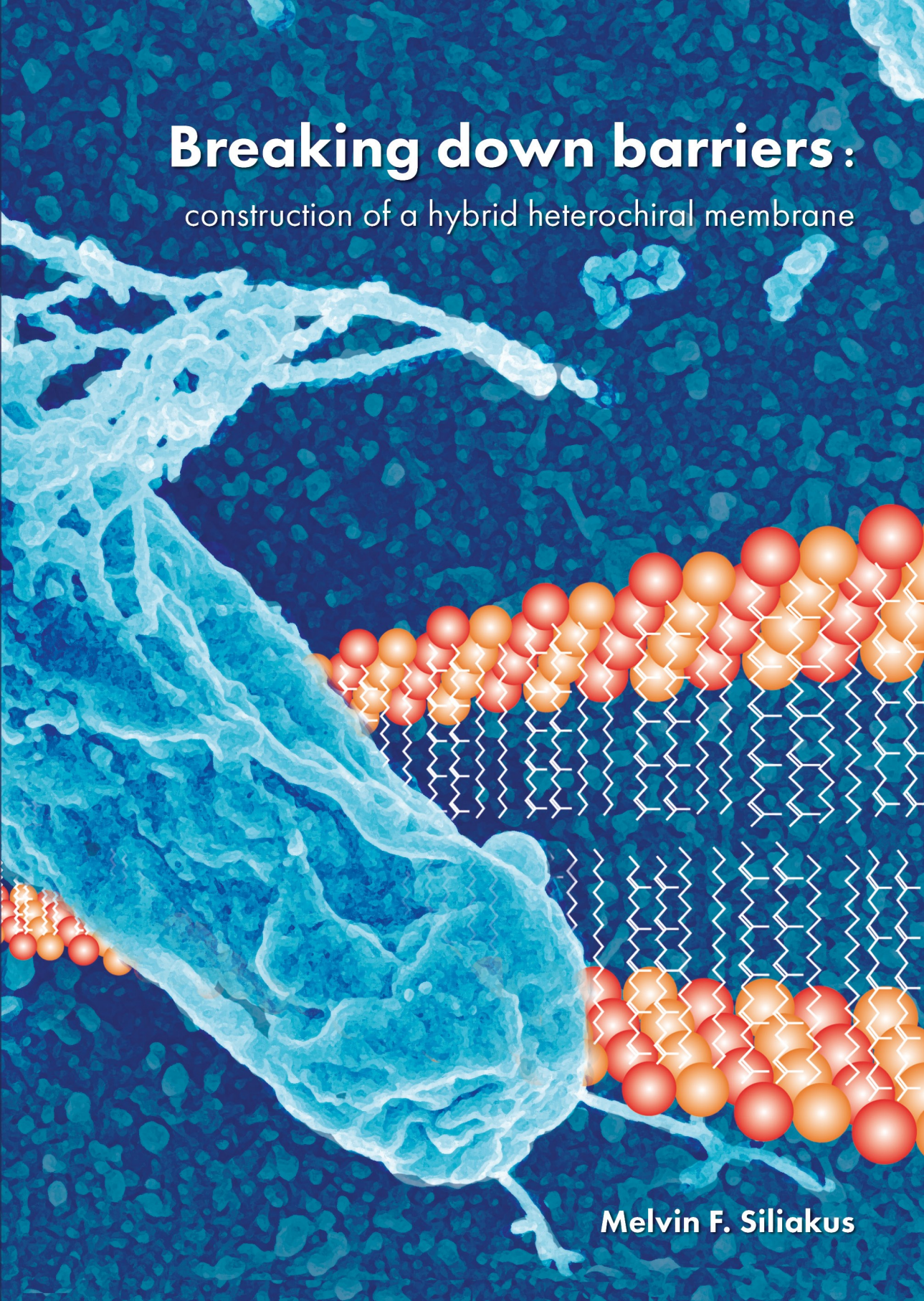
Breaking down barriers:

construction of a hybrid heterochiral membrane



Melvin F. Siliakus

Breaking down barriers:

construction of a hybrid heterochiral membrane

Melvin F. Siliakus

Thesis committee

Promotor
Prof. Dr John van der Oost
Personal chair at the Laboratory of Microbiology
Wageningen University

Co-promotor
Dr Servé W.M. Kengen
Assistant professor, Laboratory of Microbiology
Wageningen University

Other members
Prof. Dr Jaap Sinninghe Damsté, NIOZ Royal Netherlands Institute for Sea Research, Den Burg
Prof. Dr Willem J.H. van Berkel, Wageningen University
Dr Toon de Kroon, Utrecht University
Dr Ruud A. Weusthuis, Wageningen University

This research was conducted under the auspices of the Graduate School VLAG (Advanced studies in Food Technology, Agrobiotechnology, Nutrition and Health Sciences).

Breaking down barriers:

construction of a hybrid heterochiral membrane

Melvin F. Siliakus

Thesis

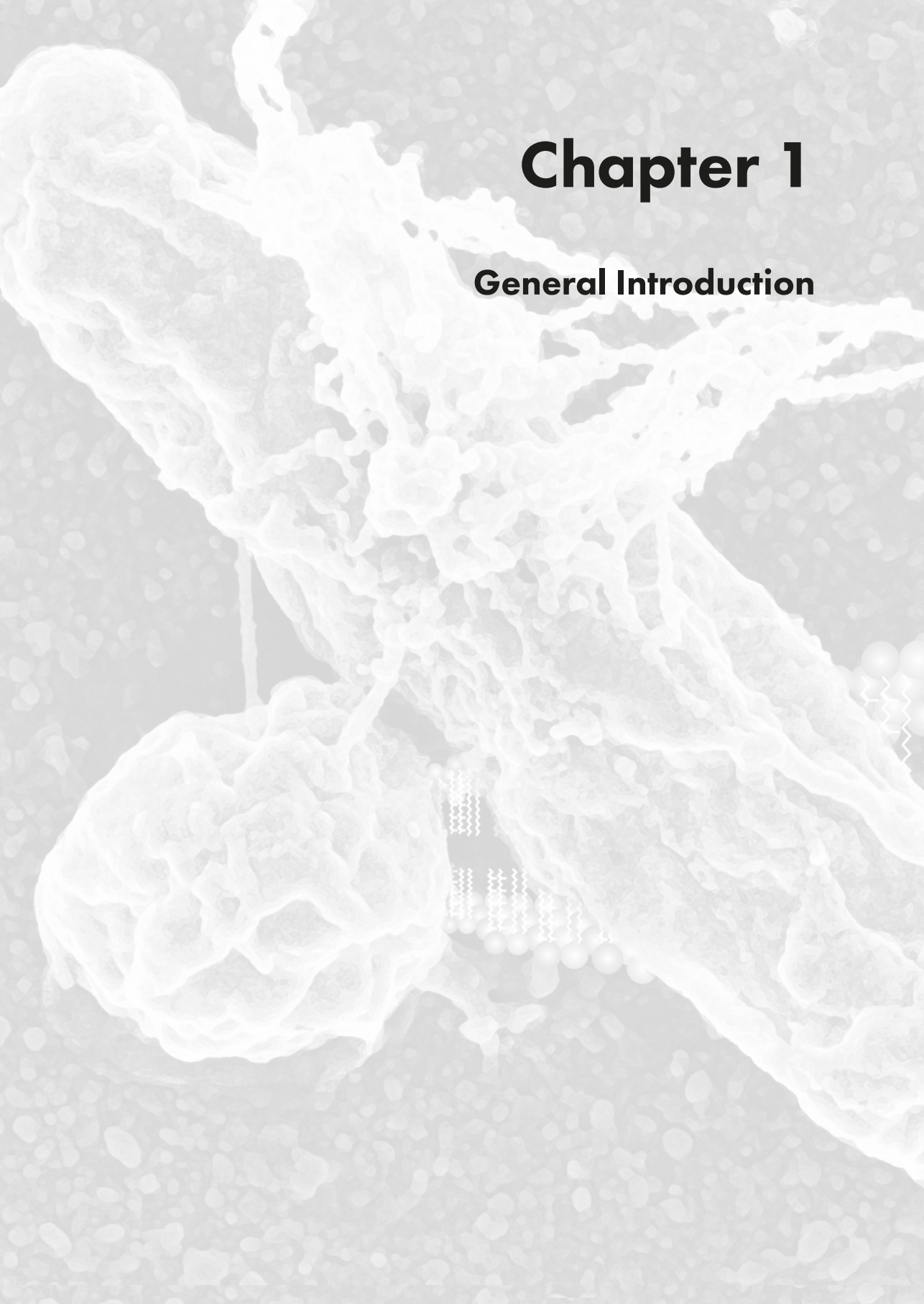
submitted in fulfillment of the requirements for the degree of doctor
at Wageningen University
by the authority of the Rector Magnificus
Prof. Dr A.P.J. Mol,
in the presence of the
Thesis Committee appointed by the Academic Board
to be defended in public
on Friday 4 November 2016
at 11 a.m. in the Aula.

Melvin F. Siliakus
Breaking down barriers: construction of a hybrid heterochiral membrane
238 pages.

PhD thesis, Wageningen University, Wageningen, NL (2016)
With references, with summary in English

ISBN 978-94-6257-929-3
DOI 10.18174/390044

Table of Contents



Chapter 1

General Introduction

Discovery of the third domain of life

In the early 20th century, bacterial classification and taxonomy were based on metabolic, physiological and morphological features. Newly discovered microbes were grouped into the domain of prokaryotes depending on the absence of a cell nucleus, presence of cell walls and the absence of organelles. Species and genus were then chosen based on biochemical and physiological characteristics. In 1977, Carl R. Woese and George E. Fox performed a phylogenetic analysis on prokaryotic and eukaryotic lineages in an attempt to develop an unbiased method for phylogenetic classification [1]. They recognized that genomes are the footprint of evolutionary descent and thus a more reliable proxy for phylogenetic classification than metabolic, physiological and morphological characteristics. The broadly distributed and highly conserved 16S ribosomal RNA sequences were therefore used to develop a general classification of all prokaryotes. Within this comparative analysis they discovered that the previously classified order of the 'archaebacteria' (methanogens) should be grouped to a separate domain because of their sequence dissimilarity with other bacteria. Following research then confirmed the existence of a third domain/kingdom of life which were termed 'Archaea' in 1990 [2]. Soon after this revelation, the multitude of novel archaea was shown to have more differences with the bacterial domain and more resemblance to the eukarya. Characteristic differences with bacteria are the pseudo-peptidoglycan, a transcription and translation machinery more similar to that of eukaryotes, a series of unique metabolic pathways and a plasma-membrane lipid composition that is fundamentally different in chemical architecture.

The 'lipid divide' and the pre-cell hypothesis

The earliest identified archaea were isolated from environments that were previously considered too extreme to support microbial life *viz*. high acidity/alkalinity, high temperatures, salinity etc. The newly discovered domain of life was therefore termed 'archaea' after the ancient Greek word (αρχαία = ancient) to reflect these organisms extreme habitats comparable to earths primordial conditions. In addition to their habitats, archaeal physiology is equally diverse as they include autotrophs, methanogens, phototrophs and sulphur reducers. Although archaea are indeed dominating the extremes of earths physicochemical challenges, we now know that milder habitats are also copiously colonized. Archaea are e.g. highly abundant in oceans as a main participant of plankton [3]. Due to their oftentimes extremophilic nature, archaea were extensively studied for the development of potential biotechnological applications. Research into their robustness then soon revealed that archaeal phospholipids display a chemical architecture that is completely distinct from bacterial and eukaryal phospholipids (**Fig. 1A** and **B**). The fundamental differences are found in the core lipids that serve as the framework for the fully maturated polar phospholipids.

These are represented by 'archaetidic acid' in archaea and 'phosphatidic acid' in bacteria and eukarya. The bacterial phosphatidic acid is composed of two fatty acid hydrocarbon chains esterified to the *sn*-1 and *sn*-2 position of a glycerol-3-phosphate (G-3-P) moiety. The archaeal archaetidic acid (DGGGP di-*O*-geranylgeranylglyceryl phosphate) on the other hand consists of two highly methyl-branched isoprenoids (termed 'phytanyls' in their saturated state) connected by ether-bonds to the *sn*-2 and *sn*-3 position of glycerol-1-phosphate (G-1-P). Although few exceptions to these rules apply for the hydrocarbon chains and ester/ether bonds [4-6], the stereochemical opposites of the glycerol-phosphates are an exclusive feature that separates the domains. The chemical differences have been termed "the lipid divide" and serve as a hallmark between the three domains [7]. Most interesting here is that although the core lipids are dissimilar, the polar head groups (ethanolamine, glycerol, serine, choline and glycerol) that are ultimately linked to the phosphate groups are shared by the three domains (**Fig. 2**). In addition to this, the archaea, and especially thermophilic species, use glycolipids with multiple hexoses as a main constituent of their membrane [8]. Another common theme of archaeal membranes is the formation of macrocyclic (**Fig. 1B**, left) and C80 tetraether lipids (also known as glycerol-dialkyl-glycerol tetraethers (GDGTs)) (**Fig. 1C**). Tetraether lipids are bipolar due to a putative condensation event that takes place between the isoprenoid ω20 carbon atoms of two diether lipids [9, 10]. Tetraether lipids are broadly distributed throughout archaea and display a high variability by the incorporation of cyclopentane rings in the phytanyl chains up to a maximum of 8 rings per GDGT (**Fig. 1D**) [11, 12].

The revealed separation and uniformity between bacterial, eukaryal and archaeal lipids ultimately raised the question how and why all life on earth was differentiated into these three domains. Currently, one of the most accepted hypothesis partitions the evolution of the domains into three evolutionary stages [13, 14]. The first stage envisions the origin of life into a multi-phenotypical population of pre-cells with membranes composed of a racemic or heterochiral glycerol-phosphate lipid mixture (i.e. a mixture of G-1-P and G-3-P lipids) [15]. The second stage was postulated in 2003 and hypothesizes the differentiation of this heterogeneous population of pre-cells (cenancestor or last universal common ancestor (LUCA)) into two lineages [14]. In the third and last stage the parallel specialization of all species known today is hypothesized. During this stage the eukaryotic lineage also appeared by a putative endosymbiosis event of bacteria and archaea that explains the mosaic distribution of both bacterial and archaeal features [13]. The foundation of bacteria and archaea however lies in the second stage in which a racemic membrane is thought to play a major role in a segregation event into homochiral membranes and the subsequent development of the bacterial and archaeal ancestors thereafter [14]. Contradictory, during the evolution of the eukaryotic lineage, a fused membrane is envisioned that also led to a racemic membrane. It is postulated that an inferior stability of mixed membranes was the main driving force behind this segregation [7]. Phylogenomic comparisons indeed strongly suggest that the

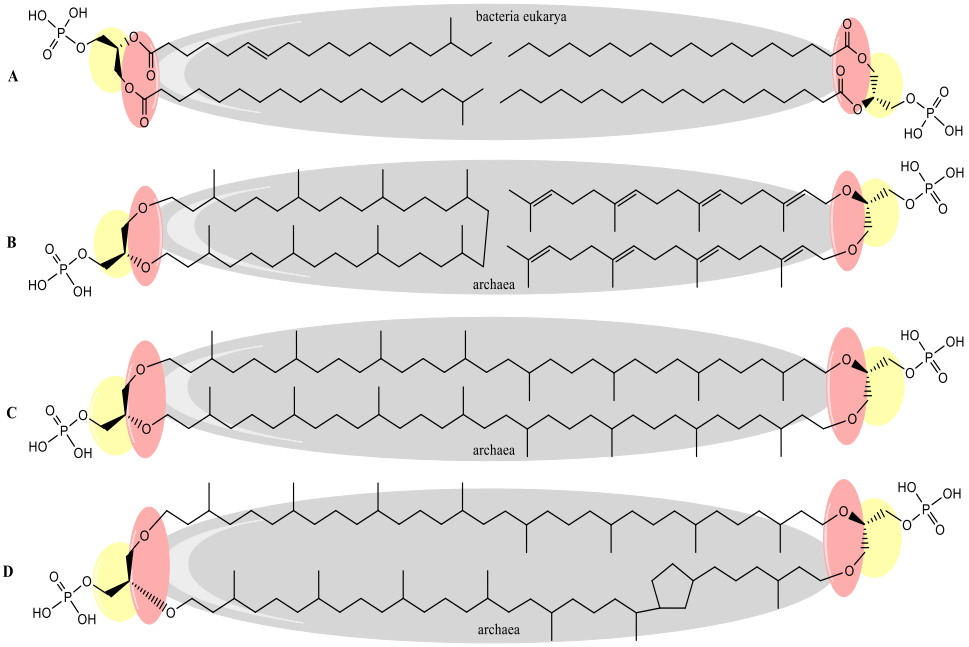


Figure 1. Bacterial and archaeal lipid variations and the lipid divide. The bacterial phosphatidic acid is represented by panel **A** right. In the left panel monounsaturated branched chain *anteiso* (top) and *iso*-fatty acid (bottom) variations are shown. The archaeal archaetidic acid or DGGGP (panel **B** right) core lipids are depicted, left lipid shows macrocyclic archaeol. The lipid divide describes the three chemical differences that are fundamentally different between the three domains of life. Grey: the hydrocarbon chains are fatty acids in bacteria and isoprenoid chains in archaea. Red: Ester-bonds link the hydrocarbon chains to the glycerol backbone in bacteria and ether-bonds in archaea. Yellow: The backbone moiety in phosphatidic acid is represented by glycerol-1-phosphate. In archaetidic acid the stereo-conformer or enantiomeric glycerol-1-phosphate is present in archaea. Archaeal bipolar glycerol dialkyl glycerol tetraether (GDGT-0) is depicted in (panel **C**) which spans the membrane to form monolayers. Panel **D** shows GDGT-1, the bipolar lipid containing 1 cyclopentane ring in the phytanyl chain.

cenancestor was equipped with genetic components of both lipid synthesis pathways [13]. Currently however there is only little data available that supports the heterochiral membrane hypothesis of pre-cells. One *in vitro* study indeed suggests that heterochiral vesicles of fatty acids are less stable than homochiral vesicles of the same compounds and exhibit reduced self-reproduction [16]. Although this data supports facilitated segregation, this study was done on fatty acids rather than glycero-phospholipids. This issue was addressed in another *in vitro* study on mixed vesicles of bacterial and archaeal lipids in which an increased temperature dependent stability was observed which proved to be highly dependent on the length of the fatty acid chains [17]. Here, both the archaeal lipid component C40 diether lipids and C80 bipolar tetraether lipids (GDGT's) showed stable vesicles with bacterial phospholipids. Although this study proofs chemical stability of mixed membranes, which is a strong argument that disputes the facilitated segregation, *in vitro* vesicles are not a correct

representation of the last universal common ancestors membrane. Experimental evidence that supports the mixed membrane of pre-cells should ideally show that heterochirality supports a sustainable situation *in vivo* as well.

Archaeal and bacterial lipid synthesis

Archaeal and bacterial lipid synthesis follows a sequence of reactions that resemble each other in analogous steps (**Fig. 2**). Synthesis of bacterial fatty acid ester-lipids starts from primary metabolites acetyl-CoA and malonyl-ACP. Fatty acid synthesis is a sequence of successive repetitive reactions that result in the formation of C14-C22 fatty acids (acyl-ACP). This acid then becomes esterified to glycerol-3-phosphate by glycerol-3-phosphate acyltransferase to form lysophosphatidic acid. A secondary acyl-ACP is attached at the *sn-2* position of glycerol-3-phosphate and catalysed by *sn*1-acyl-glycerol-3-phosphate-acyltransferase which forms phosphatidic acid. For polar head group attachments, the newly synthesized phosphatidic acid first becomes activated by CDP-diacylglycerol synthase. From this point, the phospholipid synthesis bifurcates to either the synthesis of phosphatidylserine and phosphatidylethanolamine or phosphatidylglycerol. Unlike the bacterial fatty acid synthesis cycle, the isoprenoid hydrocarbon chains of archaeal lipids are not synthesized by a cyclic reaction event. The lipid precursors geranylgeranyl-diphosphate (GGPP) are synthesized by 3 condensation steps of isopentenyl-diphosphate (IPP) with dimethylallyl-diphosphate (DMAPP). The resulting geranylgeranyl diphosphate (GGPP) is attached to the *sn*-3 position of glycerol-1-phosphate by geranylgeranylglyceryl-phosphate-synthase by forming an ether-bond. A second GGPP is attached to the *sn*-2 position by di-*O*-geranylgeranylglyceryl-phosphate synthase to form archaetidic acid (DGGGP). CDP-archaeol synthase activates DGGGP for the attachments of polar headgroups in a sequence of reactions that resembles bacterial polar head group attachments. The biochemistry that results in GDGT synthesis is currently still unknown. Some studies previously suggested that two diether molecules are the precursors of a single tetraether lipid. Here a putative tetraether synthase fuses two diether lipids by a condensation that forms two carbon-carbon bonds [10].

Natural and artificial membrane modifications

Bacteria and archaea evolved the ability to extensively remodel their membrane lipid composition in response to changing environmental conditions. The theory of homeoviscous adaptation was inspired by the observation that fluidity of the *E.coli* membrane remains relatively constant at various temperatures while the lipid composition changes [18]. These changes occur mostly on the core lipids and are implemented to tune the fluidity and permeability of the membrane to retain homeostasis. A commonly described fluidity adaptation example is the incorporation of methyl branches in the fatty acids of bacterial

phospholipids in response to lowering temperature (**Fig. 1A** left)[19]. In addition to natural environments, homeoviscous adaptation has also been reported in various applied industrial settings. Kanno and Kamagata showed that several firmicutes and actinobacteria exposed to 2.0% butanol increase their cyclopropane fatty acid (CFA) content [20]. CFA synthesis has also been recognized as main fortification adaptation against acid and alcohol exposure in *Escherichia coli* [21]. High levels of propionic acid induces an increase in long-chain saturated fatty acids and a decrease in unsaturated fatty acids in *Propionibacterium acidipropionici* [22]. These various recognized homeoviscous modifications have recently also been exploited for the improvement of industrially applied production hosts. These modifications were implemented by membrane engineering strategies to increase the yield of industrial chemicals, or improve the tolerance of the production strain against demanding processing conditions. The introduction of the desaturase DesA from *Synechocystis* to *Anacystis nidulans* transformed the cold sensitive strain more tolerant to cold temperatures [23]. Another successful effort is illustrated by *Escherichia coli* in which *cis-trans* isomerase was introduced to convert *cis*-unsaturated fatty acids to the conical shaped *trans*-unsaturated fatty acids. As a result a more rigid and less fluid membrane was formed with reduced permeability. As such, this modification provides the cells with increased tolerance to alcohols, organic acids and aromatic compounds and enhances viability to adverse physicochemical conditions [24]. Another noteworthy example of membrane engineering was shown in *Clostridium acetobutylicum* where overproduction of CFAs caused an increased butanol tolerance [25].These examples illustrate the high gain that lies in rational design of biological membranes. Despite this, membrane engineering is a relatively novel discipline for which the full potential is yet to be explored.

Heterochiral membrane engineering in *Escherichia coli*

Studies on extremophiles show that both bacteria and archaea tolerate comparable impeding conditions, which suggests an equal flexibility of these domains. Both archaea and bacteria are known to thrive at extreme habitats like hydrothermal vents, hypersaline or boiling acid lakes, or high temperature deep sea oil reservoirs. Nonetheless, subtle differences are recognized. Archaea are dominating the upper extreme temperatures, a trait which they might owe to their unique lipid compositions. Biophysical studies indeed show that archaeal ether-lipids are more robust against hydrolysis than ester-lipids due to their phytanyl hydrocarbon chains and ether bonds [26-30]. E.g. upon boiling acid hydrolysis, the archaeal ether-bonds generally remain intact, whereas ester-bonds are fully disintegrated [31]. These and other findings support that archaea are more resistant to challenging physicochemical conditions than bacteria. It is however unclear, whether extremophilic habitation of archaea is facilitated because of their distinct membranes, providing them precedence to certain niches. This would imply that a fatty acid-ester lipid

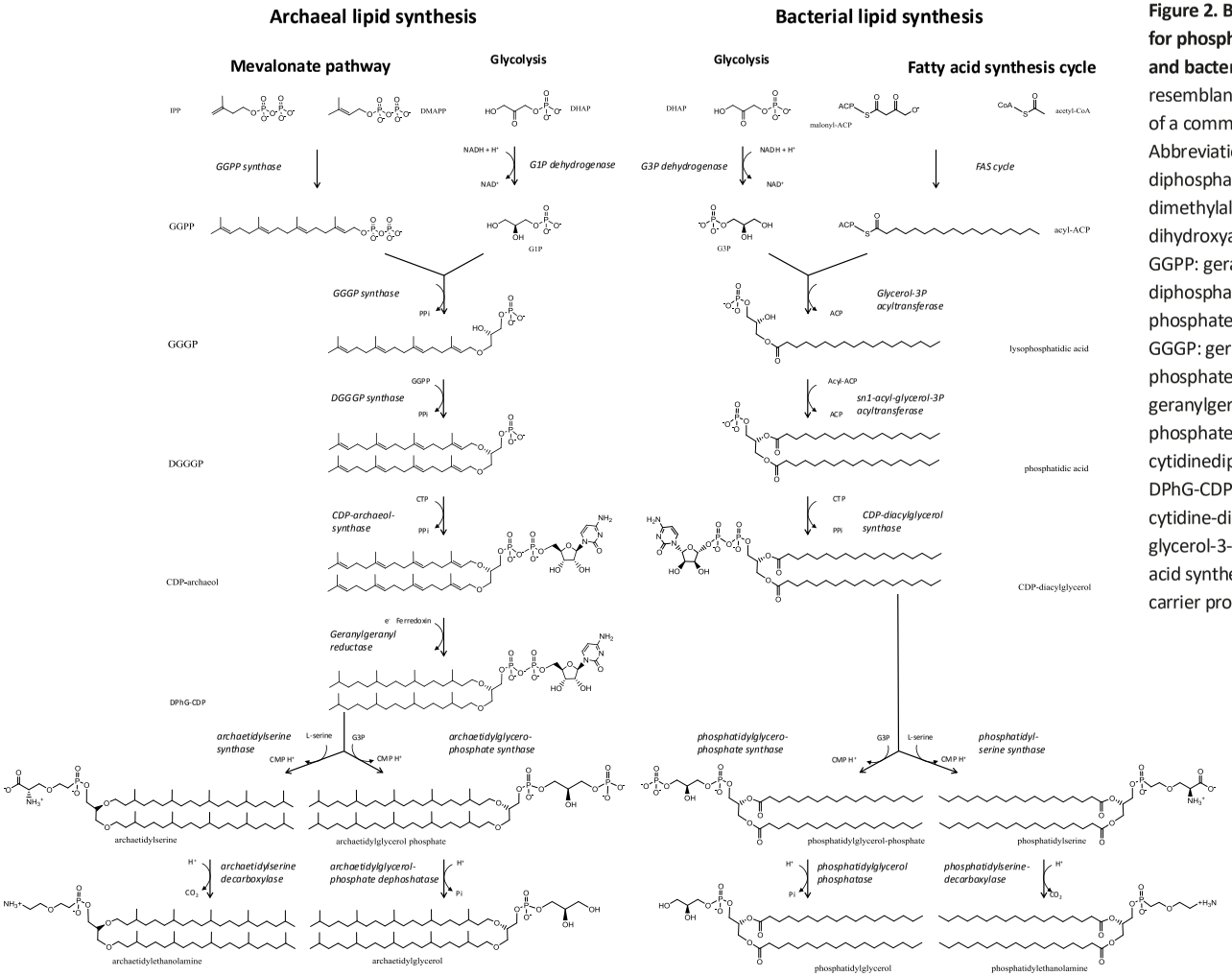


Figure 2. Biosynthetic pathways for phospholipids in archaea and bacteria. The level of resemblance is highly reminiscent of a common descent. Abbreviations: IPP: isopentenyl-diphosphate; DMAPP: dimethylallyl-diphosphate; DHAP: dihydroxyacetone phosphate; GGPP: geranylgeranyl-diphosphate; G1P: glycerol-1-phosphate; PPi: pyrophosphate; GGGP: geranylgeranylglyceryl-phosphate; DGGGP: di-*O*-geranylgeranyl-glycerylglyceryl-phosphate; CDP-archaeol: cytidinediphosphate-archaeol; DPhG-CDP: diphytanylglycerol-cytidine-diphosphate; G3P: glycerol-3-phosphate; FAS: fatty acid synthesis cycle; ACP: acyl carrier protein.

composition determines the boundaries of bacterial extremophile habitation. If true, the archaeal lipids form an interesting tool in industry that can potentially provide a protection for industrially applied bacterial organisms. The synthesis of archaeal lipids in *E.coli* has previously been accomplished in a few studies [32-34]. These studies however never exceeded 1% of total lipid content nor did they introduce both the main archaeal polar lipids archaetidylethanolamine (AE) and archaetidylglycerol (AG). These examples thus do not allow the analysis of heterochirality of a membrane as a contiguous organelle.

The aim of this thesis is therefore to metabolically engineer a complete ether lipid synthesis pathway in *Escherichia coli,* thereby evolving a bacterium with a heterochiral mixed membrane. This engineering will additionally lead to a novel bacterial organism whose membrane characteristics might partly or completely resemble the membrane composition of the putative last universal common ancestor. The strain will be analysed for *in vivo* heterochiral properties and the gained knowledge will contribute significantly to the development of robust membrane engineering approaches. Moreover can this endeavour provide insight in the membrane properties of the postulated LUCA. Based on the homeoviscous adaptation strategies observed in various physicochemical constraints, the presence of saturated diether lipids and GDGTs are correlated with thermal resistance and acid tolerance to the host strain. Conversely, unsaturated diether lipids are correlated with psychrophilicity, but also found in hyperthermophiles. The newly developed strain will be fully characterized and tested for various robustness properties. The overall synthetic biology of the strain is composed of a series of engineering targets that are described in more detail in the thesis outline section.

Thesis outline

Chapter 2: Adaptations of archaeal and bacterial membranes to variations in temperature, pH and pressure

Here we aimed to provide insight into the question whether archaea are profiting from lipids with more robust characteristics compared to bacterial lipids. The bacterial and archaeal lipid compositions and homeoviscous adaptation strategies of extremophiles are reviewed, and placed in context with the physicochemical constraint from which the extremophilic organisms were isolated. We performed a comparative analysis to uncover the regularities of lipid modifications in extreme adaptations. The distilled information was condensed and serves as a tool to predict the outcome of the implemented membrane engineering strategy.

Chapter 3: Metabolic engineering of *Escherichia coli* for high production of isoprenoids

The isoprenoid hydrocarbon chains geranylgeranyl-diphosphate make up the bulk of the archaeal lipids and is assembled from the building blocks IPP and DMAPP. For archaeal lipid production in quantities that suffice appreciable mixed membrane formation, the amount of IPP and DMAPP forms the main limitation. To address this issue, the pathway responsible for synthesis of these building blocks requires extensive upregulation. In most bacteria, the IPP and DMAPP pool are supplied by the MEP-DOXP pathway. Natively, these metabolites are essential for processing of lipopolysaccharide and peptidoglycan constituents, and synthesis of quinones required for the respiration chain. Despite these vital purposes, the available pool of IPP and DMAPP is relatively low. To deal with this shortage we metabolically engineered *E.coli* with a stable upregulation of the MEP-DOXP pathway. To accomplish elevated levels of IPP and DMAPP, a variety of synthetic operons were constructed and integrated into the chromosome. The isolated strains were then characterized for isoprenoid production.

Chapter 4: Modulated gene expression of fatty acid and phospholipid metabolism by PlsB and apo-ACP

Fatty acid and phospholipid synthesis in *E.coli* involve tightly regulated metabolic pathways for which exact details on their regulation is largely unknown. Here we describe the generation of a conditional knock-down of the endogenous pathway with the ultimate goal to further increase the ratio of the introduced archaeal lipids. The downregulation was accomplished by the incorporation of an inducible and repressible $rhaP_{BAD}$-promoter to stimulate expression of the *plsB* gene. The encoded enzyme (*sn*-glycerol-3-phosphate acyltransferase) is responsible for the esterification of acyl-ACP to the glycerol-3-phosphate backbone molecule. Characterization of the mutant revealed that additional inhibition was required by overexpression of acyl carrier protein (ACP) to completely abolish growth. The regulatory actions of PlsB and ACP were then further explored by proteome analysis and lipid complementation assays to obtain insight into the inhibitory mechanism.

Chapter 5: Formation of the ether lipids archaetidylglycerol and archaetidylethanolamine in *Escherichia coli*

CDP-archaeol is the central precursor for polar head group attachment. Here we describe the *in vitro* production of the archaeal lipids archaetidylglycerol and archaetidylethanolamine from nine purified enzymes. Various bacterial enzymes involved in the attachment of L-serine and glycerol as polar head groups were examined for substrate promiscuity for CDP-archaeol. The ether-lipid pathway constituted by a set of archaeal and bacterial enzymes was introduced in *E.coli*, which resulted in biosynthesis of AE and AG.

Chapter 6: Generation of a hybrid heterochiral membrane in *Escherichia coli*

This study describes the assembly and optimization of the archaeal lipid synthesis machinery in the upregulated MEP-DOXP strains. A variety of introduced genes was evaluated for their efficacy of archaeal lipid synthesis. This resulted in the generation of a mixed membrane and optimization of the archaeal lipid quantity. The strain was then extensively characterized for growth characteristics, cell morphology, endogenous lipid composition and its tolerance against physicochemical stressors. In the engineered strain we observed severe morphological changes by scanning electron microscopy and transmission electron microscopy. The engineered bacteria showed a high number of externally bulging lobular bodies. In addition, many cells also show the formation of intracellular vesicles. The high production of archaetidylglycerol caused a concomitant decrease in phosphatidylglycerol that suggests the functional incorporation of the foreign species. Most interestingly, the GGGPs and DGGGPs were also found to have glycerol-3-phosphate promiscuity for archaeal lipid synthesis. When analysing the resistance to physicochemical stressors, we observed a trend towards lower heat stability, but an increased cold-stability as well which confirms the expected hypothesis of mixed membranes with unsaturated archaeol. Altogether, we conclude that the incorporation of mixed membranes leads to viable cells, but with dramatic effects on their morphology and physicochemical stability.

Chapter 7: Identification of *Thermococcus kodakarensis* genes involved in GDGT synthesis

The hypothesized robustness against heat, low pH and solvents by archaeal lipids has mainly been attributed to tetraether lipids. These lipids form a dense monolayer structure which is highly impermeable to protons and solutes and therefore of main interest in membrane engineering strategies. Unfortunately, the biochemical pathway that results in tetraether lipids and the involved genes has not yet been elucidated. In this study we initiated a search for genes that are potentially involved in tetraether synthesis. A list of potential candidate genes was composed by bioinformatics analyses and the most likely candidates were subjected to disruption and overexpression in the hyperthermophilic platform archaeon *Thermococcus kodakarensis*. The implications of these genetic modifications on the membrane lipid compositions were studied by mass spectrometry. Here we demonstrate that mutants of the *TK1548*, *TK1836*, *TK1324* and *TK1957* genes greatly affect the quantity of the tetraether pool. Lipidome analysis of these mutants allowed the prediction of the most likely substrate for GDGT synthesis.

Chapter 8: Thesis summary and general discussion

Finally, the main findings of all chapters will be recapitulated in **Chapter 8**. The conclusions will be placed in context to the hypotheses presented in Chapter 1. The overall impact of successful and unsuccessful strategies on archaeal and bacterial membrane engineering are discussed. To conclude, a proposal for further research will be presented.

Chapter 2

Adaptations of Archaeal and Bacterial Membranes to Variations in Temperature, pH and Pressure

Melvin F. Siliakus[1], John van der Oost[1], Servé W. M. Kengen[1]

[1] Laboratory of Microbiology, Wageningen University, Stippeneng 4, 6708 WE, Wageningen, the Netherlands

Abstract

The cytoplasmic membrane of a prokaryotic cell is a lipid bilayer that shields its cellular content from the environment. In addition, the membrane contains proteins that are responsible for transport of proteins and metabolites as well as for signalling and energy transduction. Maintenance of the functionality of the membrane during changing environmental conditions (pH, temperature, pressure) relies on the cell's potential to rapidly adjust the lipid composition of its membrane. Despite fundamental chemical differences between bacterial ester lipids and archaeal ether lipids, both types are functional under a wide range of environmental conditions. We here provide an overview of archaeal and bacterial strategies to modulate the lipid composition of their membranes. Some molecular adjustments are unique for archaea or bacteria, whereas others are shared between the two domains. Strikingly, shared adjustments are predominantly observed near the growth boundaries of bacteria. Here we discuss the apparent relationship between lipid composition and growth boundaries that are strikingly different between archaea and bacteria.

Keywords: archaea, bacteria, membranes, adaptation, lipids

Abbreviations: AA: arachidonic acid; BCFA: branched chain fatty acid; BMP: bis-mono-acylglycero-phosphate; CL: cardiolipin; DE: diether; DHA: docosahexaenoic acid; DGGGP: di-*O*-geranylgeranylglycerylphosphate; EPA: eicosapentaenoic acid; GDGT: glyceroldialkyl-glycerol-tetraether; G-1-P: glycerol-1-phosphate; G-3-P: glycerol-3-phosphate; MUFA: monounsaturated fatty acid; PMF: proton motive force; PUFA: polyunsaturated fatty acid; SCFA: short chain fatty acid; TE: tetraether; UFA: unsaturated fatty acid

Introduction

The core lipids that serve as the framework for fully mature phospholipids are fundamentally different in bacteria and archaea. These differences are the basis of the so-called 'lipid divide' and are represented by 'phosphatidic acid' for bacteria and eukarya and 'archaetidic acid' for archaea. Phosphatidic acid is composed of two fatty acid hydrocarbon chains esterified to the *sn*-1 and *sn*-2 position of glycerol-3-phosphate (G-3-P) (**Fig. 1A**, left). Archaetidic acid (also known as di-*O*-geranylgeranylglyceryl phosphate; DGGGP) on the other hand consists of two methyl-branched isoprenoids (phytanyls once saturated) connected by ether-bonds to the *sn*-2 and *sn*-3 position of glycerol-1-phosphate (G-1-P) (**Fig. 1A**, right). Besides the latter 'archaetidic acid based' diether lipids, some archaea also contain glycerol-dialkyl-glycerol-tetraether (GDGT) lipids in their membranes (**Fig. 1C**). These lipids are bipolar and believed to be the product of a tail-to-tail condensation between two diether lipids and they form a monolayer instead of a bilayer membrane. Because of these chemical differences, bacteria and archaea also evolved domain specific adaptation mechanisms to effectively respond to different physicochemical conditions of their habitats. The common physicochemical parameters that breach the integrity of membranes are temperature, pH and hydrostatic pressure. Typical membrane characteristics that are adversely affected by these environmental changes are their permeability and fluidity. These parameters have for example a large effect on the function and mobility of membrane proteins, diffusion of nutrients and proper separation during cell division. To maintain physiological homeostasis, membrane integrity is continuously secured through a mechanism called 'homeoviscous adaptation'. This process was first demonstrated in *Escherichia coli* by the observation that fluidity of the membrane remains relatively constant at various temperatures while the lipid composition changes [18]. Cells thus actively modify their lipid composition to maintain membrane functionality at different temperatures. These modifications often cause shifts in ratios of lipid types and/or their hydrocarbon moieties, rather than complete replacement of certain species. Although the core lipids were once thought to form a sharp distinction between the two domains, recent analyses have revealed that a certain extent of overlap exists between some lipid features which in certain cases can be regarded as a form of homeoviscous adaptation.

We here provide an overview of adaptations of the core-lipids of bacterial and archaeal cells in response to changes in their physicochemical conditions. Additionally, we present a comparative analysis of the reported growth ranges of archaeal and bacterial extremophiles to reveal the boundaries of their growth. We further discuss these boundaries in relation to the encountered lipid compositions and their adaptation in the archaea and bacteria. Adaptation of cell physiology to physicochemical conditions occurs at two levels, long-term (genome evolution, defines the range within which a cell can survive) and short-term (reversible regulation of gene expression and enzyme activity to achieve optimal

functionality). A membrane adaptation can thus refer to the generally encountered lipid composition that enables a species to thrive at a particular challenging habitat within its optimum. This kind of adaptation is regarded as a native phenotype that contributes to the robustness against a particular challenge/parameter and is therefore termed 'physiological membrane adaptation'. Alternatively, an adaptation can also involve

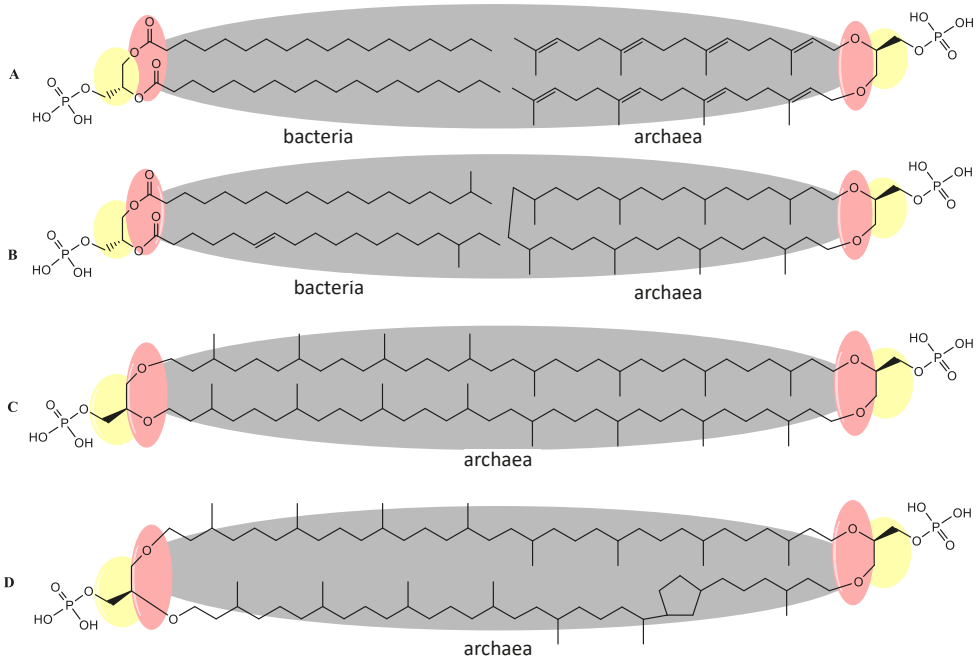


Figure 1. Common bacterial and archaeal lipid variations and the lipid divide. The bacterial 'phosphatidic acid' is represented by panel **A** left. On the right panel the archaeal 'archeatidic acid' or DGGGP core lipid is depicted. The lipid divide is presented in colours. **Grey**: the hydrocarbon chains represent fatty acids in bacteria and isoprenoid chains in archaea. **Red**: Ester-bonds link the hydrocarbon chains to the glycerol backbone in bacteria. In archaea, hydrocarbon chains are attached to the glycerol backbone by ether-bonds. **Yellow**: The backbone moiety in phosphatidic acid is represented by glycerol-3-phosphate. In archaetidic acid the backbone moiety is represented by the enantiomeric glycerol-1-phosphate. On panel **B** left, *iso*-branched chain fatty acid (top) and monounsaturated *anteiso*-branched chain fatty acid (bottom) variations are shown. Panel **B** right shows macrocyclic archaeol. Archaeal bipolar glycerol dialkyl glycerol tetraether (GDGT-0) is depicted in panel **C** and spans the membrane to form lipid monolayers. Panel **D** shows GDGT-1, the bipolar lipid containing 1 cyclopentane ring in the phytanyl chain.

the changes in the lipid composition when conditions of the natural habitat change. This kind of adaptation is regarded as a stress response to the physicochemical change beyond the organism's optimum that aids in the survival of the cell. This change will be termed here as 'membrane stress adaptation' and generally resembles the permanent physiological membrane adaptations to some extent.

Low temperature adaptation

The temperatures of environments from which microbes are isolated roughly span from the freezing point of water and below for psychrophiles to the boiling point of water, for extreme hyperthermophiles. The challenges that microbes face below or above their optimal temperature is to retain optimal functionality of their macromolecules (nucleic acids, proteins, lipids), aiming for a balance between stability (robustness) and flexibility (activity, transition states). As to membranes, the fluidity is dependent on the membrane's phase-transition temperature (T_m or transition midpoint) that expresses the temperature at which a membrane shifts from the preferred liquid crystalline phase into the rigid gel phase when the temperature drops. More specifically, at the midpoint temperature, 50% of hydrocarbon chains melt and a fluid and gel phase coincide. At temperatures below T_m, lipids are 'frozen' by alignment of the hydrocarbon chains perpendicular to the plain of the bilayer [35]. This is a result of the close ordering and side by side packing of the immobilized hydrocarbon chains and gives rise to highly impermeable membranes. Not only is the barrier function affected, but many membrane proteins only function in the liquid crystalline phase [36]. Above T_m, the phospholipid hydrocarbon chains are motile with a gradual increase of motility towards the core of the bilayer. The preferred liquid crystalline phase therefore provides a functional matrix for the many biochemical processes while being permeable to neutral molecules like H_2O, CO_2 and O_2, but impermeable to ions and solutes [37, 38].

