Inglis, R.D. Pancost, Archaeal and bacterial H-GDGTs are abundant in peat and their relative abundance is positively correlated with temperature, Geochim. Cosmochim. Acta. 227 (2018) 156–170. https://doi.org/10.1016/J.GCA.2018.02.025.

[235] D.R. Rohr, J. Toulouse, K. Pernal, Combining density-functional theory and density-matrix-functional theory, Phys. Rev. A - At. Mol. Opt. Phys. 82 (2010) 1375–1381. https://doi.org/10.1103/PhysRevA.82.052502.

[236] S.M. Jensen, V.L. Neesgaard, S.L.N. Skjoldbjerg, M. Brandl, C.S. Ejsing, A.H. Treusch, The effects of temperature and growth phase on the lipidomes of Sulfolobus islandicus and Sulfolobus tokodaii, Life. 5 (2015) 1539–1566. https://doi.org/10.3390/life5031539.

[237] J. Quehenberger, E. Pittenauer, G. Allmaier, O. Spadiut, The influence of the specific growth rate on the lipid composition of Sulfolobus acidocaldarius, Extremophiles. 24 (2020) 413–420. https://doi.org/10.1007/s00792-020-01165-1.

[238] M. Tourte, P. Schaeffer, V. Grossi, P.M. Oger, Functionalized Membrane Domains: An Ancestral Feature of Archaea?, Front. Microbiol. 11 (2020) 526. https://doi.org/10.3389/fmicb.2020.00526.

[239] E.C. Hopmans, S. Schouten, R.D. Pancost, M.T.J. Van Der Meer, J.S. Sinninghe Damsté, Analysis of intact tetraether lipids in archaeal cell material and sediments by high performance liquid chromatography/atmospheric pressure chemical ionization mass spectrometry, Rapid Commun. Mass Spectrom. 14 (2000) 585–589. https://doi.org/10.1002/(SICI)1097-0231(20000415)14:7<585::AID-RCM913>3.0.CO;2-N.

[240] J.S. Sinninghe Damsté, S. Schouten, E.C. Hopmans, A.C.T. Van Duin, J.A.J. Geenevasen, Crenarchaeol: The characteristic core glycerol dibiphytanyl glycerol tetraether membrane lipid of cosmopolitan pelagic crenarchaeota, J. Lipid Res. 43 (2002) 1641–1651. https://doi.org/10.1194/jlr.M200148-JLR200.

[241] D. Lai, J.R. Springstead, H.G. Monbouquette, Effect of growth temperature on ether lipid biochemistry in Archaeoglobus fulgidus, Extremophiles. 12 (2008) 271–278. https://doi.org/10.1007/s00792-007-0126-6.

[242] M. Sinensky, Homeoviscous adaptation: a homeostatic process that regulates the viscosity of membrane lipids in Escherichia coli, Proc. Natl. Acad. Sci. U. S. A. 71 (1974) 522–525. https://doi.org/10.1073/pnas.71.2.522.

[243] H. Shimada, N. Nemoto, Y. Shida, T. Oshima, A. Yamagishi, Effects of pH and temperature on the composition of polar lipids in Thermoplasma acidophilum HO-62, J. Bacteriol. 190 (2008) 5404–5411. https://doi.org/10.1128/JB.00415-08.

[244] I. Uda, A. Sugai, Y.H. Itoh, T. Itoh, Variation in Molecular Species of Core Lipids from the Order Thermoplasmales Strains Depends on the Growth Temperature, J. Oleo Sci. 53 (2004) 399–404. https://doi.org/10.5650/jos.53.399.

[245] I. Uda, A. Sugai, Y.H. Itoh, T. Itoh, Variation in molecular species of polar lipids from Thermoplasma acidophilum depends on growth temperature, Lipids. 36 (2001) 103–105. https://doi.org/10.1007/s11745-001-0914-2.

[246] S. Schouten, E.C. Hopmans, E. Schefuß, J.S. Sinninghe Damsté, Distributional variations in marine crenarchaeotal membrane lipids: A new tool for reconstructing ancient sea water temperatures?, Earth Planet. Sci. Lett. 204 (2002) 265–274. https://doi.org/10.1016/S0012-821X(02)00979-2.

[247] E.S. Boyd, T.L. Hamilton, J. Wang, L. He, C.L. Zhang, The role of tetraether lipid composition in the adaptation of thermophilic archaea to acidity, Front. Microbiol. 4 (2013) 62. https://doi.org/10.3389/fmicb.2013.00062.

[248] J.P. Nicolas, A molecular dynamics study of an archaeal tetraether lipid membrane: Comparison with a dipalmitoylphosphatidylcholine lipid bilayer, Lipids. 40 (2005) 1023–1030. https://doi.org/10.1007/s11745-005-1465-2.

[249] L.F. Pineda De Castro, M. Dopson, R. Friedman, Biological membranes in extreme conditions: Simulations of anionic archaeal tetraether lipid membranes, PLoS One. 11 (2016) e0155287. https://doi.org/10.1371/journal.pone.0155287.

[250] J.L. Gabriel, P. Lee Gau Chong, Molecular modeling of archaebacterial bipolar tetraether lipid membranes, Chem. Phys. Lipids. 105 (2000) 193–200. https://doi.org/10.1016/S0009-3084(00)00126-2.

[251] Z. Zeng, X.L. Liu, K.R. Farley, J.H. Wei, W.W. Metcalf, R.E. Summons, P. V. Welander, GDGT cyclization proteins identify the dominant archaeal sources of tetraether lipids in the ocean, Proc. Natl. Acad. Sci. U. S. A. 116 (2019) 22505–22511. https://doi.org/10.1073/pnas.1909306116.

[252] Z. Zeng, H. Chen, H. Yang, Y. Chen, W. Yang, X. Feng, H. Pei, P. V. Welander, Identification of a protein responsible for the synthesis of archaeal membrane-spanning GDGT lipids, Nat. Commun. 13 (2022) 1–9. https://doi.org/10.1038/s41467-022-29264-x.

[253] Z. Zeng, X.L. Liu, J.H. Wei, R.E. Summons, P. V. Welander, Calditol-linked membrane lipids are required for acid tolerance in Sulfolobus acidocaldarius, Proc. Natl. Acad. Sci. U. S. A. 115 (2018) 12932–12937. https://doi.org/10.1073/pnas.1814048115.

[254] P.A. Frey, A.D. Hegeman, F.J. Ruzicka, The radical SAM superfamily, Crit. Rev. Biochem. Mol. Biol. 43 (2008) 63–88. https://doi.org/10.1080/10409230701829169.

[255] C.D. Poulter, T. Aoki, L. Daniels, Biosynthesis of isoprenoid membranes in the methanogenic archaebacterium methanospirillum hungatei, J. Am. Chem. Soc. 110 (1988) 2620–2624. https://doi.org/10.1021/ja00216a041.

[256] T. Eguchi, Y. Nishimura, K. Kakinuma, Importance of the isopropylidene terminal of geranylgeranyl group for the formation of tetraether lipid in methanogenic archaea, Tetrahedron Lett. 44 (2003) 3275–3279. https://doi.org/10.1016/S0040-4039(03)00627-0.

[257] T.A. Langworthy, Long-chain diglycerol tetraethers from Thermoplasma acidophilum, Biochim. Biophys. Acta (BBA)/Lipids Lipid Metab. 487 (1977) 37–50. https://doi.org/10.1016/0005-2760(77)90042-X.

[258] T.A. Langworthy, Turnover of di-O-phytanylglycerol in Thermoplasma, Rev. Infect. Dis. 4 (1982) 266.

[259] N. Nemoto, Y. Shida, H. Shimada, T. Oshima, A. Yamagishi, Characterization of the precursor of tetraether lipid biosynthesis in the thermoacidophilic archaeon Thermoplasma acidophilum, Extremophiles. 7 (2003) 235–243. https://doi.org/10.1007/s00792-003-0315-x.

[260] T. Kon, N. Nemoto, T. Oshima, A. Yamagishi, Effects of a squalene epoxidase inhibitor, terbinafine, on ether lipid biosyntheses in a thermoacidophilic archaeon, Thermoplasma acidophilum, J. Bacteriol. 184 (2002) 1395–1401. https://doi.org/10.1128/JB.184.5.1395-1401.2002.

[261] N.S. Ryder, Specific inhibition of fungal sterol biosynthesis by SF 86-327, a new allylamine antimycotic agent, Antimicrob. Agents Chemother. 27 (1985) 252–256. https://doi.org/10.1128/AAC.27.2.252.

[262] G. Petranyi, N.S. Ryder, A. Stütz, Allylamine derivatives: New class of synthetic antifungal agents inhibiting fungal squalene epoxidase, Science (80-.). 224 (1984) 1239–1241. https://doi.org/10.1126/science.6547247.

[263] M. Nishihara, H. Morii, Y. Koga, Heptads of Polar Ether Lipids of an Archaebacterium, Methanobacterium thermoautotrophicum: Structure and Biosynthetic Relationship, Biochemistry. 28 (1989) 95–102. https://doi.org/10.1021/bi00427a014.

[264] T. Eguchi, M. Morita, K. Kakinuma, Multigram synthesis of mevalonolactone-d9 and its application to stereochemical analysis by 1H NMR of the saturation reaction in the biosynthesis of the 2,3-di-O-phytanyl-sn-glycerol core of the archaeal membrane lipid, J. Am. Chem. Soc. 120 (1998) 5427–5433. https://doi.org/10.1021/ja974387q.

[265] T. Eguchi, H. Takyo, M. Morita, K. Kakinuma, Y. Koga, Unusual double-bond migration as a plausible key reaction in the biosynthesis of the isoprenoidal membrane lipids of methanogenic archaea, Chem. Commun. 96 (2000) 1545–1546. https://doi.org/10.1039/b003948i.

[266] W. Fitz, D. Arigoni, Biosynthesis of 15,16-Dimethyltriacontanedioic Acid (Diabolic Acid) from [16-*H3]- and [14-*H2]-Palmitic Acids, (1992) 1533–1534. http://pubs.rsc.org/-/content/articlepdf/1992/c3/c39920001533.

[267] A. Krogh, B. Larsson, G. Von Heijne, E.L.L. Sonnhammer, Predicting transmembrane protein topology with a hidden Markov model: Application to complete genomes, J. Mol. Biol. 305 (2001) 567–580. https://doi.org/10.1006/jmbi.2000.4315.

[268] D.J. Miller, A. Jerga, C.O. Rock, S.W. White, Analysis of the Staphylococcus aureus DgkB Structure Reveals a Common Catalytic Mechanism for the Soluble Diacylglycerol Kinases, Structure. 16 (2008) 1036–1046. https://doi.org/10.1016/j.str.2008.03.019.

[269] J.E. Burke, E.A. Dennis, Phospholipase A 2 structure/function, mechanism, and signaling, J. Lipid Res. 50 (2009) S237. https://doi.org/10.1194/jlr.R800033-JLR200.

[270] A. Gulik, V. Luzzati, M. DeRosa, A. Gambacorta, Tetraether lipid components from a thermoacidophilic archaebacterium. Chemical structure and physical polymorphism, J. Mol. Biol. 201 (1988) 429–435. https://doi.org/10.1016/0022-2836(88)90149-0.

[271] J.R. De La Torre, C.B. Walker, A.E. Ingalls, M. Könneke, D.A. Stahl, Cultivation of a thermophilic ammonia oxidizing archaeon synthesizing crenarchaeol, Environ. Microbiol. 10 (2008) 810–818. https://doi.org/10.1111/j.1462-2920.2007.01506.x.

[272] F.J. Elling, M. Könneke, J.S. Lipp, K.W. Becker, E.J. Gagen, K.U. Hinrichs, Effects of growth phase on the membrane lipid composition of the thaumarchaeon Nitrosopumilus maritimus and their implications for archaeal lipid distributions in the marine environment, Geochim. Cosmochim. Acta. 141 (2014) 579–597. https://doi.org/10.1016/j.gca.2014.07.005.

[273] Y. Koga, H. Morii, Biosynthesis of Ether-Type Polar Lipids in Archaea and Evolutionary Considerations, Microbiol. Mol. Biol. Rev. 71 (2007) 97–120. https://doi.org/10.1128/mmbr.00033-06.

[274] L. Villanueva, J.S.S. Damsté, S. Schouten, A re-evaluation of the archaeal membrane lipid biosynthetic pathway, Nat. Rev. Microbiol. 12 (2014) 438–448. https://doi.org/10.1038/nrmicro3260.

[275] D.S. Nichols, M.R. Miller, N.W. Davies, A. Goodchild, M. Raftery, R. Cavicchioli, Cold adaptation in the Antarctic archaeon Methanococcoides burtonii involves membrane lipid unsaturation, J. Bacteriol. 186 (2004) 8508–8515. https://doi.org/10.1128/JB.186.24.8508-8515.2004.

[276] C.S. Knappy, B.J. Keely, Novel glycerol dialkanol triols in sediments: Transformation products of glycerol dibiphytanyl glycerol tetraether lipids or biosynthetic intermediates?, Chem. Commun. 48 (2012) 841–843. https://doi.org/10.1039/c1cc15841d.

[277] S. Schouten, M.J.L. Hoefs, M.P. Koopmans, H.J. Bosch, J.S. Sinninghe Damsté, Structural characterization, occurrence and fate of archaeal ether-bound acyclic and cyclic biphytanes and corresponding diols in sediments, in: Org. Geochem., Pergamon, 1998: pp. 1305–1319. https://doi.org/10.1016/S0146-6380(98)00131-4.

[278] X.L. Liu, J.S. Lipp, J.M. Schröder, R.E. Summons, K.U. Hinrichs, Isoprenoid glycerol dialkanol diethers: A series of novel archaeal lipids in marine sediments, Org. Geochem. 43 (2012) 50–55. https://doi.org/10.1016/j.orggeochem.2011.11.002.

[279] S.I. Ohnuma, M. Watanabe, T. Nishino, Identification and characterization of geranylgeraniol kinase and geranylgeranyl phosphate kinase from the archaebacterium Sulfolobus acidocaldarius, J. Biochem. 119 (1996) 541–547. https://doi.org/10.1093/oxfordjournals.jbchem.a021275.

[280] H. Morii, T. Eguchi, Y. Koga, In vitro biosynthesis of ether-type glycolipids in the methanoarchaeon Methanothermobacter thermautotrophicus, J. Bacteriol. 189 (2007) 4053–4061. https://doi.org/10.1128/JB.01875-06.

[281] K. Zaremba-Niedzwiedzka, E.F. Caceres, J.H. Saw, Di. Bäckström, L. Juzokaite, E. Vancaester, K.W. Seitz, K. Anantharaman, P. Starnawski, K.U. Kjeldsen, M.B. Stott, T. Nunoura, J.F. Banfield, A. Schramm, B.J. Baker, A. Spang, T.J.G. Ettema, Asgard archaea illuminate the origin of eukaryotic cellular complexity, Nature. 541 (2017) 353–358. https://doi.org/10.1038/nature21031.

[282] A. Spang, L. Eme, J.H. Saw, E.F. Caceres, K. Zaremba-Niedzwiedzka, J. Lombard, L. Guy, T.J.G. Ettema, Asgard archaea are the closest prokaryotic relatives of eukaryotes, PLoS Genet. 14 (2018) e1007080. https://doi.org/10.1371/journal.pgen.1007080.

[283] G.P. Fournier, A.M. Poole, A briefly argued case that Asgard Archaea are part of the eukaryote tree, Front. Microbiol. 9 (2018) 1896. https://doi.org/10.3389/fmicb.2018.01896.

[284] A. Spang, J.H. Saw, S.L. Jørgensen, K. Zaremba-Niedzwiedzka, J. Martijn, A.E. Lind, R. Van Eijk, C. Schleper, L. Guy, T.J.G. Ettema, Complex archaea that bridge the gap between prokaryotes and eukaryotes, Nature. 521 (2015) 173–179. https://doi.org/10.1038/nature14447.

[285] L. Manoharan, J.A. Kozlowski, R.W. Murdoch, F.E. Löffler, F.L. Sousa, C. Schleper, Metagenomes from coastal marine sediments give insights into the ecological role and cellular features of loki- and Thorarchaeota, MBio. 10 (2019). https://doi.org/10.1128/mBio.02039-19.

[286] H. Imachi, M.K. Nobu, N. Nakahara, Y. Morono, M. Ogawara, Y. Takaki, Y. Takano, K. Uematsu, T. Ikuta, M. Ito, Y. Matsui, M. Miyazaki, K. Murata, Y. Saito, S. Sakai, C. Song, E. Tasumi, Y. Yamanaka, T. Yamaguchi, Y. Kamagata, H. Tamaki, K. Takai, Isolation of an archaeon at the prokaryote–eukaryote interface, Nature. 577 (2020) 519–525. https://doi.org/10.1038/s41586-019-1916-6.

[287] T. Řezanka, L. Nedbalová, L. Procházková, K. Sigler, Lipidomic profiling of snow algae by ESI-MS and silver-LC/APCI-MS, Phytochemistry. 100 (2014) 34–42. https://doi.org/10.1016/j.phytochem.2014.01.017.

[288] M. Holčapek, R. Jirásko, M. Lísa, Recent developments in liquid chromatography-mass spectrometry and related techniques, J. Chromatogr. A. 1259 (2012) 3–15. https://doi.org/10.1016/j.chroma.2012.08.072.

[289] K. Yang, X. Han, Accurate quantification of lipid species by electrospray ionization mass spectrometry - Meets a key challenge in lipidomics, Metabolites. 1 (2011) 21–40. https://doi.org/10.3390/metabo1010021.

[290] K. Schuhmann, R. Almeida, M. Baumert, R. Herzog, S.R. Bornstein, A. Shevchenko, Shotgun lipidomics on a LTQ Orbitrap mass spectrometer by successive switching between acquisition polarity modes, J. Mass Spectrom. 47 (2012) 96–104. https://doi.org/10.1002/jms.2031.

[291] J. Wang, X. Han, Analytical challenges of shotgun lipidomics at different resolution of measurements, TrAC - Trends Anal. Chem. 121 (2019). https://doi.org/10.1016/j.trac.2019.115697.

[292] Y.H. Rustam, G.E. Reid, Analytical Challenges and Recent Advances in Mass Spectrometry Based Lipidomics, Anal. Chem. 90 (2018) 374–397. https://doi.org/10.1021/acs.analchem.7b04836.

[293] T. Züllig, M. Trötzmüller, H.C. Köfeler, Lipidomics from sample preparation to data analysis: a primer, Anal. Bioanal. Chem. 412 (2020) 2191-2209. https://doi.org/10.1007/s00216-019-02241-y.

[294] M. Lísa, M. Holčapek, Characterization of triacylglycerol enantiomers using chiral HPLC/APCI-MS and synthesis of enantiomeric triacylglycerols, Anal. Chem. 85 (2013) 1852–1859. https://doi.org/10.1021/ac303237a.

[295] M. Holčapek, H. Velínská, M. Lísa, P. Česla, Orthogonality of silver-ion and non-aqueous reversed-phase HPLC/MS in the analysis of complex natural mixtures of triacylglycerols, J. Sep. Sci. 32 (2009) 3672–3680. https://doi.org/10.1002/jssc.200900401.

[296] L. Yang, H. Nie, F. Zhao, S. Song, Y. Meng, Y. Bai, H. Liu, A novel online two-dimensional supercritical fluid chromatography/reversed phase liquid chromatography–mass spectrometry method for lipid profiling, Anal. Bioanal. Chem. 412 (2020) 2225–2235. https://doi.org/10.1007/s00216-019-02242-x.

[297] M. Lísa, M. Holčapek, High-Throughput and Comprehensive Lipidomic Analysis Using Ultrahigh-Performance Supercritical Fluid Chromatography-Mass Spectrometry, Anal. Chem. 87 (2015) 7187–7195. https://doi.org/10.1021/acs.analchem.5b01054.

[298] M. Sollai, L. Villanueva, E.C. Hopmans, R.G. Keil, J.S.S. Damsté, Archaeal sources of intact membrane lipid biomarkers in the oxygen deficient zone of the eastern tropical south Pacific, Front. Microbiol. 10 (2019) 765. https://doi.org/10.3389/fmicb.2019.00765.

[299] H.F. Sturt, R.E. Summons, K. Smith, M. Elvert, K.U. Hinrichs, Intact polar membrane lipids in prokaryotes and sediments deciphered by high-performance liquid chromatography/electrospray ionization multistage mass spectrometry - New biomarkers for biogeochemistry and microbial ecology, Rapid Commun. Mass Spectrom. 18 (2004) 617–628. https://doi.org/10.1002/rcm.1378.