To maintain sufficient membrane fluidity below their optimal growth temperatures, bacteria adopt a large variety of modifications to lower the T_m (reviewed in [39-42]). The molecular mechanisms are directed at increasing the fluidity by forming a more disordered gel phase or by prevention of the gel phase. Distinct targets of bacterial cold adaptations (stress and physiological) have been identified: (i) unsaturated fatty acids (UFAs), (ii) short chain fatty acids (SCFAs), (iii) branched chain fatty acids (BCFAs), (iv) carotenoids, and (v) glycolipids and uncommon polar lipids. The most prevalent cold-stress modification is the incorporation of mono-unsaturated fatty acids (MUFAs) [40, 42]. Bacteria actively introduce *cis*- or *trans*-double bonds (**Fig. 2**) by desaturases or synthesize UFAs *de novo* [43].

The main advantage of desaturases is the rapid response they elicit. The *cis*-unsaturated fatty acids, however, increase fluidity more efficiently than *trans*-unsaturated fatty acids. This is due to the immobile 30° kink in the acyl chain that increases the cross-sectional area of the lipid (**Fig. 2**) [45]. The incorporation of *cis*-unsaturations is a well described mechanism of cold-stress in the mesophilic *Escherichia coli* [19]. Correspondingly, the opposite conversion of *cis*- to *trans*-unsaturation fatty acids is correlated with adaptation to higher temperatures in *Pseudomonas syringae* [46]. The presence of poly-unsaturated fatty acids (PUFAs) in response to low temperatures is uncommon in mesophilic bacteria and far less effective in fluidization compared to mono-unsaturations. Nonetheless, omega-3 (ω3; EPA and DHA) and omega-6 (ω6; AA) PUFAs are abundantly detected in marine psychrophiles

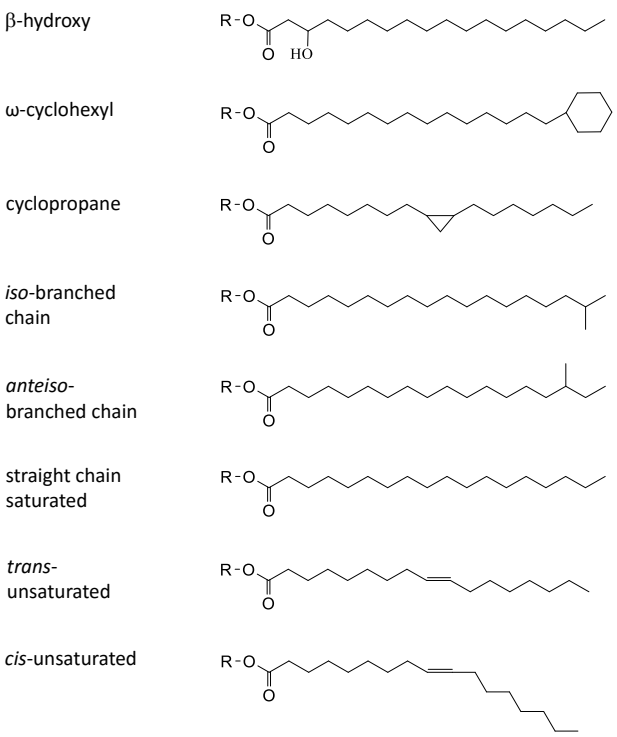


Figure 2. Variations in fatty acid chain conformation The following common fatty acid modifications are depicted from top to bottom: β-hydroxy fatty acid, ω-cyclohexyl fatty acid, cyclopropane fatty acid, *iso*-branched chain fatty acid, *anteiso*-branched chain fatty acid, saturated straight chain fatty acid, *trans* unsaturated fatty acid, *cis* unsaturated fatty acid. (adapted from [40, 44])

and cyanobacteria as an adaptation to low temperatures [39, 42, 47]. Fluidization can also be achieved by incorporation of SCFAs (<12 carbons), but can only be implemented in growing cells and are not employed as an abrupt stress adaptation [44]. It is therefore commonly accepted that SCFA formation is not a universal way of fluidity modification during cold stress. Additional membrane adaptation strategies to sub-optimal temperatures have been observed in gram-positive mesophiles (*Bacillus subtilis, Bacillus T1*) but also the gram-negative thermophile *Thermus thermophilus,* in that they exchange *iso*- for *anteiso*-BCFA (**Fig. 1B**, left and **Fig. 2**)[40, 48]. This is because *anteiso*-positioned methyl groups cause greater fluidity of the membrane due to a greater disturbance of the packing order of the hydrocarbon chains. As with incorporation of SCFAs, BCFAs are synthesized *de novo*, and do not allow a swift response to sudden temperature drops. In the psychrotolerant *Sphingobacterium antarcticus* (T_{opt}=25°) a combination of cold-specific modifications was detected [49, 50]. When cultivated at 5°C, the amount of UFAs is increased as well as the amount of BCFAs. In addition, polar and non-polar carotenoids (C40) were found to

be incorporated into the membrane like also shown for the psychrotolerant *Micrococcus roseus* (T_{opt}=20°C)[51]. These pigments are believed to insert their hydrophilic groups mainly at opposite sides of the polar regions of the bilayer and as such adopt a membrane spanning orientation [52]. Ironically, polar carotenoids like zeaxanthins are believed to decrease membrane fluidity in the liquid crystalline phase, but increase fluidity in the gel phase. Polar carotenoids are therefore assumed to balance the fluidizing effect of the fatty acid modifications while simultaneously enhancing the barrier function to ions and oxygen [50, 52, 53].

With respect to long-term physiological membrane adaptations to cold, many studies reporting on the membrane lipid composition of psychrophilic bacteria (T_{opt} < 15°C) have been performed (**Table 1**). Membrane characterization of *Clostridium psychrophilum* (T_{opt}=4°C) [54, 55], *Colwellia psychrerythraea* (T_{opt}=8.5°C) [56, 57], and *Psychromonas ingrahamii* (T_{opt}=5°C) [58, 59] revealed comparable adaptations to short-term cold stress responses. These adaptations also involve high levels of SCFAs, UFAs, polar carotenoids and glycolipids. Interestingly, although *iso*-BCFAs are detected in psychrophilic membranes, branching does not play a prominent role. Comparative analysis of BCFA percentages in psychrophiles shows a negative relationship with their growth temperature optima (**Fig. 3** and **Table S1A**). In *Psychromonas ingrahamii*, BCFAs levels of only 4.5% of the total fatty acid content are detected and in *Colwellia psychrerythraea* or *Desulfotalea psychrophila* BCFAs are present at trace amounts or completely absent. This suggests that BCFAs confer no benefit and may even be disadvantageous at low temperatures (T_{opt}≤15°C).

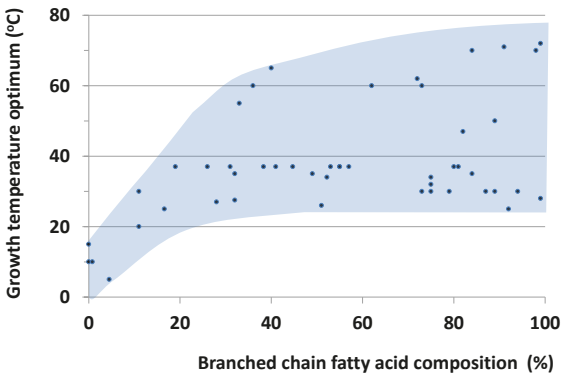


Figure 3. Bacterial temperature optima versus percentage BCFAs Documented temperature optima of 48 bacteria plotted against the percentage of BCFAs of the respective bacterium. The trend shows that at low growth temperatures, BCFA percentages are generally low, while at moderate and high growth temperatures, percentages are variable.

This trend is not only restricted to psychrophiles but also observed as stress response in psychrotolerant species like *Planococcus halocryophilus* [60]. When the psychrophilic *Clostridium psychrophilum* is cultivated below 0°C, a combination of cold-specific modifications is detected like in *Sphingobacterium antarcticus*. These bacteria use a high degree of unsaturated, cyclopropane containing fatty acids and SCFAs. Additionally, polar lipid changes are represented by a high degree of glycolipids, *sn*1-ether fatty acid plasmalogens

and cardiolipins (CLs). These comparisons indicate that long-term physiological adaptations comprise highly similar mechanisms compared to the short-term stress adaptations, except for the contribution of BCFAs.

Membrane adaptation to cold has not been studied as extensively in archaea as in bacteria. Cold-adaption in archaea was long time overlooked probably due to the late recognition of their abundance in oceans and a main interest for hyperthermophilic archaea instead. Nonetheless, it appears that bacterial psychrophiles outnumber archaea in diversity and dominate subfreezing ecosystems (**Fig. 4** and **Fig. S1B**)[61]. According to the standard terminology of psychrophilicity, (T_{opt}≤15°C, i.e. growth optimum at or below 15°C), there are at present only 2 confirmed species of archaeal psychrophiles (*Cenarchaeum symbiosum* T_{opt}=10°C [62] and *Methanogenium frigidum* T_{opt}=15°C [63]). Here we will therefore apply an adjusted terminology for archaeal psychrophilicity (T_{min}<15°C, i.e. capable of growth below 15°C). *Methanosarcinaceae* and *Methanomicrobiales* are thus far the most studied archaeal psychrophiles, but in-depth analyses of their membrane composition are scarce. The few reported lipid adaptation studies, demonstrate a couple of mechanisms to increase membrane fluidity with only one mechanism analogous to bacteria. The documented stress and physiological membrane adaptations to cold are grouped as follows: (i) unsaturated diethers, (ii) isoprenoid hydroxylation, (iii) tetraether:diether ratio, and (iv) number of pentacycli. In *Methanococcoides burtonii* (T_{opt}=23°C) [64] and *Halorubrum lacusprofundi* (T_{opt}=33°C) [65], membranes are found that completely lack GDGTs and have an increased level of unsaturated diether lipids. The physiological adaptation of unsaturation is probably a passive event in which the fully unsaturated precursor DGGGP is selectively or incompletely saturated by reductases, unlike the active introduction of double bonds by desaturases typical for bacteria. Unsaturated archaeol, however, was also found in the hyperthermophile *Methanopyrus kandleri* [66], currently the record-holder of highest growth temperature. This discrepancy seriously questions both the fluidizing effect and the absolute requirement of unsaturated diethers as psychrophilic adaptation mechanism in archaea. In addition to high levels of unsaturated archaeol, a significant amount of hydroxyarchaeol was also observed in *Methanococcoides burtonii* [64]. This modification involves the hydroxylation at the C-3 position of the *sn*-1 isoprenoid chain. The authors suggest that this modification results in an extension of the polarity of the head group, thereby shortening the core of the lipid. It is however not certain whether this type of adaptation is really cold-specific. Although archaea are also known to intercalate non-polar poly-isoprenoids like lycopene and squalene between their membrane lipids, it is not demonstrated to result in fluidity buffering like polar carotenoids do in bacterial membranes. In the extreme halophile *Halobacterium salinarum*, squalene is detected at high quantities and shown to play a major role in packing and lateral organization of the polar lipids [67]. A role of squalene in fluidity adaptation, comparable to carotenoids in bacteria, is therefore also expected in psychrophilic archaea. Squalene is probably implemented to rigidify and reduce permeability to protons and solutes.

Archaeal cold stress adaptation studies are mostly done on thermophiles grown below their optima. From these archaeal below-optimum adaptation studies a recurring theme is an increase in diether content and decrease of membrane spanning GDGTs (**Fig. 1C**). In *Thermococcus kodakarensis* (T_{opt}=85°C), a temperature drop from 85 to 60°C causes archaeol (DE) to increase from 17.7% to 49.1% at the expense of GDGTs (82.3 to 50.9%) [68]. Correspondingly, for *Archaeoglobus fulgidus* (T_{opt}=78°C) the diether content shifts from 60% to 70% when grown at a relatively lower temperature of 70°C [69]. One of the most recently described below-optimum-stress adaptation in archaea concerns a decrease in pentacycle number. Archaeal pentacyclization is a modification often encountered in hyperthermophiles. This feature involves the incorporation of cyclopentane rings along the biphytanyl chains up to 4 rings per chain (**Fig. 1D**). These rings are believed to stabilize the chain packaging and to decrease the permeability of the membrane. In *Thermoplasma acidophilum* (T_{opt}=59°C), lowering the growth temperature to 40°C results in a decrease in the average number of pentacycles per lipid. Here a change in cyclization degree from 2.1 to 1.6 cycles per lipid was observed [70]. A decrease in tetraether and pentacycle number, however, should not be regarded as universal cold-shock response. This is because membranes that are exclusively composed of tetraethers with high numbers of pentacycles per lipid have also been reported to be common in mesophilic archaea [71]. Drawing general conclusions from these changes are thus hampered by the fact that most of these studies are based on hyperthermophiles. Moreover, the observed temperature effects should be corrected for growth phase by normalizing to Optical Density, which has not always been done.

Due to the lack of reported archaea growing below -2°C, it is tempting to suggest that bacteria are better equipped to grow in extremely cold habitats. This hypothesis is supported by the difference in lowest reported temperature optimum between archaea and bacteria (**Fig. S1B**). The lowest confirmed growth optimum for archaea lies at 15°C for *Methanogenium frigidum* [72]. This is notably higher than the optimum of 5°C for the bacterium *Psychromonas ingrahamii* and whose lowest demonstrated growth temperature lies at a startling -12°C [58]. In case there is a growth temperature boundary for archaea around the freezing point of water, a plausible explanation for this threshold may lay in the methyl branched hydrocarbon chains. In the comparative analysis of BCFA percentages we showed that psychrophilic bacteria have less methyl branched lipids compared to mesophiles and thermophiles suggesting a disadvantage during cold adaptation. The methyl branched phytanyls on the other hand, are an imperative feature of archaeol. The phytanyl constituent of archaeal membranes may therefore restrict efficient adaptation mechanisms required to maintain fluidity at subfreezing temperatures. This apparent difference observed in growth temperature minima between archaea and bacteria also leads to a paradox, because the T_m of some archaeal membranes is established at -15° /-20°C [30, 73]. This average T_m thus suggests that archaeal membranes can maintain a liquid crystalline phase below 0°C without the requirement of extensive fluidity enhancing

modifications. Apparently, the isolated psychrophilic archaea and their established growth minima do not support this which may be explained by too low permeability of archaeal membranes instead.

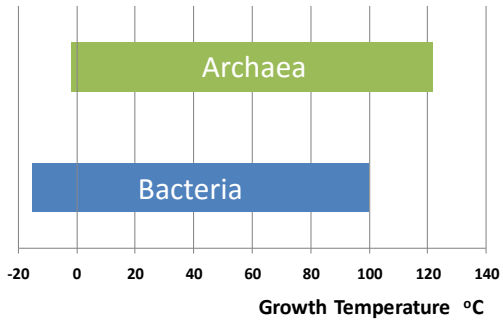


Figure 4. Growth temperature ranges of currently known archaea and bacteria The table depicts the minimum and maximum growth temperatures of well characterized psychrophilic and hyperthermophilic archaea and bacteria. Archaea survive between -2°C and 122°C, whereas bacteria survive between -12°C and 100°C.

High temperature adaptation

At higher temperatures, microbes endure increased fluidity and permeability of the membrane up to a point that lipids are too disordered to maintain a liquid crystalline phase. When the temperature rises beyond the optimum of the liquid-crystalline phase, lipids adopt a fluid phase and ultimately also a non-lamellar phase like the cubic and hexagonal structure [74]. It is therefore not surprising to find adaptation strategies which are opposite to cold adaptations, like those that induce stiffening and promote ordering of the hydrocarbon chains. Mesophilic and thermophilic bacteria mainly adjust fluidity by increasing the amount of (i) branched chain *iso*-fatty acids [18, 75, 76], (ii) saturated fatty acids [48], (iii). long-chain fatty acids and (iv) polar carotenoid content [77, 78]. Counterintuitively, the membranes of thermophilic bacteria are not devoid of *cis*-monounsaturated and *anteiso*-BCFAs [76] which are typically implemented as cold stress adaptation. What appears more important is the ratio between these fatty acids. In the highest temperature window of bacterial hyperthermophiles, some bacteria amend high amounts of BCFAs, which occasionally approach 100% of the fatty acids (**Fig. 3**). These BCFAs are for the greatest part composed of *iso*-BCFAs. Interestingly, thermophilic bacteria that are able to grow above 70°C, implement lipid species that more or less remind of typical archaeal lipids (tetraethers, diether and tetraesters). A key question that comes up here is whether these lipids are typical physiological adaptations to heat, or whether they are remnants of these bacteria's phylogenetically deeply rooted position. In *Thermotogales* species e.g. (T_{max}=90°C) diabolic-acid derived membrane spanning tetraether-lipids are detected [4]. Diether fatty acid lipids are found in *Thermodesulfotobacterium commune* T_{max}=85°C [79] and *Aquifex pyrophilus* T_{max}=95°C [5, 80]. Lastly, long chain dicarboxylic fatty acid dimethyl-esters or tetraester-lipids are found in *Thermoanaerobacter ethanolicus* T_{max}=78°C [81, 82], *Thermoanaerobacter thermosulfurigenes* T_{max}=75°C [83], *Thermosipho africanus* T_{max}=77°C [84], *Fervidobacterium pennivorans* T_{max}=80°C [4] and more. Analogous to the archaeal

tetraether lipids, these bacterial tetraether and tetraester lipids are believed to result from a tail-to-tail condensation between the regular *iso*-branched chain fatty acids that extend from opposite sides of the membrane. Contrary to the earlier dogma, these species show that membrane spanning hydrocarbon chains are not an absolute requirement for survival beyond 80°C. This became especially apparent by analysis of *Aquifex pyrophilus* (T_{max}=95°C) that does not contain membrane spanning lipids but does have ether bonds. It must be emphasized here that in this species a minority of ester-lipids was isolated, which also shows the ability of ester lipids to tolerate these temperatures *in vivo*. Nonetheless, our analysis points at a temperature-dependent boundary at ~80°C for bacterial membranes composed exclusively of ester-lipids. This is in line with the postulation of Robert Huber and Karl Stetter that ether bonds are essential for hyperthermophilic growth [5]. The highest reported bacterial growth optimum so far is demonstrated for *Geothermobacterium ferrireducens* (T_{max}=100°C) [85] as opposed to the archaeal record holder *Methanopyrus kandleri* (T_{max}=122°C) [86] (**Fig. S2**). For thermophilic archaea, membrane spanning tetraether lipids are the most abundant and frequently the only core lipid. Besides this, not much variability is observed in hyperthermophilic membranes. Tetraether lipids form monolayers that are highly stable due to a restricted motility of the hydrocarbon chains. Raising the growth temperature, therefore, elicits an increased tetraether/diether ratio like shown for *Thermococci* [68, 87]. Presence of tetraethers thus appears to be highly supportive for hyperthermophilic growth, a rule that is also followed by extreme hyperthermophilic bacteria which sometimes also use tetraether lipids. A tetraether-containing membrane, however, is not a prerequisite for heat tolerance, as *Methanopyrus kandleri, Thermococcus barophilus, Thermococcus celer* and *Aeropyrum pernix* all have no or only trace amounts of tetraethers but grow optimally between 85-105°C [88-90]. The already high stability of tetraether membranes can be further increased by the incorporation of pentacycli that cause an up-shift of the T_m. In line with this, an increase in the number of pentacycli per lipid is frequently observed, like in *Thermoplasma acidophilum* and *Sulfolobus solfataricus* [70, 91]. Finally, heat adapted membranes may involve macrocyclic archaeols. In these archaeols the isoprenoid chains are cross-linked at their tail-ends, and like tetraethers dramatically restricted in their motion, causing a reduced proton and solute permeability (**Fig. 1B**, right) [92].

Although the various reported lipid modifications suggest certain overall trends in membrane adaptations to high temperature, several exceptions to these rules have been reported as well. By far the most invalidating finding was the observed lack of GDGTs in the hyperthermophile *Methanopyrus kandleri* (T_{max}=122°C). Conflicting data were also observed for glycolipids, that were found at high levels, but also low levels under high temperature conditions [88]. The implementation of ether-bonds seems to form a high correlation with thermal adaptation in bacteria that generally do not use ether-bonds. Despite the exceptions, the boundary of survival for bacteria appears to be around 100°C (**Fig. S2**). Altogether, these exceptions to the general trends indicate that bacterial and archaeal membrane adaptations

to temperature are complex and make use of different strategies (**Table 1**). However, the question then remains what differentiates the bacteria from the archaea with respect to adaptations to low and high temperatures. Bacteria maintain a proper level of permeability and acceptable fluidity at temperatures only just above their phase transition temperatures. Permeability and fluidity of archaeal membranes on the other hand generally stay optimal throughout the entire biological temperature span between 0 and 100°C. Oftentimes, it is therefore hypothesized that the archaeal biphytanyl hydrocarbon bonds linked to glycerol by ether bonds confer higher thermal stability, increased rigidity and reduced permeability to the archaeal membranes [30, 93]. This may explain why archaeal membranes do not require extensive thermal adaptation mechanisms like bacteria do. Despite the mentioned exceptions, comparison of the lipid compositions show that a lower number of ether bonds and level of branching highly determines the lower growth temperature boundary of thermophilic bacteria compared to archaea.

High pressure adaptation

In deep-sea marine environments, microorganisms are exposed to high hydrostatic pressure. These pressures also affect the membrane fluidity, and membranes are even labelled as the most pressure-sensitive biological structures [94]. At increasing pressures, lipids become compressed and membranes consequently loose fluidity, permeability and enter gel phase similar to cold conditions [95]. Deep-sea piezophiles/barophiles are found to grow and even require pressures up to maximally 120 MPa [96, 97] (**Fig. S3**). For piezophilic bacteria and archaea, no obvious difference is observed between their survival boundaries, although this is only based on studies of the bacterium *Colwellia sp.* MT-41($pressure_{opt}$=103 MPa)[98] and of the archaeon *Pyrococcus yayanosii* CH1 ($pressure_{max}$=120 MPa) [99]. The main difference of piezophilic from psychrophilic adaptation in bacteria is the prominent presence of PUFAs [100] in addition to MUFAs. In *Alteromonas sp.* the amount of long chain PUFA 20:5 has been shown to increase with pressure [101]. In contrast, it was shown in *Photobacterium profundum* that the MUFAs and not the PUFAs are correlated with survival at high pressure in bacteria [102]. The accumulation of the polar lipids phosphatidylglycerol and phosphatidylcholine instead of phosphatidylethanolamine is also a frequently observed adaptation [103-105]. It is generally accepted that larger head groups lead to greater disruption of membrane packing and hence to enhance membrane fluidity [104, 105]. Studies in archaea on the homeoviscous adaptation to pressure are scarce and likely hampered by the fact that the identified piezophilic archaea were frequently also (hyper) thermophiles. Currently, only two studies elaborately studied the effect of increased pressure on the membrane composition of archaea. In *Methanococcus jannaschii*, an increase in the macrocyclic archaeol and concomitant decrease in archaeol and GDGTs has been observed [106]. The difficulty of studying the sole effect of pressure on archaeal membranes is exemplified in a study on the piezophilic hyperthermophilic *Thermococcus*

barophilus. Here a similar decrease of GDGTs was found, but an increase in archaeol which completely resembles cold adaptation. Interestingly, these cocci show an upregulated production of lycopene derivatives [87], but it is left unclear whether this is a temperature or pressure induced effect. The common feature however with bacteria is that increased pressure leads to adaptations that result in more fluid membranes.

Low pH adaptation

Like the discovery of microorganisms thriving at extreme temperatures, also bacteria and archaea have been found in ecosystems with extreme pH values. Bacteria and archaea have evolved various ways to cope with extreme acidity (<pH 2) or alkalinity (>pH 10). The challenge both acidophiles and alkaliphiles are facing is to retain a near neutral intracellular pH. The ΔpH ($pH_{in} - pH_{out}$) across the membrane is a major component of the proton motive force (PMF), and as such important for the energy supply of the cell. The cells actively pump out protons by means of a respiratory chain or photosystem. This outflux generates a proton motive force (PMF) that is used by ATPases to generate chemical energy in the form of ATP. Upon external down-shifts in pH, the PMF needs to be adjusted by a higher activity of the respiratory chain or other proton-pumping systems. Acidophiles evolved several mechanisms to maintain near neutral intracellular pH. They e.g. reverse their membrane potential ($\Delta\psi$) to deflect intrusion of protons. A higher expression of proton exporters and secondary transporters is therefore a common strategy. Often, enzymes and chemicals are produced that bind or buffer protons [107]. The most effective strategy however lies in reduction of proton permeability by the plasmamembrane. The neutrophilic bacterium *Escherichia coli*, remodels its membrane, based on only three strategies: (i) an increase in straight short-chain fatty acids content, (ii) a decrease in the amount of unsaturations and (iii) lowering the amount of cyclopropane-fatty acids [108] (**Fig. 2**). These modifications, however, do not follow a consistent pattern with other pH stress studies on neutrophiles. A decrease in short-chain fatty acids was observed in *Streptococcus mutans* instead [109]. In this species, another opposite effect was also observed for the number of unsaturations, which showed an increase to acid stress. Another quite contrasting finding was the fact that a cyclopropane fatty acid knockout in *Salmonella enterica typhimurium* showed a sensitivity to low pH, suggesting a positive relation with these fatty acids instead [110]. As a fourth adaptation to low pH values an overall increase of *anteiso*- and concomitant decrease in *iso*-BCFAs was observed in *Listeria monocytogenes* [111].

Obligate acidophiles synthesize a membrane that is highly impermeable to protons [112]. More consistent membrane adaptations are described in extreme acidophilic bacteria of which *Bacillus acidocaldarius* [113], *Acidiphilium sp.* [114, 115], *Alicyclobacillus acidiphilus* [116] and *Acidithiobacillus ferrooxidans* [117] are most thoroughly studied. These membranes are composed of high levels of *iso*- and *anteiso*-BCFAs, both saturated and mono-unsaturated fatty acids and uncommon β-hydroxy-, ω-cyclohexyl (**Fig. 2**) and

cyclopropane fatty acids. These permanent modifications were detected in the membranes of acidophilic bacteria which tolerate growth at ≤ pH 2.5. In peat bogs (pH 3-5) GDGT-like lipids with branched dicarboxylic fatty acids (*iso*-diabolic acid) bound by ether-bonds to G-3-P have been detected [118]. These lipids, that highly resemble the archaeal membrane spanning lipids, are clearly of bacterial origin and seem to correlate with the low pH of the habitat from which they were isolated. A widespread presence of the *iso*-diabolic acid lipids was also detected in various species of the phylum *Acidobacteria*, but do not grow below pH 3 [119]. Here we find another example for the usage of archaeal-like lipids implemented by bacteria, possibly to deal with the environmental stress. Despite their numerous adaptation strategies, bacteria have not been demonstrated to grow at pH<1 as opposed to archaea (**Fig. 5** and **Fig. S4**). As with temperature adaptations, the three typical homeoviscous adaptation strategies of archaea are also observed in acidic environments.

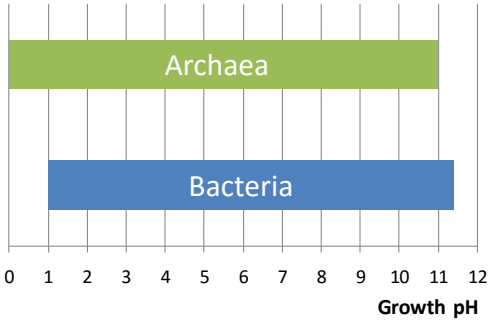


Figure 5. Growth pH range of currently known archaea and bacteria The table depicts the minimum and maximum growth pH of well characterized acidophilic and alkaliphilic archaea and bacteria. Archaea survive between pH 0 and pH 11, whereas bacteria survive between pH 1 and pH 11.4.

A key characteristic of acidophilic archaea is the presence of a membrane monolayer typically composed of nearly 100% GDGTs [71, 120]. Biophysical studies on liposomes composed of GDGTs showed that monolayer membranes are extremely impermeable to protons [121]. A common acid stress response observed in archaea is, not surprisingly, also an increase of GDGT content. Extreme acidophilic archaea, able to thrive at pH<1, are also shown to incorporate multiple cyclopentane rings, to enhance lipid packing, compressibility and membrane rigidity even more [70, 91, 120, 122]. Whether the number of cyclopentane rings is really correlated with acidophilicity, however, is still a matter of debate. Nonetheless is it highly supported by a molecular modelling study that showed a tighter packed structure compared to GDGTs without rings [123]. The relation of cyclopentane rings with thermophilicity is shown multiple times, but in the case of acidophiles hampered by the fact that they are often also thermophiles. The number of cyclopentane rings e.g. was shown to decrease in *Thermoplasma acidophilum* when grown at lower pH values [124]. In contrast, *Acidilobus sulfurireducens* exhibits an increase in the cyclopentane rings at low pH values [125]. Apparently, the pH stress response shows a species- or temperature-dependent variation in terms of the number of cyclopentane rings. Because bacteria are also known to

generate membrane spanning lipids, the lower pH boundary of archaea can be attributed to cyclopentane rings and the abundant methyl-branches. Although it is evident that the membrane-spanning lipids exhibit lower permeability compared to standard bacterial lipids, the membrane spanning lipids in bacteria do not enable them to survive equally low pH levels as archaea. In a molecular dynamics study, Chugunov *et al.* simulated the effect of methyl-groups and cyclopentane rings on membrane permeability and fluidity. A prominent finding was that the presence of methyl groups confers greater fluidization and higher permeability to the membrane than membrane spanning lipids without methyl groups. It is very likely that the inability of bacteria to include multiple methyl chains in their hydrocarbon chains explains why membrane spanning lipids are not detected in the most extreme acidophilic bacteria (pH_{min}<3).

High pH adaptation

On the other side of the pH scale, we encounter the obligate alkaliphiles. In sharp contrast to acidophiles, alkaliphilic bacteria survive comparable and even slightly higher pH values than alkaliphilic archaea, with ~pH=11.4 being the upper limit for bacteria (**Fig. 5** and **Fig. S5**). This suggests that the adaptation mechanisms of archaea and bacteria are nearly equivalent in efficiency. Although studies on the membrane composition in alkaliphilic bacteria are scarce, a few general trends can be observed. The first striking observation is the high variability of lipids present in alkaliphilic bacteria. Often, specific species like bis-mono-acylglycero-phosphate (BMP) lipids and CLs are detected in high quantities [126]. The alkaliphilic membranes are also copiously enriched in BCFAs (both *iso* as *anteiso*), and oftentimes MUFAs are very abundant [127-130]. The actual upper pH limit for life was described for a bacterium, *Bacillus pseudofirmus* (pH_{max}=11.4) [131]. In this species, high levels of CL, BMP, squalenes, and carotenoids were detected. Furthermore, the fatty acids were mainly composed of MUFAs and 92% BCFAs [132]. Altogether, it is concluded that enrichment of BCFAs is a recurring theme in extremophilic bacteria in general, except for the psychrophiles. Another alkaliphilic characteristic is the presence of squalenes, tetrahydrosqualenes or other polyisoprenes in the membrane. Hauß *et al.* showed that these neutral lipids specifically position in the center of the lipid bilayer, parallel to the plane of the membrane, thereby reducing lipid motility and thus proton leakiness [133]. Furthermore, water permeability is reduced and membrane rigidity increased which is an adaptation mechanism also suitable for low pH or high temperature. Squalene incorporation is also one of few adaptation mechanisms that are shared by both archaea and bacteria. In alkaliphilic archaea one can recognize a very distinct pattern from other archaea in which tetraethers and its derivatives are completely absent. Alkaliphilic archaeal membranes are dominated by diether lipids composed of C20:C20 or C20:C25 core lipids [71] and a complete absence of GDGTs. The same holds true for glycolipids which are found only in trace amounts or

which are completely absent. The diether C20:25 lipids unmistakeably result in more fluid and more permeable membranes, which is also reflected in the mesophilic nature of all identified alkaliphilic archaea.

Table 1. Physiological membrane adaptations by core lipid modifications typically found in bacterial and archaeal extremophiles

	Temperature				pH				Pressure	
BACTERIA	Tmin<15°C		Tmax>75°C		pH_{min}<3		pH_{max}>10		> 70 MPa	
Level of chain length		ref		ref		ref		ref		ref
shorter chain ≤C14	+	[54, 57]			+	[115, 117]	+			
longer chain ≥C18										
Level of unsaturation										
PUFA	+	[39, 42, 47]							+	[134]
MUFA-*cis*	+	[54, 57, 135]	+	[80]	+	[117]	+			
MUFA-*trans*	+	[57]								
Level of branching										
BCFA-*iso*			+	[48, 76, 136]	+	[113]	+	[130, 137-139]		
BCFA-*anteiso*			+	[48]	+	[113, 116]	+	[139]		
diabolic acid			+	[4, 140]	+	[119]				
(β)-hydroxy FA	+	[57]			+	[114, 117]				
Level of cyclization										
Ω- cyclohexyl					+	[113, 116]				
cyclopropyl	+	[54]	+	[80]	+	[114, 117]				
Level of Tetraester and etherlipids										
tetraesters			+	[84, 141]						
mono- di- tetraethers			+	[4, 5, 79, 80]	+	[118, 119]				
Level of terpenes										
polar carotenoid	+	[50, 51]	+	[77, 78]						
non-polar terpenes							+	[139]		
Other modifications										
cardiolipins	+	[54]					+	[139]		
glycolipids	+	[54]	+	[77]						
BMP							+	[139]		

	Temperature				pH				Pressure	
ARCHAEA	Tmin<15°C		Tmax> 75°C		pH_{min} <3		pH_{max} >10		>70 MPa	
Level of chain length		ref		ref		ref		ref		ref
shorter chain C20	+	[64]	+	[88-90]			+	[142-148]	+	[106, 149-151]
longer chain C25			+	[152]			+	[142-148]		
Level of saturation										
unsaturated diethers	+	[64, 65]	+	[66]						
Level of branching										
hydroxyarchaeol	+	[64]								
Level of cyclization										
pentacyclic TE			+	[70, 91]	+	[70, 91, 120, 122]				
macrocyclic			+	[153]					+	[106, 153]
Level of Tetraether lipids										
tetraethers	-	[64]	+	[68, 87]	+	[71, 120]	-	-[71]	-	[106]
Other modifications										
glycolipids			+	[66]	+	[91, 122]	-	[143, 145, 148, 154, 155]		

Abbreviations: PUFA: polyunsaturated fatty acids; MUFA-cis: *cis*-monounsaturated fatty acids; MUFA-trans: *trans*-monounsaturated fatty acids; BCFA-iso: *iso*-branched chain fatty acids; BCFA-anteiso: *anteiso*-branched chain fatty acids; BMP: bis-mono-acylglycero-phosphate; TE: tetraethers

Concluding remarks on extremophile membrane adaptations

Various studies have shown that archaeal membranes have more robust properties compared to bacterial membranes [31, 93, 156, 157]. Nevertheless, both archaea and bacteria inhabit extreme environments as extensively as they do in mild environments. To provide insight into the occurrence of both domains in extreme environments we made an inventory of maximal and minimal growth temperatures of well characterized hyperthermophiles, psychrophiles and barophiles. This analysis demonstrates an equally broad temperature growth range of ~120 degrees for both archaea and bacteria. An important difference however lies in the domination at the temperature extremes. Bacteria were found to dominate lower temperature ecosystems, whereas archaea dominate the higher temperature ecosystems. A comparable pattern also appeared when we analysed the pH scale where archaea dominate the bacteria <pH 1, and bacteria show a slight benefit at high alkalinity >pH 11. Regardless of the basal core lipid composition, both domains extensively modify their membrane composition to overcome the inhospitable conditions. Here we discussed the efficiency of homeoviscous adaptations which may correlate with the different growth boundaries. At the observed boundaries it appears that bacterial core lipids

are better accommodated to modify the membrane fluidity whereas archaeal core lipids are more efficient in rigidifying modifications. When looking into the membrane composition of the most extreme representatives of the domains, a consistent pattern appears at low pH and high temperature. Archaea form tetraether monolayer membranes that constitute an adaptation mechanism of highly efficient rigidification and reduction of permeability. Interestingly, some of these typical archaeal features are also being implemented in bacterial hyperthermophiles and acidophiles. Examples are an increased level of ether-bonds, membrane spanning lipids, and a high level of methyl-branched BCFAs. It appears that these features contribute to the fact that e.g. both *Thermotogales* and *Aquificales* can survive higher growth temperatures than other bacterial hyperthermophiles. Typical membrane spanning lipids are also found in bacteria that thrive in acidic ecosystems. However, these traits don't seem to increase their survival boundary at low pH. Despite this, it is conceivable that the cyclopentane rings do show correlation with the lowest pH survival range as comparable cyclization's are also found in bacterial lipids. In this case, it would support the hypothesized lower survival boundaries to high temperature and low pH from implementing standard bacterial lipids (fatty acid-diester-lipids). On the other hand, bacterial lipids do seem more beneficial for growth at low temperatures. It is observed that psychrophilic bacteria reduce their BCFA load as one of the main lipid modifications. This implies that at sub-freezing temperatures the effect of disturbed packing order near the interface of the bilayer caused by BCFAs, is no longer sufficient to maintain fluidity. Hence, facultative psychrophiles mainly implement unsaturations, which cause a greater cross-sectional area of the lipids and less tight packing of the hydrocarbon chains than BCFAs. The isoprenoid hydrocarbon chains of archaeol which have an imperatively high number of methyl branches may thus be a disadvantageous characteristic at low temperatures. The methyl branches maintain a too dense packing which apparently cannot easily be compensated. In sharp contrast to bacteria, imitation of bacterial lipid features by archaea is thus far not observed in psychrophilic or alkaliphilic ecosystems. Despite this, both ester- and ether-bound fatty acid lipids are also detected in archaea, but there does not seem to be a correlation with either growth temperatures or pH from the analysed archaea thus far [6]. Altogether, the deductions presented here are in line with the hypothesis that cytoplasmic membrane compositions are one of the main determining factors behind survival boundaries at low pH and high temperature.

Supplemental information

Figure S1A. Bacterial temperature optima versus branched chain fatty acid compositions

#	BCFA %	Topt (°C)	Species	BCFA ref	Temp ref
1	0	15	*Psychrobacter frigidicola*	[158]	[158]
2	0	10	*Desulfotalea psychrophila*	[135]	[135]
3	1	10	*Colwellia psychrerythraea*	[159]	[159]
4	4,5	5	*Psychromonas ingrahamii*	[160]	[161]
5	11	30	*Micrococcus roseus*	[162]	[163]
6	17	25	*Sphingobacterium antarcticus*	[164]	[49]
7	19	37	*Streptococcus faecalis*	[165]	[166]
8	26	37	*Clostridium Difficile*	[167]	[168]
9	28	27	*Pseudomonas rubescens*	[169]	[170]
10	31	37	*Flavobacterium thalpophilum*	[171]	[172]
11	32	28	*Streptosporangium album*	[173]	[174]
12	32	35	*Streptosporangium vulgare*	[173]	[174]
13	33	55	*Desulfotomaculum nigrificans*	[175]	[176]
14	38	37	*Streptococcus salivarius*	[165]	[177]
15	40	65	*Bacillus stearothermophilus*	[178]	[178]
16	41	37	*Streptococcus agalactiae*	[165]	ATCC
17	45	37	*Clostridium sordellii*	[167]	Dsmz
18	49	37	*Staphylococcus capitis*	[179]	[180]
19	51	27	*Actinoplanes philippinensis*	[173]	ATCC
20	52	34	*Desulfovibrio africanus*	[175]	[181]
21	53	37	*Desulfovibrio vulgaris*	[175]	ATCC
22	55	37	*Bacteroides vulgatus*	[173]	Dsmz
23	55	37	*Streptococcus pyogenes*	[165]	Dsmz
24	57	37	*Propionibacterium acnes*	[182]	ATCC
25	59	62.5	*Bacillus acidocaldarius*	[113]	[183]
26	62	60	*Thermus ruber*	[184]	[184]
27	72	62	*Thermomonospora curvata*	[173]	[185, 186]
28	73	30	*Bacillus pumilus*	[187]	Dsmz
29	73	60	*Clostridium thermocellum*	[188]	[189]
30	75	30	*Desulfovibrio desulfuricans*	[190]	[190]
31	75	32	*Staphylococcus simulans*	[179]	[180]
32	75	34	*Staphylococcus hominis*	[179]	[180]
33	79	30	*Bacillus cereus*	[191]	ATCC
34	80	37	*Bacillus anthracis*	[191]	[192]
35	81	37	*Staphylococcus aureus*	[179]	Dsmz
36	81	32	*Staphylococcus epidermidis*	[179]	Dsmz
37	82	47	*Arthrobacter radiotolerans*	[193]	[194]
38	84	35	*Staphylococcus haemolyticus*	[179]	Bergey's
39	84	70	*Thermodesulfotobacterium commune*	[195]	[195]
40	87	30	*Micrococcus luteus*	[196]	[197]
41	89	52.5	*Bacillus coagulans*	[173]	[198]
42	89	30	*Bacillus megaterium*	[173]	Bergey's
43	91	71	*Thermus aquaticus*	[199]	[200]
44	92	25	*Bacillus psychrophilus*	[201]	[202]
45	95	33,5	*Bacillus subtilis*	[187]	Dsmz
46	98	70	*Thermoanaerobacter thermohydrosulfuricus subsp. Carboxydovorans*	[203]	[203]
47	99	72	*Thermus thermophilus*	[204]	[205]
48	99	28	*Micrococcus conglomeratus*	[196]	Dsmz

Abbreviations: ATCC (American type culture collection) bacteriology collection, DSMZ; Deutsche Sammlung von Microorganismen und Zellkulturen, Bergey's; Bergey's manual of systematic bacteriology.

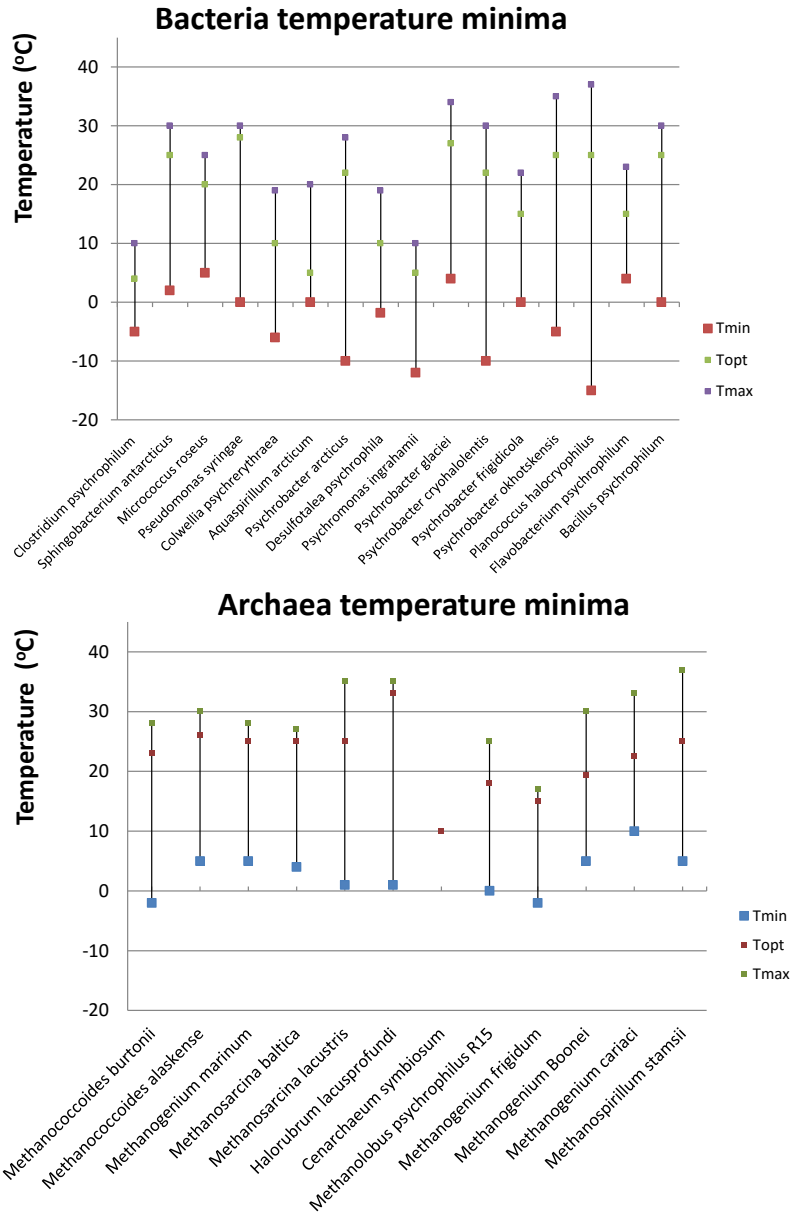


Figure S1B. Reported minimal, optimal and maximal growth temperatures of psychrophiles. Psychrophilic organisms are grouped according to the bacterial and archaeal domain.

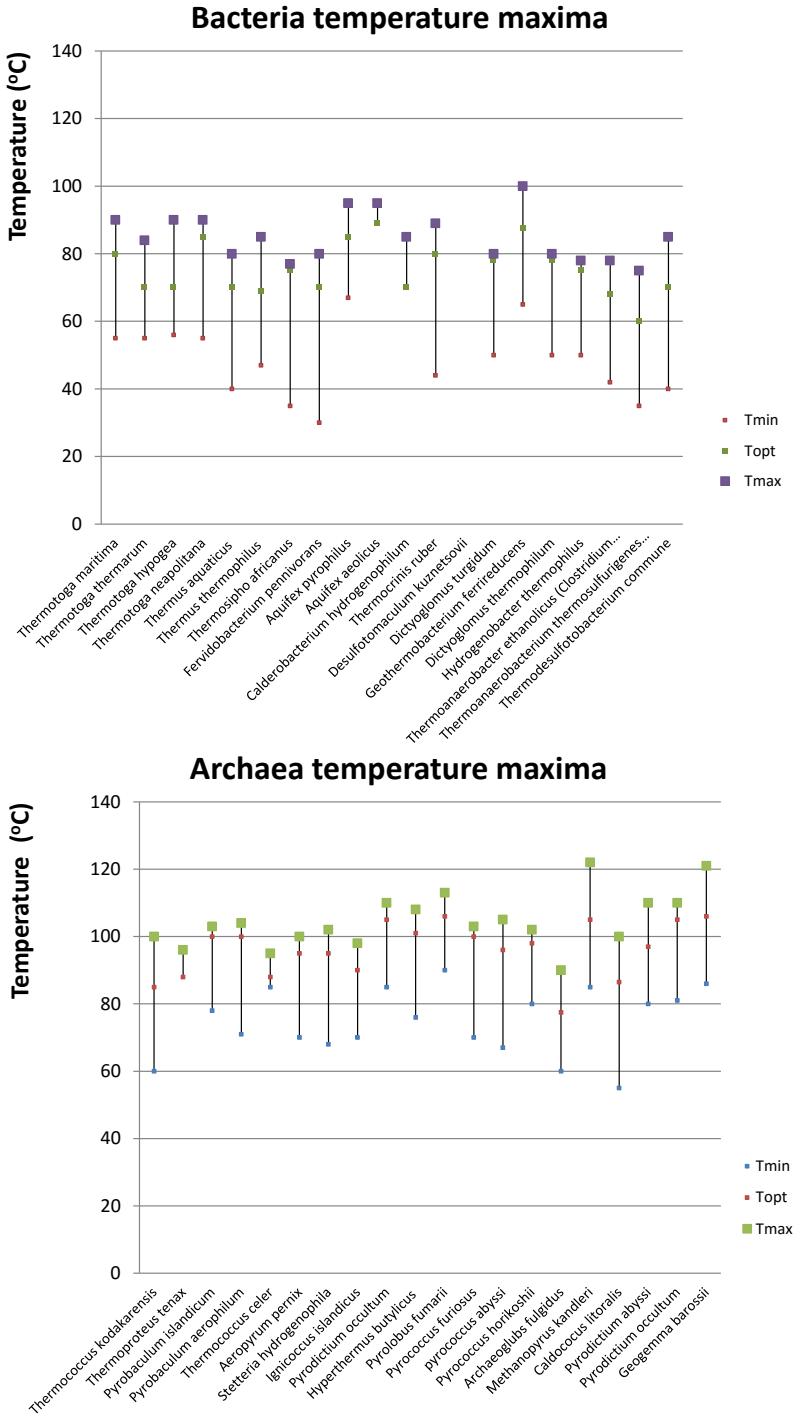


Figure S2. Reported maximal, optimal and minimal growth temperatures of hyperthermophiles. Hyperthermophilic organisms are grouped according to the bacterial and archaeal domain.

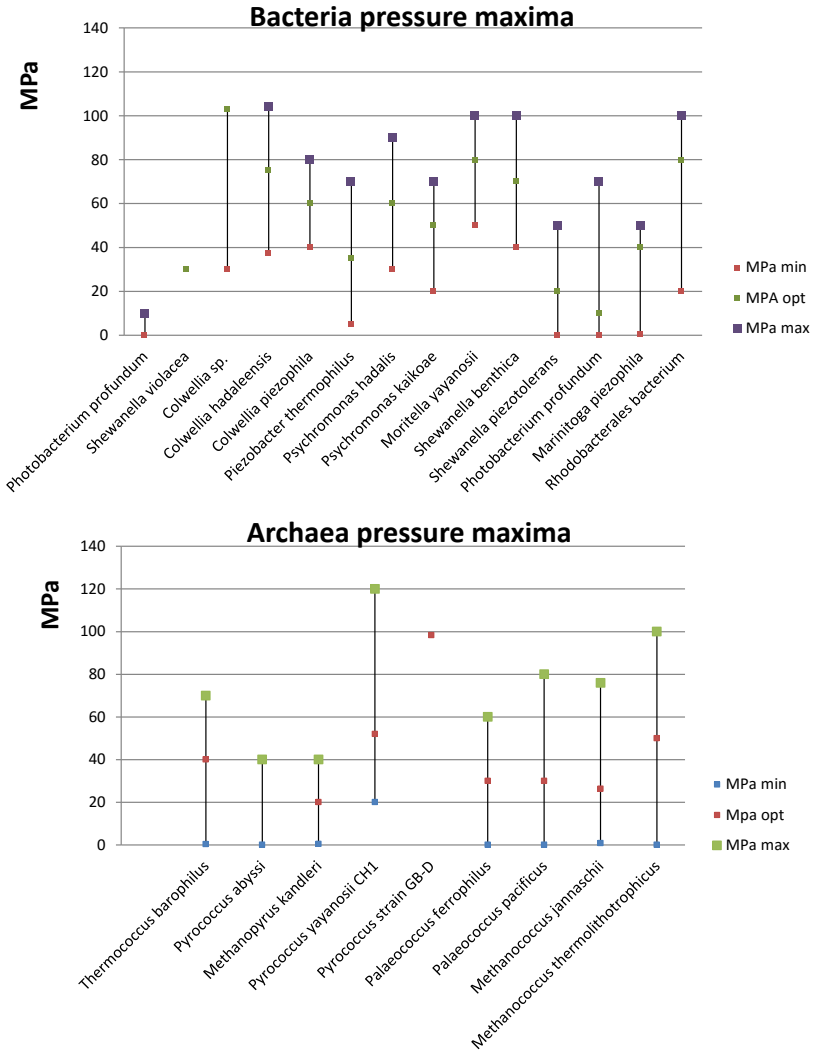


Figure S3. Reported maximal, optimal and minimal pressure-values of piezophiles/barophiles. Barophilic organisms are grouped according to the bacterial and archaeal domain.

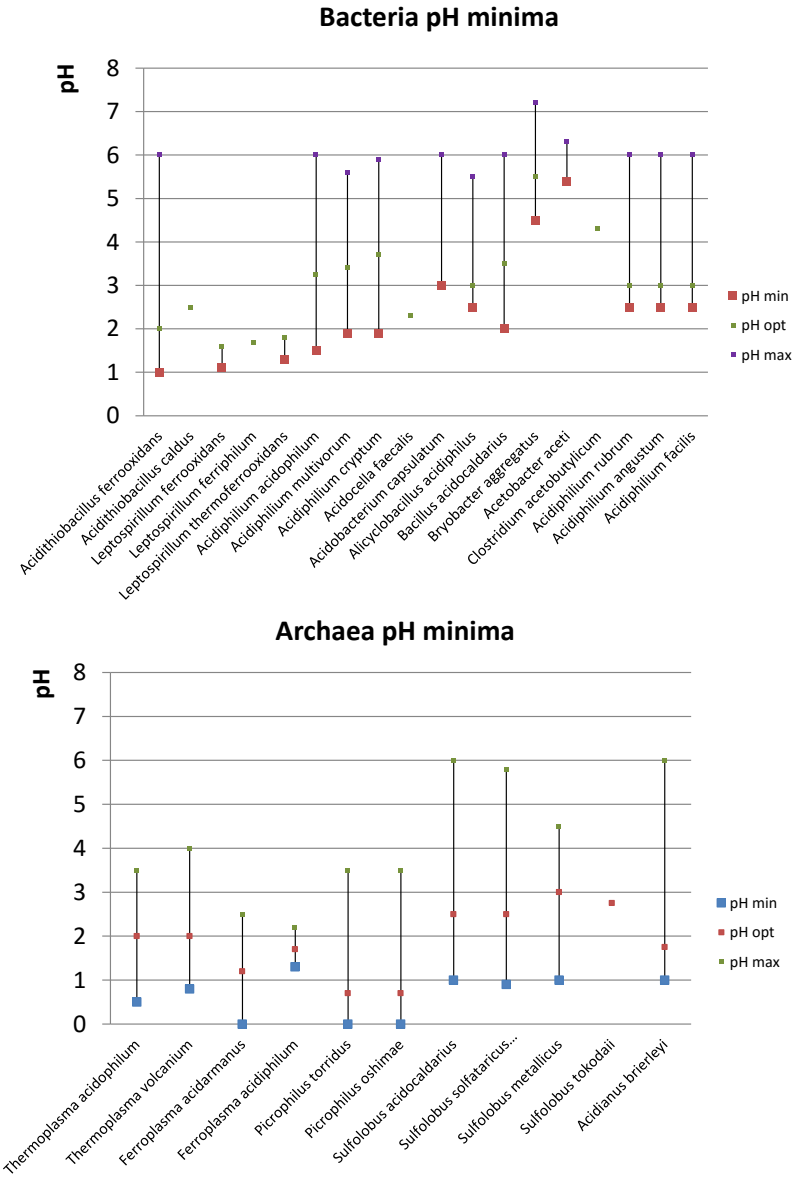


Figure S4. Reported maximal, optimal and minimal pH-values of acidophiles. Acidophilic organisms are grouped according to the bacterial and archaeal domain.

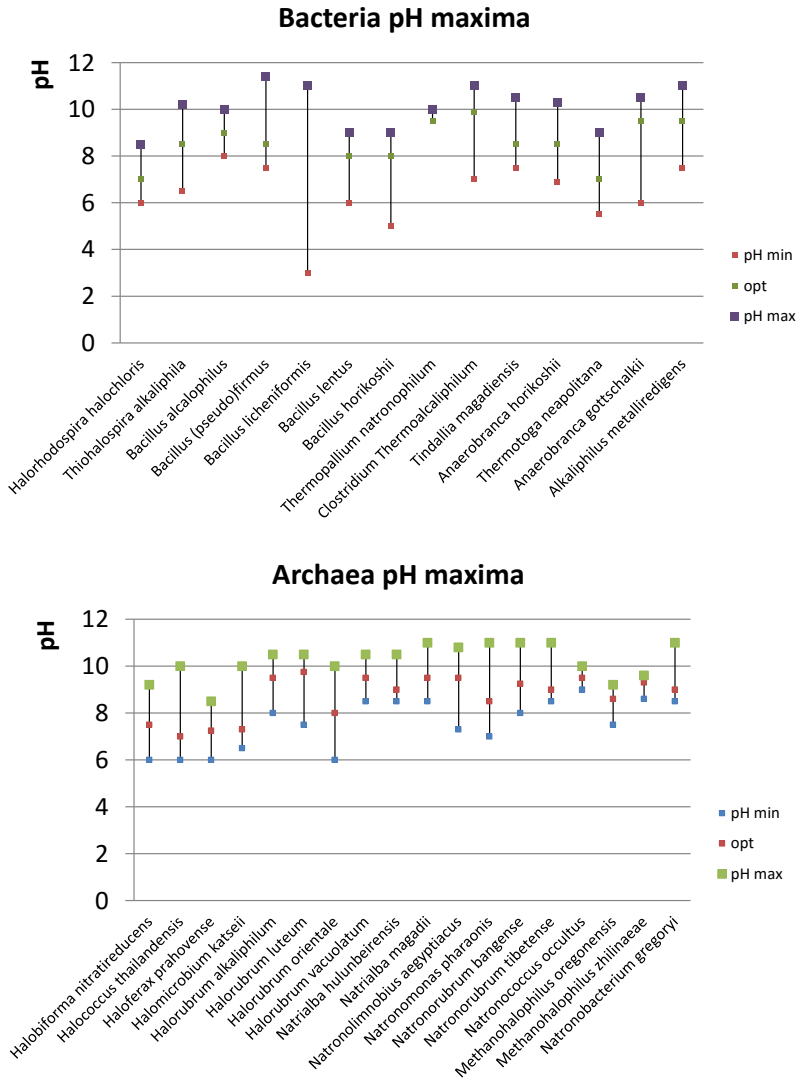


Figure S5. Reported maximal, optimal and minimal pH-values of alkaliphiles. Alkaliphilic organisms are grouped according to the bacterial and archaeal domain.

Acknowledgements

This research was carried out within the BE-Basic R&D Program, which was granted a FES (Fonds Economische Structuurversterking) subsidy from the Dutch Ministry of Economic affairs, agriculture and innovation (EL&I).

Chapter 3

Metabolic Engineering of *Escherichia coli* for High Production of Isoprenoids

Melvin F. Siliakus[1], Jessica Pasterkamp[1], Eva Kleibusch[1], Servé W.M. Kengen[1], John van der Oost[1]

[1]Laboratory of Microbiology, Wageningen University and Research Centre, Stippeneng 4, 6708 WE, Wageningen, The Netherlands.

Abstract

Heterologous expression of isoprenoid pathways has resulted in successful industrial applications. However, widespread usage of isoprenoids is still hampered by sub-optimal productivity of currently used platform organisms. A bottleneck is the intracellular supply of the two isoprenoid precursors isopentenyl-diphosphate (IPP) and dimethylallyl-diphosphate (DMAPP). In *E. coli*, these building blocks are synthesized by the MEP-DOXP pathway. To accomplish elevated levels of IPP and DMAPP, we overexpressed different combinations of genes encoding the key enzymes of the pathway (IDI, IspDF and DXS), both from plasmids and after integration from the host chromosome. By systematically screening of the ori-macrodomain of the *E.coli* chromosome, we identified a convenient insertion site at the *aceB/aceA* locus that tolerates controlled induction and repression of the integrated operons. Unlike the plasmid-encoded variants, the integrated operon resulted in significantly enhanced production of lycopene, used as a reporter of the isoprenoid production. At optimal IPTG induction, our strains increased isoprenoid production upto 5.34-fold compared to the wild-type strain. Overall, we engineered an *E.coli* strain that can serve as a platform for the balanced expression of the isoprenoid building blocks and an isoprenoid biosynthesis pathway of interest.

Abbreviations: DCW: dry cell weight; DMAPP: dimethylallyldiphosphate; DXP: 1-deoxy-D-xylulose-5-phosphate; EPOD: extensive protein occupancy domain; Fis: Factor for Inversion Stimulation; FPP: farnesyl diphosphate; GFP: green fluorescent protein; GPP: geranyl diphosphate; G3P: glyceraldehyde 3-phosphate; H-NS: histone-like nucleoid structural protein; IHF: Integration host-factor; IPP, isopentenyl diphosphate; MEP, 2C-methyl-D-erythritol; PEP, phosphoenolpyruvate; UndP: undecaprenylphosphate

Introduction

The demand for isoprenoids in industry has grown rapidly over the last years, with applications ranging from plastics and rubbers to therapeutics and solutes. Isoprenoid biosynthetic pathways from various organisms have been characterized and heterologously expressed in production strains for biotechnological applications. Oftentimes the resulting compounds are then chemically processed to a final product of interest. Successful examples are the production of the anti-malarial agent artimisinic acid in yeast [206], and the anti-cancer agent taxol by introduction of the biosynthesis pathway of taxadiene in *Escherichia coli* [207]. Although several studies aimed to optimize production levels of interesting isoprenoids in different microbial production hosts [208-213], there is still room for improvement. Generally, isoprenoid biosynthesis can be divided into two partial metabolic pathways [214]. The upstream part concerns the synthesis of the precursor's isopentenyl pyrophosphate (IPP) and dimethylallyl pyrophosphate (DMAPP), whereas the downstream part is a collection of variable pathways for the biosynthesis of the actual isoprenoid variant. The resulting compounds are then utilized in processes as diverse as respiration, protein prenylation, cell membrane maintenance, hormones or protein anchors [215].

Both archaea and bacteria have the capacity to produce isoprenoids, but they differ considerably with respect to the generation of the IPP/DMAPP building blocks. The first identified pathway is the classical 'mevalonate' or 'HMG-CoA reductase' pathway. This pathway is present in all higher eukaryotes and most archaea. In this route three acetyl-CoA molecules are joined and converted to IPP and DMAPP, with mevalonate as a key intermediate. In bacteria, an alternative pathway (non-mevalonate or MEP-DOXP) has been identified as the dominant pathway [216]. The unexpected finding that ^{14}C-labelled acetate was not directly incorporated into isoprenoids with concomitant scrambling of the label led to the identification of the alternative 'MEP-DOXP' route in bacteria. Instead of acetyl-CoA, initiation appeared to depend on the condensation of pyruvate and glyceraldehyde-3-phosphate. By now, the pathway is fully elucidated and all eight reactions are well characterized [217] (**Fig. 1**).

In bacteria, the MEP-DOXP pathway is mainly used for production of farnesyl-diphosphate (C_{15} FPP) and the subsequent production of respiratory election carriers (ubiquinone, menaquinone) and sugar carrier-lipid undecaprenylphosphate (C_{55} UndP) for processing of the cell wall [218]. The essential UndP translocates O-antigen and peptidoglycan constituents from the cytoplasmic face into the periplasmic face of the inner membrane [219, 220]. Because of this usage, the basal IPP and DMAPP production level is relatively low. Instead, archaea use the mevalonate pathway also for the synthesis of the plasma membrane lipids, and the IPP and DMAPP supply is therefore considerably higher. Despite its lower production potential, *E. coli* still remains the production organism of first choice for industrial isoprenoid production. This is mainly due to the versatility of the available genetic toolbox.

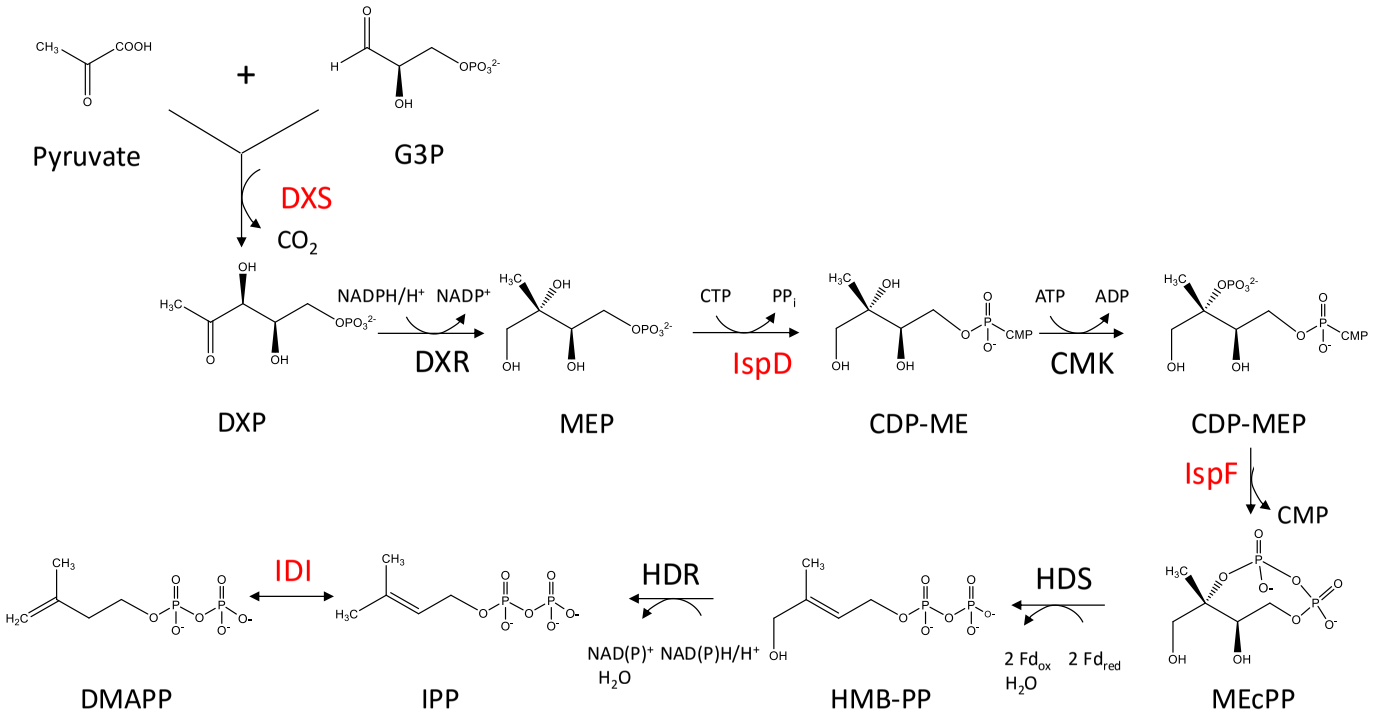


Figure 1. Non-mevalonate or MEP-DOXP pathway The MEP-DOXP pathway entails 8 catalytic steps that convert pyruvate and glyceraldehyde-3-phosphate into isopentenyl-diphosphate (IPP) and dimethylallyl-diphosphate (DMAPP). Abbreviations: G3P, glyceraldehyde-3-phosphate; DXP, 1-deoxy-D-xylulose-5-phosphate; MEP, 2C-methyl-D-erythritol 4-phosphate; CDP-ME, 4-diphospho-cytidyl-2C-methyl-D-erythritol; CDP-MEP, 4-diphosphocytidyl-2C-methyl-D-erythritol-2-phosphate; MEcPP, 2C-methyl-D-erythritol-2,4-cyclodiphosphate; HMB-PP, hydroxy-2-methyl-2-(E)-butenyl-4-diphosphate. Enzymes: DXS, 1-Deoxy-D-Xylulose-5-phosphate synthase; DXR, 1-Deoxy-D-xylulose-5-phosphate reductoisomerase; CMK, 4-diphosphocytidyl-2C-methyl-D-erythritol kinase; IspD, 2C-methyl-D-erythritol 4-phosphate cytidylyltransferase; IspF, 2C-methyl-D-erythritol-2,4-cyclodiphosphate synthase; HDS, hydroxy-2-methyl-2-(E)-butenyl-4-di-phosphate synthase; HDR, 4-hydroxy-3-methyl-2-(E)-butenyl-4-diphosphate reductase; IDI, isopentenyl-diphosphate isomerase. Enzymes targeted for overexpression in this study are depicted in red.

To establish a platform *E.coli* strain for isoprenoid production, we embarked on the construction of a strain that complies with the following characteristics. (i) The strain overproduces the isoprenoid precursors IPP and DMAPP. Overexpression studies of various MEP-DOXP genes and the resulting synthesis of the building blocks IPP and DMAPP have previously been performed to determine the individual rate limits [209]. However, more information is required on the different combinations of MEP-DOXP steps, which may optimize the established overproductions even further. (ii) The overproduction needs to be tunable to adjust to the requirements of the downstream pathway of interest. To this end, a series of artificial plasmid-borne operons with a variable combination of MEP-DOXP genes were positioned under control of an inducible *lac*-promoter. (iii) The overproduction of IPP and DMAPP needs to be stable without the requirement of an antibiotic resistance marker to retain the trait [206]. This is achieved by integration of the MEP-DOXP operon into the chromosome, which consequently also lowers a potential vector burden. Such a strain has previously been obtained, with a MEP-DOXP operon integrated into the arabinose operon region [207]. The locus of integration however (see below, iv), has never been considered as a variable that influences transcription of the construct and consequently also the production levels of IPP and DMAPP. (iv) Besides promoters and transcription factors, transcription is also dependent on the spatial chromosomal location [221]. Moreover, it has been shown that these subdomains display regions with highly heterogeneous transcription rates [221, 226, 227]. Vora *et al.* further studied these regions and revealed that highly expressed regions correlate well with high protein occupancy referred to as "*he*EPODs" (highly expressed Extensive Protein Occupancy Domains) [228]. These *he*EPODs generally show a high enrichment of RNA polymerases and Fis (Factor for Inversion Stimulation) proteins. Additionally, it was shown that distinct transcriptionally silent regions also correlate well with high protein loading and were therefore termed "*ts*EPODs" (transcriptionally silent Extensive Protein Occupancy Domains). These *ts*EPODs are especially enriched with H-NS, IHF and Fis proteins instead, and therefore thought to anchor subdomains to the central region with the multiple structural proteins that are highly loaded on these positions. Hence, *ts*EPODs are assumed to hold the structural function in chromosome organization. Here, we studied a putative subdomain based on EPOD coordinates for the integration of a set of characterized operons. The yields of these integrated operons were then compared with the plasmid-borne operons.

Results

Construction and analysis of engineered MEP-DOXP operons

To increase carbon flux through the MEP-DOXP pathway of *E.coli*, a set of synthetic plasmid-encoded operons were designed. For this purpose, we cloned 4 MEP-DOXP genes interspaced by a ribosome binding site (AGGAGAT). From the 8 catalytic steps (**Fig. 1**), the enzymes isopentenyl-

diphosphate isomerase (IDI), 2C-methyl-D-erythritol 4-phosphate cytidylyltransferase (IspD), 2C-methyl-D-erythritol-2,4-cyclodiphosphate synthase (IspF) and 1-Deoxy-D-Xylulose-5-phosphate synthase (DXS) were targeted for overproduction as these four steps have been demonstrated to be limiting the flux of the pathway [209]. The pGFPuv plasmid was chosen as a cloning vector because it encompasses a strong *lac*-promoter, a high copy origin of replication and the gene encoding the Green Fluorescent Protein (*gfp*). Native MEP-DOXP genes were cloned at the 5'-end of *gfp*. This position includes *gfp* at the C-terminus of the operon which then serves as a phenotypic marker for expression of the synthetic operons. Operon assembly was initiated by cloning of *idi* (**Fig. 2B**). This enzyme is responsible for the final catalytic step of the MEP-DOXP pathway and has been found to play an important role in controlling isoprenoid flux [209]. The next cloning concerned the *ispD* and *ispF* genes (**Fig. 2C**). A complete operon '*idi-ispDF-dxs*' was made by cloning of the *dxs* gene (**Fig. 2D**)[208, 209]. In a recent study, *in vitro* inhibition on the IspF enzyme was described requiring both the metabolites farnesyl-diphosphate (FPP), and 2C-methyl-D-erythritol 4-phosphate (MEP) [211]. This observation that suggests a regulatory role, prompted us to make two additional variations on the complete operon to study a potential inhibitory effect during upregulations of the pathway (**Fig. 2E** and **F**). First, a variant was made that lacks both intermediate steps catalyzed by IspD and IspF. By omitting both the *ispD* and *ispF* genes from the overexpression operon, a proportionally higher MEP accumulation is anticipated compared to the complete operon.

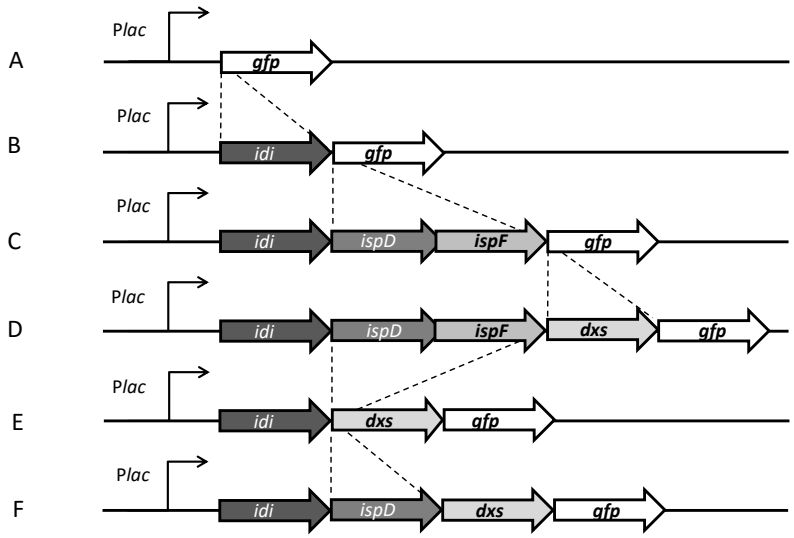


Figure 2. Artificial MEP-DOXP operons constructed in this study, A represents the unmodified pGFPuv vector. **B** represents the first operon with the insertion of *idi.* **C** represents the secondary modification to which the bicistronic genes *ispD* and *ispF* were introduced. **D** represents the complete operon in which the *dxs* gene was added. **E** and **F** represent constructs to test the effect of the addition of *ispF* and *ispD* specifically by the consecutive deletion of *ispDF* and introduction of *ispD.* Each operon is controlled by the inducible *lac*-promoter and ends with *gfp*. Dotted lines indicate the positions of the successive cloning steps.

In case of feedback-inhibition of native IspF, this is expected to result in significantly lower up-regulated production of the building blocks IPP and DMAPP. This mechanism should then be confirmed by an additional operon variant that overexpresses *ispD* but not *ispF*. Here, *ispD* overexpression might result in a lower MEP pool, comparable to the complete operon. As a consequence, this operon should in potential alleviate the inhibition. Altogether, these operons allow an estimation of the relative and cumulative contribution of the targeted MEP-DOXP enzymes on the carbon flux.

The plasmids harboring the operon variants were transformed to *E.coli* JM109 cells and induced with IPTG. These cells were then analyzed in a series of preliminary assays to first test the operon design. Observed fluorescence upon ultraviolet light exposure of cell pellets first confirmed expression of the gene encoding GFPuv, hence indicating complete expression of all the operons (**S1**). Here we also noticed a difference in cell pellet volume which is most likely reflecting a variable metabolic burden by the operon variants. We subsequently performed an IPP-isomerase activity assay of two operons and the empty plasmid control, as this could give some insight in the expression of the different operons. We observed a 32-fold and 11-fold increase of IPP-isomerase activity upon induction of the *idi* and *idi-ispDF* operons respectively compared to the induced empty plasmid control that represents native expression of *idi* (**Fig. S2**). In the operon where *ispDF* was included, a notably lower increase in IDI activity was thus detected. This relative decrease is probably caused by lower *idi* expression and suggested a ribosome competition between the operon genes. Moreover, it is shown that induction of the operons results in expression of functional IDI regardless of the addition of multiple genes. To further study expression of all genes from the complete set of operons, protein profiles were analyzed on SDS-PAGE. The gels (**Fig. 3** and **Fig. S3**) unambiguously confirmed expression of all four genes upon induction.

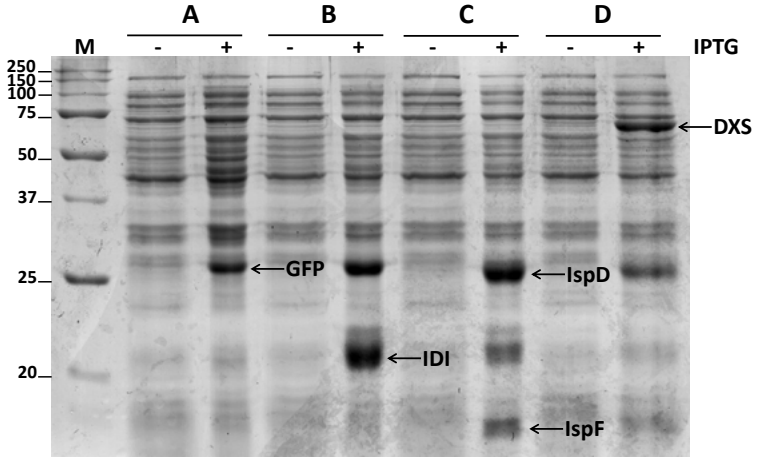


Figure 3. SDS PAGE analysis of cell-free extracts of *E.coli* with different plasmid-borne operon constructs, after IPTG (0.5 mM) induction. A: pGFPuv control **B**: *idi*-operon, **C**: *idi-ispDF*-operon, **D**: *idi-ispDF-dxs*-operon. –sign: non induced, +sign: induced at 0.5 mM IPTG. GFP; 26.8 kDa, IDI; 20.5 kDa, IspD; 25.7 kDa, IspF; 16.9 kDa, DXS 67.6 kDa.

Comparison of intensities, however, also suggests a differential gene expression between the operons. The intensity of the IDI protein for instance, gradually decreases as more genes are included in the operon. Here, the level of expression is in good agreement with the detected IPP-isomerase (IDI) activity. In the *idi-ispDF* operon, the *ispD* gene seems to result in the highest expression. And lastly, unequal gene expression is even more apparent from the *idi-ispDF-dxs* and *idi-dxs* operons (**Fig. 3D** and **Fig. S3**). Here the DXS protein forms the majority of the overexpressed products. Altogether this analysis suggests that the differential expression is determined by high translation of the bulkiest gene in the transcript.

Construction and characterization of a lycopene reporter plasmid

To establish controllable isoprenoid production, a tunable IPP and DMAPP pool is required. This not only enables broad applicability of the strain but also adjustment to a downstream pathway of interest. To enable such flexibility, the operons were placed under control of an inducible *lac*-promoter. Primarily to confirm increased MEP-DOXP flux, but also to detect its sensitivity to the *lac*-promoter, we needed a rapid and sensitive method to detect IPP and DMAPP levels. A direct quantification of IPP and DMAPP, however, is hampered by the short-life span due to conversion to FPP by endogenous farnesyl-diphosphate synthase (IspA). Conversion to an alternative stable end-product is thus required. Estimating IPP and DMAPP levels has previously been done by means of a reporter construct assembled from the carotenoid genes *crtE, crtB*, and *crtI* [212, 229-231]. This method enables a semi-quantitative determination of IPP and DMAPP levels by production of lycopene [229]. This compound is a red carotenoid pigment with an absorption optimum at 472 nm, which is readily quantified on a spectrophotometer. We have assembled a comparable inducible reporter plasmid from the *crtE, crtB* and *crtI* genes from the bacterium *Pantoea ananatis* in the compatible T7-promoter controlled pACYCduet vector. The newly obtained reporter plasmid was then transformed to JM109(DE3) cells to obtain strain WUR17.2. Induction of the reporter plasmid and the empty control plasmid was first done to test the functionality of the lycopene synthesis genes. The induced cells were pelleted and visually compared to check for red pigment development (**Fig. 4)**.

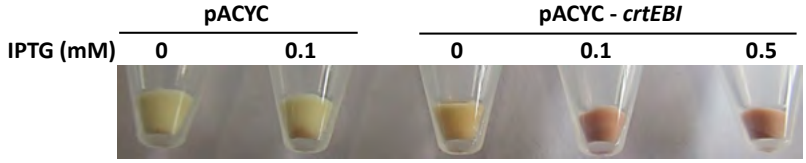


Figure 4. *E.coli* JM109(DE3) cell pellets showing lycopene production at various IPTG concentrations. JM109 (DE3) cells transformed with pACYCduet or pACYC–*crtEBI (*WUR17.2) were induced with 0, 0.1 and 0.5 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) and pelleted for lycopene visualization.

Here we established a functional reporter with only mild leaky expression and an intense color development without notable difference between 0.1 and 0.5 mM IPTG addition. It was therefore concluded that an induction >0.1 mM IPTG does not further increase the levels of lycopene and likely demarcates an optimal expression of the reporter or maximum depletion of the endogenous IPP and DMAPP pool.

Relative contributions of IDI, IspD, IspF and DXS to MEP-DOXP carbon flux.

Overproduction of IPP and DMAPP pools and the relative contribution of MEP-DOXP enzymes to carbon flux were determined by comparing the different operons. The plasmid-encoded operons were transformed to the previously generated WUR17.2 strain, and the obtained transformants were grown on LB supplemented with glycerol as carbon source. Preliminary data revealed variable growth performances in the presence of different operons. The successive addition of genes to the operons correlated with an increase in the specific growth rate (μ) for, reporter only, *idi, idi-ispDF, idi-ispDF-dxs* operons at 0.08, 0.011, 0.011, 0.016 minute^{-1}, respectively. This apparent improvement of growth rate caused by the addition of MEP-DOXP genes may suggest that overexpression of *crtE, crtB and crtI* may cause a growth limiting depletion of FPP. This effect was clearly restored during co-expression of MEP-DOXP genes which may have compensated for this depletion. After IPTG induction, the mutant strains were cultured for an additional 4 h. Total lycopene was extracted and measured along with the corresponding dry cell weights from three biological replicates as an additional preliminary test of operon functionality (**Fig. 5** and **Fig. S4**).

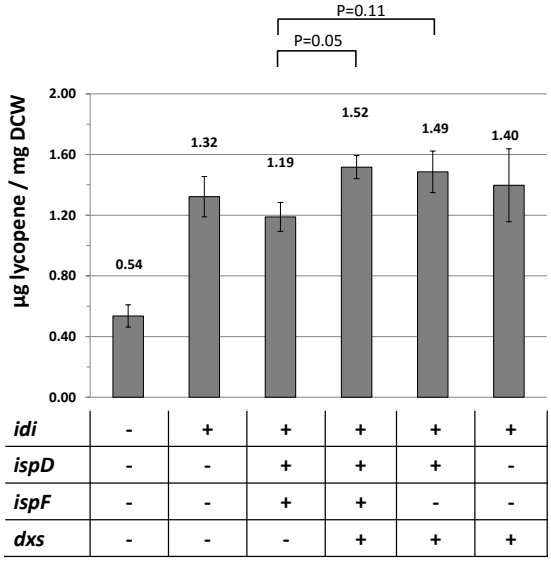


Figure 5. Average lycopene production per plasmid-encoded operon upon induction (0.1 mM IPTG). The relative lycopene production assay shows the cumulative contribution of the overexpressed MEP-DOXP genes to the carbon flux. Lane 1: reporter only, 2 *idi*, 3: *idi-ispDF*, 4*: idi-ispDF-dxs*, 5: *idi-ispD-dxs*, 6: *idi-dxs*

Here a 2.44-fold increase in lycopene by IDI alone was observed (from 0.54 to 1.32 µg lycopene /mg DCW) suggesting that the bottleneck of the MEP-DOXP flux is caused by the IPP/DMAPP isomerization [206]. Contrary to our expectations, overproduction of MEP-DOXP enzymes catalyzing earlier steps in the pathway (IspD, IspF and DXS) added to IDI did not significantly increase the production of lycopene.

Selection of EPOD dependent loci for operon integration

To generate stable strains with inducible up-regulations of IPP and DMAPP, the synthetic operons were integrated into the *E.coli* chromosome. We first exchanged the redundant *gfp* gene at the 3'-end of the plasmid operons for a selection marker. We used a kanamycin resistance gene (*kanR*) which is flanked with *lox*-sites (*lox66* and *lox71*) for its eventual replacement. These *lox* recognition sites allow for a Cre-recombinase dependent removal of the marker after integration. Recombination of the *lox66* with *lox71* sites results in formation of a desirable *lox72* scar sequence [233]. In contrast to an *frt* scar [234], the *lox72* scar tolerates a subsequent cycle of gene disruption or integration event without incidental recombination from a previously formed scar. This may become important when additional genome engineering of the here described strains is required. Downstream attachment of the *lox66-kanR-lox71* sequence yielded the plasmids pMS51, pMS52 and pMS53 (**Table 2**) which also includes the aforementioned operons *idi, idi-ispDF* and *idi-ispDF-dxs* respectively. These plasmids were then used as template for PCR-amplification of the integration cassettes. Amplifications were performed using primers elongated with ~50 bp extensions (homologous regions). These extensions were complementary to the sequences adjacent to the insertion-site at the target locus, in such a way that the orientation of the inserted operons is different from that of the neighboring chromosomal genes. Such positioning minimizes transcriptional read-through from upstream promoters.

To select integration sites which allow for IPTG induced transcription of the synthetic operons, we have used recently gained insights in the protein occupancy landscape of the *E.coli* chromosome [228]. Compelling evidence has revealed that packaging of the chromosome can significantly influence the local gene expression, meaning that the selected integration site of heterologous genes/operons has major consequences for their expression [223]. For selecting an appropriate, high expression insertion site, the key feature of this Extensive Protein Occupancy Domains (EPODs) analysis concerned the accessibility of the RNA Polymerase [228]. The EPOD analysis discriminates transcriptionally silent *ts*EPODs, probably corresponding to the core regions of the chromosome, and highly expressed *he*EPODs that have been proposed to be more accessible chromosomal loops. Based on the previously identified *ts*EPOD and *he*EPOD coordinates of the *E.coli* chromosome, we have selected potential domains as operon integration sites [228]. **Fig. 6** shows the plotted coordinates of EPODs in the selected segment of the ori-domain between nucleotides 4.000.000 and 4.500.000. The observation of two *he*EPOD regions between two *ts*EPODs from coordinates

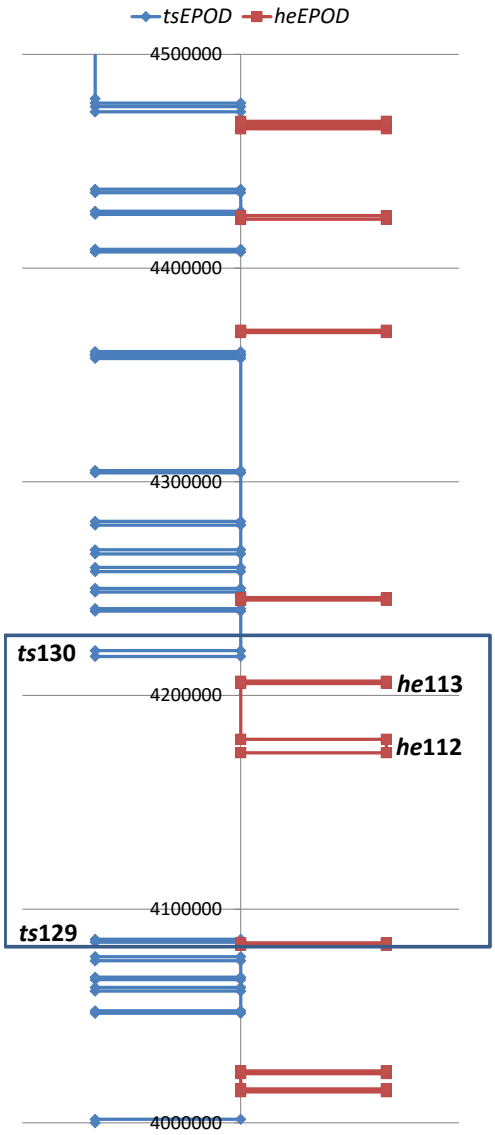


Figure 6. EPOD coordinates of a segment of the ori-macrodomain plotted on scale. Extensive protein occupancy domains between *E.coli* nucleotides 4.000.000 and 4.500.000 are shown. Within the box, two *ts*EPODs flanking two *he*EPODS are depicted that are hypothesized to form a putative subdomain. In blue: *ts*EPODs are plotted which supposedly anchor the genome subdomains to the centromere. In red, *he*EPODs. the regions localized in the subdomain loops that extend from the centromere.

4.084.552 to 4.220.964 highly suggested the presence of a relatively large loop. Here we also noticed that the *he*EPOD112 domain includes many essential genes as established by the KEIO collection [235] (e.g. the *rpl* operon encoding ribosomal proteins). In contrast, *ts*EPOD130 is exclusively composed of nonessential genes (**Fig. S5**). As these observations are in line with an expected higher accessibility of the *he*EPOD112 region and to a lower accessibility of *ts*EPOD130, we decided to integrate the synthetic operons within and between these two domains. The *ts*EPOD130 is expected to result in lower transcriptional levels and serves as a control. Additionally, a region outside an EPOD was selected between

these domains; this region (from here on referred to as 'EPOD-independent') was expected to provide a more strict control over gene expression and is included to compare the effects of the allocated EPOD regions.

A set of exclusion criteria assisting in the designation of appropriate integration sites was adopted from Bryant *et al*. [223]. To this end, promoter sequences, transcription factors, repeat sequences and possible nucleoid associated protein sequences (H-NS, Fis and IHF) were deliberately avoided (**Fig. S6**). This screen ultimately led to the selection of three loci between nonessential genes (**Fig. S5**). One locus was targeted inside *he*EPOD112 between the *rplK* and *rplA* genes. The EPOD-independent locus was selected between the genes *aceA* and *aceB*. And lastly, a locus was selected inside *ts*EPOD130, between the genes *iclR* and *arpA*. Despite several recombination attempts and screening rounds, however, integration at both the *ts*EPOD130 and *he*EPOD112 did not result in the isolation of the desired mutants. In contrast, the EPOD-independent *aceB/aceA* locus showed 71% integration efficiency for the *idi* operon from 84 screened colonies at the first recombination attempt. This sharp contrast led us to conclude that the selected EPOD-loci are highly unfavorable as target for integrations. This prompted us to continue integration of the *idi-ispDF* and *idi-ispDF-dxs* operons only at the EPOD-independent locus. For these constructs integration efficiencies of 20% and 6.6%, respectively, were observed.

Characterization of MEP-DOXP upregulation by genome integrated operons.

To estimate IPP and DMAPP pools from the integrated constructs we analyzed the mutants for their lycopene productions. The mutants were transformed with the reporter plasmid and induced with different amounts of IPTG while grown on glycerol or repressed by glucose. Glycerol serves as a readily used carbon source which does not induce catabolite repression like glucose. Moreover, as a convenient practical feature, glycerol has been reported to stimulate the heterologous production of carotenoids [213, 236]. The exact mechanism of the glycerol stimulation is unknown, but it probably relies on some control network that influences the glycerol-3P and pyruvate pools as precursor substrates of the MEP-DOXP pathway (**Fig. 1**) [213]. Comparison of lycopene color development first showed that glucose indeed establishes a very efficient repression, as no lycopene seems to be produced in these controls (**Fig. S7**). Although also without induction, the negative control with glycerol showed high levels of lycopene instead. Despite a stimulation of the isoprenoid production, glycerol thus inadvertently also promotes leaky expression of our constructs and reporter. Secondly, the initial screening indicates that the integrated single copy operons produce lycopene levels with comparable variability as their respective multicopy plasmid variants. It is therefore concluded that there is sufficient expression from the integrated constructs, and that the *aceB/aceA* insertion site provides a convenient locus for the inducible expression of synthetic gene cassettes.