[300] B.A.S. Van Mooy, H.F. Fredricks, B.E. Pedler, S.T. Dyhrman, D.M. Karl, M. Koblížek, M.W. Lomas, T.J. Mincer, L.R. Moore, T. Moutin, M.S. Rappé, E.A. Webb, Phytoplankton in the ocean use non-phosphorus lipids in response to phosphorus scarcity, Nature. 458 (2009) 69-72. https://doi.org/10.1038/nature07659.

[301] M. Scherer, G. Schmitz, G. Liebisch, Simultaneous quantification of cardiolipin, bis(monoacylglycero)phosphate and their precursors by hydrophilic interaction LC-MS/MS including correction of isotopic overlap, Anal. Chem. 82 (2010) 8794–8799. https://doi.org/10.1021/ac1021826.

[302] B.A.S. Van Mooy, G. Rocap, H.F. Fredricks, C.T. Evans, A.H. Devol, Sulfolipids dramatically decrease phosphorus demand by picocyanobacteria in oligotrophic marine environments, Proc. Natl. Acad. Sci. U. S. A. 103 (2006) 8607–8612. https://doi.org/10.1073/pnas.0600540103.

[303] R. Hájek, R. Jirásko, M. Lísa, E. Cífková, M. Holčapek, Hydrophilic Interaction Liquid Chromatography-Mass Spectrometry Characterization of Gangliosides in Biological Samples, Anal. Chem. 89 (2017) 12425–12432. https://doi.org/10.1021/acs.analchem.7b03523.

[304] M. Lísa, E. Cífková, M. Khalikova, M. Ovčačíková, M. Holčapek, Lipidomic analysis of biological samples: Comparison of liquid chromatography, supercritical fluid chromatography and direct infusion mass spectrometry methods, J. Chromatogr. A. 1525 (2017) 96–108. https://doi.org/10.1016/j.chroma.2017.10.022.

[305] Q. Shen, G. Song, H. Wang, Y. Zhang, Y. Cui, H. Xie, J. Xue, H. Wang, Isolation and lipidomics characterization of fatty acids and phospholipids in shrimp waste through GC/FID and HILIC-QTrap/MS, J. Food Compos. Anal. 95 (2021) 103668. https://doi.org/10.1016/j.jfca.2020.103668.

[306] T.H. Pham, M. Zaeem, T.A. Fillier, M. Nadeem, N.P. Vidal, C. Manful, S. Cheema, M. Cheema, R.H. Thomas, Targeting Modified Lipids during Routine Lipidomics Analysis using HILIC and C30 Reverse Phase Liquid Chromatography coupled to Mass Spectrometry, Sci. Rep. 9 (2019) 1–15. https://doi.org/10.1038/s41598-019-41556-9.

[307] K.M. Hines, J. Herron, L. Xu, Assessment of altered lipid homeostasis by HILIC-ion mobility-mass spectrometry-based lipidomics, J. Lipid Res. 58 (2017) 809–819. https://doi.org/10.1194/jlr.D074724.

[308] E. Cífková, M. Holčapek, M. Lísa, M. Ovčačíková, A. Lyčka, F. Lynen, P. Sandra, Nontargeted quantitation of lipid classes using hydrophilic interaction liquid chromatography-electrospray ionization mass spectrometry with single internal standard and response factor approach, Anal. Chem. 84 (2012) 10064–10070. https://doi.org/10.1021/ac3024476.

[309] E. Rampler, A. Criscuolo, M. Zeller, Y. El Abiead, H. Schoeny, G. Hermann, E. Sokol, K. Cook, D.A. Peake, B. Delanghe, G. Koellensperger, A Novel Lipidomics Workflow for Improved Human Plasma Identification and Quantification Using RPLC-MSn Methods and Isotope Dilution Strategies, Anal. Chem. 90 (2018) 6494–6501. https://doi.org/10.1021/acs.analchem.7b05382.

[310] T. Cajka, O. Fiehn, Increasing lipidomic coverage by selecting optimal mobile-phase modifiers in LC–MS of blood plasma, Metabolomics. 12 (2016) 1–11. https://doi.org/10.1007/s11306-015-0929-x.

[311] D. Oursel, C. Loutelier-Bourhis, N. Orange, S. Chevalier, V. Norris, C.M. Lange, Lipid composition of membranes of Escherichia coli by liquid chromatography/tandem mass spectrometry using negative electrospray ionization, Rapid Commun. Mass Spectrom. 21 (2007) 1721–1728. https://doi.org/10.1002/rcm.3013.

[312] C.W.N. Damen, G. Isaac, J. Langridge, T. Hankemeier, R.J. Vreeken, Enhanced lipid isomer separation in human plasma using reversed-phase UPLC with ion-mobility/high-resolution MS detection, J. Lipid Res. 55 (2014) 1772–1783. https://doi.org/10.1194/jlr.D047795.

[313] S.B. Breitkopf, S.J.H. Ricoult, M. Yuan, Y. Xu, D.A. Peake, B.D. Manning, J.M. Asara, A relative quantitative positive/negative ion switching method for untargeted lipidomics via high resolution LC-MS/MS from any biological source, Metabolomics. 13 (2017) 30. https://doi.org/10.1007/s11306-016-1157-8.

[314] A. Triebl, M. Trötzmüller, J. Hartler, T. Stojakovic, H.C. Köfeler, Lipidomics by ultrahigh performance liquid chromatography-high resolution mass spectrometry and its application to complex biological samples, J. Chromatogr. B Anal. Technol. Biomed. Life Sci. 1053 (2017) 72–80. https://doi.org/10.1016/j.jchromb.2017.03.027.

[315] A. Jeucken, M.R. Molenaar, C.H.A. van de Lest, J.W.A. Jansen, J.B. Helms, J.F. Brouwers, A Comprehensive Functional Characterization of Escherichia coli Lipid Genes, Cell Rep. 27 (2019) 1597-1606.e2. https://doi.org/10.1016/j.celrep.2019.04.018.

[316] M. Scherer, A. Böttcher, G. Schmitz, G. Liebisch, Sphingolipid profiling of human plasma and FPLC-separated lipoprotein fractions by hydrophilic interaction chromatography tandem mass spectrometry, Biochim. Biophys. Acta - Mol. Cell Biol. Lipids. 1811 (2011) 68–75. https://doi.org/10.1016/j.bbalip.2010.11.003.

[317] S.S. Bird, V.R. Marur, I.G. Stavrovskaya, B.S. Kristal, Separation of cis-trans phospholipid isomers using reversed phase LC with high resolution MS detection, Anal. Chem. 84 (2012) 5509–5517. https://doi.org/10.1021/ac300953j.

[318] F.D. Gunstone, I.A. Ismail, M. Lie ken jie, Fatty acids, part 16. Thin layer and gas-liquid chromatographic properties of the cis and trans methyl octadecenoates and of some acetylenic esters, Chem. Phys. Lipids. 1 (1967) 376–385. https://doi.org/10.1016/0009-3084(67)90015-1.

[319] G. Paglia, G. Astarita, Metabolomics and lipidomics using traveling-wave ion mobility mass spectrometry, Nat. Protoc. 12 (2017) 797–813. https://doi.org/10.1038/nprot.2017.013.

[320] T.P.I. Lintonen, P.R.S. Baker, M. Suoniemi, B.K. Ubhi, K.M. Koistinen, E. Duchoslav, J.L. Campbell, K. Ekroos, Differential mobility spectrometry-driven shotgun lipidomics, Anal. Chem. 86 (2014) 9662–9669. https://doi.org/10.1021/ac5021744.

[321] M. Š ala, M. Lísa, J.L. Campbell, M. Holčapek, Determination of triacylglycerol regioisomers using differential mobility spectrometry, Rapid Commun. Mass Spectrom. 30 (2016) 256–264. https://doi.org/10.1002/rcm.7430.

[322] M. Groessl, S. Graf, R. Knochenmuss, High resolution ion mobility-mass spectrometry for separation and identification of isomeric lipids, Analyst. 140 (2015) 6904–6911. https://doi.org/10.1039/c5an00838g.

[323] A.P. Bowman, R.R. Abzalimov, A.A. Shvartsburg, Broad Separation of Isomeric Lipids by High-Resolution Differential Ion Mobility Spectrometry with Tandem Mass Spectrometry, J. Am. Soc. Mass Spectrom. 28 (2017) 1552–1561. https://doi.org/10.1007/s13361-017-1675-2.

[324] C.G. Vasilopoulou, K. Sulek, A.D. Brunner, N.S. Meitei, U. Schweiger-Hufnagel, S.W. Meyer, A. Barsch, M. Mann, F. Meier, Trapped ion mobility spectrometry and PASEF enable in-depth lipidomics from minimal sample amounts, Nat. Commun. 11 (2020) 1–11. https://doi.org/10.1038/s41467-019-14044-x.

[325] H. Shimada, N. Nemoto, Y. Shida, T. Oshima, A. Yamagishi, Complete polar lipid composition of Thermoplasma acidophilum HO-62 determined by high-performance liquid chromatography with evaporative light-scattering detection, J. Bacteriol. 184 (2002) 556–563. https://doi.org/10.1128/JB.184.2.556-563.2002.

[326] M. Wagner, M. van Wolferen, A. Wagner, K. Lassak, B.H. Meyer, J. Reimann, S.V. Albers, Versatile genetic tool box for the crenarchaeote Sulfolobus acidocaldarius, Front. Microbiol. 3 (2012) 214. https://doi.org/10.3389/fmicb.2012.00214.

[327] T.D. Brock, K.M. Brock, R.T. Belly, R.L. Weiss, Sulfolobus: A new genus of sulfur-oxidizing bacteria living at low pH and high temperature, Arch. Mikrobiol. 84 (1972) 54–68. https://doi.org/10.1007/BF00408082.

[328] E.G. Bligh, W.J. Dyer, a Rapid Method of Total Lipid Extraction and Purification, Can. J. Biochem. Physiol. 37 (1959) 911–917. https://doi.org/10.1139/y59-099.

[329] M. Nishihara, Y. Koga, Extraction and composition of polar lipids from the archaebacterium, Methanobacterium thermoautotrophicum: Effective extraction of tetraether lipids by an acidified solvent, J. Biochem. 101 (1987) 997–1005. https://doi.org/10.1093/oxfordjournals.jbchem.a121969.