Quantification of lycopene then confirmed that statistically significant variable levels were indeed produced by the strains with the different operons (**Fig. 7**). Here, the *idi-ispDF-dxs* operon, as was shown for the plasmid assay, also resulted in the highest production of lycopene. A main difference with the multi-copy plasmid constructs, however, is that all integrated single copy constructs led to higher up-regulations (**Fig. 7**, grey bars). The *idi-ispDF-dxs* mutant for instance produced 2.99 instead of 1.52 μg lycopene/mg DCW when induced with 0.01 mM IPTG. This is roughly a doubling of the initially determined level from the plasmid variant and an overall 5.34-fold increase compared to the wild-type control strain (WUR17.2). It must be emphasized here that for testing of the plasmid-borne constructs 0.1 mM IPTG was used, also in the presence of glycerol. For the integrated operons, however, this IPTG concentration led to a dramatic decrease in lycopene productions (**Fig. 7**, black bars). On the basis of these observations it is suggested that the operon copy number modulates the response to IPTG either directly by the operon, or indirectly due to the reporter. Overexpression of *ispDF* in addition to *idi* (**Fig. 7**, lanes 2 and 3) was shown to contribute a little to the MEP-DOXP flux, as reflected by slightly higher lycopene levels (p=0.06). This suggests that the previously presumed inhibition (**Fig. 5**, lane 3) is most likely an artefact of the excessive overexpression [207, 237].

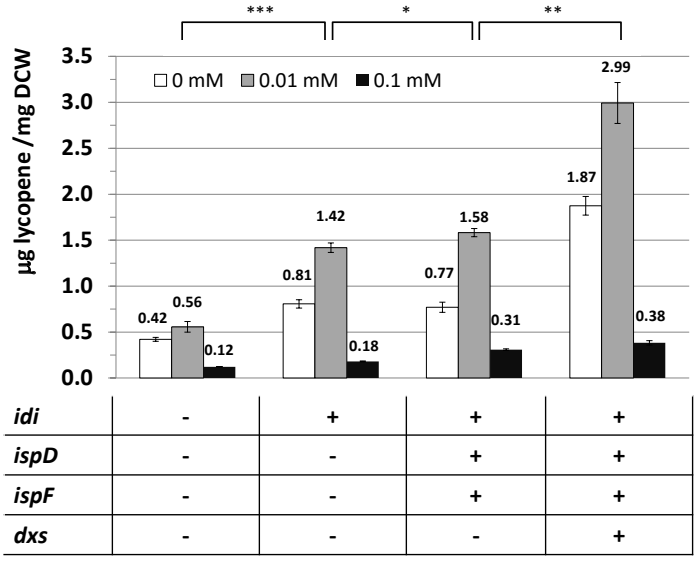


Figure 7. Lycopene production by the mutant strains with integrated MEP-DOXP operons after variable IPTG induction. White bars represent the non-induced cells. Grey bars represent 0.01 mM IPTG induced cells, and black bars represent 0.1 mM IPTG induced cells. Negative control (lane 1) represents WUR17.2 strain without operon integration. Statistical significance of the 0.01 mM inductions; * p<0.1; ** p<0.01; *** p<0.001

Like with all integrated strains, a comparable IPTG variability was observed within the control strain (WUR17.2) (**Fig. 7**, lane 1). Because of this conformity, it is more likely that the IPTG dependent variation is mainly due to induction of the lycopene reporter pathway rather than

to induction of the tested operons. To test this hypothesis, we analyzed the protein profiles of the induced cells on SDS-PAGE. The gel indeed showed a variation of expressed lycopene reporter genes *crtEBI* that corresponds with the lycopene levels (**Fig. 8**). The CrtE, CrtB and CrtI proteins which expression was controlled by T7 RNA-polymerase are abundantly present at 0.01 mM IPTG, but not visible at 0 and 0.1 mM IPTG induction. Moreover, the excessive production of the MEP-DOXP proteins, like shown for the plasmid operon variants (**Fig. 3**), is no longer observed. Because of this, it remains unclear whether the IPTG variation was really due to expression of the lycopene reporter pathway or whether the MEP-DOXP genes show a comparable variation in expression albeit below the detection limit.

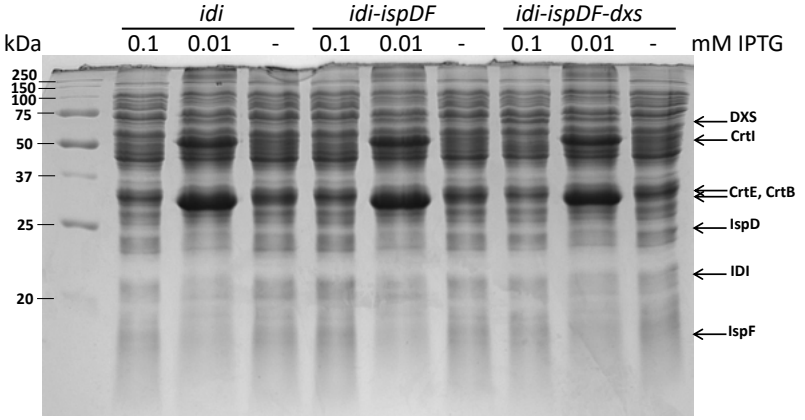


Figure 8. SDS PAGE of the differentially induced mutants. An expression optimum is shown for the lycopene reporter plasmid by the excessive presence of the CrtE: 32.6 kDa; CrtB: 34.5 kDa and CrtI: 55.0 kDa proteins. Operon gene products IDI: 20.51 kDa; IspF: 16.9 kDa; IspD: 25.7 kDa and DXS: 67.6 kDa are not distinguishable.

Discussion

The main objective of this study was to obtain a platform *Escherichia coli* strain with a stable upregulation of the MEP-DOXP pathway for isoprenoid productions. This was done by assembly of a series of synthetic operons to overexpress the crucial MEP-DOXP (*idi*, *ispD*, *ispF* and *dxs*) genes. In the course of the operon characterization we additionally studied whether these rate limiting or rate-controlling steps demonstrate a cumulative enhancing or silencing effect. This influence on carbon flux was estimated by quantification of synthesized lycopene as a representation of the available pool of IPP and DMAPP. As shown by earlier studies, we confirm that the isopentenyl-diphosphate isomerase forms a high rate-limiting step [209, 238]. High overexpression of the *idi* gene caused a 2.54-fold increase in lycopene production compared to native *idi* expression from the wild-type strain. Because of this, the *idi* gene was compared for IPP and DMAPP production upon co-expression with *ispD* and *ispF* or in a complete operon *idi-ispDF-dxs*. Here the complete operon resulted in the highest levels of lycopene of the plasmid-encoded operons at an average 2.8-fold increase.

Interestingly, in the plasmid-encoded operons, no significant differences were observed by the co-expression of *ispD, ispF or dxs* gene in any of the combinations.

To generate strains with a stable upregulation of IPP and DMAPP, we embarked on the integration of the characterized operons (*idi*, *idi-ispDF* and *idi-ispDF-dxs*) in the *E.coli* chromosome. Three targeted loci for integrations were chosen such to allow variable expression dependent on EPOD types: highly expressed, EPOD independent and transcriptionally silent. Here we observed highly efficient integrations in the EPOD independent locus as opposed to the *he*EPOD and *ts*EPOD loci that did not yield any of the desired mutant colonies. The molecular basis of these failed recombinations remains unclear, but may reflect the compact folding of the chromatin in the *ts*EPODs and the highly dense protein occupancy of EPODs in general. Because of this experimental result, we only compared the operon variants integrated from the EPOD-independent locus. Interestingly, expression from these chromosomal 'duplication variants' resulted in optimal levels of IPP and DMAPP with a fold-change of 1.07, 1.32 and 1.97 (µg lycopene/mg DCW) for the *idi*, *idi-ispDF* and *idi-ispDF-dxs* operons respectively compared to the plasmid-encoded operons. This assay confirmed that excessive expression of the targeted genes was deleterious to isoprenoid production [207]. Here the duplication of *dxs* led to the highest additional fold-change and indicates that the excessive overexpression of this enzyme was probably the cause of the low isoprenoid flux in the plasmid-assays. Likely because of this the complete *idi-ispDF-dxs* operon thus resulted in the highest detected upregulation at an 5.34-fold increase compared to a wild-type MEP-DOXP control. In addition to the single copy number improvements, a presumed inhibitory effect of the *idi-ispDF* operon in the plasmid-borne assay was no longer observed in the integrated mutant. Although a significant improvement was observed, no high upregulation was obtained. This finding led us to conclude that *ispDF* overexpression is only beneficial to the flux during co-expression with *dxs.* Despite the observed improvements it remains unclear whether duplication variants resulted in higher lycopene levels due to lower but more efficient expression levels, or because of a more beneficial expression of the reporter. Either way it seemed that much profit can be gained by IPTG calibration. In this assay we therefore also tested the influence of variable induction levels to further explore the system. Here we found that a high concentration of IPTG (i.e. 0.1 mM) severely diminishes the production of lycopene. These low levels were not observed for the plasmid-encoded operons at the same concentration of IPTG. An SDS-PAGE analysis showed that the highest levels of lycopene corresponded with high expression of the reporter enzymes and levels below detection with high IPTG concentration. Because high expression levels of the reporter did not show any correlation with lower lycopene levels, it can be concluded that excessive expression of *crtEBI* was not deleterious to the isoprenoid flux.

Regardless of the observed variabilities by IPTG, we demonstrate the construction of a series of platform strains, enabling controllable levels of isoprenoid production. Our data

show that the available IPP and DMAPP pool which is the main determining factor of the production level of isoprenoids can be further optimized. In addition to this, our data also suggests that more efficient isoprenoid production can be obtained by balancing the expression between upstream and downstream pathways which in the first place can be achieved by IPTG calibrations, but possibly even more by tuning individual expression levels of the downstream pathway.

Materials and Methods

Construction of MEP-DOXP operons

The *idi*, *ispDF*, and *dxs* genes were PCR-amplified using *E.coli* MG1655 genomic DNA as template. Cells from a 7-ml overnight culture were harvested and the chromosomal DNA was isolated according to the manufacturer's procedures "Qiagen DNeasy Blood & Tissue Kit".

PCR amplifications were performed according to the standard protocol of Phusion DNA polymerase (Thermo Fisher Scientific). Each PCR reaction contained 0.2 mM forward and reverse primer (**Table 1**), 0.2 mM dNTPs, 1x Phusion buffer, and 1U Phusion polymerase (Thermo Fisher Scientific). Each PCR cycle initiated with a 5 min incubation at 95°C. This was followed by 30-35 cycles consisting each of: 1) 30 s at 98°C, 2) 30 s annealing at variable temperatures, and 3) variable elongation at 72°C at variable duration. The cycles were finalized by an additional extension of 5 min at 72°C. The resulting DNA fragments were analyzed on a 1% agarose gel. All PCR fragments were first ligated into pCR2.1 topo vector by TOPO TA cloning according to the manufacturer's procedures. This ligation enabled sequence analysis of the isolated fragments (BaseClear and GATC). MEP-DOXP genes were cloned in consecutive steps according to standard molecular biological procedures. Restriction enzymes, ligases and gel extraction kit were provided by Thermo Fisher Scientific. The *idi* insert and pGFPuv vector (Clontech) were double digested with SphI and PstI. The digested insert and vector were excised from gel and extracted by a DNA gel extraction kit. The insert was subsequently ligated into linearized pGFPuv by T4 DNA ligase. The resulting construct pGFPuv-*idi* was transformed to competent *E.coli* DH5α cells. The subsequent fragment *ispDF*, was amplified at an annealing temperature of 58°C and an extension time of 3 min. The insert was screened for mutations in pCR2.1topo and after confirmation cloned into pGFPuv-*idi*. Cloning yielded pGFPuv-*idi-ispDF* using the XbaI and PstI restriction sites. The subsequent fragment *dxs* was amplified at an annealing temperature of 60°C and an extension time of 4 min. The screened insert was cloned into pGFPuv-*idi-ispDF* to yield the final construct pGFPuv-*idi-ispDF-dxs*. For this final cloning the XbaI and the KpnI restriction sites were used. Two variants of the pGFPuv-*idi-ispDF-dxs* construct were made by removal of the *ispF* gene. To compose the first plasmid, the *ispDF* genes were both removed by

Table 1. Oligonucleotide sequences. Restriction sites are underlined. For MEP-DOXP operon integrations, homologous regions are underlined.

Primer	TG	Sequence (5'→ 3')	RS	Primer name
	Cloning primers			
BG3606	*idi*	GGCATGCCATGCAAACGGAACACG	SphI	For-a1.1 SphI_idi
BG3599	*idi*	GCTGCAGTTATTTAAGCTGGGTAAATGCA	PstI	Rev-a1.1 idi_PstI
BG3600	*ispDF*	GCTGCAGAGGAGATATACATATGGCAACCACTCATTTGG	PstI	For-a1.2 Pst_RBS_IspDF
BG3601	*ispDF*	GTCTAGATCATTTTGTTGCCTTAATGAGTAGCG	XbaI	Rev-a1.2 IspDF_XbaI
BG3602	*dxs*	GTCTAGAGGAGATATACTGATGAGTTTTGATATTGCCAAATAC	XbaI	For-a1.3XbaI_lac_rbs_dxs
BG3603	*dxs*	GCGGTACCTTATGCCAGCCAGGCC	KpnI	Rev_a1.3dxs_KpnI
BG3600	*ispD*	GCTGCAGAGGAGATATACATATGGCAACCACTCATTTGG	PstI	For-α1.2Pst_RBS_ispDF
BG4555	*ispD*	CGCTCTAGATTACTGATGGATGGTTCGGGTG	XbaI	Rev-ispD-Xba
BG3899	*crtE*	GAACGAATTCAGCCCGAATGACGGTCTGC	EcoRI	For_EcoRI_crtE
BG3900	*crtE*	GAATCTTAAGGCGCGACCAGTTCCTGAG	AflII	Rev_AflII_crtE
BG3901	*crtBI*	GCTGAGATCTGATGAAACCAACTACGGTAATTGG	BglII	For_BglII_crtBI
BG3902	*crtBI*	CTTACTCGAGAAAGACATGGCGCTAGAG	XhoI	Rev_XhoI_crtBI
BG4429	*BB*	GACGCGTACGGTGTCTTTTTTACCTGTTTGACC	BsiWI	For-BsiWI-PI-BB
BG4430	*BB*	GACGCTTAAGCTACCTCTGGTGAAGGAGTTGG	AflII	Rev-AflII-PII-BB
BG4476	*kanR*	TTCCGTCAGCCAGTTTAGTC	-	Rev-Kan1-BB
BG5004	*aceB*	TACATCGAAGCGTGGATCTCTGG	-	For EI–scr UP
BG5005	*aceA*	ATCTGATCGGCACGACGGAAGG	-	Rev EI–scr DOWN
	Integration primers			
BG4883	MEP-DOXP operons	CTCGCTCCATCGAAGGTACTGCACGTTCCATGGGCCTGGTAGTGGAAGATTAAGCTACCTCTGGT-GAAGGAGTTG	-	For-H1A-MEP-DOXP
BG4884	MEP-DOXP operons	GCATCAACTTTCTCGCGGATAACACGCATGCGCTTGGTCAGTTTAGCCATTTCTTAGTCCTCCGGCAGTGAGC-GCAACGCAATT	-	Rev-H2A-MEP-DOXP-
BG4885	MEP-DOXP operons	CTGACCCTGCCAGGCTACCGCCTGTTAGCGTAAACCACCACATAACTATG GCTACCTCTGGTGAAGGAGTTG	-	For-H1B-P2
BG4886	MEP-DOXP operons	CTGTAATTCTTCAATTTGTTGTGTACGGGTTTTCATGTGCAGATGCTCGGCAGTGAGCGCAACGCAATT	-	Rev-H2B-Lac O3
BG4887	MEP-DOXP operons	CGCTCAATCTCATAATGCAGCCGTAAAAGTTTCGGTGGAATGAGATCTTGGCAGTGAGCGCAACGCAATT	-	For-H1C-LacO3
BG4888	MEP-DOXP operons	CTCAGCATTATTTTTAAACATCAAACCACTTAATTATTAACAAAATCGCGCTACCTCTGGTGAAGGAGTTG	-	Rev- H2C-P2

Abbreviations: RS: restriction sites; TG: Target gene; BB: 'backbone' represents the *lox71-kanR-lox66* knock-in cassette.

digestion with PstI and XbaI and the single *ispD* gene put in place to yield pMS31. The *ispD* insert was amplified at an annealing temperature of 58°C and an extension time of 30 s. The PCR fragments were cloned into the pCR2.1 cloning vector and screened for mutations. Both insert and pGFPuv-*idi-ispDF-dxs* were double digested with PstI and XbaI. The resulting insert was ligated into the linearized vector to yield pGFPuv-*idi-ispD-dxs*. To yield the second variant, the linearized plasmid after removal of *ispDF* was treated with klenow fragment (ThermoFisher Scientific). The resulting blunt-ends were self-ligated to yield the pGFPuv-*idi-dxs* (pMS33) construct.

Transformation and screening

All transformations were done by heat shock as following. Frozen chemically competent *E.coli* cells were thawed 10 min on ice. Either DH5α cells, made competent by a calcium rubidium treatment, or commercial TOP10 cells were used (Invitrogen). Ligation mixtures (2 µl) containing 20 ng DNA or 1 ng plasmid DNA were added to the cells and gently mixed. The cells were incubated on ice for 30 min. The cells were heat shocked by an incubation of 45 s at 42°C and then immediately cooled on ice to incubate 5 min for recovery. After recovery, 250 µl SOC medium was added at ambient temperature and the cells were incubated 1 h at 37°C while aerated at 170 rpm. Finally, the cells were plated on lysogeny broth (LB)-agar supplemented with appropriate antibiotics to select for transformants. The resulting colonies were screened for inserts by colony PCR and the plasmids subsequently sent for sequencing. The resulting operon constructs were then re-transformed to competent *E.coli* JM109 for expression analysis.

Assembly of a lycopene reporter construct

The *crtE* gene was PCR-amplified using *Pantoea ananatis* genomic DNA as template. Purified DNA isolate was provided by DSMZ GmbH (DSM30080). PCR amplifications were performed as described previously. Amplification of *crtE* was performed at an annealing temperature of 55°C and an extension time of 35 s. The fragment was first cloned into the pCR2.1topo vector and inserts selected by blue/white screening on LB-agar Amp^{100} X-gal plates. The T7-promoter controlled pACYCduet vector (Novagen) and the *crtE* insert were digested with EcoRI and AflII (BspTI). The vector was treated with antarctic phosphatase (NEB) to prevent self-circularization. The *crtE* insert was ligated into linearized pACYCduet by T4 ligase. The ligation mixture (2µl) was subsequently transformed to *E.coli* TOP10 cells according to the heat shock protocol. The bicistronic *crtB* and *crtI* genes were PCR-amplified as a single fragment using *Pantoea ananatis* genomic DNA as template. PCR amplifications were performed as described previously. The *crtBI* fragment was amplified at an annealing

temperature of 58°C and an extension time of 80 s. The fragment was cloned into the pCR2.1topo vector for screening. The screened insert was digested with BglII and XhoI as well as the previously constructed pACYC-*crtE* plasmid. The linearized plasmid was treated with antarctic phosphatase (NEB) to prevent self-circularization. Ligation by T4 ligase yielded the completed lycopene reporter pACYC-*crtEBI* (pMS17) (**Table 2**). After confirmation of correct mutation-free inserts, the reporter plasmid was transformed to *E.coli* JM109(DE3) cells for further processing to yield strain WUR17.2.

Table 2. *Escherichia coli* strains and plasmids used in this study

Parent strain	Annotation	Inserts	Plasmid(s)
JM109	WUR06	(-)	pGFPuv ctrl
JM109	WUR04.2	*idi*	pMS03
JM109	WUR08.2	*idi–ispDF*	pMS08
JM109	WUR11.2	*idi–ispDF-dxs*	pMS11
DH5α	WUR17	*crtEBI*	pMS17
JM109	WUR31	*idi–ispD-dxs*	pMS31
JM109	WUR33	*idi–dxs*	pMS33
JM109(DE3)	WUR17.2	*crtE-B-I*	pMS17
JM109(DE3)	WUR17.3	*crtE-B-I* (-)	pMS17 and pMS01
JM109(DE3)	WUR17.4	*crtE-B-I (idi)*	pMS17 and pMS03
JM109(DE3)	WUR17.5	*crtE-B-I (idi-ispDF)*	pMS17 and pMS08
JM109(DE3)	WUR17.6	*crtE-B-I (idi-ispDF-dxs)*	pMS17 and pMS11
JM109(DE3)	WUR31.3	*crtE-B-I (idi-ispD-dxs)*	pMS17 and pMS31
JM109(DE3)	WUR33.3	*crtE-B-I (idi-dxs)*	pMS17 and pMS33
TOP10	WUR46	BsiWI - *lox71-kanR–lox66* - AflII	pMS46
TOP10	WUR51	*idi–lox71-kanR-lox66*	pMS51
TOP10	WUR52	*idi–ispDF - lox71-kanR-lox66*	pMS52
TOP10	WUR53	*idi–ispDF–dxs – lox71-kanR-lox66*	pMS53
JM109(DE3)	WUR61	lambda RED *(exo, bet, gam)*	pKD46
		Genome modifications	
JM109(DE3)	WUR63	*idi - lox71-kanR-lox66 EI*	-
JM109(DE3)	WUR70	*idi-ispDF - lox71-kanR-lox66 EI*	-
JM109(DE3)	WUR71	*Idi-ispDF-dxs- lox71-kanR-lox66 EI*	-
JM109(DE3)	WUR89	*idi EI*	-
JM109(DE3)	WUR90	*idi-ispDF EI*	-
JM109(DE3)	WUR91	*idi-ispDF-dxs EI*	-
JM109(DE3)	WUR92	*idi EI*	pMS17
JM109(DE3)	WUR93	*idi-ispDF EI*	pMS17
JM109(DE3)	WUR94	*idi-ispDF-dxs EI*	pMS17

Abbreviations: EI: EPOD Independent

Strains and induction assay

The newly obtained *E.coli* JM109(DE3) strain carrying the lycopene reporter construct (WUR17.2) was used to analyze the various plasmid-borne MEP-DOXP operons. This strain was transformed with constructs: pMS01, pMS03, pMS08, pMS11, pMS31 and pMS33 to

yield WUR17.3, WUR17.4, WUR17.5, WUR17.6, WUR31.3 and WUR33.3 by electrophoresis respectively. LB-agar plates supplemented with 34µg/ml chloramphenicol (Cm^{34}) and 100µg/ml ampicillin (Amp^{100}) were inoculated from glycerol stocks. Antibiotics were used as selection marker to retain both the reporter pMS17 and MEP-DOXP operons, respectively. The plates were incubated overnight at 37°C. Ten ml LB Cm^{34} Amp^{100} was supplemented with 10 mM $MgSO_4$ and 0.4% glycerol and inoculated with a fresh single colony to start a pre-culture. The pre-cultures were incubated overnight at 37°C and stirring at 180 rpm. Three biological replicates per strain were inoculated in 25 ml LB media with 1/50 volume pre-culture. The cells were grown to logarithmic phase by incubating the culture at 37°C and stirring at 180 rpm. IPTG was added to the cultures to a final concentration of 0.1 mM or 0.01 mM at an OD_{600} of 0.55. The cultures were than incubated for an additional 4 h. Samples were cooled on ice for 15 min. Two 10-ml aliquots were transferred consecutively to pre-cooled 15-ml Greiner tubes. One aliquot was destined for lycopene extraction and 1 aliquot was destined for dry cell weight (DCM) determination.

Sample preparation and SDS-PAGE

To analyze protein profiles, pelleted cells from 4 ml culture were resuspended in 1 ml cold lysis buffer (Tris-HCl 50 mM, pH=8, 1 mM EDTA, 150 mM NaCl, 1 mM DTT and 0.1 mM PMSF) The cells were lysed by four consecutive sonication rounds (20 s pulse and 30 s on ice). To obtain cleared lysates, the crude extracts were centrifuged 10 min at 10,000 x*g* at 4°C. A 12% or 13.5% running gel and a 4% stacking gel were made to visualize protein profiles. Per sample 15 µg protein was applied and a "BioRad Precision plus Protein standard" was included as a reference. To visualize the proteins, the gel was stained overnight with coomassie brilliant blue according to standard procedures of Sambrook and Maniatis [239].

IDI activity assay

Isopentenyl-diphosphate isomerase (IDI) activity from cleared lysate was determined by measuring the formation of DMAPP according to Satterwhite *et al.* [240]. The reaction mixture contained: 10 mM Tris-HCl, 0.1 mM DTT, 2 mM $MgCl_2$, 0.04 mM ^{14}C-isopentenyl diphosphate (10 Ci/mol) (Perkin Elmer). The reaction was initiated by addition of 3 µg cleared lysate protein to 50 µl reaction mixture. The reaction was incubated for 10 min at 37°C and stopped by the addition of 200 µl HCl/methanol (3/1 v/v) followed by 0.5 ml MQ. Samples were vortexed, and incubated 10 min at 37°C. This incubation was followed by the addition of 20 µl saturated NaCl solution. Acidified DMAPP was extracted by addition of 10 ml toluene, vortexed, centrifuged (5 min, 1,000 x*g*), and the toluene phase removed twice. The extracts were pooled, $MgSO_4$ added until complete desiccation, and 1 ml from the toluene and the water phase transferred to scintillation vials. Radioactivity was measured in both aqueous and organic phases by scintillation counting to quantify the percentage of substrate conversion.

Lycopene quantification

Aliquots (10 ml) from biological triplicates were centrifuged at 4,700 x*g* for 10 min. The cell pellet was resuspended in 1 ml MQ, and transferred to a clean 2-ml screw-cap vial and centrifuged (12,000 x*g*, 5 min). The resulting cell pellet was suspended in 500 µl acetone and incubated 30 min at 55°C to promote extraction and centrifuged again (20 min, 16,000 x*g*, 4°C). The supernatant containing lycopene was transferred to a clean screw-cap vial. The extracts were kept in the dark to prevent light induced break-down of lycopene. This extraction procedure was repeated twice and the three lycopene extracts pooled to a final volume of 1.5 ml. The extract was centrifuged additionally at 16,000 x*g* for 2 min to pellet incidental cell-debris. A 250-µl sample was removed and diluted with 750 µl acetone from which the absorbance was measured at 472 nm in an acetone-resistant cuvette (Brand).The lycopene concentration was determined by interpolation of the measured absorbance to a standard calibration curve of lycopene (Sigma-Aldrich). The calibration curve of lycopene dissolved in acetone resulted in the following equation (y=0.2254x-0.0077) established at an R^2 =0.9998. The total lycopene concentration was normalized to the dry cell weights (DCW) per 10 ml induced culture. For DCW determinations, a 10-ml culture aliquot was centrifuged (4,700 x*g*, 2 min), and the pellet was resuspended in 1 ml MQ and transferred to a pre-weighed aluminum dish. The cells were dried for 24 h at 105°C and weighed. The lycopene production was expressed as "µg lycopene/mg DCW". All values are derived from biological triplicates.

Chromosome integration of MEP-DOXP operons

A *lox71-kanR-lox66* gene cassette as a selection marker was cloned to pCR2.1 topo by Topo-TA cloning to yield pMS46. The purpose of the cloning was to extend the cassette with an N-terminal BsiWI and a C-terminal AflII restriction site. These sites allow the insertion of the *lox71-kanR-lox66* cassette downstream of the previously constructed MEP-DOXP operons. The *lox71-KanR-lox66* insert was amplified with Phusion polymerase. The PCR reaction contained: 1 µl *lox71-kanR-lox66* template, 0.2 uM forward primer (BG4429), 02 µM reverse primer (BG4430), 0.2 mM dNTPs, 1x Phusion buffer, and 1U Phusion polymerase. The PCR cycle initiated with a 5 min incubation at 98°C. This was followed by 30 cycles consisting each of: 1) 30 s denaturation at 98°C, 2) 30 s annealing at 52°C, and 3) 40 s elongation at 72°C. The cycles were finalized by an additional extension of 5 min at 72°C. The *lox71-kanR-lox66* PCR fragment was cloned in pCR2.1 topo vector and transformed to *E.coli* TOP10 cells according to the manufacturer's procedures. Correct insertion was verified by sequence analysis. The main purpose of the following cloning procedure was the exchange of the *gfp* gene from the previously cloned constructs pMS03, pMS08, pMS11 with the *lox71-kanR-lox66* selection marker. A BsiWI compatible Acc65I and BspTI (AflII) double digestion on 2 µg pMS03, pMS08 and pMS11 plasmid DNA was performed to generate linear vector.

To generate a *lox71-kanR-lox66* insert, 4 µg pMS46 was similarly digested. Digests were applied on 1% agarose gel and separated for 30 min at 100V. Correct fragments were excised from gel and purified on a DNA clean and concentrator-5 column (ZYMO Research). The *lox71-kanR-lox66* inserts were ligated to linearized vectors to generate pMS51, pMS52 and pMS53 (**Table 2**).

To identify appropriate sites for integration, *he*EPOD112, *ts*EPOD130 and the region between these EPODs 'EPOD independent' were screened for *cis*-regulatory elements. Promoter sequences, transcription factors and repeat sequences were identified by means of the regulonDB database (http://regulondb.ccg.unam.mx/). We used (anti)sense consensus sequences to screen for Histone-like Nucleoid Structural proteins (H-NS), TCGWTWAATT [241, 242], Factor for Inversion Stimulation (Fis) GNNNRNNWNNYNNNC[243], Integration host-factor (IHF), WATCAANNNNTTR[244, 245], sites. Sequences with less than 20, 10 and 0% deviation from the consensus sequences respectively depending on their abundance were deliberately avoided for integration. Interestingly, we observed a relatively high abundance of possible H-NS sites within *ts*EPOD130 (19 sites / 2717 nt's) comparison to the *he*EPOD (4 sites / 6317 nt's). This comes down to an average H-NS-site frequency of 1 per 1579 nt's in *he*EPOD112 and 1 site per 143 nt's in *ts*EPOD130. This 11-fold difference strongly supports the hypothesis of a structural role for *ts*EPOD130 and the presence of a putative subdomain.

Homologous end extension PCR was performed by Phusion polymerase PCR (ThermoFisher Scientific). The following primers were used to extend the operons with homologous ends (**Table 1)**. For *he*EPOD locus insertion BG4883 and BG4884 primers were used and annealed at 55°C. For the EPOD-independent locus insertion BG4885 and BG4886 primers were used and annealed at 56°C. For the *ts*EPOD locus insertion BG4887 and BG4888 primers were used and annealed at 59°C. A standard PCR protocol was performed to amplify the knock-in cassettes for the *idi*, *idi-ispDF* and *idi-ispDF-dxs* operons using pMS51, pMS52 and pMS53 as template DNA, respectively. The resulting PCR-fragments were purified on a clean and concentrator 5 column (ZYMO research).

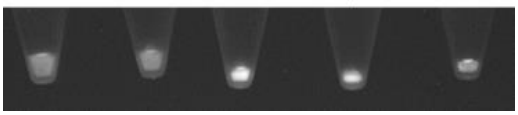
The integration and clearing of the selection marker was performed as described previously in the gene disruption strategy by Datsenko and Wanner [234]. Briefly, competent JM109(DE3) cells with induced pKD46 for phage λ Red-recombinase (WUR61) were electroporated with 375 ng purified cassette DNA at 2500V, 200Ω, 25 µF and 5 ms. After 1.5 h recovery on SOC medium at 37°C, cells were plated on LB-agar supplemented with 50 µg/ml kanamycin. Mutant colonies were screened by colony PCR using primers BG5004 and BG5005 complementary to the *aceB* and *aceA* gene respectively at the EPOD-independent locus according to standard PCR conditions. Positive clones were annotated as WUR63, WUR70 and WUR71 for the *idi, idi-ispDF* and *idi-ispDF-dxs* operons in the EPOD-independent region, respectively. Correct mutation-free insertions were confirmed

by sequence analysis of the amplified knock-in region with primers BG4476 and BG4886. Novel strains were transformed with pJW168 harboring the Cre-recombinase to remove the kanamycin resistance gene leaving a *lox72* scar sequence. The cells were spread on LB-agar plates supplemented with 100 μg/ml ampicillin and 0.5 mM IPTG to induce Cre-recombinase and grown at 30°C. To finalize the recombination, purification of the cells and clearance of the plasmids, the colonies were re-streaked on LB-agar plates without additives and kept overnight at 37°C to grow. The novel strains were annotated as WUR89, WUR90 and WUR91 for the integrated operons *idi, idi-ispDF* and *idi-ispDF-dxs,* respectively. The mutant strains were transformed with the lycopene reporter construct pMS17 to yield strains WUR92, WUR93 and WUR94.

Acknowledgements

This project was carried out within the BE-Basic R&D Program, which was granted a FES (Fonds Economische Structuurversterking) subsidy from the Dutch Ministry of Economic affairs, agriculture and innovation (EL&I).

Supplemental Information



idi	-	+	+	+	+
ispD	-	+	+	+	-
ispF	-	+	-	+	-
dxs	-	-	+	+	+

Figure S1. GFP fluorescence of cell pellets of upon ultraviolet light exposure GFPuv fluorescence of harvested cultures (10 ml) confirms complete transcription of the entire operons. A steady decrease in the pellet volume is observed by the successive addition of MEP-DOXP genes. Negative control contains the unmodified pGFPuv plasmid.

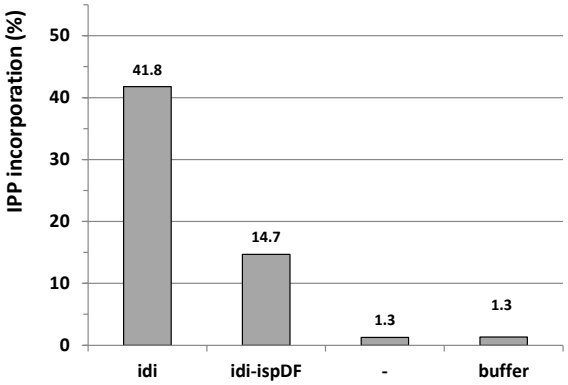
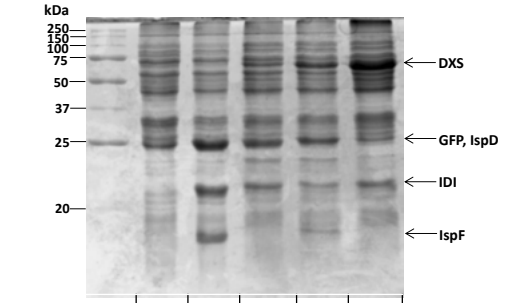


Figure S2. IDI activity assay of cell-free extracts of *E.coli* with different operon constructs after IPTG (0.1 mM) induction Assays show an efficient upregulation of IDI activity compared to wild type (-). A decrease in activity is observed when *IspDF* genes are included in the operon. Negative controls: (-) contains unmodified pGFPuv plasmid only; (buffer) control without cell-free extract.



idi	-	+	+	+	+
ispD	-	+	+	+	-
ispF	-	+	-	+	-
dxs	-	-	+	+	+

Figure S3. SDS-PAGE of cell-free extracts of *E.coli* with different operon constructs, after IPTG (0.1 mM) induction. Lane 1: pGFPuv control; 2: *idi-ispDF*-operon; 3: *idi-ispD-dxs*-operon; 4: *idi-ispDF-dxs*, 5: *idi-dxs*. Operon gene products GFP: 26.8 kDa; IDI: 20.51 kDa; IspD: 25.7 kDa; IspF: 16.9 kDa; DXS 67.6 kDa.

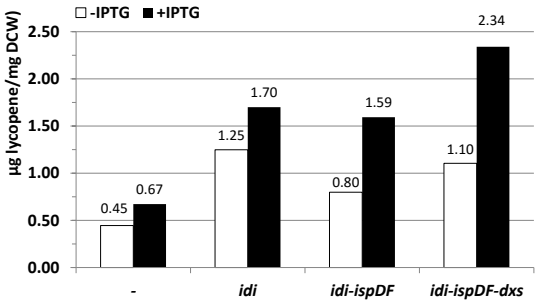


Figure S4. Relative lycopene production comparison per plasmid-encoded operon Induced vs. non induced lycopene yields.

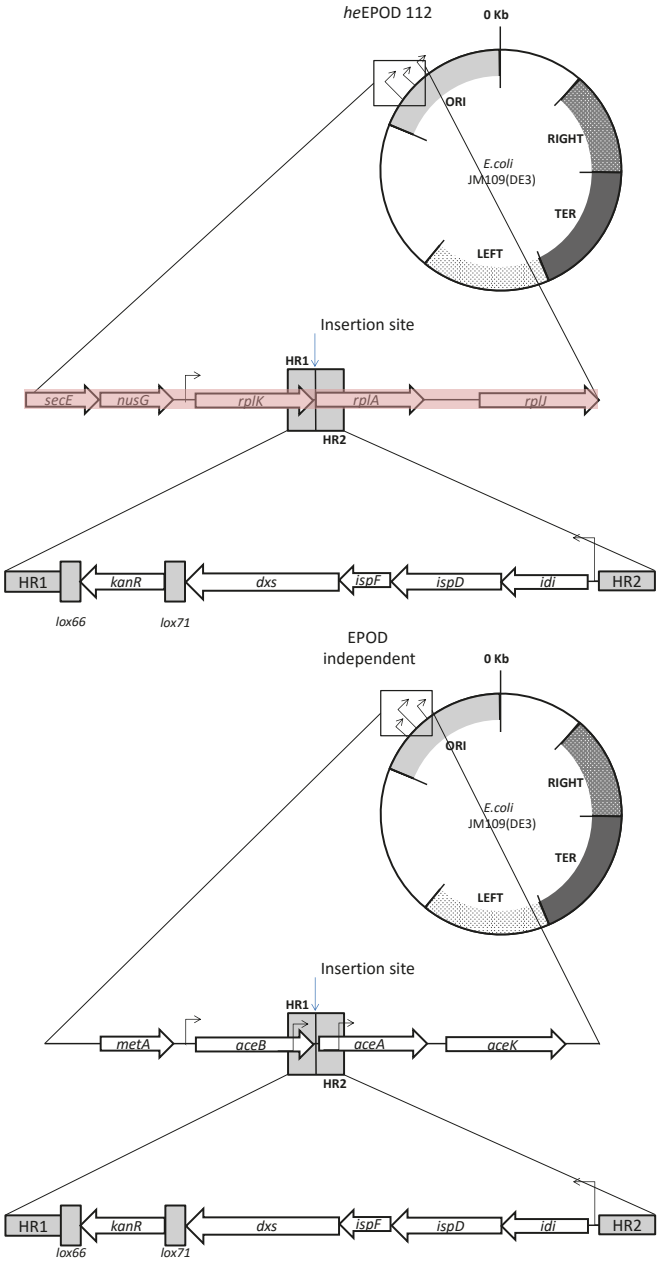


Figure S5. Integration loci according to identified EPOD regions In red: *he*EPOD112, In blanc: EPOD independent, In blue *ts*EPOD130. For all three chosen loci the integration strategies are depicted for the complete operon. For the *he*EPOD112 region, a locus between the *rplK* and *rplA* genes encoding nonessential: 50S ribosomal subunit protein L1 and L11 respectively was selected for integration. For the EPOD independent region, a locus between the *aceB* and *aceA* genes encoding nonessential malate synthase A and isocitrate lyase respectively was selected for integration. For the *ts*EPOD130 region, a locus between the nonessential *arpA* and *iclR* genes encoding ankyrin repeat protein and DNA-binding transcriptional repressor respectively was selected for integration. (Adapted from Bryant *et al.*[223]).

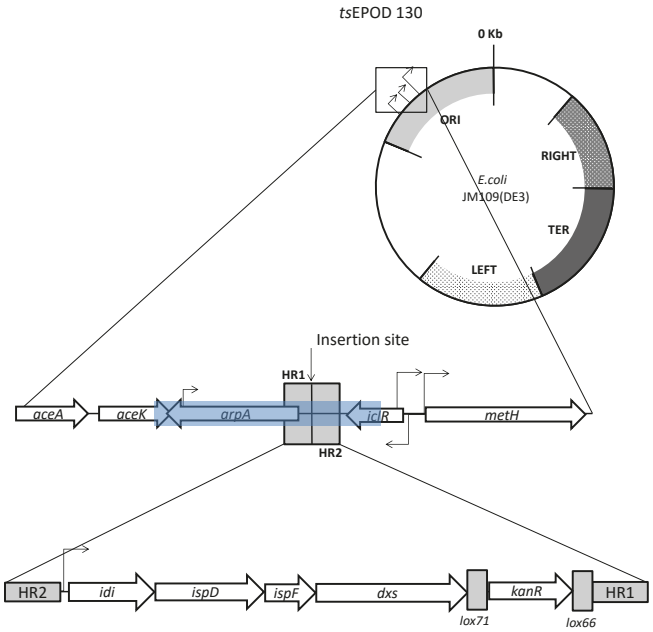


Figure S5. Continued

Figure S6. Relative positioning and abundance of nucleoid associated proteins dispersed in *he*EPOD112, *ts*EPOD130 and the EPOD-independent region between the EPODS, Fis, Factor for inversion stimulation; IHF, integration host factor; H-NS, Histone-like Nucleoid Structural proteins. Relatively high predicted abundance of H-NS sites are detected in *ts*EPOD130**.**

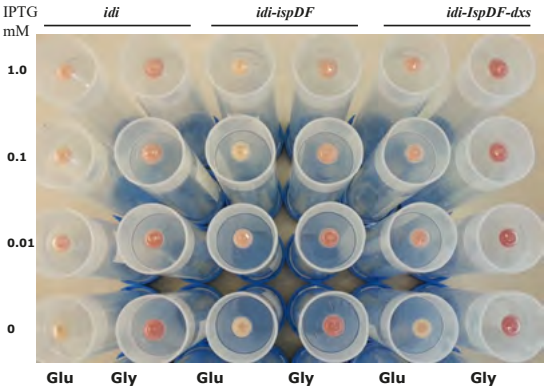


Figure S7. Lycopene production assay in *E.coli* JM109(DE3) Three different MEP-DOXP operons were integrated encompassing (i) *idi*, (ii) *idi-ispDF* and (iii) *idi-ispDF-dxs*. The mutants were grown on LB supplemented with glucose (Glu) or glycerol (Gly) to show the effect of catabolite repression. Cultures were induced with Isopropyl β-D-1-thiogalactopyranoside (IPTG) at variable concentrations to promote expression of the integrated operons and lycopene reporter

Chapter 4

Modulated Gene Expression of Fatty acid and Phospholipid Metabolism by PlsB and apo-ACP

Melvin F. Siliakus[1], Miquel Rozas[1], Teunke van Rossum[1], John van der Oost[1], Servé W. Kengen[1]

[1]Laboratory of Microbiology, Wageningen University and Research Centre, Stippeneng 4, 6708 WE, Wageningen, The Netherlands

Fatty acid metabolism and phospholipid synthesis in *Escherichia coli* have previously been studied in detail. However, the balanced regulation between fatty acid biosynthesis (FAB), fatty acid degradation (FAD) and phospholipid (PL) synthesis is still not fully understood. Glycerol-3-phosphate acyltransferase (PlsB) is the key enzyme that is responsible for the coupling between the biosynthesis of fatty acids and phospholipids. To specifically study FAB and FAD expression, PlsB transcription and activity was varied, and the alternative phospholipid synthesis route via the PlsX disrupted. The level of PlsB was manipulated by exchanging its native promoters by a repressible and inducible promoter. Down-regulation of the expression of *plsB* indeed resulted in decreased growth. As reported previously, additional inhibition of PlsB activity was obtained at the enzyme level by overexpression of the acyl carrier protein (ACP), causing partial growth arrest in wild-type *plsB* cells. Unexpectedly, when *plsB* was overexpressed in the engineered strain this further intensified the apo-ACP-induced growth inhibition. To explain these differences and study the balanced regulation between FAD, FAB and PL synthesis, proteome analysis was performed. This revealed that ACP overproduction induces PlsB-dependent responses on the fatty acid biosynthesis and degradation pathways. Not only does apo-ACP inhibit PlsB activity, it also results in decreased concentrations of PlsB and an increase of 1-acyl-*sn*-glycerol-3-phosphate acyltransferase (PlsC). Furthermore we show that a simultaneous increase of both apo-ACP and PlsB results in down-regulation of fatty acid and phospholipid biosynthesis but stimulates expression of FAD.

Keywords: Beta-oxidation, Fatty acid/metabolism, Lycophopholipid, Membranes, phospholipids/phosphatidic acid

Abstract

Introduction

The membrane of *Escherichia coli* is probably the most extensively studied membrane of any biological system [246, 247]. The gained insights have been instrumental in understanding the lipid synthesis pathways in both prokaryotes and eukaryotes. Biochemical research has elucidated each step in the phospholipid biosynthesis pathways. Based on this knowledge, efforts are on-going that aim at efficient synthesis of lipids and fatty acids as potential sources of biofuels. Fatty acids (acyl chains) are hydrocarbon chains with a terminal carboxylate group, that vary in length from short chains (C4-C6), medium chains (C7-C11), to long chains (C12-C22). Multiple enzymes involved in distinct fatty acid-dependent pathways have previously been reconstructed by demonstrating fatty acid delivery through the interaction of the acyl-carrier-protein (ACP) and partner proteins [248-253]. This has revealed that fatty acids contribute to four different physiological functions. Firstly, fatty acids are the biggest constituent of glycero-phospholipids, sphingolipids and lipopolysaccharides, and as such the main components of biological membranes. Secondly, some proteins are post-translationally modified by acylation or lipidation. In that case, the covalent attachment of acyl chains to membrane proteins act as anchor to the membrane like shown for lipoprotein Lpp[254] and hemolysin HlyA[255]. Thirdly, fatty acids may be fuel molecules that are stored as triglycerides in eukaryotes. Bacteria can also internalize and utilize exogenous fatty acids either as carbon source or as membrane constituent. Upon uptake of the fatty acid by FadL and FadD, the acyl chain is oxidized to acetyl-CoA units by beta-oxidative fatty acid degradation (FAD) enzymes FadA, FadB and FadE (**Fig. 1**). The fourth function of fatty acids is that of precursor molecules for cofactors. Fatty acid octanoic acid serves as a building block of lipoic acid, a cofactor essential for the functioning of several key enzymes in oxidative decarboxylation reactions [256]. The fatty acid pimelic acid serves as the precursor of biotin, another important co-factor involved in metabolic fixation of carbon dioxide [257].

Fatty acid biosynthesis

The bacterial fatty acid biosynthesis (FAB) cycle starts with the carboxylation of 2 acetyl-CoA to malonyl-CoA by the multi-subunit complex accABCD [258] (**Fig. 1**). The malonyl moiety is subsequently attached as thio-ester to the phosphopantetheine prosthetic group of holo-ACP by FabD to form malonyl-ACP [259]. Three 3-ketoacyl-ACP synthases then catalyse the condensation of acetyl-CoA to the growing acyl-ACP. FabH initiates the first elongation by condensing acetyl-CoA to malonyl-ACP [260-262]. Subsequent elongations are catalysed either by FabF or FabB [263]. Each successive condensation is followed by a reduction to 3-hydroxyacyl-ACP by FabG [264], a dehydration to enoyl-ACP by FabA or FabZ [265-267] and a final reduction to acyl-ACP by FabI. The latter enzyme has a key role in the completion of each elongation cycle [268, 269].

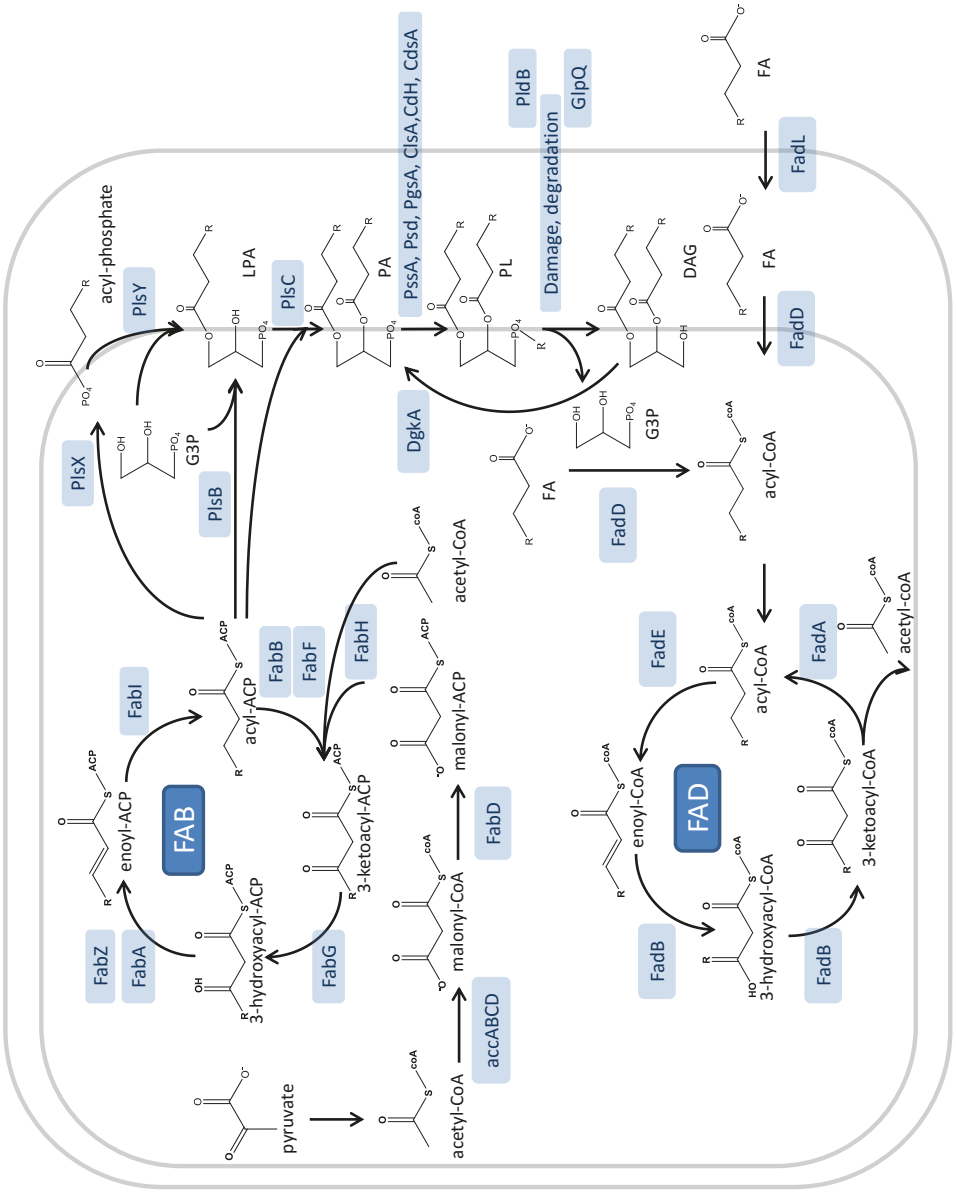


Figure 1. Fatty acid and phospholipid metabolism Fatty acid and phospholipid synthesis and degradation pathways (adapted from Janßen and Steinbüchel 2014 [247]

Phospholipid biosynthesis

The first step in the biosynthesis of phospholipids is the esterification of a fatty acid to glycerol-3-phosphate (G3P) to form lysophosphatidic acid (LPA). The incorporated fatty acids can be utilized either as acyl-ACP or acyl-CoA. The reaction is catalysed by acyltransferases of either the PlsX-PlsY pathway or the PlsB pathway (**Fig. 1**). The PlsX-PlsY pathway is the most abundant system encountered in bacteria. First, the membrane-associated phosphate acyltransferase (PlsX) phosphorylates the fatty acid from acyl-ACP to form the intermediate acyl-phosphate. This activated acyl-chain is then transferred to the *sn*-1 position of G3P by glycerol-3-phosphate acyltransferase (PlsY). LPA synthesis may also proceed via the alternative PlsB pathway. The PlsB pathway, however, is essential to *E. coli* and cannot be substituted by the dispensable PlsX-PlsY pathway [270]. Glycerol-3-phosphate acyltransferase (PlsB) is restricted to mycobacteria, alpha- and gamma-proteobacteria like *E. coli*, and requires no phosphorylated intermediate [271]. In contrast to PlsX, PlsB is an integral membrane protein which acylates G3P by utilizing either acyl-ACP or acyl-CoA as substrate. For this esterification there is a clear distinction between acyl-ACP from the cytosol and acyl-CoA imported from the environment. The 1-acyl-sn-glycerol-3-phosphate acyltransferase (PlsC) subsequently transfers a secondary fatty acid to LPA at the *sn*-2 position to form phosphatidic acid (PA). Analogous to PlsB, PlsC can also utilize either acyl-CoA or acyl-ACP as the fatty acid donor. Prior to polar head group attachment CdsA activates PA to form CDP-diacylglycerol (CDP-DAG). Finally, functional phospholipids are being formed, from which CDP-DAG bifurcates to two routes; either phosphatidyl-L-serine and phosphatidyl-ethanolamine or phosphatidyl-glycerol and cardiolipin which are synthesized by PssA and Psd, or by PgsA, Pgp and ClsA respectively.

Regulation of FAB and PL synthesis

Although a coordinated regulation between FAB, PL synthesis and FAD is evident, this process is still not fully understood. The fact that the phospholipid content of the cell remains constant regardless of the growth phase, suggests a coordinated regulation between all three pathways. A better understanding of this regulation might aid in designing metabolic engineering strategies for lipid biofuel productions. According to Zhang and Rock [246], the long-chain acyl-ACP end-products of FAB are the key regulators of the PL biosynthetic pathway. Cellular acyl-ACP levels in turn are controlled by the glycerol-3P *sn*-1-acyltransferase (PlsB) activity. Additionally, acyl-ACP levels also rely on ACP availability and the rate of the FAB cycle. Acyl-ACP has been shown to regulate Acc, FabH and FabI activity, which are responsible for the rate of FA synthesis [272, 273]. Heath and Rock *et al.* revealed that an acyl-ACP accumulation inhibits fatty acid biosynthesis. This regulation was especially evident during the stringent response. Following amino acid depletion, the (pppGpp synthetase (RelA) causes the alarmone (p)ppGpp to accumulate, which effectively

silences PlsB activity and concomitantly causes acyl-ACP to accumulate resulting in FAB inhibition [274]. Although these studies indeed show a feed-back inhibition, it is not clear whether this is the only mechanism that silences phospholipid synthesis when resources are depleting. In addition, there is only partial understanding of the responses triggered on a transcriptional level.

To shed more light on this regulation, we induced an inhibition of phospholipid synthesis. A barrier was established at the first committed step of phospholipid synthesis (PlsB) and the effects on growth and the proteome analysed. Because *plsB* is essential to *E.coli* [270], we exchanged the *plsB* promoter region for a tuneable and repressible promoter to generate a conditional *plsB* knock-down strain. To effectively control the rate of LPA formation, a *plsB* knock-down requires an additional knock-out of the classical PlsX-PlsY route. Complete abolishment of LPA formation was further explored by modulating the critical pool of ACP which allowed us to study cellular responses to the ACP isoforms: apo-ACP and acyl-ACP. The inactive apo-ACP, which is devoid in the attachment of the prosthetic phosphopantetheine group by holo-ACP synthase [275, 276], is shown to inhibit PlsB and PlsC *in vitro* [277]. The introduced modifications to inhibit PL synthesis were followed by analysis of the proteome to get more insight into the regulation of FAB, FAD and PL synthesis. This analysis revealed a modulated gene expression between FAB and FAD during induced silencing of membrane synthesis.

Results

Overproduction of PlsB and ACP causes growth arrest

To gain more insight in control mechanisms of phospholipid synthesis we constructed mutant strains that allowed for analysis of the first step of the main phospholipid synthesis route, i.e. the step catalysed by PlsB (**Fig. 1**). Because we were specifically interested in the effects of ACP and acyl-ACP on PlsB functionality, we used an *E.coli* JM109(DE3) *plsX* knockout strain as basis (**Fig. S1**). To enable modulation of *plsB* expression, we replaced the native promoters *plsBp2* and *plsBp* with the rhamnose promoter based on the sequence of a pRham–*plsB*-Chis plasmid, the functionality of which was demonstrated beforehand by anti-His Western blotting (**Fig. S2**). Instead of using this plasmid-based *plsB* system, a chromosomal promoter exchange of native *plsB* was generated (strain rhaP$_{BAD}$-*plsB*) as it was anticipated to provide a more accurate control. The growth characteristics of the *plsB*-promoter mutant strain rhaP$_{BAD}$-*plsB* was analysed upon induction and repression of *plsB* expression. Unexpectedly, growth was not blocked in the absence of rhamnose, but continued up to an OD of ~2.5. Even in the presence of glucose, which further represses *plsB* expression, this initial growth was observed, suggesting that the LB-medium contains available phospholipids or that cells can cope with a minimal amount of leaky expression of *plsB*. As expected, addition of rhamnose (0.1 % (w/v)) caused a continuation of growth albeit at rhamnose concentrations of >0.1% (w/v) (**Fig. 2**, panel 3). These findings prompted us to determine PlsB activity in the repressed and induced cultures. The PlsB activity

assays confirmed that rhamnose induces expression of *plsB*, and that glucose efficiently silences expression of *plsB* by catabolite repression (**Fig. 3**). These findings proof that the constructed mutants allow a controllable expression of *plsB*, but that (unexpectedly) growth does not correlate well with different PlsB levels, indicating some compensatory mechanism.

We subsequently transformed wild-type *plsB* and the rhaP$_{BAD}$-*plsB* mutant strain with a pACYC-*acpP* plasmid to study the effects of Acyl Carrier Protein (ACP). The inactive apo-ACP is known to competitively inhibit PlsB activity *in vitro* [277], and therefore is also expected to inhibit PlsB activity *in vivo*. We compared the growth characteristics of the wild-type (*plsB* under control of its native promoters) and the rhaP$_{BAD}$-*plsB* mutant at a fixed level of PlsB expression (0.5% (w/v) rhamnose). Here we found that variable induction of *acpP* by different IPTG concentrations had clear effects (**Fig. S3**). The fact that overexpressed ACP co-localizes with an S36T-ACP mutant with disabled pantetheinylation on urea-PAGE (**Fig. S4C**) confirms that apo-ACP is responsible for this inhibition. Although growth in the ACP-overexpressed cells was reduced severely (**Fig. 2**, panel 2), growth was never reduced completely. The rhaP$_{BAD}$-*plsB* strain, on the other hand, was expected to overcome the *in vivo* inhibition by apo-ACP by compensation with the overexpressed PlsB pool; although a 6-fold increase in PlsB activity was detected in the rhaP$_{BAD}$-*plsB* strain (**Fig. 3**), this situation did not counteract the growth inhibition, but instead it further enhanced the inhibition and led to almost complete growth arrest (**Fig. 2**, panel 4). Remarkably, introduction of the pACYC-*acpP* plasmid to the rhaP$_{BAD}$-*plsB* strain (**Fig. S3**, panel 1) also reduces cell growth significantly. This growth reduction by apparent leaky expression of *acpP* highly contrasts growth of the mutant before transformation with pACYC-*acpP.* In this strain growth was only mildly affected by the increased PlsB pool (**Fig. 2**, panel 3). Based on the *in vitro* inhibition of PlsB by apo-ACP [277] and on the growth arrest during simultaneous overexpression of ACP and PlsB, it is concluded that growth inhibition by apo-ACP can not only be explained by PlsB inactivation.

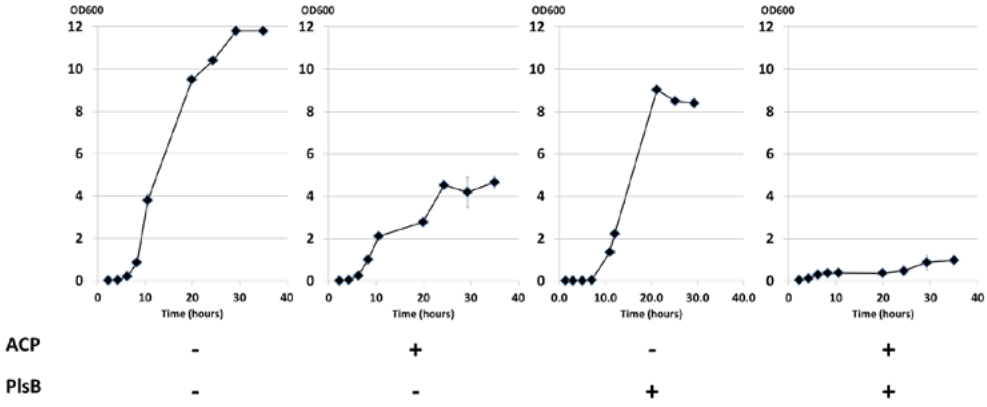


Figure 2. Growth curves of a*cpP* and *plsB* overexpressed cells *AcpP* overexpression is elicited by induction of the pACYC-*acpP* plasmid (50 µM IPTG) and *plsB* overexpression by induction of the rhaP$_{BAD}$-*plsB* strain (0.5% (w/v) L-rhamnose). Lower panel shows induction (+) and no induction (-) of *acpP* and *plsB* expression.

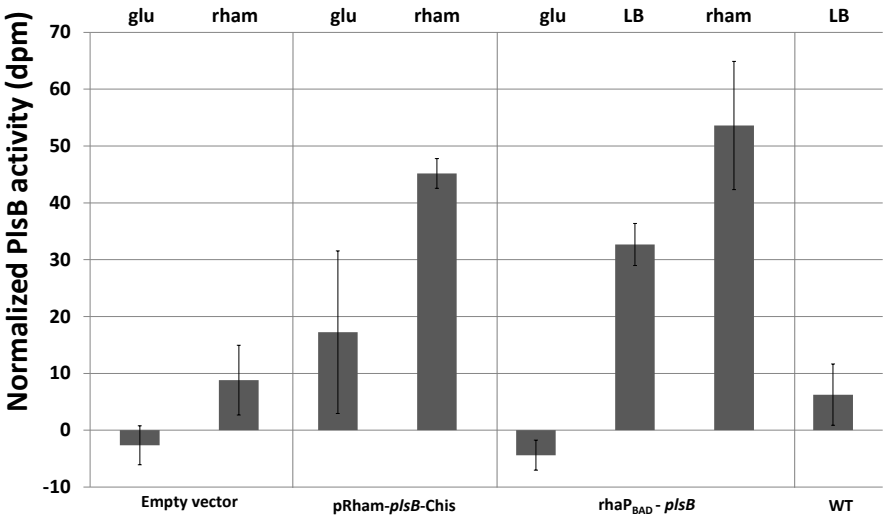


Figure 3. Activity assay upon overexpression and repression of *plsB* The rhaP$_{BAD}$-*plsB* strain shows induction, repression, and no stimulation of *plsB* by the rhamnose promoter substitution on the addition of respectively 0.5% (w/v) rhamnose, 20 mM glucose, and LB only. The mutant strain is compared to an *E.cloni* strain transformed with empty vector and pRham-*plsB*-Chis plasmid as positive control. Additionally, the mutant is compared to an unmodified *E.coli* JM109(DE3) strain (WT). All error bars are based on 4 biological replicates.

4

FAB and PL synthesis are regulated by ACP and PlsB

plsB overexpression

To assess the expression profile during inhibited membrane synthesis and understand the observed changes in growth behaviour, we further analysed ACP and ACP - PlsB overexpressed cells. To obtain milder growth inhibitions, we resigned from using the pACYC-*acpP* plasmid which relies on T7 RNAP expression. Instead, wild-type *plsB* and rhaP$_{BAD}$-*plsB* cells were transformed with the pCA24N–*acpP* plasmid that relies on the T5-lac promoter from the ASKA collection (strain JW1080) [278]. The novel strains were cultured on LB medium supplemented with 0.5% (w/v) rhamnose and 0, 0.1 and 0.5 mM IPTG. We performed a proteome analysis on cleared lysates and membrane fractions to obtain insight into changed levels of proteins involved in fatty acid and phospholipid synthesis or degradation (raw data in supplementary excel file **S12**). We first looked into the changes elicited by *plsB* overexpression alone (**Fig. S5**) to distinguish between effects of *plsB* and *acpP* overexpression. Overproduction of PlsB (28-fold, **Fig. S6**) promotes expression of fatty acid synthesis cycle proteins FabB, FabG, FabI and a repression of FadD whose peptides were no longer detected. These changes correspond with transcriptionally regulated responses by FadR and (p)ppGpp by RelA [247]. As FadR was only sparingly detected during PlsB overproduction (**Fig. S7**, 0 mM IPTG), (p)ppGpp is a plausible regulator responsible for these collateral changes in an attempt to silence PlsB activity. We also observed a large reduction

of phospholipid head group enzymes PssA, Psd, which suggests an overproduction of phosphatidylglycerol and reduction of phosphatidylethanolamine. Together these changes may contribute to the formation of elongated cells observed during *plsB* overexpression (**Fig. S8**).

acpP overexpression in wild-type *plsB* cells

We then looked into proteome changes caused by enhanced levels of ACP. The effect of *acpP* induction was analysed in both the wild-type and mutant *plsB* strain. In contrast to an expected activity increase, the increased ACP pool causes a down-regulation of PlsB in wild-type *plsB* cells (**Fig. 4** and **Fig. S9**). Furthermore, no major changes were observed for the proteins involved in fatty acid synthesis and beta-oxidation cycle that could explain the toxicity of apo-ACP. Instead, the reduced PlsB pool is accompanied by an up-regulation of the fatty-acid synthesis genes, FabA, FabF, FabG, FabH and FabI. In contrast, proteins of the beta-oxidation cycle did not change markedly. The FadD protein (acyl-CoA synthetase) however showed a down-regulation at low IPTG induction. Interestingly, at high induction, the FadD pool increased, which facilitates coupled import and activation of fatty acids. The increased apo-ACP pool also causes an up-regulation of PlsC. Taken together, the inhibition and down-regulation of PlsB is accompanied by higher expression at the level of both acyl-ACP and phosphatidic acid synthesis. Therefore, we speculate that growth continuation is facilitated by expediting flux through PlsB by increasing the acyl-ACP pool or by increasing the internalization of acyl-CoA. These observed responses correspond with a FadR-mediated control [279, 280] for which an up-regulation was indeed detected at high induction of *acpP*. Altogether these findings support an accumulation of acyl-ACP due to reduced PlsB activity.

Overproduction of ACP and PlsB in rhaP_{BAD}–*plsB* cells

The shapes of the aforementioned growth curves (**Fig. 2** and **Fig. S3**) indicate that an inhibitory mechanism takes place that (directly or indirectly) depends on the size of both the ACP and the PlsB pool. Our proteome analysis shows that large ACP and PlsB pools affect several of the involved enzymes, and that this appears to be the opposite after overexpression of ACP alone (**Fig. 4**, right panel). Our earlier discussed responses elicited by *plsB* overexpression alone also appear to differ markedly. During *acpP* and *plsB* co-expression, we observed a down-regulation of the FAB cycle enzymes FabB, FabG and FabI (**Fig. 4**, and **Fig. S10**). Conversely, 3 proteins that make up the FAD pathway (FadB, FadD and FadE) are highly induced. In strain rhaP_{BAD}-*plsB*, the level of FadE was already increased 2-fold in response to *plsB* overexpression. Additional ACP however causes even higher titres of FadE (2.6 and 3.2-fold). This is in sharp contrast to the wild-type *plsB* strain in which *acpP* overexpression alone does not change the FadE pool. A simultaneous increase in PlsB and ACP levels apparently stimulates a shift in expression of fatty acid metabolism from anabolic (FAB) to catabolic (FAD) (**Fig. 4**). In contrast to wild-type *plsB* cells where sufficient acyl-ACP is expected, a reduced synthesis of acyl-ACP causing LPA shortage is a likely explanation

for the growth arrest. Based on these observations we hypothesize a shortage of *de novo* synthesized fatty acids in the rhaP$_{BAD}$-*plsB* strain but not in the wild-type *plsB* strain. Both strains were therefore expected to allow growth rescue by free lysophosphatidic acid (LPA) addition, but only growth rescue by free fatty acid addition in the wild-type *plsB* strain.

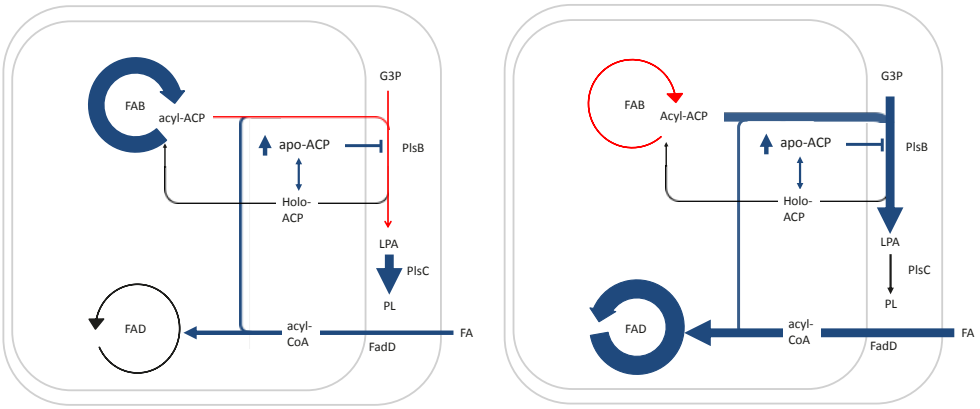


Figure 4. Summarized proteome changes Proteome changes are shown between ACP overproduction (wild-type *plsB)* in the left panel and ACP/PlsB overproduction (rhaP$_{BAD}$-*plsB*) in the right panel. Arrow thicknesses are not provided on absolute scale, but illustrated in **Fig. S9** and **Fig. S10**. Blue arrows: up-regulations; black arrows: unchanged; red arrows: down-regulations in reference to non-induced cells of the respective strain.

4

Transcription factors do not correlate with observed changes

We further analysed the transcriptional regulators involved in the expression of the analysed pathways[247] to find trends that may explain the observed differences in the proteome. In both wild-type *plsB* and rhaP$_{BAD}$-*plsB* cells a down-regulation of FabR and an up-regulation of FadR were observed compared to non-induced cells (**Fig. S7**). The FabR and FadR changes correlate well with a stimulation of the FAB cycle. Although this relation is evident in the wild-type *plsB* cells, the changes observed in the rhaP$_{BAD}$-*plsB* strain cannot be attributed to these transcription factors. In addition to FadR and FabR we therefore also looked at the transcription factors RpoS, CRP, (p)ppGpp by RelA and RpoE. During overexpression of *plsB*, we observed accompanied FAD up-regulations that were only observed at simultaneous *acpP* induction. The stimulation of FAD proteins in these cells is most likely regulated by the observed increase in RpoS. This finding however highly contrasts the increased FadR up-regulation which is known to cause a down-regulation of this pathway. Likewise did we observe an increased pool of RelA that can account for the changes in the FAB cycle. Conversely, the changes in FadR and FabR highly contradict this outcome. Hence, this data clearly shows that the transcriptional regulation profile does not allow a prediction of the metabolic status of the rhaP$_{BAD}$-*plsB* cells. This suggests that ACP apparently exerts effects both by allosteric and (in)direct transcriptional control.

PlsB and apo-ACP overproduction diminishes availability of fatty acids for PL synthesis

Based on the proteome analysis, we predicted an inhibition of membrane synthesis in wild-type *plsB* cells upon *acpP* induction as a result of decreased lysophosphatidic acid (LPA) synthesis. For the rhaP$_{BAD}$-*plsB* cells, overexpression of *plsB* and *acpP* was expected to cause a depletion of LPA by acyl-ACP shortage instead. These expectations were based on the down-regulation of FAB and up-regulation of FAD, which were not observed in the wild-type *plsB* cells. To test this hypothesis, we attempted to restore growth by the addition of LPA or free fatty acids to the growth medium during *acpP* induction.

Our rescue assays demonstrate that diminished growth by *acpP* overexpression in the wild-type *plsB* strain can indeed be fully restored by the addition of LPA (**Fig. 5**, left panel). This observation thus confirms that apo-ACP inhibits acyl-ACP utilization by PlsB *in vivo*. This, however, also entails that sufficient acyl-ACP was present for PlsC to synthesize PA, indicating that the FAB cycle is not fully inhibited by the accumulating acyl-ACP. This finding also agrees well with the proteome data that demonstrates an up-regulation of FAB proteins. Growth recovery also occurs upon addition of fatty acids, which suggests that the inhibition of PlsB by apo-ACP can be overcome by competing with imported acyl-CoA.

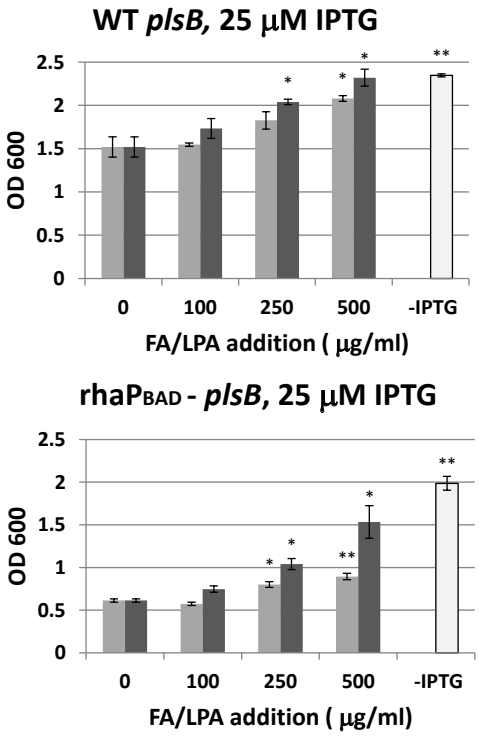


Figure 5. LPA and FA growth rescue assays. Effect of LPA and fatty acid addition on cell growth in wild-type *plsB* and rhaP$_{BAD}$-*plsB* cells during *acpP* or *acpP* and *plsB* overexpression respectively. Grey: oleic and palmitic acid administration, dark grey: LPA (14:0 and 18:1) administration, white, no IPTG control.*significantly different from 0 μg/ml control $P<0.05$, **significantly different from 0 μg/ml control $P<0.01$.

A similar LPA administration however, could not fully restore the growth arrest in the rhaP$_{BAD}$-*plsB* mutant (**Fig. 5**, right panel). This suggests that unlike the wild-type situation, the level of acyl-ACP is now insufficient for PlsC to give complete recovery. This apparent lower acyl-ACP level agrees well with the down-regulation of FAB proteins in the mutant strain. Thus, the observed aggravated growth inhibition in rhaP$_{BAD}$-*plsB* cells can be attributed to a reduced availability of acyl-ACP rather than a reduced activity of PlsB alone. In contrast to wild-type *plsB*, internalization of fatty acids showed only minor growth restoration despite a down-regulation of FAB that suggests a high need for fatty acids. This difference might be explained by activity of the increased beta-oxidation cycle that competes with PlsB for the internalized fatty acids. Altogether, a high level of PlsB and apo-ACP is shown to result in the silencing of FAB and PL synthesis by a combination of transcriptional and allosteric control.

Discussion

Only few studies have explored the balanced regulation between fatty acid and phospholipid synthesis. Previous studies report that long-chain acyl-ACP is the main regulator of membrane synthesis by feedback inhibition on FAB and requires proper balancing [246]. Here, we set out to further investigate the regulatory mechanisms of acyl-ACP on membrane synthesis, underlying the observed growth inhibition. An acyl-ACP accumulation implies that formation by FAB and consumption by PlsB, PlsX and PlsC are unbalanced. Additionally, long chain acyl-ACP can only accumulate upon obstruction of PlsB, PlsX and PlsC activity or a G3P shortage. To provide clarity in this mechanism, we investigated whether there are other regulatory systems involved that change expression of FAB, PL synthesis and FAD upon induced inhibition of membrane synthesis.

Wild-type *plsB*

To study changes in expression levels of the FAB cycle and PL synthesis pathway, we inhibited phospholipid synthesis by overexpression of ACP in wild-type *plsB* cells. As apo-ACP did not abolish growth completely, these cells were expected to respond by up-regulating their PlsB level. Proteome analysis in these cells however demonstrated a down-regulation of PlsB instead. To explain this down-regulation, we also analysed changes in transcription factors. Our data suggests that the observed down-regulation is likely elicited by the alarmone (p)ppGpp caused by RelA up-regulation. A PlsB down-regulation and blocked activity by apo-ACP should lead to an accumulation of long-chain-acyl-ACPs and the concomitant cessation of fatty acid synthesis [272-274, 281-283]. This, however, does not agree with the observed up-regulated expression of fatty acid synthesis proteins FabH, FabG, FabF, FabB, FabI, as well as of PlsC shown in this study. Feedback inhibition of FabH and FabI by acyl-ACP (allosteric) is apparently compensated by up-regulation of these genes. This questions the main role of acyl-

ACP as a controller of the rate of fatty acid synthesis. To further study the mechanism of the growth inhibition, we performed a FA and LPA rescue assay. Here we observed that growth was fully restored by the addition of LPA, which confirms that PlsB inhibition by apo-ACP was indeed the causative obstruction. In addition, this restored growth implies that sufficient FAB activity is maintained, as for PA formation from LPA, acyl-ACP is also required. Apparently, LPA addition relieves the acyl-ACP accumulation by activity of PlsC which utilizes acyl-ACP as well. Fatty acid addition to the medium also led to growth rescue which implies that PlsB inhibition by apo-ACP tolerates acyl-CoA utilization in contrast to earlier observations [277].

In the absence of exogenously added fatty acids or LPA, PlsB inhibition is evidently accompanied by increased acyl-ACP and PA synthesis as shown by the up-regulated expression of FAB and PlsC enzymes. Altogether, it is shown that an accumulation of acyl-ACP is not an exclusive condition for silencing of phospholipid synthesis, and not necessarily accompanied by down-regulation of FAB proteins.

rhaP$_{BAD}$-*plsB* mutant

By replacing the native promoters of *plsB* with a rhamnose promoter we showed that repression of *plsB* inhibits growth, but did not suffice to cause a complete growth arrest. Therefore we attempted to enhance the inhibition of membrane synthesis by also increasing the apo-ACP pool. Overexpression of ACP in combination with a repressed *plsB* expression proved successful in abolishing growth. We expected that the inhibitory effect of apo-ACP could be relieved by causing an overexpression of *plsB* as was already demonstrated for ppGpp induced growth inhibition [274]. However, overexpression of PlsB (28-fold) failed to rescue growth at increased ACP pools. In contrast, growth inhibition was even aggravated. This finding is a first indication that apo-ACP has additional inhibitory effects on growth, apart from merely PlsB inhibition. During simultaneously high PlsB and ACP pools, proteome data showed notable changes which were not observed when either PlsB or ACP are overexpressed separately. Here we observed a down-regulation of fatty acid synthesis proteins FabI and FabB in contrast to an up-regulation which was observed when ACP alone was overexpressed. Concomitantly, a significant increase in fatty acid degradation enzymes FadB, FadE, FadD and FadL was induced. Altogether these changes suggest a decreased acyl-ACP pool and an increased beta-oxidation activity. Therefore, we hypothesized that a shortage of available fatty acids for membrane synthesis was the detrimental deficiency instead of an inhibition of PlsB alone. To test this, a rescue assay by FA and LPA supplementation to the medium was also performed on these cells. This assay demonstrated that growth inhibition in this strain is not recovered by fatty acid uptake and only poorly by LPA which validates the hypothesis. This data shows that high apo-ACP and PlsB pools induces a metabolic shift towards degradation of fatty acids. Another contributing player is (p)ppGpp, that is expected to accumulate by the observed up-regulation of RelA. This alarmone plays a crucial role in the stringent response which modulates transcription

of many genes under conditions of amino acid limitation. Thus, (p)ppGpp may have been formed in response to high PlsB expression in an attempt to suppress the activity and gene transcription. An additional effect of the stringent response might therefore also be found by the down-regulation of FAB and increase of FAD. The accumulation of (p)ppGpp, however, cannot explain all differences between the strains, as RelA was also found to be highly expressed in the wild-type *plsB* cells.

In conclusion, our data shows that a model in which acyl-ACP accumulation always triggers silenced membrane synthesis is too simple. We postulate that FAB down-regulation and FAD up-regulation are transcriptionally-controlled phenomena in response to simultaneously elevated pools of both PlsB and apo-ACP.

Materials and Methods

Generation of a *plsX* knock-out strain

A *plsX* knock-out was obtained according to a modified gene disruption protocol by Datsenko and Wanner [234]. Briefly, *E.coli* JM109(DE3) cells were transformed with pKD46 by heat shock. Electrocompetent cells were made by growing a culture on 2x Yeast Peptone ampicillin 100µg/ml supplemented with 10 mM L-arabinose to induce the λ Red recombinase. Primers to amplify the gene disruption cassette are listed in **Table 1**. The 5'-end of For-H1-plsX-P1 contains a 39 nucleotides homologous end, corresponding to the upstream region of *plsX*. The 5'-end of Rev-H2-plsX-P2 contains a 40 nucleotides homologous end that corresponds to the upstream region of *fabHp* which is located at the 3'-end of *plsX*. A new disruption cassette was developed based on the cassette of Westra E.R. *et al.*[284], but with *lox71*(left)/*lox66*(right) sites [285] instead of *frt* sites flanking *kanR* (synthesized and cloned SfiI/SfiI in pMA-RQ by Geneart AG, see supplementary information (**S11**) for description and sequence). The primers For-H1-plsX-P1 and Rev-H2-plsX-P2 were used to amplify the *plsX* knockout cassette from the plasmid template pMA-RQ *lox71-kanR-lox66*. The amplified fragments were applied on agarose gel and extracted. The purified knockout cassettes were transformed at 100 ng DNA by electroporation 2.5 kV, 200Ω, 25 µF at 5 ms. Cells were plated at 37°C on lysogeny broth (LB)-agar plates with 50 µg/ml kanamycin. For subsequent elimination of *kanR* by Cre-recombinase, parts of the protocol from Datsenko and Wanner [234] were replaced by components of the protocol from Palmeros *et al.* [286]. The integrated kanamycin resistance gene was cleared by transformation with pJW168 and subsequent IPTG induction of Cre-recombination of the flanking *lox*-sites by plating on LB-agar Amp^{100} plates supplemented with 0.5 mM IPTG. Recombination events were confirmed by colony-PCR amplification using primers *plsX*-k.o-For and *plsX*-k.o-Rev specific for the flanking *rpmF* and *fabH* genes respectively. The newly obtained *plsX* knock-out strain was cleared from plasmids by growth at 37°C and designated WUR45.

Cloning and characterization of the pRham-*plsB* plasmid

The *plsB* gene was amplified from *E. coli* MG1655 genomic DNA by using the primer pair For-prham-*plsB* and Rev-prham-*plsB*-Chis. *E.cloni* cells (Lucigen corp.) were co-transformed with 75 ng *plsB* insert and 25 ng linear pRham™ vector according to the manufacturer's procedure. The obtained colonies were screened by colony PCR using the pRham forward and pETite reversed primer. The newly obtained plasmid was designated pMS32 (pRham *plsB*-Chis). A single colony of *E.cloni* pRham-plsB-Chis was used to inoculate fresh LB medium supplemented with 50 µg/ml kanamycin to obtain a pre-culture. Fresh LB-Kan[50] medium (9x10ml) was inoculated with a 100x diluted pre-culture and supplemented with glucose or rhamnose at an OD_{600}=0.8. The induced cultures were grown an additional 4h at 37°C. Cells were subsequently harvested by centrifugation. The cell pellets were resuspended in 1.5 ml ice cold lysis buffer: 50 mM Tris HCl, pH=8, EDTA 1 mM, NaCl 150 mM, DTT 1 mM, PMSF 0.1 mM. Induced *E.cloni* cells were homogenized by 4 pulses of 20 s ultrasonication. Cleared lysates were obtained by 10 min centrifugation of the crude extracts at 10,500 x*g* at 4°C. The protein content of cleared lysates was quantified by the Roti Nanoquant method according to Bio-RAD. Twenty µg protein per sample was applied on 8.5% SDS-PAGE. Proteins were transferred to nitrocellulose membrane for 1.5 h at 130 mA. The nitrocellulose membrane was blocked with bovine serum albumin (BSA). PlsB was visualized with primary antibody mouse anti-His6 IgG and secondary antibody Alexa Fluor488 goat anti-mouse IgG.

Table 1. Oligonucleotides, strains and plasmids used in this study

Primer	Oligo Sequence (5'→ 3')
Construction of mutants	
For-H1-*plsX*-P1	GGCAACTGGGGAAAGACCAAACCGGGCGGCGACGATACCGGTGTCTTTTTTACCTGTTTGACC
Rev-H2-*plsX*-P2	CCCTTCACCCTGAGATTTCAGCAGAGAAAGGAACATCCTGCTACCTCTGGTGAAGGAGTTG
plsX-k.o.-For	TCCATGGCCGTACAACAG
plsX-k.o.-Rev	AAAGCGTGCGTAGCAGAAGTC
For-*plsB*-KI-S3	ATTTCCAGCGGCTCTAACG
For-*plsB*-exch-Nter	TTCAGGCGCTTAACGGTAAC
Construction of plasmids	
For-prham-*plsB*	GAAGGAGATATACATATGTCCGGCTGGCCACGAATTTAC
Rev-prham-*plsB*-Chis	GTGATGGTGGTGATGATGCCCTTCGCCCTGCGTCGCACT
pRHam forward	GCTTTTTAGACTGGTCGTAGGGAG
pETite reversed	CTCAAGACCCGTTTAGAGGC
For-NcoI-*acpP*	GCGCCATGGGCACTATCGAAGAACGCGTTAAG
Rev-BamHI-*acpP*	GCCGGATCCTTACGCCTGGTGGCCGTTGATGTA
ACYCDuetUP1	GGATCTCGACGCTCTCCCT
DuetDOWN1	GATTATGCGGCCGTGTACAA
For-SDM-ACP-S36T	CGTTGAAGACCTGGGCGCGGATACTCTTGACACCGTTGAGCTGG
Rev-SDM-ACP-S36T	CCAGCTCAACGGTGTCAAGAGTATCCGCGCCCAGGTCTTCAACG

Homologous ends and restriction sites are underlined

Strains	Plasmids	Genes / Characteristics
WUR45		JM109(DE3) Δ*plsX*
WUR46	pMS46	pCR2.1topo *lox71-kanR-lox66*
WUR61	pKD46	JM109(DE3) Δ*plsX,* pKD46
WUR102		JM109(DE3) Δ*plsX*, Δ*dgkA*, rhaP$_{BAD}$-*plsB*
WUR32	pMS32	pRham *plsB*-Chis
WUR57	pMS57	pRham -
WUR132	pMS132	pACYCduet – *acpP*
WUR133	pMS133	pACYCduet – *acpP* (S36T)
WUR105	pUC18	JM109(DE3) Δ*plsX*, Δ*dgkA*, rhaP$_{BAD}$-*plsB,* pUC18
WUR106	pCA24N *acpP*	JM109(DE3) Δ*plsX*, Δ*dgkA*, rhaP$_{BAD}$-*plsB,* pCA24N *acpP*
WUR113	pUC18	JM109(DE3) pUC18
WUR115	pCA24N *acpP*	*acpP*
WUR139	pMS132	JM109(DE3) Δ*plsX,* pACYCduet – *acpP*
WUR140	pMS133	JM109(DE3) Δ*plsX,* pACYCduet – *acpP* (S36T)
WUR141	pMS132	JM109(DE3) Δ*plsX*, Δ*dgkA*, rhaP$_{BAD}$-*plsB,* pACYCduet – *acpP*
WUR142	pMS133	JM109(DE3) Δ*plsX*, Δ*dgkA*, rhaP$_{BAD}$-*plsB,* pACYCduet – *acpP* (S36T)

4

plsB promoter exchange with rhaP$_{BAD}$ and strain characterization

The sequence of the rhaP$_{BAD}$ -promoter exchange cassette was designed according to the sequence of the rhaP$_{BAD}$-*plsB* promoter-gene sequence of the previously characterized pMS32 plasmid. A stretch of 288 nucleotides upstream of the *plsB* startcodon which entails the rhaP$_{BAD}$ promoter, was attached to the 5'-end of the *lox71-kanR-lox66* cassette. This resulting construct was additionally flanked at the 5'-end with 227 nucleotides of the *plsB* gene which represents the 5'-homologous region H1. At the 3'-end of the *lox71-kanR-lox66* cassette, 292 nucleotides of the *lexA* gene including the upstream *lexA* repressor binding sites 1, 2, 3 and *lexAp* promoter sequence were attached and represent the 3'-homologous region H2. The 5'-terminus of the homologous region H1 was extended with nucleotides 5'-GGC to form a unique SfoI site which upon digestion leaves a native blunt-end terminus. The 3'-termini of the homologous region H2 was likewise extended with nucleotides 5'-GCC to form a secondary SfoI site which upon digestion also leaves another native blunt-end terminus. The cassette was artificially synthesized by GeneArt (Life Technologies) and subcloned in a pMA-RQ plasmid. To obtain linear promoter exchange cassettes, 13 µg of pMA-RQ entailing the synthesized construct was double digested with AvaII (Eco47I) and SfoI (EheI) Thermo Fisher Scientific Inc. The digest was subsequently applied on a 1% (w/v) agarose gel and the 1887 bp fragment excised from gel. Competent WUR61 (JM109(DE3) Δ*plsX*, pKD46) cells were transformed with 300 ng gel-extracted and purified cassette DNA

by electroporation. The transformed cells were subsequently plated on LB-agar plates supplemented with 50 µg/ml kanamycin and 0.2% (w/v) rhamnose. The newly obtained colonies were screened for insert by colony PCR using the primer pair: For-*plsB*-KI-S3 and For-*plsB*-exch-Nter. An integration efficiency of 40% was established for the screened colonies. The *kanR* gene was cleared from the genome by Cre-recombinase as described previously. The newly obtained *plsB* mutant strain was designated as WUR102 (JM109(DE3) Δ*plsX*, Δ*dgkA*, rhaP$_{BAD}$-*plsB*). The *plsB* mutant and wild-type were additionally transformed with pUC18 plasmid to include an independent antibiotic selection marker and are designated as WUR105 and WUR113 respectively. A single colony was used to inoculate 25 ml LB 100 µg/ml ampicillin, rhamnose 0.05% (w/v) as a pre-culture. Second pre-cultures were started in 10 ml LB-Amp^{100} from 1/1000 volume pre-culture 1. Secondary pre-cultures were supplemented with 20 mM glucose, and 0, 0.01, 0.02, 0.05, 0.1, 0.2, 0.5, 1 and 2% (w/v) rhamnose and grown o/n at 37°C. Twenty-five ml LB Amp^{100} supplemented with appropriate substrates were inoculated at 1/1000 volume from the secondary pre-cultures in triplicates and grown at 37°C. Growth curves were obtained by measuring turbidity at 600 nm. Cells were subsequently checked under a light microscope at a 1000x magnification to examine their morphology.