[330] G.M. Cabrera, M.L. Fern´andez Murga, G. Font de Valdez, A.M. Seldes, Direct analysis by electrospray ionization tandem mass spectrometry of mixtures of phosphatidyldiacylglycerols from Lactobacillus, J. Mass Spectrom. 35 (2000) 1452–1459. https://doi.org/10.1002/1096-9888(200012)35:12<1452::AID-JMS81>3.0.CO;2-N.

[331] T. Houjou, K. Yamatani, H. Nakanishi, M. Imagawa, T. Shimizu, R. Taguchi, Rapid and selective identification of molecular species in phosphatidylcholine and sphingomyelin by conditional neutral loss scanning and MS3, Rapid Commun. Mass Spectrom. 18 (2004) 3123–3130. https://doi.org/10.1002/rcm.1737.

[332] A. Criscuolo, M. Zeller, K. Cook, G. Angelidou, M. Fedorova, Rational selection of reverse phase columns for high throughput LC–MS lipidomics, Chem. Phys. Lipids. 221 (2019) 120–127. https://doi.org/10.1016/j.chemphyslip.2019.03.006.

[333] A. Plückthun, E.A. Dennis, Acyl and Phosphoryl Migration in Lysophospholipids: Importance in Phospholipid Synthesis and Phospholipase Specificity, Biochemistry. 21 (1982) 1743–1750. https://doi.org/10.1021/bi00537a007.

[334] J.E. O'Gara, K.D. Wyndham, Porous hybrid organic-inorganic particles in reversed-phase liquid chromatography, J. Liq. Chromatogr. Relat. Technol. 29 (2006) 1025–1045. https://doi.org/10.1080/10826070600574747.

[335] H. Ogiso, T. Suzuki, R. Taguchi, Development of a reverse-phase liquid chromatography electrospray ionization mass spectrometry method for lipidomics, improving detection of phosphatidic acid and phosphatidylserine, Anal. Biochem. 375 (2008) 124–131. https://doi.org/10.1016/j.ab.2007.12.027.

[336] H. Nakanishi, Y. Iida, T. Shimizu, R. Taguchi, Separation and quantification of sn-1 and sn-2 fatty acid positional isomers in phosphatidylcholine by RPLC-ESIMS/MS, J. Biochem. 147 (2010) 245–256. https://doi.org/10.1093/jb/mvp171.

[337] M.A. Besseling, E.C. Hopmans, R. Christine Boschman, J.S. Sinninghe Damsté, L. Villanueva, Benthic archaea as potential sources of tetraether membrane lipids in sediments across an oxygen minimum zone, Biogeosciences. 15 (2018) 4047-4064. https://doi.org/10.5194/bg-15-4047-2018.

[338] F.J. Elling, M. Könneke, G.W. Nicol, M. Stieglmeier, B. Bayer, E. Spieck, J.R. de la Torre, K.W. Becker, M. Thomm, J.I. Prosser, G.J. Herndl, C. Schleper, K.U. Hinrichs, Chemotaxonomic characterisation of the thaumarchaeal lipidome, Environ. Microbiol. 19 (2017) 2681–2700. https://doi.org/10.1111/1462-2920.13759.

[339] D. Benhaim, E. Grushka, Effect of n-octanol in the mobile phase on lipophilicity determination by reversed-phase high-performance liquid chromatography on a modified silica column, J. Chromatogr. A. 1209 (2008) 111–119. https://doi.org/10.1016/j.chroma.2008.08.118.

[340] B.K. Lavine, J.P. Ritter, S. Peterson, Enhancement of selectivity in reversed-phase liquid chromatography, J. Chromatogr. A. 946 (2002) 83–90. https://doi.org/10.1016/S0021-9673(01)01520-5.

[341] E. Cífková, R. Hájek, M. Lísa, M. Holčapek, Hydrophilic interaction liquid chromatography-mass spectrometry of (lyso)phosphatidic acids, (lyso)phosphatidylserines and other lipid classes, J. Chromatogr. A. 1439 (2016) 65–73. https://doi.org/10.1016/j.chroma.2016.01.064.

[342] F.L. Ndlandla, V. Ejoh, A.C. Stoltz, B. Naicker, A.D. Cromarty, S. van Wyngaardt, M. Khati, L.S. Rotherham, Y. Lemmer, J. Niebuhr, C.R. Baumeister, J.R. Al Dulayymi, H. Swai, M.S. Baird, J.A. Verschoor, Standardization of natural mycolic acid antigen composition and production for use in biomarker antibody detection to diagnose active tuberculosis, J. Immunol. Methods. 435 (2016) 50–59. https://doi.org/10.1016/j.jim.2016.05.010.

[343] M. Lísa, E. Cífková, M. Holčapek, Lipidomic profiling of biological tissues using off-line two-dimensional high-performance liquid chromatography-mass spectrometry, J. Chromatogr. A. 1218 (2011) 5146–5156. https://doi.org/10.1016/j.chroma.2011.05.081.

[344] N. Parrish, G. Osterhout, K. Dionne, A. Sweeney, N. Kwiatkowski, K. Carroll, K.C. Jost, J. Dick, Rapid, standardized method for determination of Mycobacterium tuberculosis drug susceptibility by use of mycolic acid analysis, J. Clin. Microbiol. 45 (2007) 3915–3920. https://doi.org/10.1128/JCM.02528-06.

[345] P. Dugo, T. Kumm, M.L. Crupi, A. Cotroneo, L. Mondello, Comprehensive two-dimensional liquid chromatography combined with mass spectrometric detection in the analyses of triacylglycerols in natural lipidic matrixes, J. Chromatogr. A. 1112 (2006) 269–275. https://doi.org/10.1016/j.chroma.2005.10.070.

[346] M. Lísa, M. Holčapek, T. Řezanka, N. Kabátová, High-performance liquid chromatography-atmospheric pressure chemical ionization mass spectrometry and gas chromatography-flame ionization detection characterization of Δ5-polyenoic fatty acids in triacylglycerols from conifer seed oils, J. Chromatogr. A. 1146 (2007) 67–77. https://doi.org/10.1016/j.chroma.2007.01.122.

[347] N. Tahiri, P. Fodran, D. Jayaraman, J. Buter, M.D. Witte, T.A. Ocampo, D.B. Moody, I. Van Rhijn, A.J. Minnaard, Total Synthesis of a Mycolic Acid from Mycobacterium tuberculosis, Angew. Chemie - Int. Ed. 59 (2020) 7555–7560. https://doi.org/10.1002/anie.202000523.

[348] EMEA, European Medicines Agency: An unacceptable choice, 2011. https://www.ema.europa.eu/en/documents/scientific-guideline/ich-q-2-r1-validation-analytical-procedures-text-methodology-step-5_en.pdf.

[349] G. Van Meer, D.R. Voelker, G.W. Feigenson, Membrane lipids: Where they are and how they behave, Nat. Rev. Mol. Cell Biol. 9 (2008) 112–124. https://doi.org/10.1038/nrm2330.

[350] M.T. Nakamura, B.E. Yudell, J.J. Loor, Regulation of energy metabolism by long-chain fatty acids, Prog. Lipid Res. 53 (2014) 124–144. https://doi.org/10.1016/j.plipres.2013.12.001.

[351] T. Shimizu, Lipid mediators in health and disease: Enzymes and receptors as therapeutic targets for the regulation of immunity and inflammation, Annu. Rev. Pharmacol. Toxicol. 49 (2009) 123–150. https://doi.org/10.1146/annurev.pharmtox.011008.145616.

[352] A.E. Saliba, I. Vonkova, A.C. Gavin, The systematic analysis of protein-lipid interactions comes of age, Nat. Rev. Mol. Cell Biol. 16 (2015) 753–761. https://doi.org/10.1038/nrm4080.

[353] M.D. Resh, Fatty acylation of proteins: The long and the short of it, Prog. Lipid Res. 63 (2016) 120–131. https://doi.org/10.1016/j.plipres.2016.05.002.

[354] S. Grösch, S. Schiffmann, G. Geisslinger, Chain length-specific properties of ceramides, Prog. Lipid Res. 51 (2012) 50-62. https://doi.org/10.1016/j.plipres.2011.11.001.

[355] B. Antonny, S. Vanni, H. Shindou, T. Ferreira, From zero to six double bonds: Phospholipid unsaturation and organelle function, Trends Cell Biol. 25 (2015) 427–436. https://doi.org/10.1016/j.tcb.2015.03.004.

[356] E. Sezgin, I. Levental, S. Mayor, C. Eggeling, The mystery of membrane organization: Composition, regulation and roles of lipid rafts, Nat. Rev. Mol. Cell Biol. 18 (2017) 361–374. https://doi.org/10.1038/nrm.2017.16.

[357] T. Harayama, H. Riezman, Understanding the diversity of membrane lipid composition, Nat. Rev. Mol. Cell Biol. 19 (2018) 281–296. https://doi.org/10.1038/nrm.2017.138.

[358] M. van Wolferen, A. Orell, S.V. Albers, Archaeal biofilm formation, Nat. Rev. Microbiol. 16 (2018) 699–713. https://doi.org/10.1038/s41579-018-0058-4.

[359] J.C.M. Holthuis, A.K. Menon, Lipid landscapes and pipelines in membrane homeostasis, Nature. 510 (2014) 48–57. https://doi.org/10.1038/nature13474.

[360] A. Yamashita, Y. Hayashi, Y. Nemoto-Sasaki, M. Ito, S. Oka, T. Tanikawa, K. Waku, T. Sugiura, Acyltransferases and transacylases that determine the fatty acid composition of glycerolipids and the metabolism of bioactive lipid mediators in mammalian cells and model organisms, Prog. Lipid Res. 53 (2014) 18–81. https://doi.org/10.1016/j.plipres.2013.10.001.

[361] L. Doyle, O.G. Ovchinnikova, K. Myler, E. Mallette, B.S. Huang, T.L. Lowary, M.S. Kimber, C. Whitfield, Biosynthesis of a conserved glycolipid anchor for Gram-negative bacterial capsules, Nat. Chem. Biol. 15 (2019) 632–640. https://doi.org/10.1038/s41589-019-0276-8.