G3P acyltransferase (PlsB) activity assay

WUR57 (empty vector) WUR32 (pRham-*plsB*-Chis plasmid) and WUR102 (rhaP$_{BAD}$-*plsB*) were grown in 100 ml LB supplemented with 20 mM glucose or 0.5% (w/v) rhamnose and appropriate antibiotics. Likewise, WUR115 (rhaP$_{BAD}$-*plsB* with pCA24N-*acpP*) cells were grown at 0.5% (w/v) rhamnose and 1 mM IPTG to test PlsB activity upon *acpP* induction. The cells were harvested after 16 and 42 h growth resp. at 37°C by a 10 min centrifugation at 4,700 rpm and 4°C. Cell pellets were subsequently resuspended in 8 ml cold lysis buffer (Tris HCl 50 mM pH=8, 1 mM EDTA, 150 mM NaCl, 1 mM DTT and 0.1 mM PMSF). The cells were lysed by two passages through a french press cell at 16,000 psi. To obtain cleared lysates, the crude extracts were centrifuged 10 min at 10,000 x*g* at 4°C. The cleared lysates were then ultracentrifuged at 120,000 x*g* at 4°C for 70 min to obtain membrane fractions. The pelleted membranes were covered with 200 µl lysis buffer and resided at 4°C >48 h. The membranes were resuspended in the lysis buffer by gently pipetting. The protein content was quantified by Roti nanoquant (Roth). The G3P acyltransferase reaction mixture was composed of: 100 mM Tris-HCl (pH=8.4), 0.3 M NaCl, 1 mg/ml BSA, 3.5 mM MgCl2, 1 mM ^{14}C-glycerol-3P (10 mCi/mmol) (Perkin Elmer), 50 µM palmitoyl-CoA (Sigma-Aldrich) in a final volume of 200 µl. The reaction was initiated by addition of 10 µg membrane fraction to the reaction mixture and incubated at 24°C for 10 min. The reaction was stopped by addition of 0.6 ml of 1% $HClO_4$. Lysophosphatidic acid product was extracted by addition of 3 ml $CHCl_3$:MeOH (1:2 v/v) and vortexing. This was followed by addition of 1 ml $CHCl_3$ and 1 ml 1% $HClO_4$. Two phases were reconstituted by centrifugation (5 min at 1,000 x*g*). The

upper aqueous phase was removed by aspiration, and the organic phase was washed twice with 2 ml 1% $HClO_4$ to remove residual substrate. The product was measured by counting disintegrations on a scintillation counter.

Cloning of *acpP* and *acpP* S36T in pACYCduet

The *acpP* gene was amplified by PCR using the primer pair For-NcoI-*acpP* and Rev-BamHI-*acpP* from *E.coli* MG1655 genomic DNA template. The resulting PCR product was digested with NcoI and BamHI, and then ligated into the corresponding restriction sites of pACYC-duet (Novagen). This construct was designated pMS132 and transformed to competent NEB5α cells (New England Biolabs Inc.) by heat shock according to the manufacturer's procedure. The newly obtained colonies were screened for insert by colony PCR using the provided ACYCDuetUP1 and DuetDOWN1 primers. To mutate the serine at amino acid position 36 to a threonine, a single nucleotide mutation thymine109 was exchanged for an adenine. Quikchange site directed mutagenesis was performed on the newly obtained pMS132 plasmid using the primer pair For-SDM-ACP-S36T and Rev-SDM-ACP-S36T according to Stratagene. The mutated plasmid pMS133 was obtained by performing a Phusion polymerase (Life Technologies) reaction followed by a DpnI reaction. The resulting plasmid was transformed to NEB5α cells and screened by sequence analysis.

Strain and ACP pool characterization

The newly obtained plasmids pMS132 and pMS133 were transformed by electroporation to the previously generated WUR45 and WUR102. The newly obtained strains were designated WUR139 and WUR140 from WUR45 respectively and WUR141 and WUR142 from WUR102 respectively. A single colony was used to inoculate 10 ml LB supplemented with 34 µg/ml chloramphenicol and grown overnight at 37°C to obtain a pre-culture. For each induction, 25 ml LB-Cm^{34} supplemented with 0.5% (w/v) rhamnose with appropriate IPTG addition was inoculated at 1/100 volume from the obtained pre-culture. Each induction was performed in triplicate. Growth curves were obtained by measuring turbidity at 600 nm. For ACP detection on urea-PAGE, fresh cultures were obtained and incubated on ice after 28.5 h. of growth. The cells were harvested by centrifugation for 10 min at 4700 rpm on a Sorvall centrifuge at 4°C. The obtained cell pellets were resuspended in 1 ml cold lysis buffer. Cells were homogenized by 4 pulses of 20 sec ultrasonication. Cleared lysates were obtained by centrifugation of the crude extracts (10 min at 10,500 x*g* at 4°C). The protein content was subsequently quantified by the Roti-Nanoquant method (Roth). ACP species were analysed on a 13% PAGE containing 2.5M urea. Twenty µg of protein per sample was diluted in sample buffer (3x) (150 mM Tris-HCL pH=6.8, DTT 300 mM, 0.3% bromophenol blue, 30% glycerol, urea 2.5M). Samples were boiled 10 min at 95°C and loaded on 13% PAGE containing 2.5 M urea. Proteins were subsequently stained by PAGE-blue (Bio-RAD) according to the manufacturers procedures.

Proteome analysis of ACP overproduction

WUR102 and WUR45 were transformed with pCA24N-*acpP* isolated from the ASKA collection strain JW1080 to obtain strains WUR106 and WUR115 respectively. A pre-culture was obtained by inoculation of 10 ml LB Cm^{34} from a single colony and growing overnight at 37°C. 100 ml LB Cm^{34} supplemented with 0.5% (w/v) rhamnose was inoculated at 1/1000 volume from the pre-culture. The samples were induced at 0, 0.1 and 0.5 mM IPTG and grown approximately 42 h at 37°C to obtain a stable growth plateau. Cleared lysates and membrane fractions were generated as described previously. Membrane fractions were resuspended in a solution containing 0.5 M 6-aminocaproic acid and 5 g/L n-dodecyl-beta-D-maltoside to solubilize membrane proteins. Membrane fractions were solubilized for 1 h. at 20°C at vigorous stirring. After 15 min sonication in a sonicator bath, the fractions were centrifuged at 14,000 rpm for 30 min to pellet insolubilized aggregates. Protein content of the obtained cleared lysate samples and solubilized membrane proteins was subsequently quantified by Roti-nanoquant. Twenty-five µg protein from both cleared lysate and membrane fraction was loaded on a pre-cast 10% acrylamide gel Mini-PROTEAN® TGX (BioRad) and electrophoresed 12-15 min at 15 mA to migrate 1-1.5 cm. The gel was washed with demiwater and stained with PAGEblue (Thermo Fisher Scientific). The stained proteins were excised from gel. Trypsin digestion and peptide clean-up were performed according to the standard procedures of the Radboud UMC Centre for Proteomics, Metabolomics and Glycomics, (Nijmegen, the Netherlands). LC-MS/MS profiling of the peptides was performed by nanoflow C18 reversed phase liquid chromatography coupled to a Maxis 4G +ETD Q-TOF instrument (Bruker Daltonics) The samples were measured with CID (Collision Induced Dissociation) as fragmentation technique and chromatographic separation was achieved via a linear gradient of 7 to 35% acetonitrile using 0.1% formic acid as ion pair reagent. Data analysis was performed using the MaxQuant software package.

FA and LPA growth rescue assay

Induced cultures of WUR139 and WUR140 were obtained as described previously for the strain characterization. Both strains were induced with 25 µM IPTG. Three stock solutions of fatty acids and LPA were prepared at 2, 5 and 10 mg/ml in LB Tween-20 5%. Fatty acid stock solutions were prepared from palmitic acid and oleic acid (Sigma Aldrich) at a (1.63 : 1.00) ratio as shown for *E.coli* W3110 fatty acid composition at 30°C [287, 288]. The LPA stock solutions were prepared from 1-myristoyl-2-hydroxy-sn-glycero-3-phosphate (14:0) and 1-oleoyl-2-hydroxy-sn-glycero-3-phosphate (18:1) provided by Avanti Polar Lipids Inc. by solubilizing in a (1 : 4.26) ratio as determined at physiological compositions at 30°C [287, 288] in LB Tween-20 5%. Fatty acid and LPA mixtures were added directly after induction and inoculation. Triplicate cultures were grown in 10 ml LB from each FA and LPA containing medium at 0, 100, 250 and 500 µg/ml. Control cultures were subjected to similar volumes

of LB Tween-20 5% without fatty acids or LPA. Growth performance was determined by measuring turbidity at 600 nm after 20 h. incubation at 37°C.

Acknowledgements

This work was carried out within the BE-Basic R&D Program, which was granted a FES (Fonds Economische Structuurversterking) subsidy from the Dutch Ministry of Economic affairs, agriculture and innovation (EL&I).

Supplemental information

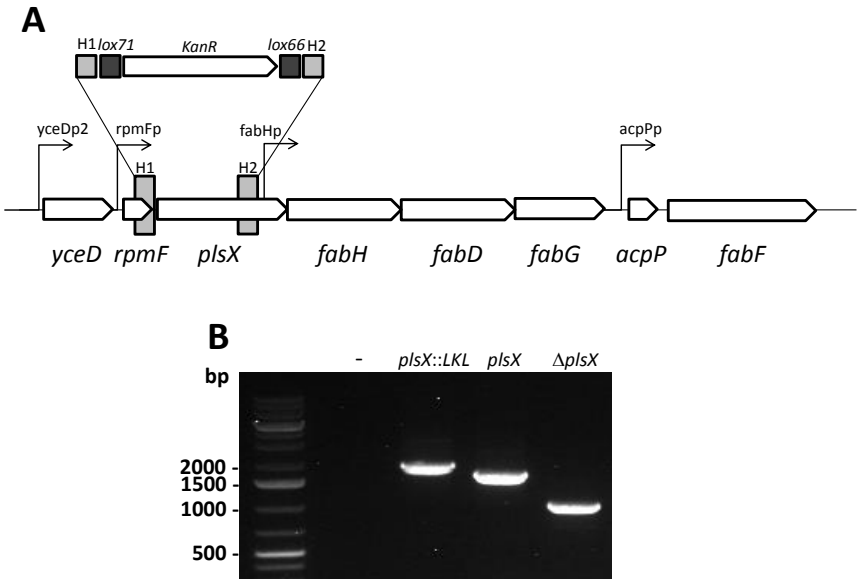


Figure S1. Generation of a *plsX* knock-out. A: Fab gene cluster *plsX-fabHDG* showing the *fabHp* promoter at the 3'-end of *plsX*. The knock-out cassette shows the corresponding homologous regions positioned on the *rpmF* and *plsX* gene. Although the *fabHp*-promoter is compensated by the upstream *YceDp1* and *YceDp2* promoters, transcriptional alteration of the fatty acid synthesis cycle was avoided by excluding this region from the disruption. **Fig. S1B**: 1% agarose gel showing PCR amplification of the knock-out strain *plsX* region at 1011 bp, wild-type *plsX* at 1655 bp and the gene disruption harbouring the *lox* flanked kanamycin resistance gene (LKL) at 1999 bp.

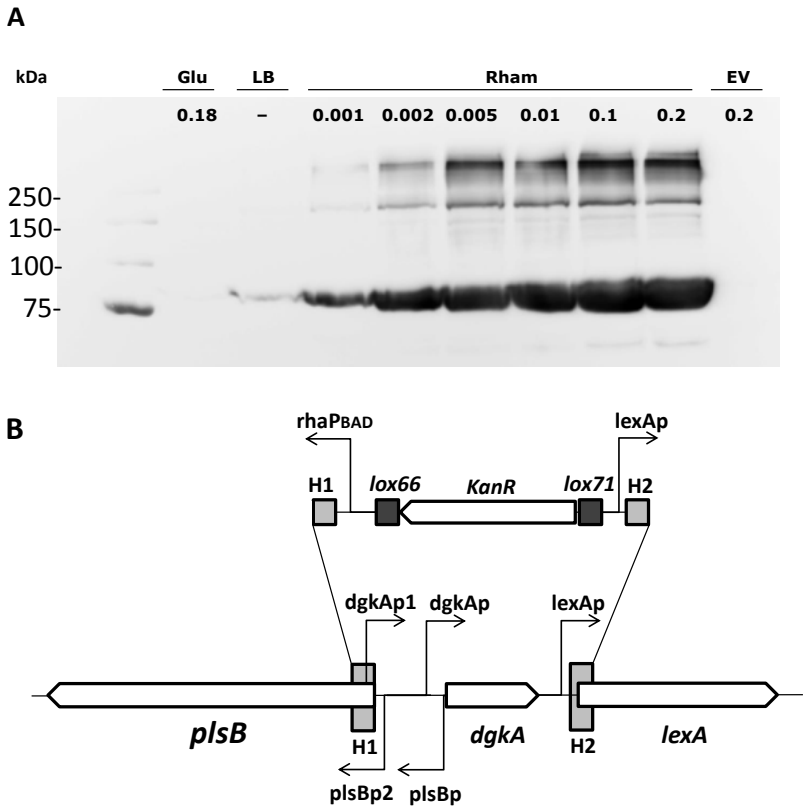


Figure S2. Exchanging the native promoters *plsBp2* and *plsBp* with a rhaP$_{BAD}$ promoter. A: Inducible control of *plsB* was tested in a pRham-*plsB*-Chis plasmid. Western Blot analysis on the C-terminal His-tagged PlsB was performed by mouse-anti-His IgG, and visualised with Alexa-fluor 488. Effective catabolite repression by glucose and induction by rhamnose addition is confirmed. In the absence of glucose and rhamnose 'LB', leaky production of PlsB (91.4 kDa) is visible. Tuneable control of *plsB* expression was shown by adding different amounts of rhamnose (%). EV: empty vector **S2B:** A knockout cassette was employed that replaces *plsBp2* and *plsBp* with the rhamnose promoter based on the acquired plasmid sequence of pRham-*plsB*-Chis. Because of partial overlap of *plsBp* and some putative transcription factor sites with the 5'-end of *dgkA*, it was decided to disrupt *dgkA* simultaneously with the promoter exchange. The *dgkA* gene is shown essential only during hypo-osmotic circumstances, and therefore assumed to be of no relevant influence in further analyses.

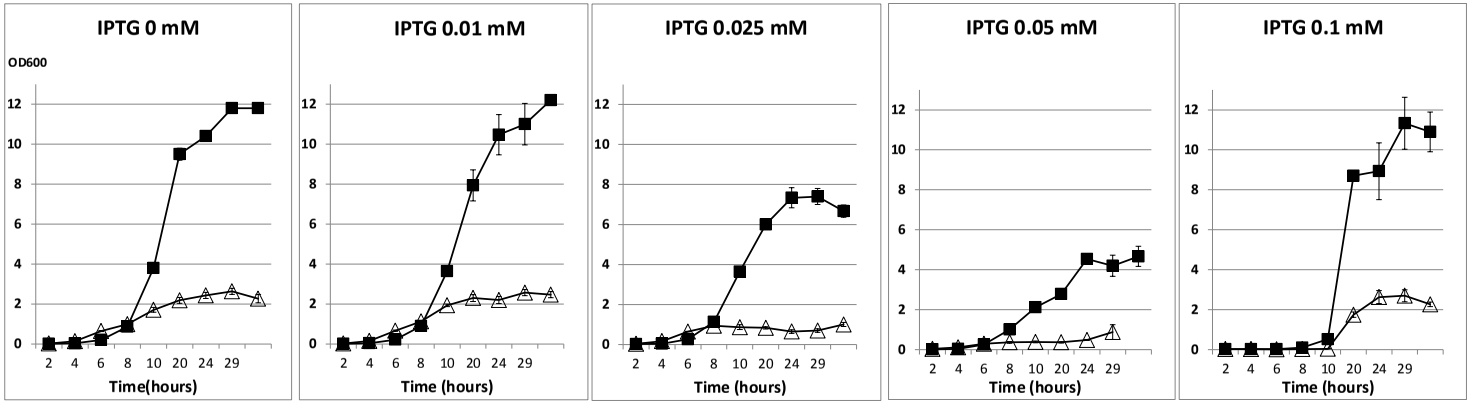


Figure S3. IPTG-induction profiles of *acpP* showing clear effects on growth of wild-type *plsB* and rhaPBAD-*plsB* cells during 0.5% rhamnose induction of *plsB*. ■ : Wild-type *plsB*, and Δ: $rhaP_{BAD}$ - *plsB* at different *acpP* induction levels. Cells were grown on LB supplemented with 0.5% rhamnose in four-fold replicates. Unexpectedly both strains showed a full growth recovery at high induction of ACP (0.1 mM IPTG) (**S3**: panel 5), urea-PAGE analysis showed that at this IPTG concentrations ACP levels in the cleared lysates and insoluble fractions had decreased (**S4A & B**), confirming a correlation between the level of ACP expression and growth inhibition.

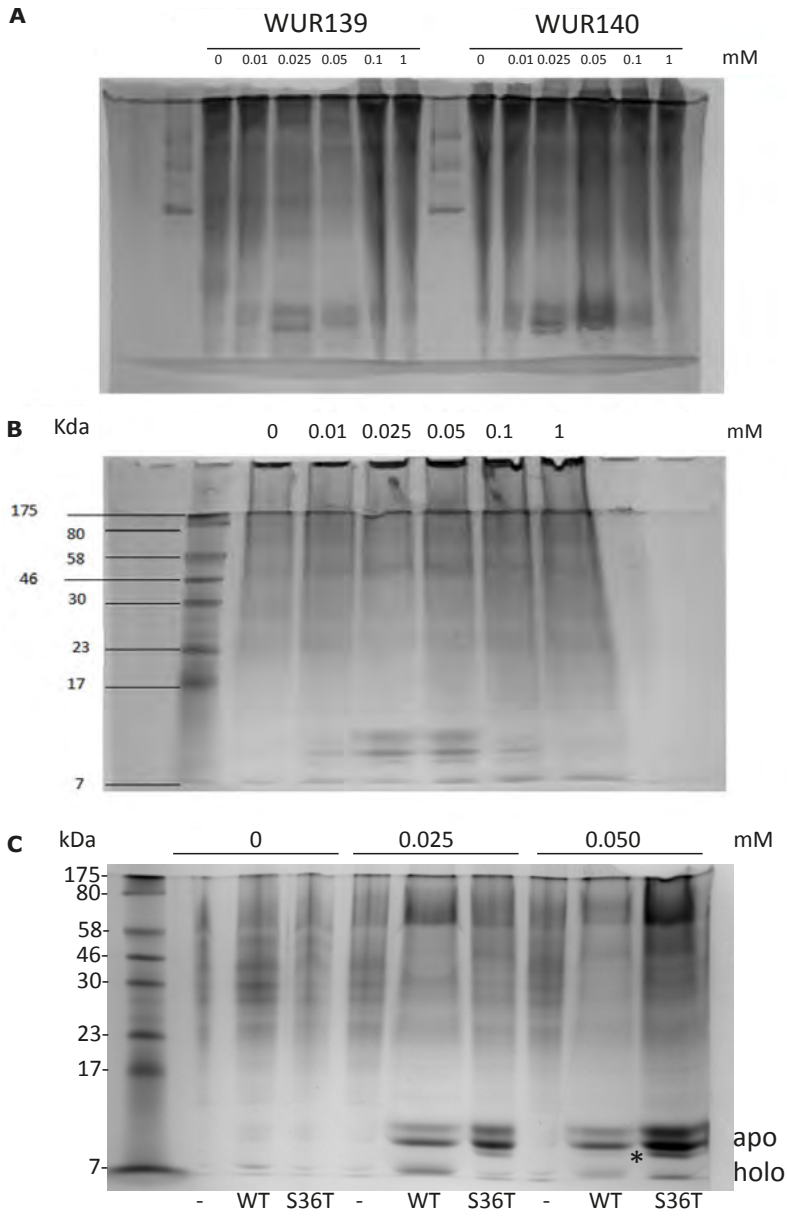


Figure S4. Urea-PAGE analysis of the ACP pool A: Urea PAGE showing cleared lysate fractions from WUR139 (ΔplsX, pACYC - acpP) and WUR140 (ΔplsX, pACYC – acpP S36T) at various IPTG concentrations. *acpP* S36T is a mutant in which the serine-36 pantetheinylation is disabled causing an apo-ACP accumulation. **S4B**: Urea PAGE showing WUR140 inclusion body aggregates resolubilized in lysis buffer and by sonication. ACP aggregates are abundant at concentrations of 0.025 and 0.05 mM IPTG. Less ACP fragments (soluble and insoluble) are present at concentrations >0.1 mM IPTG indicating a correlation between growth inhibition and ACP overexpression. **S4C**: Urea-PAGE showing overexpression of ACP and ACP-S36T at 0, 0.025 and 0.05 mM IPTG induction. Protein samples were normalized to 20 µg. An increased production of apo-ACP is shown. Holo-ACP is likely caused by endogenous expression. The asterisk indicates the position of a partially (un)-folded or variant of undescribed nature.

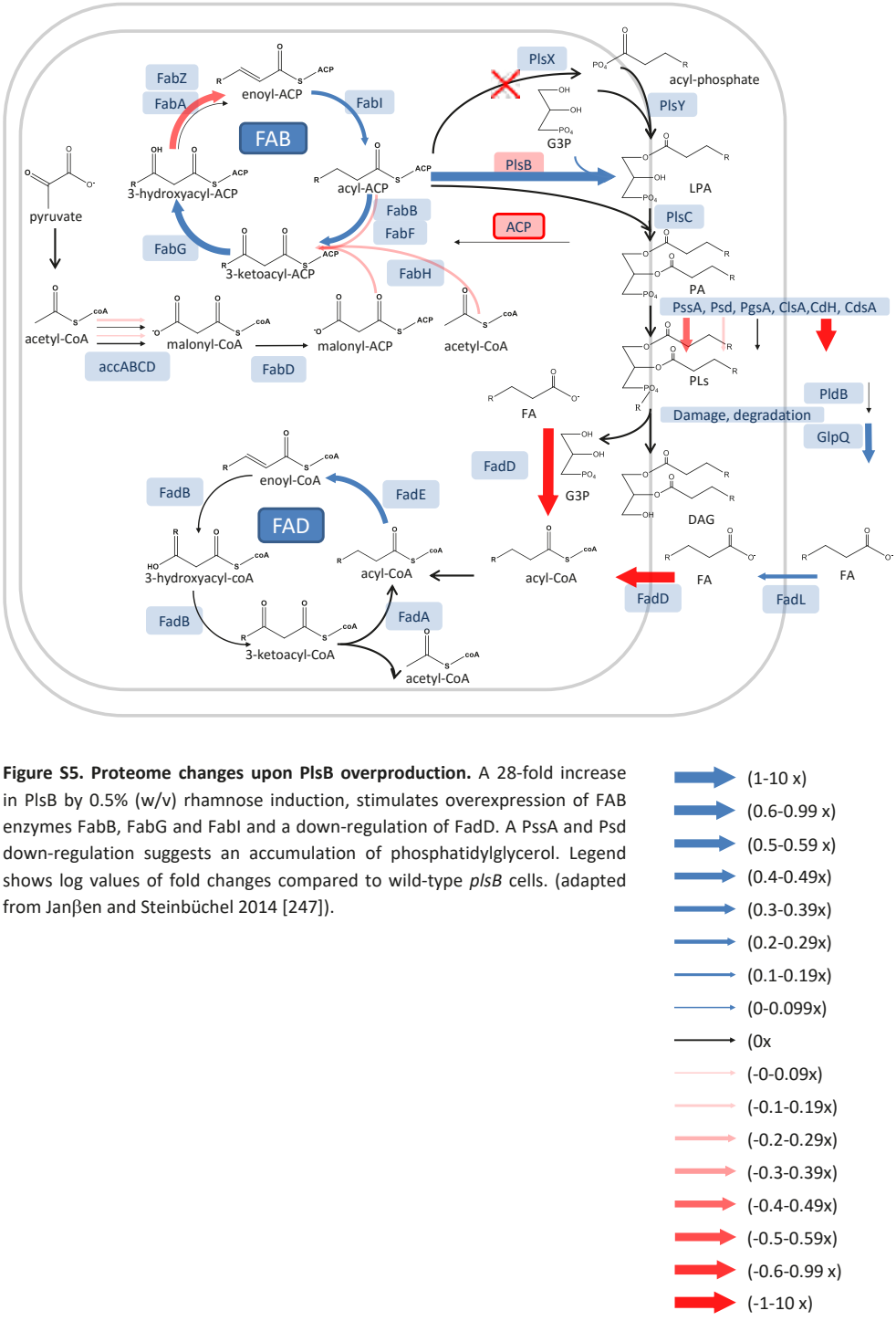


Figure S5. Proteome changes upon PlsB overproduction. A 28-fold increase in PlsB by 0.5% (w/v) rhamnose induction, stimulates overexpression of FAB enzymes FabB, FabG and FabI and a down-regulation of FadD. A PssA and Psd down-regulation suggests an accumulation of phosphatidylglycerol. Legend shows log values of fold changes compared to wild-type *plsB* cells. (adapted from Janβen and Steinbüchel 2014 [247]).

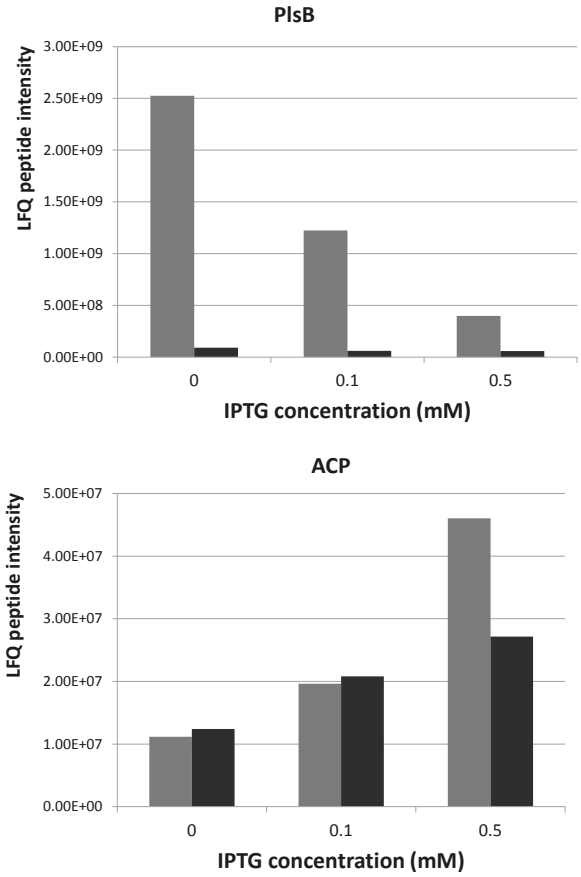


Figure S6. Proteome values (LFQ peptide intensities) of ACP and PlsB. PlsB overexpression is shown by rhamnose addition at a maximum of 28-fold increase. Conversely, PlsB pools are decreased by the presence of high ACP pools despite equal rhamnose induction. ACP levels are produced in higher quantities in the rhaP$_{BAD}$- *plsB* strain at 0.5 mM IPTG induction. Grey bars: rhaP$_{BAD}$ - *plsB*, black bars: wild-type *plsB*.

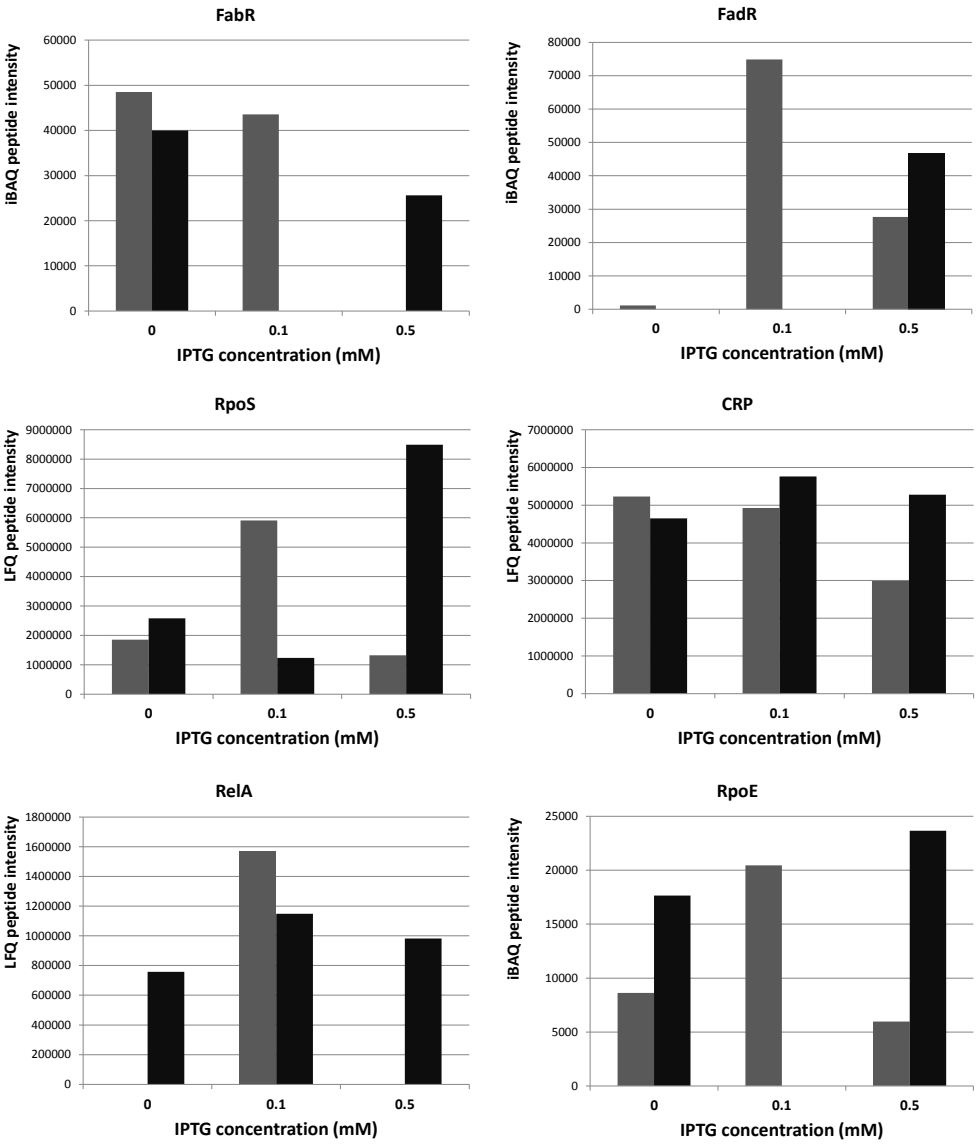
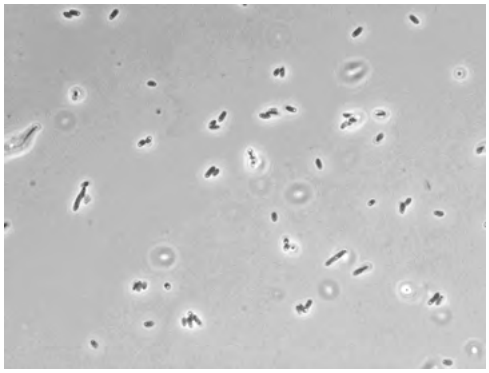


Figure S7. Proteome values of transcription factors FabR, FadR, RpoS, CRP, RelA and RpoE. In grey: rhaP$_{BAD}$-*plsB*, in black: wild-type *plsB.*

Δ*plsX, ΔDgkA,* rhaP$_{BAD}$- *plsB*, at 0.5% rhamnose



Δ*plsX, ΔDgkA,* rhaP$_{BAD}$- *plsB*, at 20 mM Glucose

Figure S8. Light microscopy pictures of *E.coli* cultures. PlsB overexpression by 0.5% (w/v) rhamnose addition changes the cell morphology to elongated cells. PlsB repression by 20 mM glucose addition changes cell morphology to coccoid shaped cells. Cells shown were taken from the same magnification of 1000 x.

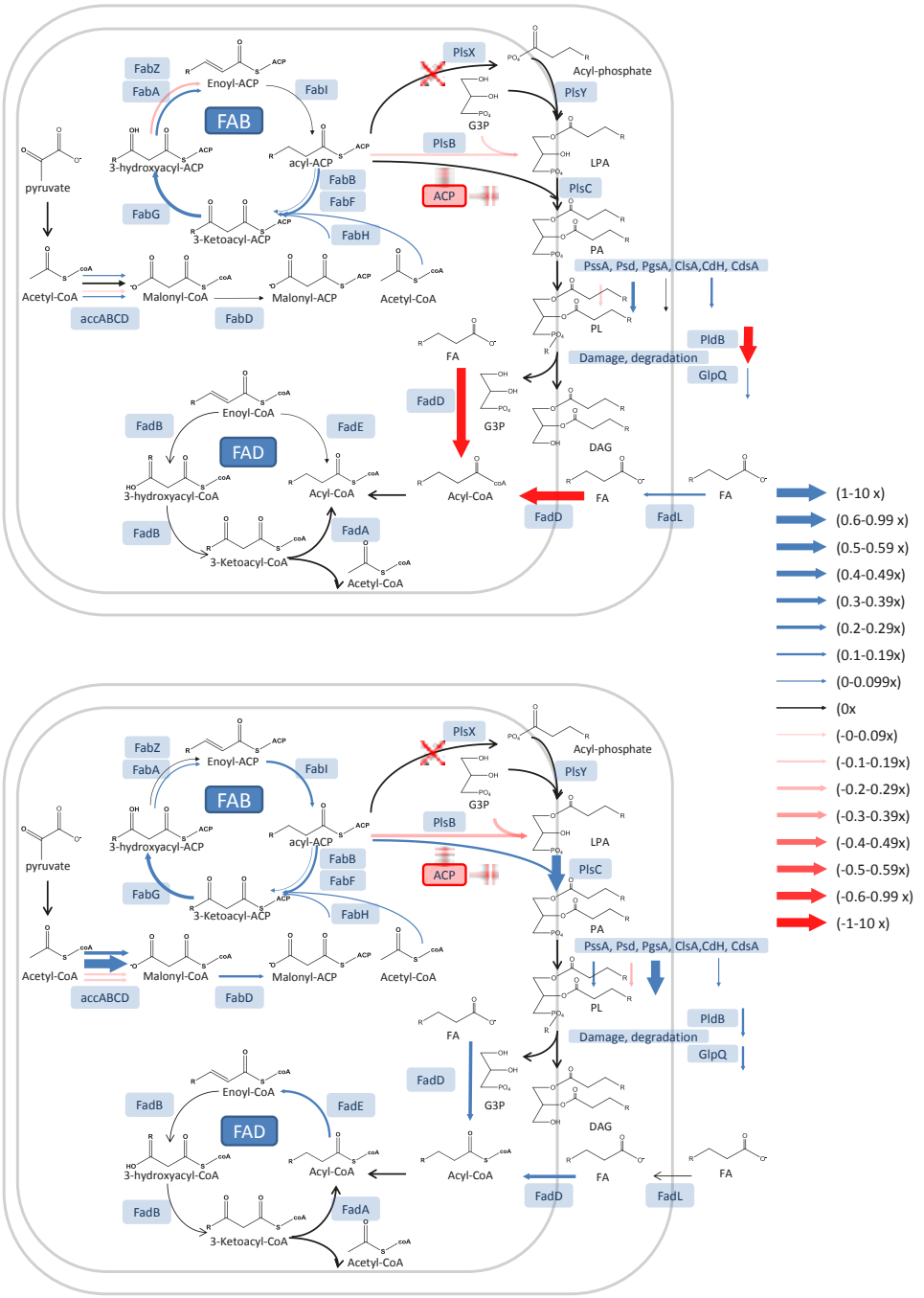


Figure S9. Wild-type plsB strain proteome Representation of fatty acid and phospholipid synthesis and degradation pathways at 0.1 mM (upper panel) and 0.5 mM (lower panel) IPTG induction of *acpP*. Legend shows log values of fold changes compared to non-induced sample. Blue: upregulation, Red: downregulation. (adapted from Janβen and Steinbüchel 2014 [247])

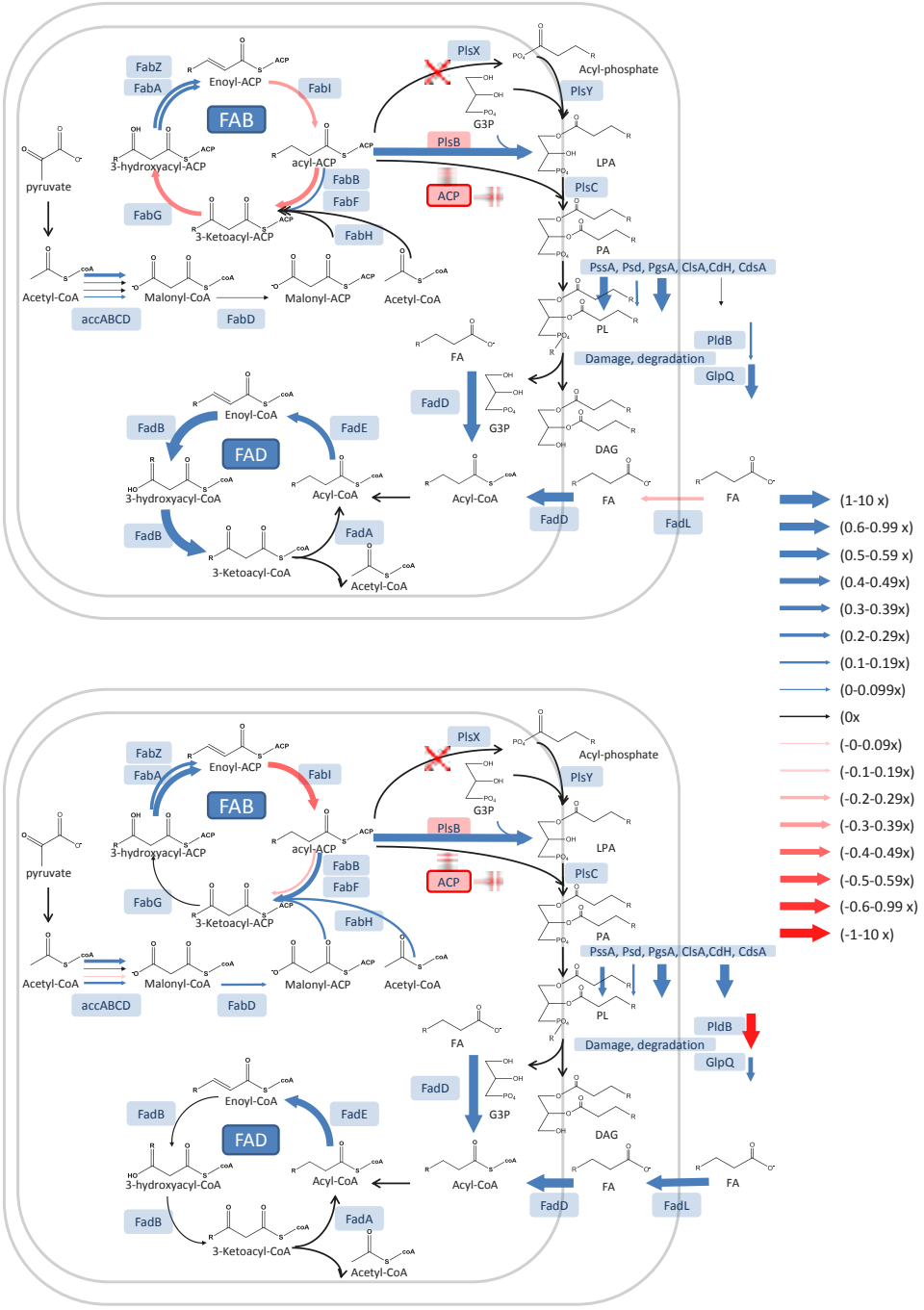


Figure S10. RhaP$_{BAD}$ - *plsB* strain proteome. Representation of fatty acid and phospholipid synthesis and degradation pathways at 0.1 mM (upper panel) and 0.5 mM (lower panel) IPTG induction of acpP. PlsB overproduction by 0.5% (w/v) rhamnose induction. Legend shows log values of fold changes compared to non-induced sample. Blue: upregulation, Red: downregulation. (adapted from Janβen and Steinbüchel 2014 [247])

GGCCGTCAAGGCCGCATAAGCTTGGTGTCTTTTTTACCTGTTTGACCCTGCAG*TACCGTTC*
*GTATAATGTATGCTATACGAAGTTAT*AGATCTCTATTTGTTTATTTTTCTAAATACATTCAAAT
ATGTATCCGCTCATGAGACAATAACCCTGATAAATGCTTCAATAATATTGAAAAAGGAAG
AGTATGAGCCATATTCAACGGGAAACGTCTTGCTCTAGGCCGCGATTAAATTCCAACATG
GATGCTGATTTATATGGGTATAAATGGGCTCGCGATAATGTCGGGCAATCAGGTGCGACA
ATCTATCGATTGTATGGGAAGCCCGATGCGCCAGAGTTGTTTCTGAAACATGGCAAAGGT
AGCGTTGCCAATGATGTTACAGATGAGATGGTCAGACTAAACTGGCTGACGGAATTTATG
CCTCTTCCGACCATCAAGCATTTTATCCGTACTCCTGATGACGCATGGTTACTCACCACTG
CGATCCCCGGGAAAACAGCATTCCAGGTATTAGAAGAATATCCTGATTCAGGTGAAAATA
TTGTTGATGCGCTGGCAGTGTTCCTGCGCCGGTTGCATTCGATTCCTGTTTGTAATTGTCCT
TTTAACAGCGACCGCGTATTTCGTCTCGCTCAGGCGCAATCACGAATGAATAACGGTTTG
GTTGATGCGAGTGATTTTGATGACGAGCGTAATGGCTGGCCTGTTGAACAAGTCTGGAAA
GAAATGCACAAACTTTTGCCATTCTCACCGGATTCAGTCGTCACTCATGGTGATTTCTCAC
TTGATAACCTTATTTTTGACGAGGGGAAATTAATAGGTTGTATTGATGTTGGACGAGTCG
GAATCGCAGACCGATACCAGGATCTTGCCATCCTATGGAACTGCCTCGGTGAGTTTTCTC
CTTCATTACAGAAACGGCTTTTTCAAAAATATGGTATTGATAATCCTGATATGAATAAATT
GCAGTTTCATTTGATGCTCGATGAGTTTTTCTAAGTCGAC*ATAACTTCGTATAATGTATGCTA*
*TACGAACGGTA*GCGGCCGCCAACTCCTTCACCAGAGGTAGGAATTCCTGGGCCTCATGGGC
C

Figure S11. Disruption cassette sequence and design. *lox71-kanR-lox66* disruption cassette. Underlined and italic nucleotides indicate restriction sites and lox sites, respectively. The cassette contains from 5' to 3' a SfiI site (used for cloning by Geneart), a HindIII site, a site for primer annealing, a PstI site, *lox71*, a BglII site, *kanR*, a SalI site, *lox66*, a NotI site, a site for primer annealing, an EcoRI site, and a SfiI site (used for cloning by Geneart).

Figure S12. Proteomics raw data excel file. Quantification of 1255 proteins in reference to E.coli K12 proteome by Maxquant version 1.5.0.0. software package

Chapter 5

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

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

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

[2]Stratingh Institute for Chemistry, University of Groningen, Nijenborgh 7, 9747 AG Groningen, The Netherlands

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

[4]Current address: Department of Medicine, Section of Infectious Diseases, Boston University School of Medicine, 02118 Boston, Massachusetts, USA

Abstract

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

Keywords: archaea; ether lipid biosynthesis; membrane proteins; Liquid chromatography mass spectrometry

Introduction

The cytoplasmic membrane is an essential constituent of cells. It forms a barrier that separates the cytosol from the external milieu. In conjunction with energy transducing complexes and transporter proteins, the phospholipid bilayer allows a homeostasis of the intracellular concentration of nutrients and other metabolites within the cell[289, 290]. The lipid composition of the cytoplasmic membrane differs between organisms, and is one of the elements that distinguishes Archaea from Bacteria and Eukarya. This marked diversion of the lipid composition is also termed "the lipid divide" that may finds its origin in the early stages of the evolution of life[14, 17, 291, 292].

Archaeal lipids are composed of highly branched isoprenoid chains ether-linked to a glycerol-1-phosphate (G1P) backbone, compared to fatty acid chains ester-linked to a glycerol-3-phosphate (G3P) backbone as typically found in bacteria and eukarya. Besides these main characteristics, archaeal membranes display a further diversity in their lipid composition consisting of different modifications of the two major structures: *sn*-2,3-diphytanylglycerol diether, called archaeol and *sn*-2,3-diphytanylglycerol tetraether lipid, known as caldarchaeol [292-294]. The biosynthetic pathway leading to the formation of archaeal lipids has been studied in some detail [9, 33, 215, 295-297] and most of the enzymes involved in the biosynthesis have been identified and characterized. However, the entire pathway is not completely understood, nor is it clear how caldarchaeol is formed. The isoprenoid building blocks isopentenyl pyrophosphate (IPP) and dimethyallyl pyrophosphate (DMAPP) are synthesized via the mevalonate pathway in Archaea [215] and are combined through sequential condensation reactions catalyzed by geranylgeranyl diphosphate (GGPP) synthase and farnesylgeranyl diphosphate (FGPP) synthase, depending on the length of the isoprenoid chain product [298, 299]. Glycerol-1-phosphate (G1P) in archaea and glycerol-3-phosphate (G3P) in Bacteria and Eukarya are synthesized by similar reactions although the enzymes involved, i.e., glycerol-1-phosphate dehydrogenase (G1PDH)[300, 301] and glycerol-3-phosphate dehydrogenase (G3PDH), are not evolutionary related and belong to different protein families [295]. The elongated isoprenoid chain and G1P are subsequently linked together through ether linkages by two prenyltransferases. A cytosolic protein geranylgeranylglyceryl phosphate synthase (GGGP synthase) selectively attaches the isoprenoid chain to the G1P[302, 303] leading to the first ether bond formation. Next, the di-*O*-geranylgeranylglyceryl phosphate synthase (DGGGP synthase)[304-306] catalyzes the second ether bond formation step linking another isoprenoid chain to the lipid precursor yielding DGGGP. The following step concerns the activation of DGGGP via a CTP-transferring reaction by CDP-archaeol synthase (CarS) that was recently discovered[306]. CDP-archaeol is an important intermediate for the successive steps in lipid biosynthesis where the CDP group is replaced by a polar head group. Serine, ethanolamine, glycerol and *myo*-inositol are common polar head groups found throughout the three domains of life.

The enzymes involved in the initial replacement of CMP from a CDP-alcohol with different types of polar head groups share mostly a common mode of action among Eukarya, Bacteria and Archaea (**Fig. 1**). A different mechanism to produce phosphatidylethanolamine (PE) has been described for Eukarya that involves a direct replacement of the CDP group with ethanolamine[307, 308].

The *E. coli* membrane is composed by 70-80% of phosphatidylethanolamine (PE), 20-25% of phosphatidylglycerol (PG) and 5% or less of cardiolipin as the major phospholipid components[289, 290, 309]. PG represent the main anionic phospholipid and is important for various cellular processes such as the initiation of DNA synthesis[290, 310-312] and protein translocation[313-315], whereas PE is the main non-bilayer lipid. In order to re-program *E. coli* for the synthesis of archaeal ether lipids, it will be essential to produce the two archaeal lipids archaetidylethanolamine (AE) and archaetidylglycerol (AG) to accommodate at least the polar head group composition. For polar head group modification, the branch point in bacteria lies with CDP-diacylglycerol (CDP-DAG), produced by CDP-diacylglycerol synthase encoded by the *cdsA* gene[316]. CDP-DAG acts as a substrate for two sets of enzymes. For the formation of PE, the CDP group of CDP-DAG is replaced with L-serine by phosphatidylserine synthase (Pss) leading to the production of phosphatidylserine (PS). Next, PS is decarboxylated by phosphatidylserine decarboxylase (Psd) converting it into PE. In the other biosynthetic pathway the CDP group of CDP-DAG is initially replaced by a glycerol-3-phosphate (G3P) moiety by phosphatidylglycerol synthase (Pgs) leading to the formation of phosphatidylglycerol phosphate (PGP). The enzyme phosphatidylglycerol phosphatase (Pgp) removes the phosphate resulting in the formation of PG[246, 289, 290, 309]. In Archaea, the formation of the corresponding archaeal lipids, archaetidylethanolamine (AE) and archaetidylglycerol (AG) seems to take place via very similar mechanisms (**Fig. 1**). The two biosynthetic branches diverge from the CDP-archaeol towards the formation of AE or AG. The former is produced by the sequential action of two enzymes, archaetidylserine synthase (Ass)[317] and archaetidylserine decarboxylase (Asd) which replace the CDP group with L-serine whereupon a decarboxylation reaction yields the AE. Likewise, AG synthesis involves the enzyme archaetidylglycerol phosphate synthase (Ags) which attaches a G3P to CDP-archaeol with the consequent formation of archaetidylglycerol phosphate (AGP) which is subsequently dephosphorylated into archaetidyglycerol (AG) by the action of archaetidylglycerol phosphatase (Agp)[318, 319].

The bacterial and archaeal synthase enzymes involved in polar head group attachment all contain a well conserved domain, identified as D-G-x(3)-D-x(3)-D in the PROSITE database that classify these proteins as members of CDP-alcohol phosphatidyltransferase family [7, 318]. Previous bioinformatics analysis [7, 318, 320] revealed a wide distribution of these enzymes in Bacteria and Archaea suggesting the existence of an ancestral enzyme in the last universal common ancestor (LUCA)[7] able to produce both archaeal- and bacterial-like lipids.

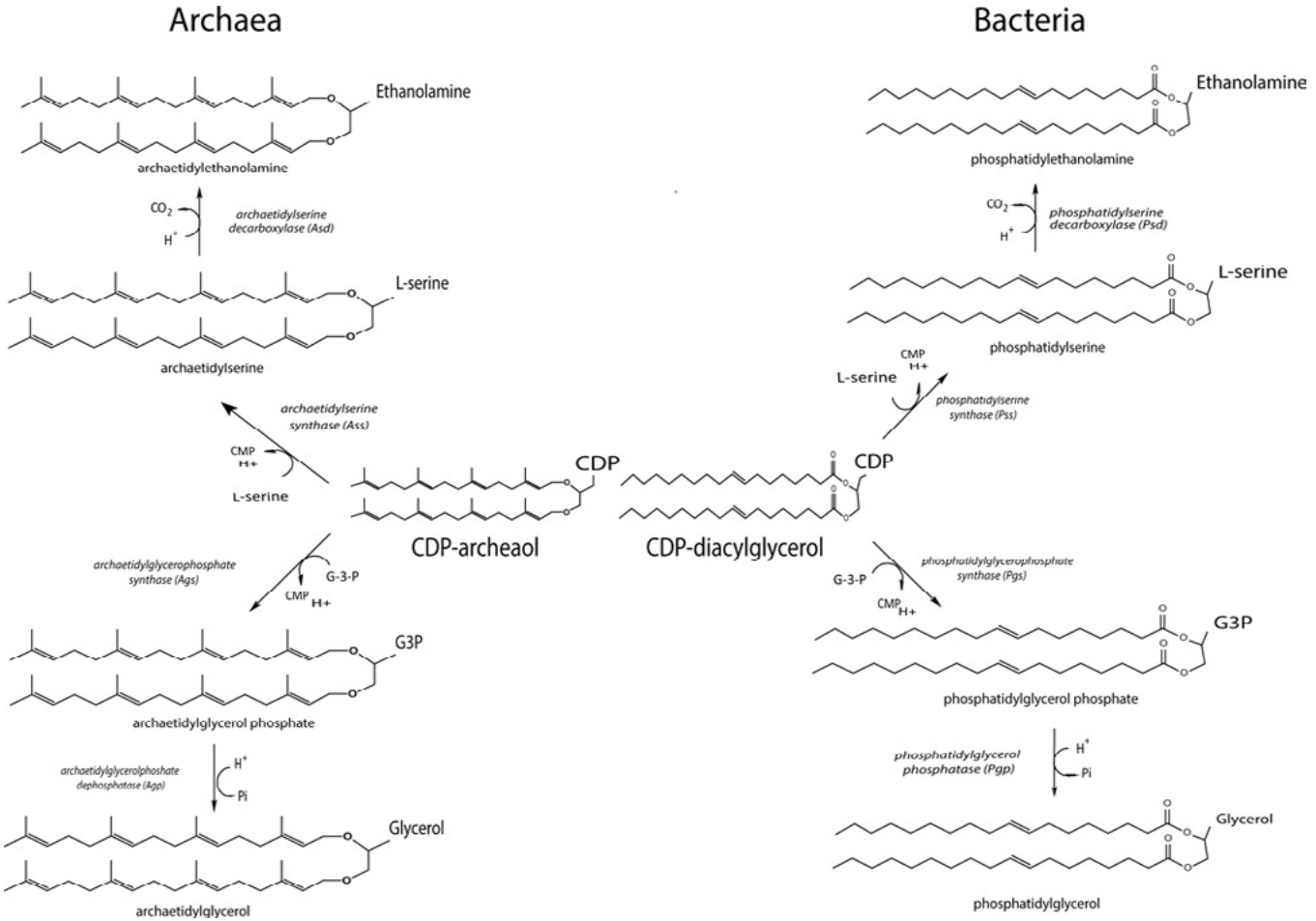


Figure 1. Polar head group attachment in bacteria and archaea. The scheme represents the enzymatic steps of the ether lipid biosynthetic pathway for polar head group attachment in Archaea and Bacteria. The archaeal enzymes involved in the replacement of CDP-group with L-serine (green) or glycerol (blue) from CDP-archaeol are colored in light blue whereas the corresponding bacterial enzymes are in violet.

Therefore, this could indicate substrates promiscuity among these enzymes and possibly bacterial enzymes are able to catalyze polar head group attachment to the CDP-archaeol. In this study, we have tested various bacterial enzymes for the formation of AE and AG from CDP-archaeol and *in vitro* reconstituted the entire pathways employing simple building blocks and up to nine purified enzymes from archaeal and bacterial origin. In addition, we have introduced the archaeal ether lipid biosynthetic pathway for AE and AG synthesis into *E. coli* making use of the substrate promiscuity of bacterial enzymes catalyzing the final polar head group conversion steps.

Results

Archaetidylglycerol formation by bacterial PgsA and PgpA enzymes

PG synthesis in *E. coli* proceeds via two enzymatic steps. The first reaction is catalyzed by phosphatidylglycerophosphate synthase. *E. coli* contains two individual genes that encode this activity [321, 322]. The *pgsA* gene exclusively functions in phosphatidylglycerol formation [323], whereas the *pgsB* gene is also involved in the synthesis of the Lipid A core of Lipopolysaccharide (LPS)[324]. The resulting phosphatidylglycerol phosphate (PGP) is dephosphorylated to phosphatidylglycerol. In *E. coli* three main phosphatidylglycerol phosphatases are found, encoded by the *pgpA*, *pgpB* and *pgpC* genes [325]. PgpA exhibits a narrow substrate specificity dephosphorylating only PGP[326] while PgpB also shows phosphatase activity towards diacylglycerol pyrophosphate [327], phosphatidic acid and lysophosphatidic acid [328, 329] revealed the occurrence of Pgs homologs in Archaea that belong to CDP-alcohol phosphatidyltransferase family, suggesting functional and structural conservation of these enzymes among bacteria and archaea. The same analysis revealed a weak sequence conservation of Pgp in Archaea, limited to PgpA and PgpB only. To investigate the ability of bacterial enzymes to accept CDP-archaeol as a substrate for archaetidylglycerol (AG) formation, the *pgsA* and *pgpA* genes of *E. coli*, were cloned and overexpressed in *E. coli* BL21. Both overexpressed proteins localized to the membrane fraction after cell fractionation. Upon membrane solubilization with the detergent n-dodecyl-β-D-maltopyranoside (DDM), PgsA and PgpA could be purified by Ni-NTA affinity chromatography (**Fig. 2A**). The activity of the enzymes was analyzed in *in vitro* coupled reactions in detergent solution using LC-MS for detection (**Fig. 2D**). The activity of the *E. coli* PgsA towards the archaeal substrates CDP-archaeol (CDP-ol) was tested using chemically synthesized DGGGP as a substrate that was converted into CDP-ol in the presence of the *A. fulgidus* CarS and CTP as described previously [306] (**Fig. 2B**, lane 2). Upon the addition of the *E. coli* PgsA and G3P, the formation of AGP could be demonstrated (m/z= 869.51 $[M-H]^-$) (**Fig. 2B**, lane 3). In the subsequent reaction the conversion of AGP into AG (m/z= 789.55 $[M-H]^-$) was observed when also the *E. coli* PgpA was included in the reaction (**Fig.**

2B, lane 4). The *E. coli* PgsA was unable to use archaeal DGGGP as substrate (**Fig. 2B**, lane 5). As control, the activity of PgsA and PgpA was also tested towards the bacterial substrate CDP-diacylglycerol (CDP-DAG). The latter was produced by incubating the purified *E. coli* CdsA protein with its substrates phosphatidic acid (PA, 18:1) and CTP (**Fig. 2C**, lane 2). The formation of PGP (m/z= 853.50 $[M-H]^-$) was detected only in presence of the *E. coli* PgsA and G3P (**Fig. 2C**, lane3), which was further converted into the final product PG (m/z= 773.54 $[M-H]^-$) upon the addition of the *E. coli* PgpA (**Fig. 2C**, lane 4). These results demonstrate that the *E. coli* PgsA and PgpA recognize and convert the archaeal substrate CDP-archaeol into archaetidylglycerol phosphate and archaetidylglycerol, respectively.

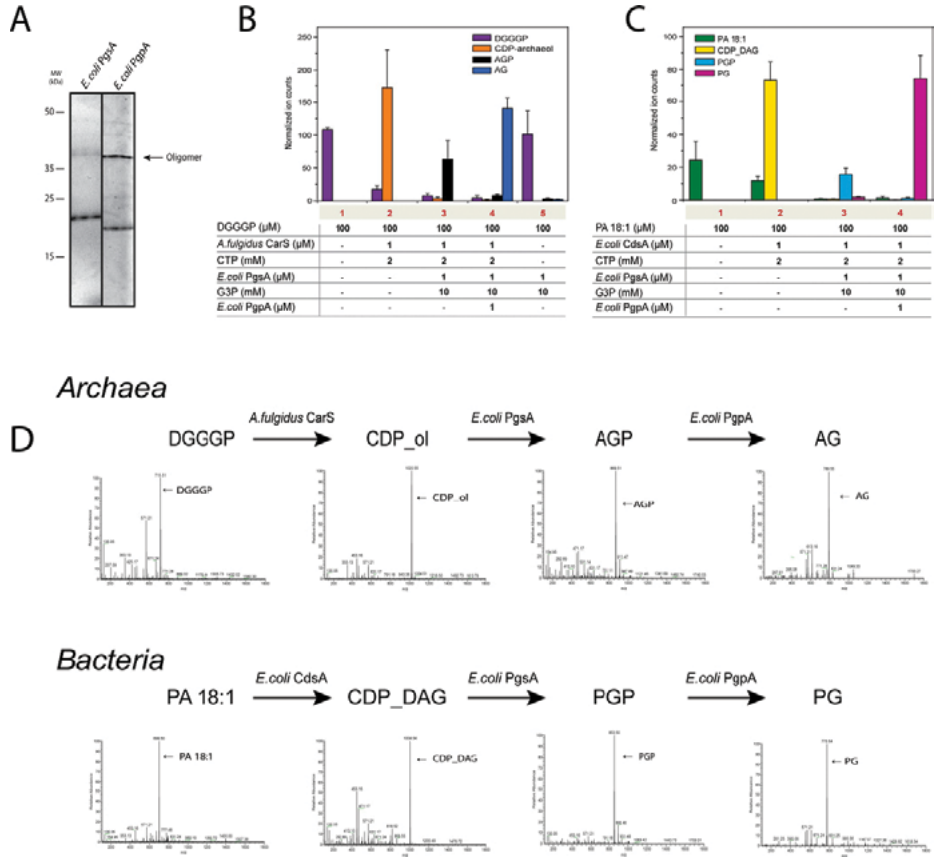


Figure 2. ***In vitro*** **demonstration of AG biosynthesis involving the** ***E. coli*** **PgsA and PgpA. (A)** Coomassie stained SDS-PAGE gels showing the Ni-NTA purified proteins PgsA (21 kDa) and PgpA (20.5 kDa) from the bacteria *E. coli*. *In vitro* reactions using **(B)** DGGGP or **(C)** PA (C18:1,C18:1) and the purified proteins as specified, to test the substrate specificity of the *E. coli* PgsA and PgpA. Total ions counts from LC-MS data were normalized using DDM as internal standard. The data are the average of three experiments ± SE. **(D)** Schematic representation of the *in vitro* reactions. The purified enzymes used in the experiments are highlighted in blue and the mass spectra from the LC-MS runs of the corresponding products are depicted in the red boxes: DGGGP (m/z= 715.51 [M-H]-), CDP_ol (m/z= 1020.54 [M-H]-), AGP (m/z= 869.51 [M-H]-) and AG (m/z= 789.55 [M-H]-) in panel B and PA 18:1 (m/z= 699.50 [M-H]-), CDP_DAG (m/z= 1004.54 [M-H]-), PGP (m/z= 853.50 [M-H]-) and PG (m/z= 773.54 [M-H]-) in figure C.

Archaetidylethanolamine formation by bacterial PssA and Psd enzymes

Like PG synthesis, phosphatidylethanolamine (PE) synthesis also requires the action of two enzymes: phosphatidylserine synthase (Pss) for the synthesis of phosphatidylserine (PS) which is further decarboxylated to PE by phosphatidylserine decarboxylase (Psd)[330]. Two different subclasses of phosphatidylserine synthases exist: Pss-I, a cytoplasmic protein present mainly in Gram-negative bacteria such as *E. coli*, and Pss-II, a membrane protein that is found in Gram-positive bacteria, yeast[331] and archaea[332]. The Pss-II enzymes contain a highly conserved domain present in CDP- alcohol phosphatidyltranferases[332]. Previously, archaetidylserine synthase (Ass) activity was identified in the membrane fraction of *Methanothermobacter thermautotrophicus*[317]. Secondary structure analysis of the archaeal members of the CDP-alcohol phosphatidyltransferase family indicates the presence of eight conserved transmembrane domains (TMDs) comparable to the corresponding bacterial protein family (**Fig. 3A**).

Interestingly, the PssA sequence of *B. subtilis* is substantially smaller than the other members of this family (only 177 amino acids instead of 451 amino acids of the *E. coli* Pss-II), but it shows the core of five TMDs (**Fig. 3B**) but lacking two C- terminal TMDs. A membrane fraction of *B. subtilis* incubated with the different archaeal-like substrates showed archaetidylserine (AS) formation whereas such activity could not be demonstrated in *E. coli* [317] suggesting a more narrow substrate specificity of the *E. coli* Pss-I and Pss-II enzymes. Therefore, the *PssA* gene of *B. subtilis* was cloned and overexpressed in *E. coli* BL21 strain under the control of T7 promoter. The protein was solubilized from the membrane with DDM and purified by Ni-NTA affinity chromatography (**Fig. 3C**). Coupled *in vitro* reactions were performed using DGGGP as initial substrate along with the *A. fulgidus* CarS and CTP. Products were isolated through *n*-butanol extraction and analyzed by LC-MS (**Fig. 3F**). In the presence of the *B. subtilis* PssA, AS (m/z= 802.53 [M-H]-) formation was observed in the presence of L-serine and Mg^{2+} (**Fig. 3D**, lane 3). PssA showed no activity towards DGGGP indicating the strict requirement of a CDP-activated intermediate for the reaction (**Fig. 3D**, lane 4). Moreover, no AS was formed in presence of EDTA (**Fig. 3D**, lane 5) which chelates divalent ions, consistent with a requirement of Mg^{2+} for enzymatic activity[333]. As a control, formation of phosphatidylserine by *B. subtilis* PssA was demonstrated in the presence of phosphatidic acid (PA, 18:1), CTP and the *E. coli* CdsA (**Fig. 3D**, lane 6). To examine the activity of the *E. coli* PssA towards CDP-archaeol, the enzyme was also overexpressed and purified (**Fig. 3C**). In the presence of CDP-archaeol as substrate, no AS formation was observed demonstrating that the *E. coli* PssA (the Pss-II enzyme) indeed does not recognize the archaeal precursor (**Fig. 3E**, lane 1). The enzyme, however, converted CDP-DAG to phosphatidylserine (m/z= 786.53 [M-H]-) in a coupled enzyme assay using phosphatidic acid as substrate (**Fig. 3E**, lane 4).

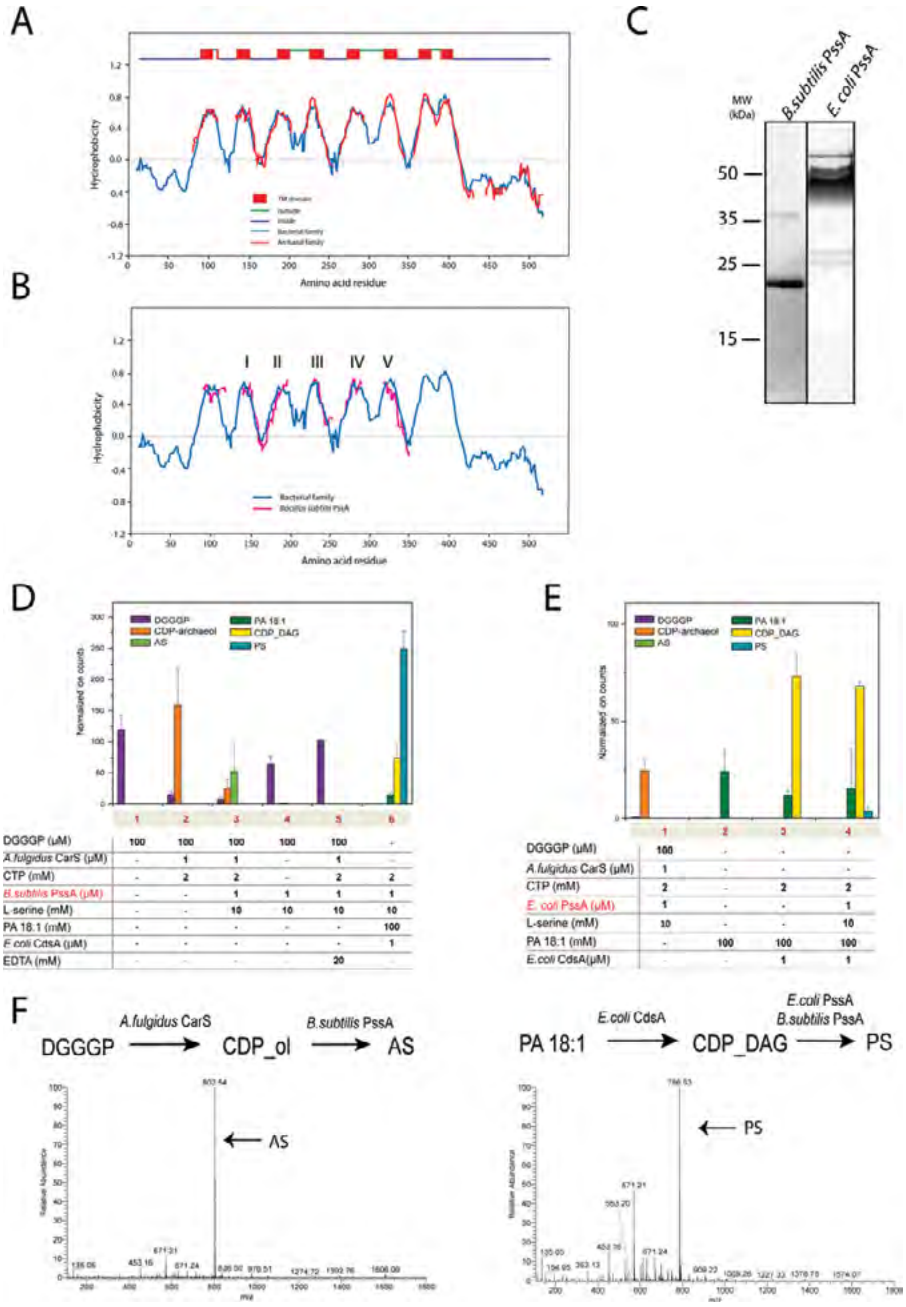
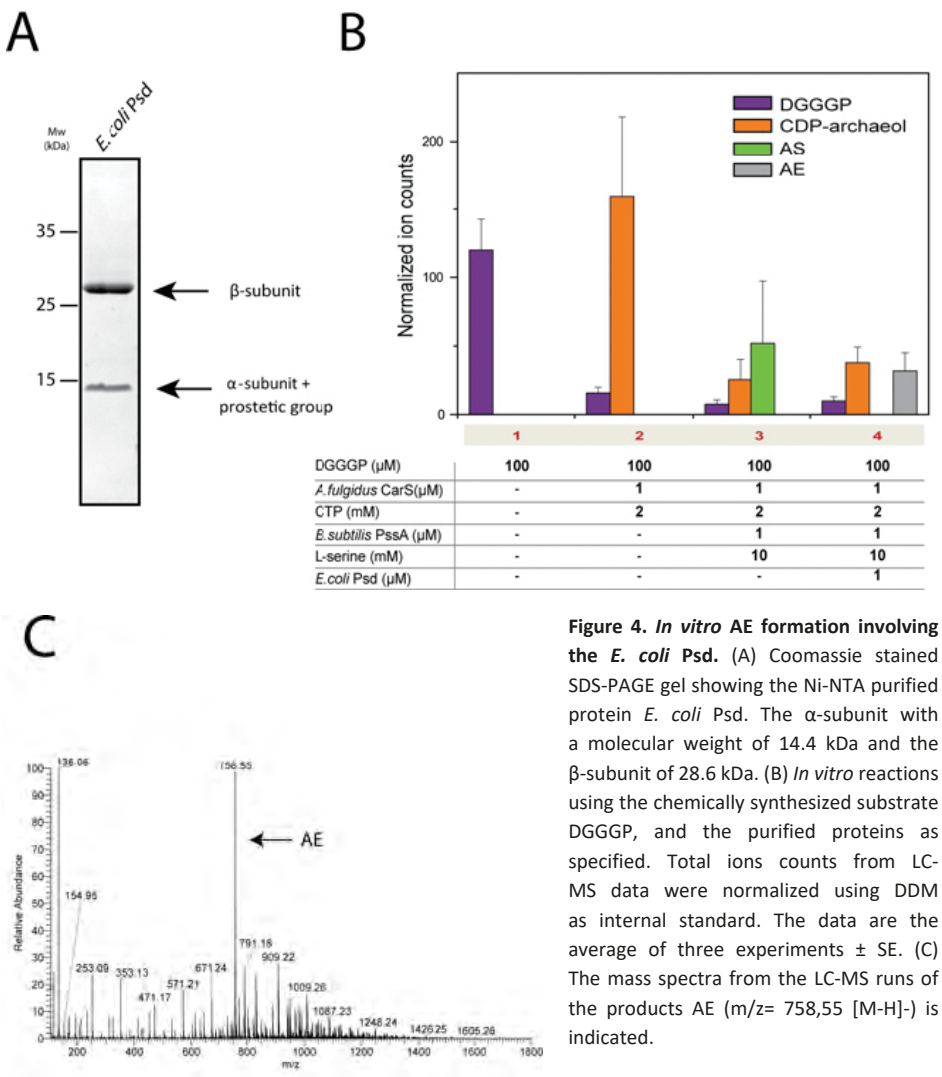


Figure 3. ***In vitro*** **AS synthesis by the** ***B. subtilis*** **PssA**. **(A)** Averaged hydropathy profile of the bacterial Pss proteins (blue line) and archaeal ones (red line) is aligned highlighting their conservations in the two kingdoms. The membrane topology prediction is depicted above the plot. **(B)** Hydropathy profile alignment of *B. subtilis* PssA (purple line) and the averaged hydropathy profile of the bacterial Pss protein family (blue line). The conserved transmembrane domains (TMDs) are indicated by Roman numbers. **(C)** Coomassie stained SDS-PAGE gels showing the Ni-NTA purified proteins *B. subtilis* PssA (21 kDa) and *E. coli* PssA (53.6 kDa). Specificity of the bacterial *B.*

subtilis PssA **(D)** and *E. coli* PssA **(E)** towards DGGGP and PA (C18:1,C18:1) as assessed by means of an *in vitro* assay using the purified enzymes. Total ions counts from LC-MS data were normalized using DDM as internal standard. The data are the average of three experiments ± SE. **(F)** Schematic representation of the performed *in vitro* reactions. The mass spectra from the LC-MS runs of the two products AS (m/z= 802.54 [M-H]-) and PS (m/z= 786.53 [M-H]-) are indicated in the boxes.



DGGGP (μM)	100	100	100	100
A.fulgidus CarS(μM)	-	1	1	1
CTP (mM)	-	2	2	2
B.subtilis PssA (μM)	-	-	1	1
L-serine (mM)	-	-	10	10
E.coli Psd (μM)	-	-	-	1

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

In Archaea and Bacteria, AE and PE are produced by a decarboxylation reaction that removes the carboxylic acid group of L-serine. In *E. coli*, the phosphatidylserine decarboxylase is encoded by the *psd* gene that specifies a membrane-associated pro-enzyme which undergoes an autocatalytic internal cleavage [334] leading to two subunits – the α subunit containing a pyruvoyl prosthetic group and a β subunit [335, 336]. Previous bioinformatics

analysis [318]. However, the archaeal Asd has not been biochemically characterized. Because of the general mechanism of the decarboxylation reaction [307], the possibility exists that the endogenous phosphatidylserine decarboxylase of *E. coli* is able to recognize AS. Therefore, the *E. coli psd* gene was overexpressed in *E. coli* BL21 strain under the control of T7 promoter. Membranes bearing overexpressed levels of Psd were solubilized with DDM and the enzyme was purified by Ni-NTA affinity chromatography (**Fig. 4A**).

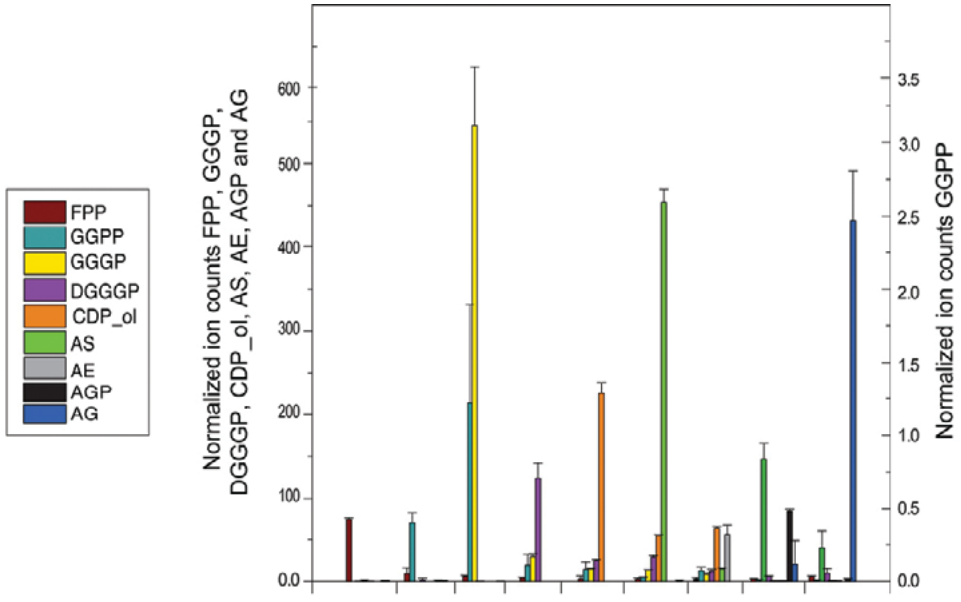
The α-subunit showed a slower migration on SDS-PAGE, as expected on the basis of its predicted molecular mass. This is likely due to the presence of the pyruvoyl prosthetic group attached. [335] The enzymatic activity of the *E. coli* Psd was tested by *in vitro* coupled reactions as described above and product formation was detected by LC-MS (**Fig. 4C**). Conversion of AS to AE (m/z= 758.55 $[M-H]^-$) by the *E. coli* Psd was observed in the presence of Mg^{2+}, L-serine, *B. subtilis* PssA, *A. fulgidus* CarS, CTP and DGGGP (**Fig. 4B**, lane 4).

In vitro reconstitution of archaetidylethanolamine and archaetidylglycerol formation

Previously, we have described the *in vitro* reconstitution of CDP-archaeol formation starting from the precursors IPP and DMAPP/FPP [306] using two bacterial enzymes and three archaeal enzymes that were overexpressed and purified from *E. coli* upon codon optimization. Due to the wider substrate

specificity described above, the extension of this pathway for the reconstitution of the formation of AE and AG requires four additional bacterial enzymes described in the previous section. The ether lipid biosynthetic pathway reconstitution [306] starts from the two isoprenoid building blocks IPP and FPP which undergo several cycles of condensation leading to isoprenoid chains with the required C_{20} length [337]. GGPP (m/z= 449.19 $[M-H]^-$) formation was detected by LC-MS with an *E. coli* farnesyl diphosphate synthase (IspA) [338] mutant as described previously (**Fig. 5**, lane 2).[306] Despite the unique feature of archaeal G1P dehydrogenases to synthesize the G1P, some bacterial enzymes are also able to perform this reaction [339]. The enzyme AraM (G1PDH) from *B. subtilis* [340] produces the required glycerophosphate backbone which can be attached to the GGPP via the first ether bond leading to the synthesis of GGGP (m/z= 443.26 $[M-H]^-$). This latter reaction is catalyzed by the *M. maripaludis* GGGP synthase (MmarC7_1004) [302] which in combination with IspA and G1PDH and the substrates IPP, FPP, DHAP and NADH leads to the conversion of GGPP into GGGP (**Fig. 5**, lane 3). DGGGP production (m/z= 715.51 $[M-H]^-$) was observed in a subsequent reaction with the *A. fulgidus* DGGGP synthase (AF0404) [304] (**Fig. 5**, lane 4). Next, DGGGP is converted into CDP-archaeol (m/z= 1020.54 $[M-H]^-$) by the *A. fulgidus* CarS (AF1740) in the presence of CTP [306] (**Fig. 5**, lane 5). CDP-archaeol is the precursor for the formation of AS (m/z= 802.51 $[M-H]^-$) in the presence of the *B. subtilis* PssA and L-serine (**Fig. 5**, lane 6), which is further converted into AE (m/z= 758.54 $[M-H]^-$) (**Fig. 5**, lane 7) by the *E. coli* Psd. In another reaction, the CDP-archaeol is converted into AGP (m/z= 869.51 $[M-H]^-$

) by the *E. coli* PgsA in the presence of G3P (**Fig. 5**, lane 8). AGP is then converted into AG (m/z= 789.54 $[M-H]^-$) by the addition of the *E. coli* PgpA to the *in vitro* reactions (**Fig. 5**, lane 9). Taken together, the reactions described here employing purified enzymes represent the *in vitro* reconstitution of the entire archaeal lipid pathway using a set of archaeal and bacterial enzymes.



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

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

Archaetidylglycerol and archaetidylethanolamine formation in *E. coli*

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

Table 1. Expression vectors used in this study.

Plasmid	Description	Reference
pRSF-Duet-1	Cloning and expression vector (Kan^R), T7 promoter	Novagen
pET-Duet-1	Cloning and expression vector (Amp^R), T7 promoter	Novagen
pCDF-Duet-1	Cloning and expression vector (Str^R), T7 promoter	Novagen
pACYC-Duet-1	Cloning and expression vector (Cm^R), T7 promoter	Novagen
pSJ122	Synthetic gene encoding codon optimized DGGGP synthase from *A. fulgidus* with N-terminal His-tag and redesigned ribosome binding site AGGACGTTAACAT cloned into pRSF-Duet vector using the primers 41 and 42	Jain *et al.*[306]
pSJ135	PCR product of *ispA* gene with N-terminal His-tag from *E. coli K12* genomic DNA containing a double mutation Y79H and S140T. PCR product of *idi* gene with His-tag at the N-terminus from *E. coli K12* genomic DNA. Both genes were cloned into pCDF-Duet vector using the primers 62, 63, 24 and 57	This study
pSJ138	Synthetic gene encoding codon optimized GGGP synthase from *M. maripaludis* with N-terminal His-tag. PCR product of *araM* with C-terminal His-tag from *B. subtilis* genomic DNA. Both genes were cloned into pET-Duet vector using the primers 70, 71, 11 and 12	This study
pSJ140	Synthetic gene encoding codon optimized DGGGP synthase from *A. fulgidus* with N-terminal His-tag and redesigned ribosome binding site AGGACGTTAACAT. Synthetic gene encoding codon optimized CDP-archaeol synthase from *A. fulgidus* with C-terminal His-tag. Both synthetic genes are cloned into pRSF-Duet vector using the primers 32, 20, 84 and 86	This study
pSJ148	PCR product of *cdsA* with N-terminal His-tag from *E. coli* genomic DNA cloned into pACYC-Duet vector using the primers 103 and 106	This study
pAC004	PCR product of *pss* with C-terminus His-tag from *B. subtilis* genomic DNA cloned into pACYC-Duet vector using the primers 89 and 90	This study
pAC008	PCR product of *psd* with C-terminus His-tag from *E. coli* genomic DNA cloned into pACYC-Duet vector using the primers 533 and 534	This study
pAC011	PCR product of *pss* with C-terminus His-tag from *E. coli* genomic DNA cloned into pET-Duet vector using the primers 542 and 543	This study
pAC015	PCR product of *pgsA* with C-terminus His-tag from *E. coli* genomic DNA cloned into pRSF-Duet vector using the primers 551 and 552	This study
pAC017	PCR product of *pgpA* with C-terminus His-tag from *E. coli* genomic DNA cloned into pET-Duet vector using the primers 562and 563	This study

B. subtilis PssA). All genes were expressed under the control of the T7 promoter and induced with 0.25 mM of IPTG for 3.5 h. Total lipids were extracted from the *E. coli* membrane fraction using the Bligh and Dyer method [341] and analyzed by LC-MS. Four different engineered *E. coli* strains were compared, containing a different combination of archaeal lipid enzymes: (i) control harboring only the empty vectors; (ii) five ether lipid enzymes (*E. coli* IDI, *E. coli* mutant IspA, *B. subtilis* AraM, *M. maripaludis* GGGPS and *A. fulgidus* DGGGPS); (iii) six ether lipid enzymes (*E. coli* IDI, *E. coli* mutant IspA, *B. subtilis* AraM, *M. maripaludis* GGGPS, *A. fulgidus* DGGGPS and *A. fulgidus* CarS) and (iv) seven ether lipid enzymes (*E. coli* IDI, *E. coli* mutant IspA, *B. subtilis* AraM, *M. maripaludis* GGGPS, *A. fulgidus* DGGGPS, *A. fulgidus* CarS and *B. subtilis* PssA). When the total phospholipid composition of *E. coli* was compared for the different strains, no major differences were observed for the four differently engineered *E. coli* strains (**Fig. 6A**). However, a detailed analysis shows the presence of a peak corresponding to AE in *E. coli* strain 4 (**Fig. 6B**, lane 4) which contains all the ether lipid genes required for synthesis compared to the other strains that lack the *B. subtilis* PssA (**Fig. 6B**, lane 3) or both the *A. fulgidus* CarS and *B. subtilis* PssA (**Fig. 6B**, lane 2) or the control (**Fig. 6B**, lane 1). Another archaeal specific product DGGGP-Gro was detected in the engineered *E. coli* strains, which is absent in the control strain (**Fig. 6B**, lane 1). This compound was also detected in a previous study [32]. The formation of DGGGP-Gro, that hereafter is termed unsaturated AG, was observed in the *E. coli* strains that contain the basic set of five ether lipid genes up to the genes required for the formation of DGGGP (**Fig. 6B**, lane 2). However, there is an enhanced production of AG when the strain also contains the *A. fulgidus* CarS (**Fig. 6B**, lane 3). Upon the introduction of the AS producing *B. subtilis* PssA enzyme (strain 4), the amount of AG decreased and instead AE is produced (**Fig. 6B**, lane 4).

Formation of unsaturated AG has been attributed to an endogenous reaction in *E. coli* which directly attaches the glycerol moiety to DGGGP [32], as it was assumed that the *E. coli* PgsA and PgpA enzymes are not able to accept the archaeal intermediated as a substrate. As shown in the previous section, these two enzymes do recognize CDP-archaeol and convert it into AG. Therefore, the presence of unsaturated AG in the *E. coli* strain expressing the ether lipid genes is due to the low, but significant activity of the endogenous CDP-diacylglycerol synthase towards DGGGP. To confirm this hypothesis an *in vitro* reaction was performed where DGGGP was incubated in presence of the purified CDP-diacylglycerol synthase (*E. coli* CdsA) in presence of CTP. Under those conditions, a significant fraction of the DGGGP was converted into CDP-archaeol (**Fig. 6C**, lane 2**)** demonstrating that CdsA is active with the archaeal substrate. When the *E. coli* PgsA and PgpA and the substrate G3P were added to the reaction, the CDP-archaeol was converted into AG (**Fig. 6C**, lane 3). Taken together, these data demonstrate that the low activity of the endogenous *E. coli* CdsA (*in vitro* less than 1% compared to CarS in our previous work [306]) towards DGGGP results in substantial AG production *in vivo*. It is concluded that with a limited set of genes the archaeal lipids AG and AE can be produced in *E. coli*.

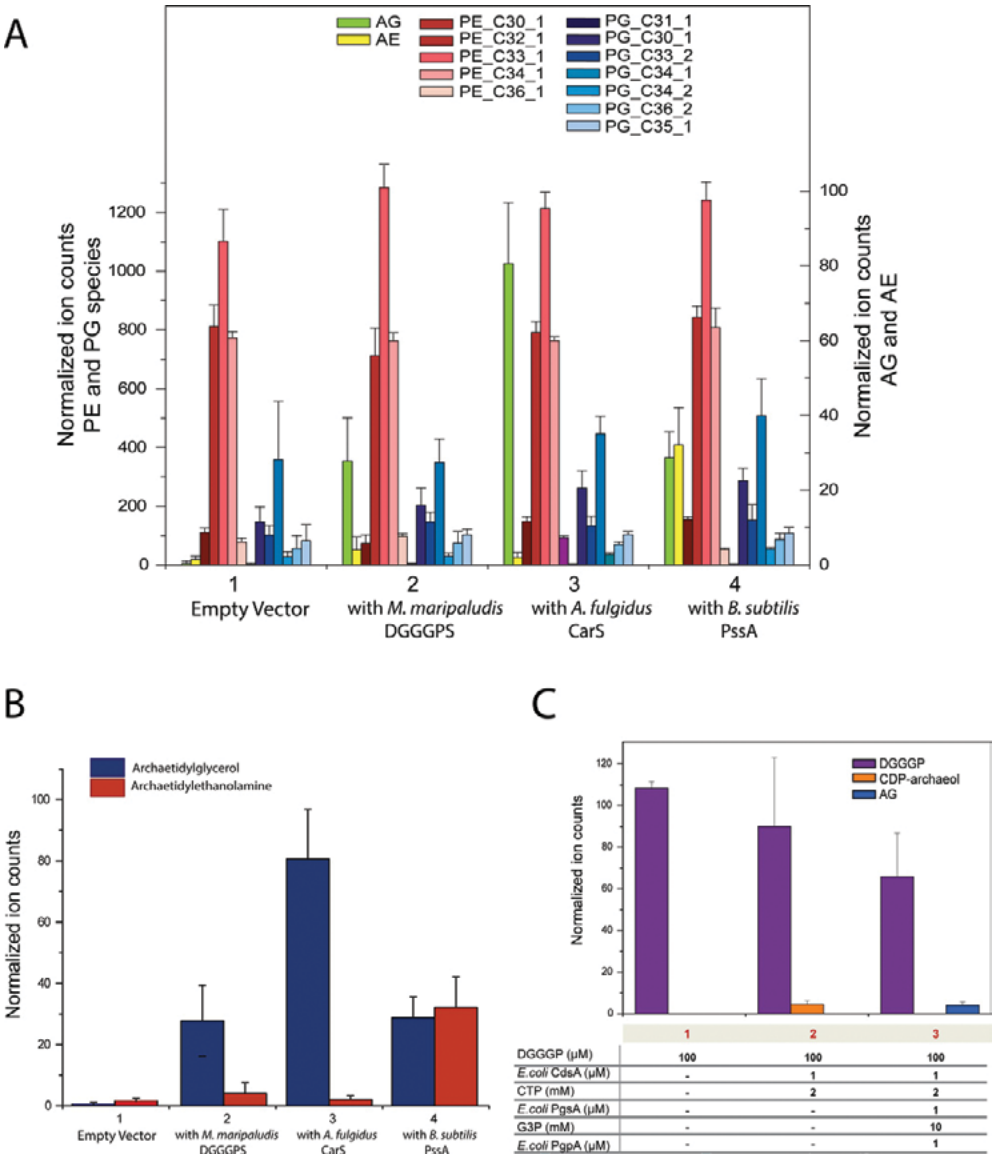


Figure 6. Archaeal lipid production in *E. coli*. (A) Total lipid analysis of four different engineered *E. coli* strains containing a different combination of ether lipid enzymes. The lipids were extracted from the membrane fraction and analyzed by LC-MS. The total ion counts of the several PE and PG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

Discussion

Since several decades *E. coli* has been used for metabolic engineering such as the improvement of isoprenoids or carotenoids production [208, 237, 238, 342]. Several attempts have been made to reconstruct the archaeal ether lipid biosynthetic pathway into *E. coli*. Lai *et al.* [33] demonstrated production of DGGGP upon the overproduction of the endogenous *E. coli* IDI and production of four enzymes (G1PDH, GGPPS, GGGPS and DGGGPS) from the hyperthermophilic archaeon *Archaeoglobus fulgidus*. Likewise,Yokoi *et al.* [32] employed ether lipid genes from the mesophilic archaeon *Methanosarcina acetivorans* to produce DGGGP and another compound that was identified as the phosphatidylglycerol-type derivate of DGGGP, named DGGGP-Gro. Isobe *et al*. [34] In addition expressed the *M. acetivorans* geranylgeranyl reductase for double bonds reduction in conjunction with the four aforementioned archaeal enzymes [32]. They observed the *in vivo* production of archaeal lipids with a fully saturated isoprenoid chain only when the cofactor *M. acetivorans* ferredoxin was co-expressed.