[362] S. Jain, A. Caforio, A.J.M. Driessen, Biosynthesis of archaeal membrane ether lipids, Front. Microbiol. 5 (2014) 641. https://doi.org/10.3389/fmicb.2014.00641.

[363] B.C. Chung, E.H. Mashalidis, T. Tanino, M. Kim, A. Matsuda, J. Hong, S. Ichikawa, S.Y. Lee, Structural insights into inhibition of lipid i production in bacterial cell wall synthesis, Nature. 533 (2016) 557–560. https://doi.org/10.1038/nature17636.

[364] C. Zhang, A.P.R. Phillips, R.L. Wipfler, G.J. Olsen, R.J. Whitaker, The essential genome of the crenarchaeal model Sulfolobus islandicus, Nat. Commun. 9 (2018) 4908. https://doi.org/10.1038/s41467-018-07379-4.

[365] R.L.H. Andringa, Synthetic studies of archaeal lipids and related long-chain isoprenoids, University of Groninngen, 2022.

[366] N.A.W. de Kok, M. Exterkate, R.L.H. Andringa, A.J. Minnaard, A.J.M. Driessen, A versatile method to separate complex lipid mixtures using 1-butanol as eluent in a reverse-phase UHPLC-ESI-MS system, Chem. Phys. Lipids. 240 (2021) 105125. https://doi.org/10.1016/j.chemphyslip.2021.105125.

[367] K. Ohara, A. Muroya, N. Fukushima, K. Yazaki, Functional characterization of LePGT1, a membrane-bound prenyltransferase involved in the geranylation of p-hydroxybenzoic acid, Biochem. J. 421 (2009) 231–241. https://doi.org/10.1042/BJ20081968.

[368] M. Melzer, L. Heide, Characterization of Polyprenyldiphosphate: 4-Hydroxybenzoate Polyprenyltransferase from Escherichia coli, Biochim. Biophys. Acta (BBA)/Lipids Lipid Metab. 1212 (1994) 93–102. https://doi.org/10.1016/0005-2760(94)90193-7.

[369] I.G. Young, R.A. Leppik, J.A. Hamilton, F. Gibson, Biochemical and genetic studies on ubiquinone biosynthesis in Escherichia coli K-12:4-hydroxybenzoate octaprenyltransferase., J. Bacteriol. 110 (1972) 18–25. https://doi.org/10.1128/jb.110.1.18-25.1972.

[370] L. Bräuer, W. Brandt, D. Schulze, S. Zakharova, L. Wessjohann, A structural model of the membrane-bound aromatic prenyltransferase UbiA from *E. coli*, ChemBioChem. 9 (2008) 982-992. https://doi.org/10.1002/cbic.200700575.

[371] B.A. Kellogg, C.D. Poulter, Chain elongation in the isoprenoid biosynthetic pathway, Curr. Opin. Chem. Biol. 1 (1997) 570–578. https://doi.org/10.1016/S1367-5931(97)80054-3.

[372] F.H. Wallrapp, J.J. Pan, G. Ramamoorthy, D.E. Almonacid, B.S. Hillerich, R. Seidel, Y. Patskovsky, P.C. Babbitt, S.C. Almo, M.P. Jacobson, C.D. Poulter, Prediction of function for the polyprenyl transferase subgroup in the isoprenoid synthase superfamily, Proc. Natl. Acad. Sci. U. S. A. 110 (2013) E1196-202. https://doi.org/10.1073/pnas.1300632110.

[373] K. Ohara, K. Mito, K. Yazaki, Homogeneous purification and characterization of LePGT1 - A membrane-bound aromatic substrate prenyltransferase involved in secondary metabolism of Lithospermum erythrorhizon, FEBS J. 280 (2013) 2572–2580. https://doi.org/10.1111/febs.12239.

[374] The Multiple-System Atrophy Research Collaboration, Mutations in COQ2 in Familial and Sporadic Multiple-System Atrophy, N. Engl. J. Med. 369 (2013) 233–244. https://doi.org/10.1056/nejmoa1212115.

[375] W. Kabsch, IUCr, XDS, 66 (2010) 125–132. https://doi.org/10.1107/S0907444909047337.

[376] G.M. Sheldrick, A short history of SHELX, Acta Crystallogr. Sect. A Found. Crystallogr. 64 (2008) 112–122. https://doi.org/10.1107/S0108767307043930.

[377] T.C. Terwilliger, J. Berendzen, Automated MAD and MIR structure solution, Acta Crystallogr. Sect. D Biol. Crystallogr. 55 (1999) 849–861. https://doi.org/10.1107/S0907444999000839.

[378] P. Emsley, B. Lohkamp, W.G. Scott, K. Cowtan, Features and development of Coot, Acta Crystallogr. Sect. D Biol. Crystallogr. 66 (2010) 486–501. https://doi.org/10.1107/S0907444910007493.

[379] P. V. Afonine, R.W. Grosse-Kunstleve, N. Echols, J.J. Headd, N.W. Moriarty, M. Mustyakimov, T.C. Terwilliger, A. Urzhumtsev, P.H. Zwart, P.D. Adams, Towards automated crystallographic structure refinement with phenix.refine, Acta Crystallogr. Sect. D Biol. Crystallogr. 68 (2012) 352–367. https://doi.org/10.1107/S0907444912001308.

[380] G. Jones, P. Willett, R.C. Glen, Molecular recognition of receptor sites using a genetic algorithm with a description of desolvation, J. Mol. Biol. 245 (1995) 43–53. https://doi.org/10.1016/S0022-2836(95)80037-9.

[381] R.B. Best, X. Zhu, J. Shim, P.E.M. Lopes, J. Mittal, M. Feig, A.D. MacKerell, Optimization of the additive CHARMM all-atom protein force field targeting improved sampling of the backbone φ, ψ and side-chain χ1 and χ2 Dihedral Angles, J. Chem. Theory Comput. 8 (2012) 3257–3273. https://doi.org/10.1021/ct300400x.

[382] K. Vanommeslaeghe, A.D. MacKerell, Automation of the CHARMM general force field (CGenFF) I: Bond perception and atom typing, J. Chem. Inf. Model. 52 (2012) 3144–3154. https://doi.org/10.1021/ci300363c.

[383] M.J. Abraham, T. Murtola, R. Schulz, S. Páll, J.C. Smith, B. Hess, E. Lindah, Gromacs: High performance molecular simulations through multi-level parallelism from laptops to supercomputers, SoftwareX. 1-2 (2015) 19-25. https://doi.org/10.1016/j.softx.2015.06.001.

[384] S. Dowler, G. Kular, D.R. Alessi, Protein lipid overlay assay., Sci. STKE. 2002 (2002) pl6. https://doi.org/10.1126/stke.2002.129.pl6.

[385] L. Wessjohann, B. Sontag, Prenylation of Benzoic Acid Derivatives Catalyzed by a Transferase from Escherichia coli Overproduction: Method Development and Substrate Specificity, Angew. Chemie (International Ed. English). 35 (1996) 1697–1699. https://doi.org/10.1002/anie.199616971.

[386] X. Robert, P. Gouet, Deciphering key features in protein structures with the new ENDscript server, Nucleic Acids Res. 42 (2014) W320-4. https://doi.org/10.1093/nar/gku316.

[387] S.F. Altschul, T.L. Madden, A.A. Schäffer, J. Zhang, Z. Zhang, W. Miller, D.J. Lipman, Gapped BLAST and PSI-BLAST: A new generation of protein database search programs, Nucleic Acids Res. 25 (1997) 3389–3402. https://doi.org/10.1093/nar/25.17.3389.

[388] T. Frickey, A. Lupas, CLANS: A Java application for visualizing protein families based on pairwise similarity, Bioinformatics. 20 (2004) 3702–3704. https://doi.org/10.1093/bioinformatics/bth444.

[389] R.C. Edgar, MUSCLE: A multiple sequence alignment method with reduced time and space complexity, BMC Bioinformatics. 5 (2004) 1–19. https://doi.org/10.1186/1471-2105-5-113.

[390] I. Letunic, P. Bork, Interactive Tree Of Life (iTOL): An online tool for phylogenetic tree display and annotation, Bioinformatics. 23 (2007) 127–128. https://doi.org/10.1093/bioinformatics/btl529.

[391] A. Musatov, E. Sedlák, Role of cardiolipin in stability of integral membrane proteins, Biochimie. 142 (2017) 102–111. https://doi.org/10.1016/j.biochi.2017.08.013.

[392] R.N.A.H. Lewis, R.N. McElhaney, The physicochemical properties of cardiolipin bilayers and cardiolipin-containing lipid membranes, Biochim. Biophys. Acta - Biomembr. 1788 (2009) 2069–2079. https://doi.org/10.1016/j.bbamem.2009.03.014.

[393] W. Dowhan, M. Bogdanov, E. Mileykovskaya, Functional roles of lipids in membranes, in: D.E. Vance, J.E. Vance (Eds.), Biochem. Lipids, Lipoproteins Membr., Fifth Edit, Elsevier, San Diego, 2008: pp. 1–37. https://doi.org/10.1016/B978-044453219-0.50003-9.

[394] F. Kawai, M. Shoda, R. Harashima, Y. Sadaie, H. Hara, K. Matsumoto, Cardiolipin Domains in Bacillus subtilis Marburg Membranes, J. Bacteriol. 186 (2004) 1475–1483. https://doi.org/10.1128/JB.186.5.1475-1483.2004.

[395] E. Mileykovskaya, W. Dowhan, Visualization of phospholipid domains in Escherichia coli by using the cardiolipin-specific fluorescent dye 10-N-nonyl acridine orange, J. Bacteriol. 182 (2000) 1172–1175. https://doi.org/10.1128/JB.182.4.1172-1175.2000.

[396] J.M. Wood, Perspective: challenges and opportunities for the study of cardiolipin, a key player in bacterial cell structure and function, Curr. Genet. 64 (2018) 795–798. https://doi.org/10.1007/s00294-018-0811-2.

[397] E. Mileykovskaya, W. Dowhan, Cardiolipin membrane domains in prokaryotes and eukaryotes, Biochim. Biophys. Acta - Biomembr. 1788 (2009) 2084–2091. https://doi.org/10.1016/j.bbamem.2009.04.003.

[398] W. Basu Ball, J.K. Neff, V.M. Gohil, The role of nonbilayer phospholipids in mitochondrial structure and function, FEBS Lett. 592 (2018) 1273–1290. https://doi.org/10.1002/1873-3468.12887.

[399] N. Ikon, R.O. Ryan, Cardiolipin and mitochondrial cristae organization, Biochim. Biophys. Acta - Biomembr. 1859 (2017) 1156–1163. https://doi.org/10.1016/j.bbamem.2017.03.013.

[400] R. Arias-Cartin, S. Grimaldi, J. Pommier, P. Lanciano, C. Schaefer, P. Arnoux, G. Giordano, B. Guigliarelli, A. Magalon, Cardiolipin-based respiratory complex activation in bacteria, Proc. Natl. Acad. Sci. U. S. A. 108 (2011) 7781–7786. https://doi.org/10.1073/pnas.1010427108.

[401] A. Corcelli, S. Lobasso, L.L. Palese, M.S. Saponetti, S. Papa, Cardiolipin is associated with the terminal oxidase of an extremely halophilic archaeon, Biochem. Biophys. Res. Commun. 354 (2007) 795–801. https://doi.org/10.1016/j.bbrc.2007.01.060.

[402] M. Zhang, E. Mileykovskaya, W. Dowhan, Gluing the respiratory chain together: Cardiolipin is required for supercomplex formation in the inner mitochondrial membrane, J. Biol. Chem. 277 (2002) 43553–43556. https://doi.org/10.1074/jbc.C200551200.

[403] S. Hiraoka, H. Matsuzaki, I. Shibuya, Active increase in cardiolipin synthesis in the stationary growth phase and its physiological significance in Escherichia coli, FEBS Lett. 336 (1993) 221–224. https://doi.org/10.1016/0014-5793(93)80807-7.

[404] L.A. Luévano-Martínez, A.J. Kowaltowski, Phosphatidylglycerol-derived phospholipids have a universal, domain-crossing role in stress responses, Arch. Biochem. Biophys. 585 (2015) 90–97. https://doi.org/10.1016/j.abb.2015.09.015.

[405] M. Schlame, M.L. Greenberg, Biosynthesis, remodeling and turnover of mitochondrial cardiolipin, Biochim. Biophys. Acta - Mol. Cell Biol. Lipids. 1862 (2017) 3–7. https://doi.org/10.1016/j.bbalip.2016.08.010.

[406] H.F. Tian, J.M. Feng, J.F. Wen, The evolution of cardiolipin biosynthesis and maturation pathways and its implications for the evolution of eukaryotes, BMC Evol. Biol. 12 (2012) 32. https://doi.org/10.1186/1471-2148-12-32.

[407] F. Noguchi, G. Tanifuji, M.W. Brown, K. Fujikura, K. Takishita, Complex evolution of two types of cardiolipin synthase in the eukaryotic lineage stramenopiles, Mol. Phylogenet. Evol. 101 (2016) 133–141. https://doi.org/10.1016/j.ympev.2016.05.011.

[408] B.E. Tropp, Cardiolipin synthase from Escherichic coli, Biochim. Biophys. Acta - Lipids Lipid Metab. 1348 (1997) 192–200. https://doi.org/10.1016/S0005-2760(97)00100-8.

[409] C. Sohlenkamp, O. Geiger, Bacterial membrane lipids: Diversity in structures and pathways, FEMS Microbiol. Rev. 40 (2015) 133–159. https://doi.org/10.1093/femsre/fuv008.

[410] F. Krieger, A. Möglich, T. Kiefhaber, Effect of proline and glycine residues on dynamics and barriers of loop formation in polypeptide chains, J. Am. Chem. Soc. 127 (2005) 3346–3352. https://doi.org/10.1021/ja042798i.

[411] J. Jacob, H. Duclohier, D.S. Cafiso, The role of proline and glycine in determining the backbone flexibility of a channel-forming peptide, Biophys. J. 76 (1999) 1367–1376. https://doi.org/10.1016/S0006-3495(99)77298-X.

[412] R. Matsumi, H. Atomi, A.J.M. Driessen, J. van der Oost, Isoprenoid biosynthesis in Archaea - Biochemical and evolutionary implications, Res. Microbiol. 162 (2011) 39–52. https://doi.org/10.1016/j.resmic.2010.10.003.

[413] J. Peretó, P. López-García, D. Moreira, Ancestral lipid biosynthesis and early membrane evolution, Trends Biochem. Sci. 29 (2004) 469-477. https://doi.org/10.1016/j.tibs.2004.07.002.

[414] M. Exterkate, A.J.M. Driessen, Synthetic Minimal Cell: Self-Reproduction of the Boundary Layer, ACS Omega. 4 (2019) 5293–5303. https://doi.org/10.1021/acsomega.8b02955.

[415] C.R.H. Raetz, W. Dowhan, Biosynthesis and function of phospholipids in Escherichia coli, J. Biol. Chem. 265 (1990) 1235–1238. https://doi.org/10.1016/s0021-9258(19)40001-x.

[416] A. Caforio, A.J.M. Driessen, Archaeal phospholipids: Structural properties and biosynthesis, Biochim. Biophys. Acta - Mol. Cell Biol. Lipids. 1862 (2017) 1325–1339. https://doi.org/10.1016/j.bbalip.2016.12.006.

[417] S.C. Kushwaha, M. Kates, G.D. Sprott, I.C.P. Smith, Novel polar lipids from the methanogen Methanospirillum hungatei GP1, Biochim. Biophys. Acta (BBA)/Lipids Lipid Metab. 664 (1981) 156–173. https://doi.org/10.1016/0005-2760(81)90038-2.

[418] G.D. Sprott, G. Ferrante, I. Ekiel, Tetraether lipids of Methanospirillum hungatei with head groups consisting of phospho-N,N-dimethylaminopentanetetrol, phospho-N,N,N-trimethylaminopentanetetrol, and carbohydrates, Biochim. Biophys. Acta (BBA)/Lipids Lipid Metab. 1214 (1994) 234–242. https://doi.org/10.1016/0005-2760(94)90069-8.

[419] C.A.P. Hidalgo-Ahumada, M.K. Nobu, T. Narihiro, H. Tamaki, W.T. Liu, Y. Kamagata, A.J.M. Stams, H. Imachi, D.Z. Sousa, Novel energy conservation strategies and behaviour of Pelotomaculum schinkii driving syntrophic propionate catabolism, Environ. Microbiol. 20 (2018) 4503–4511. https://doi.org/10.1111/1462-2920.14388.

[420] B.G. Hall, Building phylogenetic trees from molecular data with MEGA, Mol. Biol. Evol. 30 (2013) 1229–1235. https://doi.org/10.1093/molbev/mst012.

[421] L.R. Sita, Convenient Highly Stereoselective Syntheses of (3R,7R,11R)- and (3S,7R,11R)-3,7,11,15-Tetramethylhexadecanoic Acid (Phytanic Acid) and the Corresponding 3,7,11,15-Tetramethylhexadecan-1-ols, J. Org. Chem. 58 (1993) 5285–5287. https://doi.org/10.1021/jo00071a049.

[422] H.C. Aspinall, N. Greeves, W.M. Lee, E.G. McIver, P.M. Smith, An improved williamson etherification of hindered alcohols promoted by 15-crown-5 and sodium hydride, Tetrahedron Lett. 38 (1997) 4679–4682. https://doi.org/10.1016/S0040-4039(97)00965-9.

[423] E.A.H. Yeh, E. Kumli, K. Damodaran, D.P. Curran, Bare-minimum fluorous mixture synthesis of a stereoisomer library of 4,8,12-trimethylnonadecanols and predictions of NMR spectra of saturated oligoisoprenoid stereoisomers, J. Am. Chem. Soc. 135 (2013) 1577–1584. https://doi.org/10.1021/ja311606u.

[424] A.B. Smith, I.G. Safonov, Total synthesis of (+)-dactylolide, Org. Lett. 4 (2002) 635–637. https://doi.org/10.1021/ol017271e.

[425] S.E. Schaus, B.D. Brandes, J.F. Larrow, M. Tokunaga, K.B. Hansen, A.E. Gould, M.E. Furrow, E.N. Jacobsen, Highly selective hydrolytic kinetic resolution of terminal epoxides catalyzed by chiral (salen)CoIII complexes. Practical synthesis of enantioenriched terminal epoxides and 1,2-diols, J. Am. Chem. Soc. 124 (2002) 1307–1315. https://doi.org/10.1021/ja016737l.

[426] M.L. Alcaraz, L. Peng, P. Klotz, M. Goeldner, Synthesis and properties of photoactivatable phospholipid derivatives designed to probe the membrane-associate domains of proteins, J. Org. Chem. 61 (1996) 192–201. https://doi.org/10.1021/jo951350k.

[427] D.J. Keith, S.D. Townsend, Total Synthesis of the Congested, Bisphosphorylated Morganella morganii Zwitterionic Trisaccharide Repeating Unit, J. Am. Chem. Soc. 141 (2019) 12939–12945. https://doi.org/10.1021/jacs.9b06830.

[428] F. Sung'hwa, A. Strik, H. Regeling, B. Zwanenburg, G.J.F. Chittenden, Synthesis of 2(R),3-dihydroxypropyl and 2(R),3(R)-dihydroxybutyl β-d-fructopyranosides and some derivatives, Carbohydr. Res. 341 (2006) 846–854. https://doi.org/10.1016/j.carres.2006.02.005.

[429] M. Iwashita, K. Makide, T. Nonomura, Y. Misumi, Y. Otani, M. Ishida, R. Taguchi, M. Tsujimoto, J. Aoki, H. Arai, T. Ohwada, Synthesis and evaluation of lysophosphatidylserine analogues as inducers of mast cell degranulation. Potent activities of lysophosphatidylthreonine and its 2-deoxy derivative, J. Med. Chem. 52 (2009) 5837–5863. https://doi.org/10.1021/jm900598m.

[430] K. Takai, K. Nakamura, T. Toki, U. Tsunogai, M. Miyazaki, J. Miyazaki, H. Hirayama, S. Nakagawa, T. Nunoura, K. Horikoshi, Cell proliferation at 122°C and isotopically heavy CH4 production by a hyperthermophilic methanogen under high-pressure cultivation, Proc. Natl. Acad. Sci. U. S. A. 105 (2008) 10949. https://doi.org/10.1073/PNAS.0712334105.

[431] Y. Koga, Thermal adaptation of the archaeal and bacterial lipid membranes, Archaea. 2012 (2012). https://doi.org/10.1155/2012/789652.

[432] M. De Rosa, A. Gambacorta, B. Nicolaus, W.D. Grant, A C_{25}, C_{25} diether core lipid from archaebacterial haloalkaliphiles, J. Gen. Microbiol. 129 (1983) 2333–2337. https://doi.org/10.1099/00221287-129-8-2333.

[433] H. Morii, H. Yagi, H. Akutsu, N. Nomura, Y. Sako, Y. Koga, A novel phosphoglycolipid archaetidyl(glucosyl)inositol with two sesterterpanyl chains from the aerobic hyperthermophilic archaeon Aeropyrum pernix K1, Biochim. Biophys. Acta - Mol. Cell Biol. Lipids. 1436 (1999) 426–436. https://doi.org/10.1016/S0005-2760(98)00157-X.

[434] T.A. Langworthy, W.R. Mayberry, P.F. Smith, Long chain glycerol diether and polyol dialkyl glycerol triether lipids of Sulfolobus acidocaldarius, J. Bacteriol. 119 (1974) 106–116. https://doi.org/10.1128/jb.119.1.106-116.1974.

[435] G.D. Sprott, M. Meloche, J.C. Richards, Proportions of diether, macrocyclic diether, and tetraether lipids in Methanococcus jannaschii grown at different temperatures, J. Bacteriol. 173 (1991) 3907–3910. https://doi.org/10.1128/jb.173.12.3907-3910.1991.

[436] T.A. Langworthy, T.G. Tornabene, G. Holzer, Lipids of archaebacteria, Zentralblatt Fur Bakteriol. Angew. Und Okol. Microbiol. Abt.1 Orig.C Hyg. 3 (1982) 228–244. https://doi.org/10.1016/S0721-9571(82)80036-7.

Appendix

A

Acknowledgements

I have lived in Groningen for approximately nine years now, but it seems like time flew by. Although there definitely were moments where days seemed to go on forever. I experienced my time as a PhD candidate as a rollercoaster; I have experienced both the best and worst times of my life. The friends I made were a huge uplifting factor during the best of times, and supporting during the worst. During these past five years, my friends helped me to grow as a scientist and as a person, and I want to sincerely thank them, and the other people I met during my PhD.

First and foremost, I would like to thank my supervisor **Arnold Driessen** for offering me this position. I can clearly remember that I asked if you had any positions available after I finished my masters' colloquium, and that we later discussed the opportunity you had for me. Right from the start it was clear this would be a challenging project; but with your guidance we managed to publish high-quality work that I can be proud of. Without this opportunity I wouldn't have met the people I met and developed the way I did. I always enjoyed the nice casual chats and I especially had a great dinner at Sabrina's defense with you and your lovely wife Wil. Moreover, I especially remember the amazing Christmas get-together with the alcohol-free beer! Now that the pandemic-disruptions seems to calm down and that measures are being relaxed, I hope that you will organize more of such gatherings for the current MolMic members. The new ones do not know what they are missing!

Secondly, I wanted to thank **Adriaan Minnaard** for his co-supervision of my PhD. I always appreciated your input in our work discussions, as you often could provide a fresh point of view and valuable chemistry insights on how certain enzymatic reactions could happen or which compounds could be synthesized in your laboratory. Your valuable scientific feedback and kindness was always appreciated and is something I will definitely remember.

A very special thanks to my paranymphs **Max** and **Riccardo**. **Max** you were one of the first people I newly met when I started my PhD. We shared an office for a few months, and when you returned a while later to start your own PhD, we ended up sharing an office for the better part of three years! During this time I got to know you better, and learned we have several common interests. I was glad you also came along for the Corsica summer school and hope you feel the same way. You are also the person that introduced me to the best pizza I have ever had, baked by your girlfriend **Eleonora**. And although disturbed by the pandemic, and thesis writing, I enjoy our irregular squash sessions and hope we can pick this back up sooner rather than later. **Riccardo**, I still remember when I met you during my master's internship in Dirk-Jan's group. However, we did not really get in touch until a while after I joined the lab and started my PhD, and even then we weren't too close. However, that changed about two years ago as we gradually started to hang out more and I noticed we share similar opinions on a lot of stuff, particularly professional discussions. You have proven yourself a good friend, someone I can talk with and someone really enjoy hanging out with. The little "road-trips" we made are something I really enjoyed as well. I hope we can do that more often before one of us heads off to start a new adventure somewhere else. And not unimportantly, you were of great help with my thesis, especially when it came to making figures.

The honorary paranymphs, **László and Marten** are definitely not forgotten. Together with **Riccardo** we were the GANG. Guys, what a long, strange, trip it has been. When I started my PhD at MolMic I already recognized you three from my time as a master student but never thought we would become such close friends. In my experience, you guys are the main reason I felt accepted in the social group at MolMic and for a large part are responsible for the good time I had during the first half of my PhD. From **Lazi's** hurricane-force sneezes that I could feel in the back of my neck several offices over and **Ricca's** random squirt gun assassinations, to a very memorable trip to CHAINS with **Marten**; these are only some of the memories I cherish. I really enjoyed the time we spent together during lunches, board game evenings, online game nights, movies, drinks (+online), BBQs, Greece, Berlin, mini-road trips and various other outings. However it should go without saying that our friendship is not based on some particular experiences I recite here, but rather the accumulation of the many experiences we shared. You all know I like to go into the tiniest of details, but that is definitely too much to list here! I sincerely hope we manage to stay in touch despite geographical challenges and hopefully we can make and keep the yearly camping trip something of a tradition. And since I share scientific interests with all of you, who knows, we might professionally meet or collaborate in the future!

Ruben, I would like to thank you for your expertise and outstanding collaborative qualities that enabled me to work on this PhD. You always had time to discuss various topics on short notice, you were open to hear new ideas and answer my chemistry questions. I definitely look back on this collaboration with good memories. I am sure you will do well in the future, but I still wish you the best for your future career!

Greetje, you deserve a special mention in my acknowledgements as your assistance in the DGGGPS project cannot be understated. And yet it has been, as I noticed you weren't mentioned in the DGGGPS paper. However, I would like to personally rectify this. As I cannot speak for the other authors, I thank you for your efforts here. Thank you very much!

I started my PhD during a period not many people joining the lab. So For me there is a clear distinction between pre-me MolMic members and post-me MolMic members. I want to thank all **pre-me MolMic members** for providing such a welcome atmosphere! **Aleksandra**, together with **Max** you were my first officemates at MolMic. I would like to thank you for the many nice discussions and chats we have had, and still occasionally have. Hopefully this summer we can organize a meetup once again with a BBQ and drinks! **Carsten,** I will always remember that your defense was the first defense I paranymphed for. Thank you, it was an honor! **Annarita, Fernando and Carsten** many thanks for all the sweetness your office brought to the group! **André, Stephanie, Sabrina and Zsofi**, the other members of "la famiglia pequena", thank you for making me feel so welcome and for the awesome times we had. We made great memories together during our hangouts in the park, game nights, Halloween party and trips to Berlin and Greece! **André,** the first of the gentlemen's club to leave MolMic. I highly appreciated your presence at MolMic. You always managed to color the atmosphere wherever you went and the topics we discussed at the lunch table were often as random as they came! You are a great guy and I hope you are doing well in Germany! **Anmara,** thank you for taking care of various administrative aspects during my PhD and guiding me through the steps required for the defence of this PhD thesis. **Dirk-Jan**, I would like to thank you for your interesting questions and valuable comments on my work discussion talks. Furthermore I enjoyed and highly value the talks we had during parties, drinks and lunches.

The **post-me MolMic members** aren't forgotten either. However due to circumstances we didn't hang out as much. **Lúcia**, we met as officemates when I started my PhD. We got along well and I was glad to hear you were returning for a PhD later! Thank you for the get-togethers you organized at your place. I will remember those mini-burgers and brigadeiros you made for a long time to come (especially the banana flavored ones!). Your flamboyant style and openness is definitely something I would miss. I hope we stay in touch and meet up to play some more boardgames in the future! **Xiaohuan**, I want to specially thank you for the food-exchanges we have had and you cooking me amazing dumplings (With **Lin's** help for the dough and **your boyfriend** for the stuffing!) and inviting me over for the dinner get-togethers and the moon festival celebration. **Pedro**, I still remember the nice dinner we had together in the food-court shortly after you joined! You always managed to infuse me with some positivity and brighten my day. I had hoped to spend some more time with you after the pandemic, but unfortunately you weren't around for very long. I am glad you found a position you really like in Texas and I wish you the best of luck for the future! **Mirthe**, the fellow outdoor-gardener in MolMic, thank you for inviting me over to your place last summer. I enjoyed our scientific discussions and the teas we shared! **Sonja**, thank you for your vibrant attitude in MolMic and the nice talk we had at your party shortly after you started in MolMic! I would like to thank **Rianne, Alka, Parichita, Jeroen, Jiayi, Janny, Maria, Fabiola, Alicia** and other MolMic members for creating and maintaining a generally pleasant atmosphere at MolMic and wish you all the best!

Once upon a time, there were three very different girls who did their masters' studies at the RUG. These students all three had particular properties: They all had different hair colors, they were all being trained as scientists in MolMic and they all worked for me. My name is not Charlie, but these were ***my Angels: Kim, Noëlle and Aarti.*** I am glad you were in our lab and learned a lot from all of you, and hope you have learned a lot from me too! I have fond memories of a nice "kapsalon" late in the evening because of a busy day, a great Christmas borrel at Bax and some of you even joined in for my birthday! That said, I think you all know that I am sorry that I couldn't always supervise all of you as much as you (and I) would have liked. Although I hope that in the end that all of you ended up in a place you feel good at and wish you all the best of luck in your future endeavors!

Lieve **Paula**, Ik kan me onze ontmoeting in Australië nog herinneren als de dag van gisteren. Het blijft een gebeurtenis ten gevolge van een aparte samenloop van verschillende factoren. Als dat niet gebeurd was zou ik nooit in Groningen terecht gekomen zijn en zou mijn carrière er mogelijk vrij anders uit gezien hebben. Ik wil je bedanken voor al je steun gedurende mijn academische loopbaan en de minder mooie tijden die daarmee gepaard gingen. We hebben veel dingen doorgemaakt de afgelopen jaren. Maar Ik ben blij dat we deze tijd hebben kunnen delen en zal hier hoofdzakelijk met goede herinneringen op terug kijken.

Lieve **Papa en Mama**, zonder jullie steun was deze ervaring voor mij buiten bereik gebleven. Al vanaf de basisschool stonden jullie achter me. Dit bewezen jullie wederom tijdens het middelbare school traject, en nogmaals toen ik naar de HBO mocht gaan. Zonder jullie ondersteuning had ik dan ook nooit naar Australië kunnen gaan en zou ik dus vervolgens ook niet naar Groningen gegaan zijn. Ik wil jullie bedanken voor de mooie tijden die jullie mogelijk hebben gemaakt, en voor alle steun en toeverlaat tijdens de minder mooie tijden. Als laatste, maar niet het minste, hoop ik dan ook dat jullie trots zijn op wat ik nu uiteindelijk met al jullie steun bereikt heb en met wat ik in de toekomst nog ga doen. Ik houd van jullie en hoop nog vele jaren met jullie te kunnen genieten.

About the author

Niels Alexander Willem de Kok was born on the 3rd of February 1992 in the city of Helmond, the Netherlands. After graduating from the St Joseph IVO-MAVO school, and subsequently the Carolus Borromeus College HAVO school; in 2010 he enrolled in the bachelor programme of Applied science, Science and life at Fontys hogeschool in Eindhoven. During the third year he got the opportunity to travel to Adelaide, Australia, to conduct research on nanoparticle cytotoxicity during a bachelors internship. This was followed by an internship in Molecular Genetics group at the University of Groningen where he worked on the development of a heterologous lantibiotic expression system in *Bacillus subtilis*. After an early graduation, he pursued a master's degree in molecular biology and biotechnology at the University of Groningen. During this master's, he further developed his research skills and learned various techniques. This knowledge was used to aid in the characterization of glucanotransferases from *Lactobacillus aviarius* with novel substrate- and product-specificities in the Microbial Physiology group, and to study cell wall synthesis in *Bacillus subtilis* with fluorescent amino acid analogs and fluorescence microscopy in the Molecular Microbiology group at the University of Groningen. After graduation in 2017, having developed a keen interest in the construction of cell barriers, antibiotic biosynthesis and the effect of antibiotics on said barriers; he joined the Department of Molecular Microbiology in 2017 as a PhD candidate under supervision of prof. dr. Arnold Driessen. His research focused on the biosynthesis of complex phospholipids in Archaea and was completed in spring 2022. The results of that work are described in this thesis.

List of publications

Working title: Lateral membrane organization is the main target of the antimicrobial activity of the peptidomimetic AMC-109
Adéla Melcrová, Josef Melcr, Sourav Maity, Mariella Gabler, Jonne van der Eyden, Niels A. W. de Kok, Wenche Stensen, John S. M. Svendsen, Arnold J. M. Driessen, Siewert J. Marrink, Wouter H. Roos
In preparation

The catalytic and structural basis of archaeal glycerophospholipid biosynthesis
Niels A. W. de Kok and Arnold J. M. Driessen**
in submission

A versatile method to separate complex lipid mixtures using 1-butanol as eluent in a reverse-phase UHPLC-ESI-MS system
Niels A. W. de Kok, Marten Exterkate, Ruben L. H. Andringa, Adriaan J. Minnaard, Arnold J.M. Driessen**
Chem. Phys. Lipids (2021), 240:105125, https://doi.org/10.1016/j.chemphyslip.2021.105125

A Unified Approach for the Total Synthesis of cyclo-Archaeol, iso-Caldarchaeol, Caldarchaeol, and Mycoketide
Ruben L. H. Andringa, Niels A. W. de Kok, Arnold J. M. Driessen and Adriaan J. Minnaard**
Angewandte Chemie (2021), 133: 2-9, https://doi.org/10.1002/ange.202104759

A promiscuous archaeal cardiolipin synthase enables construction of diverse natural and unnatural phospholipids
Marten Exterkate, Niels A. W. de Kok, Ruben L. H. Andringa, Niels H. J. Wolbert, Adriaan J. Minnaard, Arnold J. M. Driessen**
J. of Biol. Chem. (2021), 296, 100691, https://doi.org/10.1016/j.jbc.2021.100691

CRISPR-based transcriptional activation tool for silent genes in filamentous fungi
László Mózsik*, Mirthe Hoekzema*, Niels A. W. de Kok, Roel A. L. Bovenberg, Yvonne Nygård & Arnold J. M. Driessen**
Scientific reports (2021), 11, 1118, https://doi.org/10.1038/s41598-020-80864-3

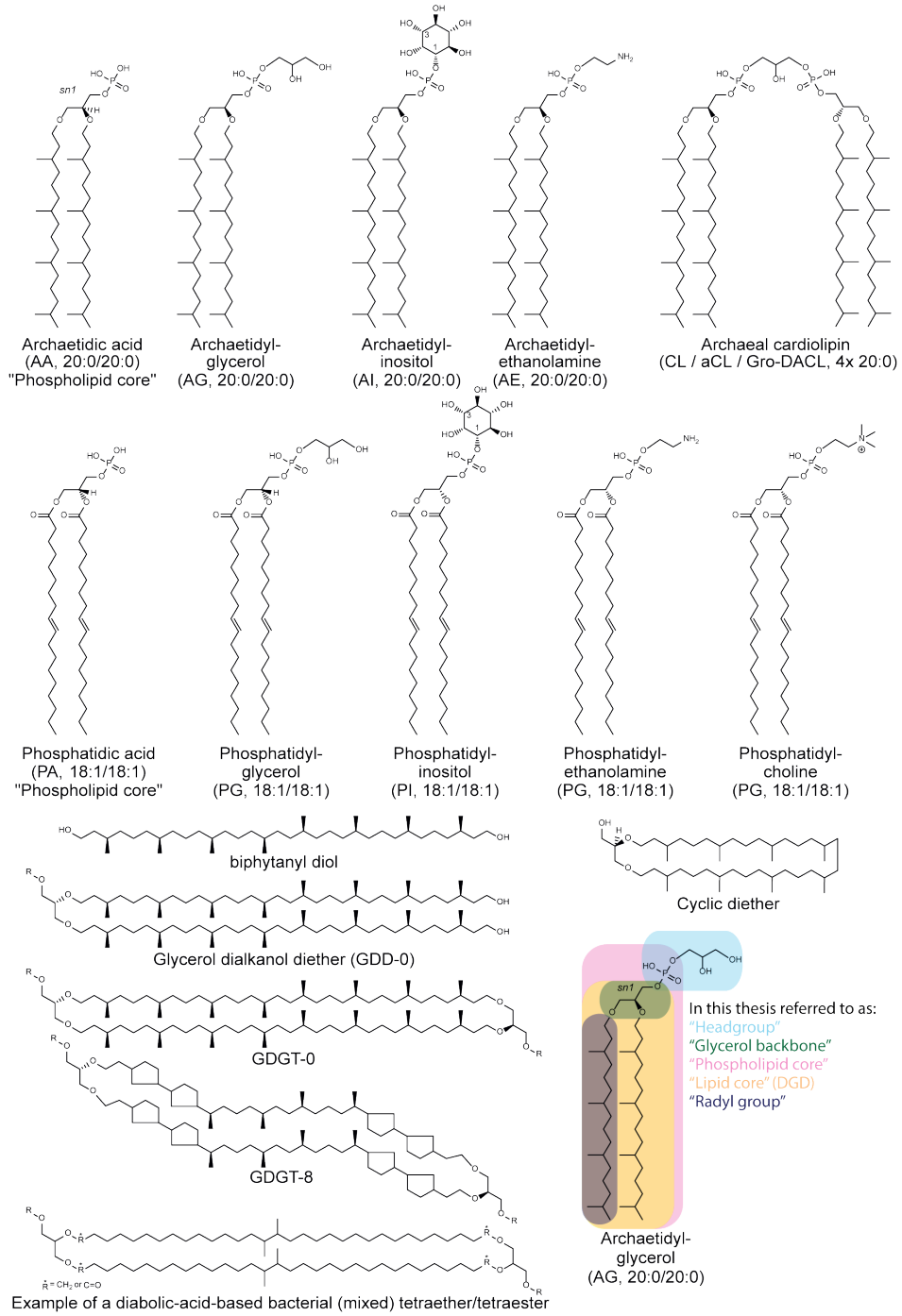
Structural and Functional Insights into an Archaeal Lipid Synthase
Sixue Ren*, Niels A. W. de Kok*, Yijun Gu*, Weizhu Yan, Qiu Sun, Yunying Chen, Jun He, Lejin Tian, Ruben L. H. Andringa, Xiaofeng Zhu, Mei Tang, Shiqian Qi, Heng Xu, Haiyan Ren, Xianghui Fu, Adriaan J. Minnaard, Shengyong Yang, Wanjiang Zhang, Weimin Li, Yuquan Wei, Arnold J. M. Driessen** and Wei Cheng**
Cell Reports (2020), 33, 108294, https://doi.org/10.1016/j.celrep.2020.108294

MiniBacillus PG10 as a Convenient and Effective Production Host for Lantibiotics
Amanda Y. van Tilburg, Auke J. van Heel, Jörg Stülke, Niels A. W. de Kok, Anne-Stéphanie Rueff, and Oscar P. Kuipers**
ACS Synthetic Biology (2020), 9 (7), 1833-1842, https://doi.org/10.1021/acssynbio.0c00194

Biochemical characterization of two GH70 family 4,6-α-glucanotransferases with distinct product specificity from Lactobacillus aviarius subsp. aviarius DSM 20655
Xiangfeng Meng, Joana Gangoiti, Niels de Kok, Sander S. van Leeuwen, Tjaard Pijning, Lubbert Dijkhuizen**
Food Chemistry (2018), 253: 236-246, https://doi.org/10.1016/j.foodchem.2018.01.154

Pentapeptide-rich peptidoglycan at the Bacillus subtilis cell-division site
Danae Morales Angeles, Yun Liu, Alwin M. Hartman, Marina Borisova, Anabela de Sousa Borges, Niels de Kok, Katrin Beilharz, Jan-Willem Veening, Christoph Mayer, Anna K. H. Hirsch, Dirk-Jan Scheffers**
Molecular Microbiology (2017), 104: 319-333, https://doi.org/10.1111/mmi.13629

* These authors contributed equally
** Corresponding author



A