In the present study we aimed to produce the two major archaeal-like lipids archaetidylethanolamine (AE) and archaetidylglycerol (AG) in *E. coli* as these lipids have the same polar head group as the major phospholipids phosphatidylethanolamine (PE) and phosphatidylglycerol (PG) present in the *E. coli* membrane. However, this required a further understanding of the enzymatic steps required for polar head group attachment. Importantly, we have previously shown the production of CDP-archaeol *in vitro* by also including the CDP-archaeol synthase (CarS) into the pathway that yielded DGGGP. Expression of CarS provides a means to substantially increase the production of the endogenously produced DGGGP-Gro that corresponds to archaetidylglycerol (**Fig. 6B**). During polar head group attachment, the CDP group from the CDP-activated precursor is replaced by a different polar head group. In Bacteria and in Archaea, the reactions involved in this process are very similar and mediated by enzymes belonging to the CDP-alcohol phosphatidyltransferase superfamily. Given the high sequence homology among archaeal and bacterial phosphatidyltransferase [318, 320], we investigated the substrate promiscuity of the enzymes PgsA and PgpA that are involved in PG formation, and their ability to recognize CDP-archaeol. Using *in vitro* reactions, purified *E. coli* PgsA was able to produce AGP from CDP-archaeol which in turn was produced by the *A. fulgidus* CarS from DGGGP, CTP and G3P. By means of the *E. coli* phosphatase PgpA, the AGP was readily converted into AG. This demonstrates a high substrate promiscuity of these bacterial enzymes and alleviates the need to introduce archaeal enzymes into *E. coli* to perform these reactions steps. Indeed, when the ether lipid biosynthesis pathway up to the formation of CDP-archaeol is introduced into *E. coli*, AG formation is observed. However, CarS is not essential, as even in its absence, some AG can be formed. The origin of this AG was previously unknown, but we now show that this is due to a low activity of the endogenous *E. coli* CdsA for DGGGP resulting in the formation of CDP-archaeol that is further converted by *E. coli* PgsA and PgpA into AG.

In contrast to AG formation, AE formation has not yet been observed previously in *E. coli* which must imply that the *E. coli* phosphatidylserine synthases are unable to recognize CDP-archaeol in order to produce the intermediate AS that by decarboxylation should be further converted into AE. Indeed, the purified *E. coli* PssA was inactive with this substrate. Structural analysis of the bacterial and archaeal members of CDP-alcohol phosphatidyltransferase family indicated a high level of secondary structure conservation of these enzymes with eight transmembrane segments. The *B. subtilis* PssA is a truncated version of these enzymes, being significantly shorter but still containing the highly conserved protein core. Importantly, this enzyme was previously shown to be active with the archaeal-like substrates [343]. Indeed, the purified *B. subtilis* PssA catalyzed the production of AS from CDP-archaeol and L-serine. The next step, the decarboxylation of the serine moiety seems less specific, as purified *E. coli* Psd mediated the decarboxylation of AS with the concomitant formation of AE. Thus, for the production of AE in *E. coli* cells only the *B. subtilis* PssA needs to be introduced (**Fig. 6B**).

Having established the exact sequence of reactions needed for the archaeal ether lipid biosynthetic pathway [306] including the polar head group attachment, the reconstitution of AE and AG formation *in vitro* could be realized. Using a combination of archaeal and bacterial enzymes and a breakdown in individual reactions, the synthesis of the unsaturated archaeal lipids AE and AG from the initial building blocks IPP, FPP and DHAP was achieved with nine purified enzymes. This defined the conditions needed for the reconstitution of AE and AG formation *in vivo*, using *E. coli* as a host. Since three endogenous *E. coli* enzymes (Psd, PgsA and PgpA) recognize the archaeal precursors, the *in vivo* reconstitution depends on the expression of seven ether lipid genes. This includes overexpression of the *E. coli idi* to boost IPP formation, and expression of a mutant IspA to generate GGPP; the *B. subtilis* AraM for G1P formation, and *B. subtilis* PssA for AS formation, the key ether lipid biosynthetic proteins, the *M. maripaludis* GGGPS, the *A. fulgidus* DGGGPS and the *A. fulgidus* CarS. Although the conversion reactions appear efficient *in vitro*, the amounts of AE and AG produced *in vivo* are still low and compared to the total *E. coli* lipid content, they are probably less than 1%. However, the expression of CarS elevated the levels three-fold as compared to a previous report on AG formation [32].

The work described here represents a unique strategy to synthesize archaeal ether lipids in bacteria. Although the levels are still low, it is important to realize that the production of AE and AG as reported in this study, is performed in the presence of a fully functional phospholipid biosynthetic pathway. High-level AE and AG production likely also requires the upregulation of the entire pathway leading to IPP production. In addition, the pathway resulting in the reduction of the isoprenoid chains needs to be introduced to produce the saturated archaeatidyl compounds. Future studies should be directed towards a gradual down-regulation of the endogenous ester-bonded phospholipid biosynthetic pathway so that the endogenous lipids can be replaced by ether lipids. Such a bacterial strain could be used as an experimental model to examine the impact of the 'lipid divide' on the physiology and robustness of bacteria.

Materials and Methods

Bacterial strain and cloning procedures

Escherichia coli and *Bacillus subtilis* genomic DNA was used as template for the amplification of genes encoding the bacterial enzymes. *Escherichia coli* DH5α (Invitrogen) was used for the cloning steps. The plasmids and the primers used in this study are listed in **Table 1** and **2**. *Escherichia coli* BL21 (DE3) or Lemo21 (DE3) [344] was used as protein over-expression host strain and grown in aerobic condition at 37°C in Luria-Bertani (LB) medium supplemented with the required antibiotics. Kanamycin (50 µg/ml), chloramphenicol (34 µg/ml), streptomycin (50 µg/ml) and ampicillin (50 µg/ml) in conjunction with 0.2% glucose were added to the medium when necessary.

Table 2. Oligonucleotide primers used in the study.

Primers name	Primer sequence 5'→3'	Restriction enzyme site
11	GCGCGAATTCATGCATCACCACCACC	EcoRI
12	GCGCAAGCTTTCATTTTTTGGACAGC	HindIII
20	GCGCCTCGAGGACAGGTTTCCCGACTGGAAAG	XhoI
24	GATATACCATGGGCAGCCATCACCATC	NcoI
32	GGCGCCATATGCTGGATCTGATTCTGAA	NdeI
41	GACCAAGCTTGCGGCCGCATAATGC	HindIII
42	GATGCTCGAGTTAGAATGCACCGGCG	XhoI
57	GCGCGAATTCTTATTTATTACGCTGGATGATGTAG	EcoRI
62	CACTCATTAATTCATGATGATTTACCGGCAATGG	blunt
63	AGCGTGGATACACTCAACGGC	blunt
70	GCGCCATATGAATCGTATCGCAGCTGAC	NdeI
71	GCGCCTCGAGTTAGTGATGATGGTGGTGATGTTCATATAGACCATGGTTGATCAGCG	XhoI
84	GCCGCCATGGGTAGTCATCATCACCACCATC	NcoI
86	GCGCGAATTCTTAGAATGCACCGGCGA	EcoRI
89	GCGCCATATGAATTACATCCCCTGTATGATTACG	NdeI
90	GCGCCTCGAGTTAGTGATGGTGATGGTGGTGATGATGATTCCATCTCCCAGACTCCAG	XhoI
103	GCGCCTCGAGTTAGTGATGGTGATGGTGGTGATGATGAAGCGTCCTGAATACCAGTAAC	XhoI
106	GCCGCCATGGGCAGCCATCACCATCATCACCACAGCCTGAAGTATCGCCTGATATCTGC	NcoI
533	GCGCCTCGAGAAAACAATGGCCTGGAGGCTACCTTGTTAAATTCATTTAAACTTTCGCTAC	XhoI
534	GCGCTTAATTAATTAGTGATGGTGATGGTGGTGATGATGGACCTGGTCTTTTTTGTCGTCAAC	PacI
542	CGGCCATATGCTGTCAAAATTTAAGCGTAATAAAC	NdeI
543	GCGCCTCGAGTTAGTGATGATGGTGGTGATGCAGGATGCGGCTAATTAATC	XhoI
551	GCGCCATATGCAATTTAATATCCCTACGTTGC	NdeI
552	CGCGCTCGAGTCAGTGATGGTGATGGTGGTGATGATGCTGATCAAGCAAATCTGCACGC	XhoI
562	CGCGGAATTCATGACCATTTTGCCACGCCATAAAG	EcoRI
563	CGGCGCGGCCGCCTAGTGATGGTGATGGTGGTGATGATGCGACAGAATACCCAGCG	NotI

Expressions and purification of ether lipid enzymes

The bacterial proteins *B. subtilis* PssA and *E. coli* Psd were expressed in *E. coli* BL21 strain and induced with 1 mM of isopropyl β-D-1-thiogalactopyranoside (IPTG). *E. coli* PssA was induced with 0.5 mM of IPTG in the same over-expression strain *E. coli* BL21, whereas *E. coli* PgsA and PgpA were expressed in *E. coli* Lemo strain and induced with 0.4 mM of IPTG and 0.5 mM of L-rhamnose. After 2,5 h of induction the cytoplasmic and membrane fractions were separated as described in a previous study [306]. The inner membrane vesicles (IMVs) of *E. coli* expressing the membrane proteins (*B. subtilis* PssA and *E. coli* Psd, PgsA and PgpA) were isolated as previously described [345]. The *E. coli* IMVs harboring the *B. subtilis* PssA and *E. coli* Psd were resuspended in buffer A (50 mM Tris-HCl pH 7.5, 300 mM NaCl and 10% glycerol) and 0.5 mg/ml of IMVs were solubilized in 2% of DDM detergent for 1 h at 4°C. A centrifugation (173,400x*g*) step of 30 min at 4°C removed the insolubilized materials and the supernatant was incubated with Ni-NTA beads (Sigma) for 1 h at 4°C. The Ni-NTA beads were washed ten times with 40 column volumes (CV) of buffer B (50 mM Tris-HCl pH 7.5, 300 mM NaCl, 10% glycerol and 0.2% DDM) supplemented with 20 mM Imidazole and the proteins were eluted three times with 0.5 CV of buffer B supplemented with 250 mM Imidazole. The *E. coli* IMVs (1 mg/ml) containing the *E. coli* PgsA and PgpA were resuspended in buffer C (50 mM Tris-HCl pH 7.5, 150 mM NaCl and 10% glycerol). The solubilization, the removal of the insolubilized materials and the Ni-NTA beads incubation was performed in the same manner as described above. The *E. coli* PgsA-bound beads were washed five times with 40 CV of buffer D (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 10% glycerol and 0.2% DDM) supplemented with 10 mM Imidazole and eluted three times with 0.5 CV of buffer D supplemented with 250 mM Imidazole. The Ni-NTA beads containing *E. coli* PgpA were washed ten times with 40 CV of buffer D supplemented with 10 mM Imidazole and the protein was eluted in one single wash of 0.5 CV of buffer D supplemented with 300 mM Imidazole.

For the purification of the soluble protein *E. coli* PssA, the cytoplasmic fraction was incubated with Ni-NTA beads in buffer C overnight at 4°C. The beads were washed three times with 40 CV of buffer C supplemented with 10 mM Imidazole, one times with 40 CV of buffer C supplemented with 60 mM Imidazole end eluted with one wash of 2 CV of buffer C supplemented with 300 mM Imidazole. The purified proteins purity was assessed on SDS-PAGE (12%) stained with Coomassie Brilliant Blue and the purified protein concentration was determined by measurement of the absorbance at 280 nm. The other cytosolic proteins (*E. coli* IDI, *E. coli* mutant IspA, *B. subtilis* AraM and *M. maripaludis* GGGPS) and the membranes proteins (*A. fulgidus* DGGGPS, *A. fulgidus* CarS and *E. coli* CdsA) used in this study were expressed and purified as described previously [306].

In vitro assays for archaeal lipids production

In vitro reactions were performed in 100 µl of assay buffer containing a final concentration of 50 mM Tris-HCl pH7.5, 10 mM $MgCl_2$, 52.5 mM NaCl, 87.5 mM Imidazole, 0.07% DDM and 3.5% glycerol. Where specified 100 µM synthetic DGGGP, 100 µM PA (18:1), 2 mM CTP, 10 mM G3P, 10 mM L-serine 20 mM EDTA and the indicated amount of purified enzymes were added to the reaction mixture. For the *in vitro* AE and AG reconstitution, 100 µl of reaction volume was used containing the following assay buffer: 50 mM Tris-HCl pH 7.5, 10 mM $MgCl_2$, 15 mM NaCl, 25 mM Imidazole, 0.02% DDM, 1% glycerol, 0.5 mM dithiothreitol (DTT), 0.1 mM FPP, 0.1 mM IPP, 2 mM NADH and 2 mM DHAP in. Reactions were incubated at 37°C for 1 h as described previously [306], and the products were extracted with two times 0.3 ml of *n*-butanol. Extracted lipids were evaporated under a stream of nitrogen gas and resuspended in 50 µl of methanol for the LC-MS analysis.

In vivo archaeal lipids synthesis

Engineered *E. coli* strains were aerobically grown at 37°C in 200ml of LB medium supplemented with half antibiotic concentration (kanamycin (25 µg/ml), chloramphenicol (17 µg/ml), streptomycin (25 µg/ml) and ampicillin (25 µg/ml), 0.2% of glucose and 1 mM $NiCl_2$. The cells were induced with 0.25 mM IPTG and after 3.5 h, the total membrane fractions were isolated as previously described [306]. The internal standard eicosane (20 µM) was added to the total membrane fractions (8 mg/ml) and lipids were extracted by means of the Bligh and Dyer method [341]. The chloroform extractable lipid fraction was washed with the aqueous phase from a blank Bligh and Dyer extraction and evaporated under a stream of nitrogen gas. The evaporated samples were resuspended with 0.3 ml 1:2 chloroform:MeOH, evaporated and finally resuspended in 100 µl MeOH to be analyzed in LC-MS.

LC-MS analysis

The *in vitro* reaction samples were analyzed using an Accela1250 high-performance liquid chromatography system coupled with an electrospray ionization mass spectrometry (ESI-MS) Orbitrap Exactive (Thermo Fisher Scientific). A volume of 5µl of each sample was used for the analysis. The column, the LC method and the MS parameters used in this study have been already described in details in our previous work [306]. The capillary and in the tube lens voltage were set to -75V and -190V, respectively. For the *in vivo* experiment samples different method was used for the analysis. A volume sample of 5µl was injected into a Shim-pack XR-ODS/C8/Phenyl column with dimension 3.0 mm I.D. x 75 mm (Shimadzu) operating at 55°C with a flow rate of 400 µl/min. Mobile phase A (10mM ammonium formate with 0.1% formic acid in water:acetonitrile 40:60 v/v) and mobile phase B (10 mM ammonium formate with 0.1% formic acid in acetonitrile:isopropanol 10:90 v/v) were used

as follows: initial condition started with 40% B, a linear gradient was started in 2 min from 40% to 43% B, gradient from 43% to 50% B in 0.1 min, 54% B isocratic for the next 9.9 min, linear gradient from 54% to 70% B in 0.1 min, 99% B isocratic for the following 5.9 min, gradient from 99% to 40% B in 0.1 min and 40% B isocratic for the last 1.9 min [346]. The MS settings used for this analysis were the same as described above. The thermo XCalibur processing software was used for the data analysis and Genesis algorithm for automated peak detection and integration was applied to this analysis.

Acknowledgements

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

Chapter 6

Generation of a Hybrid Heterochiral Membrane in *Escherichia coli*

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

[1]Laboratory of Microbiology, Wageningen University and Research Centre, Stippeneng 4, 6708 WE, Wageningen, The Netherlands
[2]Department of Molecular Microbiology, Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, 9747 AG Groningen, The Netherlands
[3]Stratingh institute for Chemistry, University of Groningen, Nijenborgh 7, 9747 AG Groningen
[4]Current address: Department of Medicine, Section of Infectious Diseases, Boston University School of Medicine, 02118 Boston, Massachusetts, USA

†authors contributed equally

Abstract

Archaea and Bacteria share many genetic and biochemical features, yet their phospholipids are fundamentally different. Whereas bacterial lipids are composed of glycerol-3-phosphate (G3P) with ester-linked fatty acid side chains, archaeal lipids consist of glycerol-1-phosphate (G1P) with ether-linked isoprenoids with distinct chirality. This phenomenon is referred to as 'the lipid divide'. It has been speculated that the last universal common ancestor (LUCA) cell could have possessed a mixed membrane that at some stage has segregated, resulting in the bacterial and archaeal lineages. However, a mixed membrane has never been found, possibly because it would not tolerate a viable situation due to instability. In the present study, we tested this hypothesis by constructing and characterizing a mixed membrane in *Escherichia coli.* By combining metabolic engineering of the MEP-DOXP pathway with an isoprenoid-based lipid synthesis pathway we engineered a mixed membrane consisting of up to 28% archaeal lipids. G1P-based archaeal lipids were formed both in the presence and in the absence of a heterologous G1P-dehydrogenase, showing that *E.coli* can synthesize G1P endogenously. Despite normal growth of the engineered cells, it was observed that the production of archaeal lipids resulted in marked changes in their endogenous lipid composition as well as in dramatic morphological deformities. Remarkably, robustness assays also demonstrated an elevated cell survival when exposed to a cryo-treatment and to butanol exposure. From these results, we conclude that mixed membranes are readily tolerated *in vivo*, which would agree with the existence of natural cells with hybrid bacterial/archaeal membranes.

Keywords: LUCA, heterochirality, archaetidylglycerol, archaetidylethanolamine, di-*O*-geranylgeranylglyceryl-phosphate, metabolic engineering, glycerol-1-phosphate, glycerol-3-phosphate

Introduction

Archaeal lipids are fundamentally different from those of bacteria. Archaeal lipids are composed of isoprenoid chains attached to *sn*-glycerol-1-phosphate by ether bonds, whereas bacterial lipids are composed of fatty acid chains esterified to *sn*-glycerol-3-phosphate. Although this intrinsic difference is commonly used to distinguish bacteria and archaea ('Lipid Divide'), there are bacterial lipids that contain fatty acids attached by diether- or tetraether-bonds [4, 5, 79, 81, 84, 119], as well as archaeal lipids containing fatty acid based ester-lipids [6]; in addition, polar head groups are shared between archaeal and bacterial lipids. Despite these few discrepancies, there are, as yet, no examples of exchange of the glycerophosphate enantiomers between lipids of the two domains. Over the years, this lipid divide, has given rise to a range of different scenarios for the evolution of the domain-specific lipid types [7, 13-15, 347].

In 1994, Kandler postulated the pre-cell hypothesis to explain the distinct and uniform features across the three domains of life [15]. A heterogeneous population of pre-cells or last universal common ancestor (LUCA) cells was proposed which ultimately diverged into the bacterial and archaeal lineages. These pre-cells must have had some basic features shared by all descending lineages, including the systems responsible for processing of the genetic information (transcription, translation and replication). In case a reproducing pre-cell with complete inheritable characteristics existed, the genetic material must have been confined within a closed membrane. At this point however, the pre-cell hypothesis fails to explain the lipid divide. According to Wächtershäuser [14] there are two possibilities to account for the evolution of extant membrane compositions. A pre-cell may have existed with a uniform membrane based on either ether-G1P or on ester-G3P lipids. From this ancestor, separate lineages evolved with the concomitant emergence of new lipids and subsequent loss of the original lipid. In a more widely accepted alternative scenario, a pre-cell membrane evolved with mixed ether-G1P/ester-G3P lipids, and through selection of one lipid over the other segregated when the domain divergence occurred [14]. Both scenarios, however, inevitably lead to an hypothesized appearance and disappearance of hybrid membranes from which also cross-hybridized diasteriomers and enantiomer lipids (isoprenoid-G1P/isoprenoid-G3P vs. fatty acyl-G1P/fatty acyl-G3P resp.) were proposed to drive the segregation [7]. The mixed membrane theory has caused heated debates over the last years. In his review [7], Koga shows that a wide range of specificities of CDP-alcohol phosphatidyltransferases on unnatural lipid substrates (isoprenoid-G3P and fatty acyl-G1P) occur [317]. This observation supports the idea that the four mentioned lipid types already existed in pre-cells. Besides that, it was emphasized that lipids with different hydrocarbon chains tend to segregate, but not because of the different glycerophosphate backbones. The biochemical pathways for synthesis of the core lipids phosphatidic acid and archaetidic acid are not related, but do show a high degree of analogy (**Fig. S1**). In line with this, shared

enzyme families and substrate promiscuities of polar head-group specific enzymes may, to some extent, support the mixed membrane hypothesis. *In vitro* biophysical and modelling studies in turn report some contradictory findings on mixed membranes. On the one hand, it has been demonstrated that heterochiral vesicles of fatty acids are unstable and exhibit reduced self-reproduction compared to homochiral vesicles [16]. This claim found some support in a mathematical model on membrane bioenergetics that postulated that the LUCA membrane required a highly impermeable membrane made of fatty acids and isoprenes only [347]. On the other hand, arguments have been put forward that would be in favour of the existence of mixed membranes. The enzyme families of polar head group attachment (phospholipid synthases) are all homologs, which would agree with the model that LUCA contained mixed lipid membranes. This is supported by an *in vitro* study in which mixed membranes from bacterial and archaeal lipids exhibited a decreased permeability at elevated temperatures and were therefore postulated to be highly stable [17]. However, there is as yet no consensus on the early-evolution mixed-membrane hypothesis.

The pathway for diether lipid synthesis is completely elucidated (**Fig. S1**) [100]. The reconstruction of the archaeal core lipid di-*O*-geranylgeranylglyceryl-phosphate (DGGGP) was first successfully demonstrated in *Escherichia coli* [33]. For the biosynthesis of DGGGP four genes are required. Firstly, the isoprenoid geranylgeranyldiphosphate is assembled from its building blocks isopentenyl-diphosphate (IPP) and dimethylallyl-diphosphate (DMAPP) by geranylgeranyl diphosphate synthase (GGPPs). Secondly, the other lipid building block concerns the backbone molecule *sn*-glycerol-1-phosphate (G1P) that is being generated by G1P-dehydrogenase. Thirdly, two enzymes subsequently attach two GGGP isoprenoid chains to G1P by formation of ether bonds: *sn*-3-*O*-geranylgeranylglyceryl phosphate synthase (GGGPs) attaches GGPP to G1P at the *sn*-3 position to form GGGP, after which attachment of the second GGPP to the *sn*-2 position is catalysed by a membrane protein called *sn*-2,3 di-*O*-geranylgeranyl-glyceryl phosphate synthase (DGGGPs). For the final stage of polar head group attachment, the required enzymes of archaea and bacteria are orthologs. Yokoi *et al.* have demonstrated that in *E.coli* DGGGP is converted to archaetidylglycerol (AG). The promiscuous activity of *E.coli*'s native phosphatidylglycerol phosphate synthase (PgsA) and phosphatidylglycerol phosphatase (PgpA) have recently been shown to be responsible for these conversions [348]. The final synthesis step to the full diether has been described by Isobe *et al.* who established a complete reduction of the lipid isoprenoid chains to phytanyls by identification of a geranylgeranyl reductase (GGr) along with its accompanying ferredoxin [34]. The increase in biochemical knowledge on both lipid synthesis routes has stimulated researchers to construct a mixed membrane *in vivo,* however, with limited success as the contribution of archaeal lipids never exceeded 1% [32-34, 348].

In the present study we engineered an upregulation of the genes involved in the synthesis of the building blocks IPP and DMAPP, and co-expressed a composite archaeal lipid synthesis pathway in *E. coli*. The main objective of this endeavour was to evaluate whether a hybrid

heterochiral membrane is tolerated *in vivo* and what the implications are on growth, cell morphology and membrane biochemistry. This is the first report of a mixed archaeal/ bacterial membrane in a biological system and the first characterization of a hybrid glycerophosphate heterochiral membrane *in vivo*.

Results

Metabolic engineering of a hybrid membrane in *Escherichia coli*

1. Enhanced supply of IPP and DMAPP

To screen for higher *in vivo* diether-lipid production different variants of the archaeal lipid synthesis pathway were constructed in *E.coli* JM109(DE3) (**Table 1**). The main limitation to the production of isoprenoid compounds and thus also for archaeal lipids in bacteria is the supply of its building blocks isopentenyl-diphosphate (IPP) and dimethylallyl-diphosphate (DMAPP). To construct a hybrid membrane, the MEP-DOXP pathway which is responsible for synthesis of IPP and DMAPP in *E.coli* must be upregulated. We addressed this constraint by the chromosomal integration of two synthetic fragments composed of the native *idi* gene or a combination of the *idi, ispD, ispF* and *dxs* genes. These constructs were integrated by a modified gene disruption protocol from Datsenko and Wanner [234] (**Fig**. **S2**). The *aceA-aceB* locus was chosen for integration (Chapter 3). The two obtained mutant strains **WUR89** and **WUR91** were tested for inducible MEP-DOXP upregulation by means of a reporter plasmid that encodes enzymes that catalyse the conversion of IPP and DMAPP to the red carotenoid lycopene. Here we observed a 2.5- and 5.3-fold increase in lycopene production, respectively, confirming that upon induction the pool of IPP and DMAPP had increased (**Fig**. **S3)**.

2. Production of DGGGP and archaetidylglycerol

To enable di-*O*-geranylgeranylglyceryl-phosphate (DGGGP) production, the MEP-DOXP integrated strains were complemented with 2 compatible plasmids. The first plasmid (pMS148) harbours the previously characterized *crtE* (GGPP synthase) gene from *Pantoea ananatis* [229] and *araM* (*sn*-glycerol-1-phosphate dehydrogenase, G1PDH) gene from *Bacillus subtilis* [340] for synthesis of lipid precursors geranylgeranyl-diphosphate (GGPP) and *sn*-glycerol-1-phosphate (G1P), respectively. The initial purpose of the secondary plasmid was to test for optimal heterologous functionality and coupled activity of *Methanococcus maripaludis* GGGP-synthase with DGGGP-synthases.

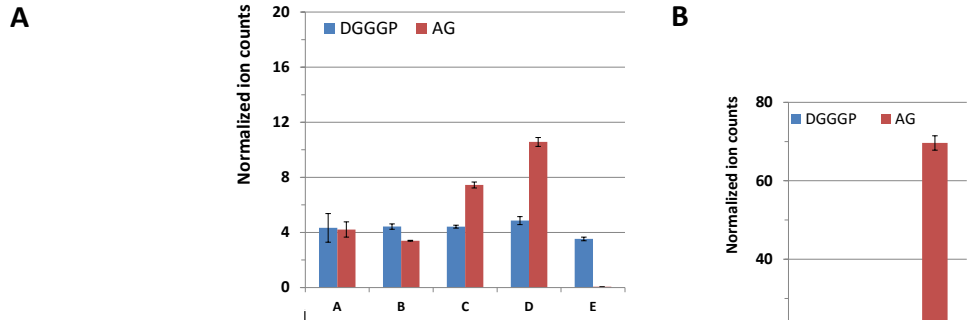
We first compared two different DGGGP synthase genes, either from *Archaeoglobus fulgidus* (*AF0404*) to yield pSJ139 or from *M. maripaludis* to yield pSP001. Introduction of these plasmids provided the following 5 mutant strains (**Table 1**): (i) Wild type JM109(DE3) strain

harbouring pMS148 and pSJ139 annotated as **WUR150**; (ii) integrated *idi* gene, pMS148 and pSJ139 annotated as **WUR151**; (iii) integrated *idi-ispDF-dxs* operon, pMS148 and pSJ139 annotated as **WUR152**; (iv) integrated *idi-ispDF-dxs* operon, pMS148 and pSP001 annotated as **WUR153**; v) and lastly a negative control strain harbouring *idi-ispDF-dxs*, complemented with pSJ139 only, and annotated as **WUR168**. As a final variable during the lipid production assays, the mutant strains were grown either on LB medium or defined minimal medium (OPT1) which was optimized for isoprenoid production [213]. This OPT1 medium was modified to our specific requirements by decreasing the $MgSO_4$ content to 0,012 mg/L and glycerol to 1%, addition of 1 mM $NiCl_2$ as a co-factor for G1PDH, and replacing the composed vitamins for a combined MEM-vitamin mixture. Total lipid extracts from isolated membrane fractions were obtained using the Bligh and Dyer method [341] and analysed on LC-MS for production of DGGGP (m/z=715.51 [M-H$^+$]), archaetidylglycerol (AG) (m/z=789.54 [M-H$^+$]), and the native *E.coli* lipids phosphatidylglycerol (PG) (m/z=773.54 [M-H$^+$]) and phosphatidylethanolamine (PE) (m/z=775.61 [M-H$^+$]) which together account for 95% of *E.coli*'s lipids [309]. Comparison of the strains having both introduced plasmids showed variable levels of AG in the membrane fractions (**Fig**. **1A,** lanes A-D). In accordance with previous reports [32, 348], heterologous production of AG was obtained without the special requirement for specific CDP-archaeol synthase (CarS), archaetidyl glycerophosphate synthase (Ags) and phosphatase (Agp) activity. The CdsA, PgsA and PgpA enzymes native to *E.coli* recognize and convert the archaeal DGGGP substrate to AG; these enzymes are promiscuous in that they accept substrates derived from both G1P and G3P, and with fatty acid and isoprenoids [348]. The metabolic engineering of the MEP-DOXP pathway caused an overall increased production of DGGGP and AG compared to the control strain with native IPP and DMAPP pools. This upregulation was especially apparent for the **WUR153** strain (**Fig 1A**, lane D) which contains the complete *idi-ispDF-dxs* operon and the *M. maripaludis* DGGGPs. Here we observed a 1.7-fold upregulation of DGGGP leading to 6.1% diether lipids of the cumulative lipid areas. The most radical improvement, however, was observed when **WUR152** was grown on the statistically optimized OPT1 medium which resulted in diether ratios up to 15.6% (**Fig. 1B**). This finding suggested possibly even higher upregulations for **WUR153** where the *M. maripaludis* DGGGP-synthase seemed to be more efficient than the *A. fulgidus* DGGGP-synthase.

Table 1 Strains, integrations and plasmids used for DGGGP and AG production.

Strain	Integrated operon	Plasmid 1	Plasmid 2	DGGGPs
WUR150	-	pMS148	pSJ139	*A.fulg*
WUR151	*idi*	pMS148	pSJ139	*A.fulg*
WUR152	*idi-ispDF-dxs*	pMS148	pSJ139	*A.fulg*
WUR153	*idi-ispDF-dxs*	pMS148	pSP001	*M.marp*
WUR168	*idi-ispDF-dxs*	-	pSJ139	*A.fulg*

Abbreviations: *A.fulg*, *Archaeoglobus fulgidus; M.marp, Methanococcus maripaludis.*



	A	B	C	D	E
idi-int	-	+	-	-	-
idi-ispDF-dxs-int	-	-	+	+	+
P1: *crtE-araM*	+	+	+	+	-
P2A: *GGGPs/DGGGPs Afulg*	+	+	+	-	+
P2B: *GGGPs/DGGGPs Marp*	-	-	-	+	-
% diethers	3,7	2,9	4,5	6,1	1,0

	C	C_opt
LB	+	-
OPT1	-	+
% diethers	4,5	15,6

Figure 1. Level of produced DGGGP and AG in different genetic variants or different growth media. A. The total amount of diethers (pink bar) is expressed as a percentage of the cumulative ion counts. A: WUR150; B: WUR151; C: WUR152; D: WUR153; E: WUR168. **B.** Level of DGGGP and AG in WUR152 on different growth medium, LB and OPT1. DGGGP = di-*O*-geranylgeranylglyceryl-phosphate; AG = archaetidylglycerol.

When both strains were grown on OPT1 medium, an increase in AG production was again observed (**Fig. 2A,** lanes B and C). A very high increase however was not detected without integration of the MEP-DOXP operon in **WUR150**. Like anticipated, **WUR153** on OPT1 medium showed the highest productions of AG. Cumulative ether lipid contents up to 37.8% of the total lipid extract were found.

3. Production of archaetidylethanolamine

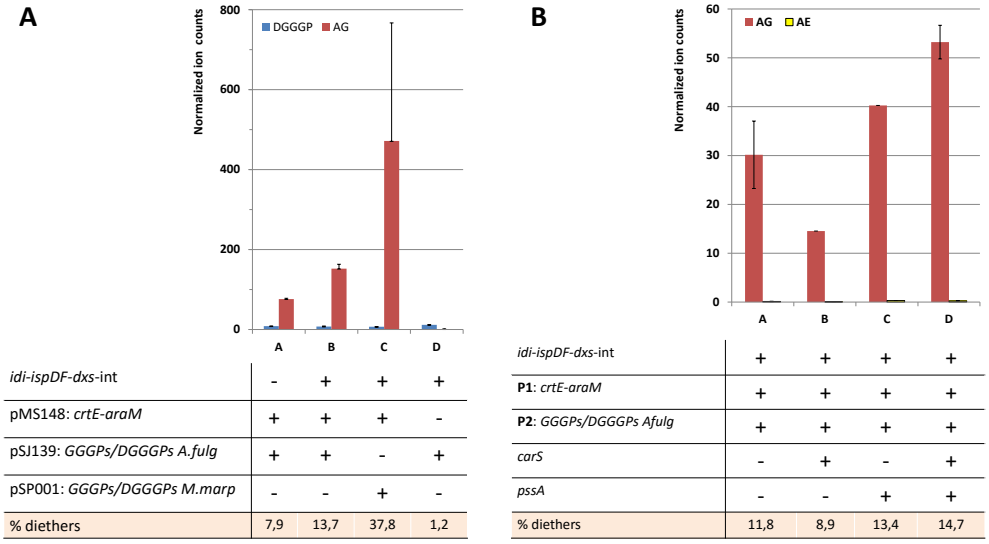
Archaetidylethanolamine (AE) is the main non-bilayer lipid constituent of the archaeal membrane which is essential for the realization of a functional membrane and the proteins therein [349]. Previously, production of AE was established in *E.coli* by the heterologous expression of the *B. subtilis* phosphatidylserine synthase gene (*pssA*). By *in vitro* coupled reactions, this enzyme, unlike the counterpart of *E.coli,* was found to display promiscuous activity towards CDP-archaeol (in addition to its native substrate CDP-diacylglycerol). This central core lipid was converted to archaetidylserine (AS) by the exchange with L-serine by PssA. When *E.coli* phosphatidylserine decarboxylase (Psd) was also added to the reaction, *in vitro* formation of AE was confirmed [348]. To also include production of phosphatidylethanolamine (PE) in the mixed membrane, three variations of the **WUR153** strain were constructed by the addition of *A. fulgidus carS* and *B. subtilis pssA* to the secondary plasmid. These additions provided 3 additional strains (**Table 2**): (i) WUR153 with

carS annotated as **WUR177** (ii) WUR153 with *pssA* annotated as **WUR178** and finally (iii) WUR153 with both *carS* and *pssA* annotated as **WUR182**. Induction and lipid analysis including archaetidylethanolamine (AE) (m/z=758.54 [M-H$^+$]), on these strains confirmed production of AE (**Fig. 2B**). In accordance to our previous findings, the *carS* gene also significantly contributed to the production of AG, but only in addition to *pssA*. Interestingly, AE was detected in all strains with *pssA,* albeit at very low amounts. Increased expression of *pssA* by optimizing RBS sequences could not compensate for this low production (*data not shown*). Despite persistently low levels of AE, **WUR182** which has the complete pathway present, showed reproducibly high levels of AG. Further experiments were therefore continued with this strain.

Table 2 Strains tested for archaetidylethanolamine production. Variations of the inserts in plasmid 2 are indicated with pSJ139 as negative control.

Strain	P1	P2	Insert 1	Insert 2	Insert 3	Insert 4
WUR153	pMS148	pSJ139	DGGGPs *M. marp*	GGGPs *M.marp*	-	-
WUR177	pMS148	pAC027	DGGGPs *M. marp*	GGGPs *M.marp*	*carS*	-
WUR178	pMS148	pAC028	DGGGPs *M. marp*	GGGPs *M.marp*	*pssA*	-
WUR182	pMS148	pAC029	DGGGPs *M. marp*	GGGPs *M.marp*	*carS*	*pssA*

Abbreviations: P1, plasmid 1; P2, plasmid 2; *M.marp, Methanococcus maripaludis*



	A	B	C	D
idi-ispDF-dxs-int	-	+	+	+
pMS148: *crtE-araM*	+	+	+	-
pSJ139: *GGGPs/DGGGPs A.fulg*	+	+	-	+
pSP001: *GGGPs/DGGGPs M.marp*	-	-	+	-
% diethers	7,9	13,7	37,8	1,2

	A	B	C	D
idi-ispDF-dxs-int	+	+	+	+
P1: *crtE-araM*	+	+	+	+
P2: *GGGPs/DGGGPs Afulg*	+	+	+	+
carS	-	+	-	+
pssA	-	-	+	+
% diethers	11,8	8,9	13,4	14,7

Figure 2. Level of produced DGGGP, AG and AE in different genetic variants A. Calculated ratio of diethers (pink bar) is expressed as a percentage of the cumulative ion counts. A: WUR150; B: WUR152; C: WUR153; D: WUR127; **B.** Level of produced AG and AE expressed as normalized ion counts. The total amount of diethers (pink bar) is expressed as a percentage of the total ion counts. A: WUR153; B: WUR177; C: WUR178; D: WUR182. DGGGP = di-*O*-geranylgeranylglyceryl-phosphate; AG = archaetidylglycerol; AE = archaetidylethanol-amine

4. Induction optimization

Heterologous isoprenoid production may benefit from a proper balancing of the expression between the upstream MEP-DOXP pathway, and the downstream pathway. This effect was also apparent when we analysed the different integrated operons by lycopene productions. Variable levels of lycopene were produced by varying the time of induction and the IPTG concentration. Here, a variable IPTG concentration was found to correlate well with the level of expression of the *crtEBI* reporter genes (*data not shown*). This IPTG-dependent expression level strongly supported a potential for improved diether production, which was also tested in the **WUR182** strain. To this end, we looked for better induction by varying the onset of induction or the IPTG concentration. We observed that IPTG supplementation immediately before inoculation resulted in higher amounts of archaeal lipids compared to postponed induction (**Fig. 3A**, lane 1 vs. lane 2 and 3). Significantly lower levels of AG were obtained after late induction at OD_{600}=0.3 or OD_{600}=0.6. Interestingly, a high AG production was also observed without the addition of IPTG. These levels, attributed to leaky expression, even exceeded the levels obtained at postponed inductions [350]. For the early induced cells a highest average content of 28% archaeal lipids (consisting of AG, AE and DGGGP) was demonstrated. Together with the earlier obtained diether levels, this shows that *E.coli* cells can tolerate high levels of archaeal lipids in their membrane and are perfectly able grow.

5. Substrate promiscuity of archaeal enzymes

To be able to attribute potential phenotypical changes to the presence of archaeal lipids and not to an overall high abundance of heterologous membrane proteins, we constructed a control strain (**WUR187**) with a disrupted *araM* gene. By deleting a large segment of the *araM* gene, responsible for the synthesis of *sn*-glycerol-1-phosphate, we aimed to disrupt archaeal lipid synthesis, but maintain the heterologous enzyme load. With the initial purpose to calibrate for optimal IPTG concentrations, we performed a serial dilution of IPTG on the **WUR182** strain and compared it with **WUR187**. This analysis showed that optimal AG production occurs at ~0.05 mM IPTG (**Fig. 3B**, lane 4). Unexpectedly, the **WUR187** Δ*araM* control strain appeared to still contain a substantial amount of archaeal lipids (up to 10%), albeit at a lower level than **WUR182** (**Fig. 3B**, lanes 6 and 7). This observation suggested that GGGPs and DGGGPs display substantial substrate promiscuity towards *sn*-glycerol-3-phosphate.

The glycerophosphate configuration, however, is traditionally recognized as one of the most striking differences between archaeal and bacterial lipids. The finding of AG in the lipidome of *E. coli,* a bacterium in which G1P synthesis has never been reported, thus raises questions about the molecular basis of this observation. The first enzyme of the ether lipid biosynthesis pathway committed in the recognition of G1P or G3P is GGGP-synthase. This enzyme typically attaches the isoprenoid GGPP to the glycerol-1-phosphate backbone by an ether-bond. The crystal structures of the GGPPs enzymes from *A. fulgidus* [302] and *Methanothermobacter thermoautotrophicus* have been elucidated, and the catalytic site

and the binding site for G1P have been identified [303]. When activity of this enzyme was tested towards G1P or G3P it was found to have an activity of only 5% with G3P as substrate compared to the activity with G1P [351]. We therefore also investigated the activity of *M. maripaludis* GGGP-synthase towards G1P and G3P in an *in vitro* reaction using GGPP as initial substrate. Surprisingly, GGGP formation was observed both in the presence of G1P and G3P at equal levels (**Fig. 4A**, lanes 2 and 3) compared to the negative control where no glycerophosphate was added (**Fig. 4A**, lane 1). As a control, the specificity of purified PlsB from *E. coli* towards G3P and G1P was tested in another coupled *in vitro* reaction. As expected the synthesis of lysophosphatidic acid (LPA) was observed in the presence of G3P (**Fig. 4B**, lane 2) but not with G1P (**Fig. 4B**, lane 3). Kinetic analysis of GGGP-synthase towards G1P and G3P showed, however, an 8-fold higher affinity of the enzyme for G1P (Km = 5.75 ± 1.57 µM for G1P, Km = 46.67 ± 6 µM for G3P) (**Fig. 4C**). Consequently, substrate specificity of all the downstream enzymes of the ether lipid pathway were also tested for G3P specificity. Coupled *in vitro* reactions were used to test the stereoselectivity of the second enzyme (DGGGP-synthase) involved in the ether bonds formation towards G1P and G3P.

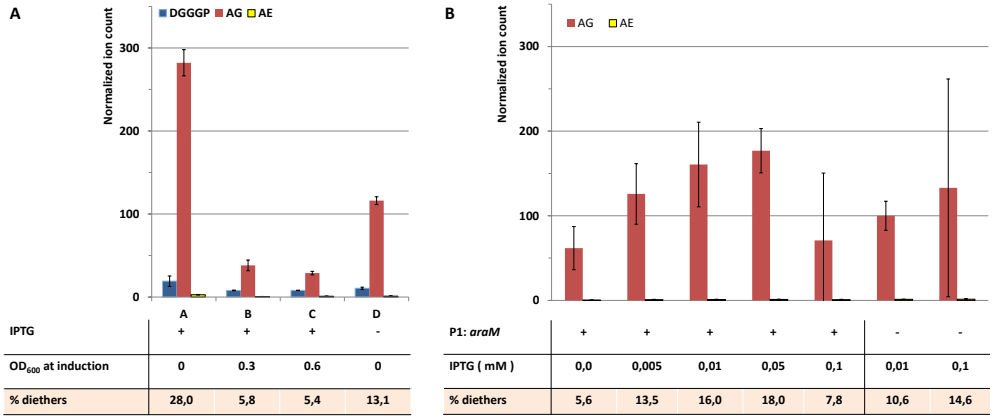
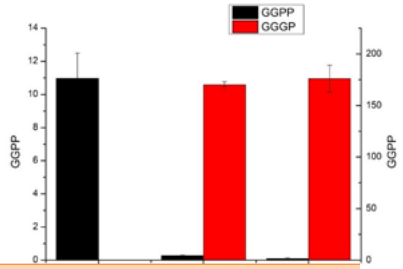


Figure 3. Optimization of IPTG induction in WUR182 A. DGGGP, AG and AE production expressed as normalized ion counts. The total amount of diethers (pink bar) is expressed as a percentage of the total ion counts. WUR182 was assayed for optimal onset of induction (IPTG = 0.1 mM). **B.** An IPTG calibration curve was tested on WUR182, Additionally, the Δ*araM* control strain WUR187 was tested for AG and AE production.

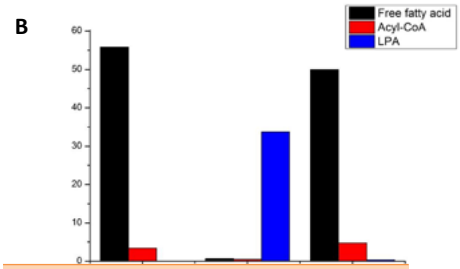
By using purified GGGPs from *M. maripaludis* and DGGGPS from *A. fulgidus*, DGGGP synthesis was observed also when using G3P (**Fig. 4D**, lane 2 vs. **Fig. 4E,** lane 4). Remarkable, however, is the low amount of DGGGP produced in the presence of GGGP in G3P configuration compared to GGGP in G1P configuration. This result implies that DGGGPs plays a more important role in the selection of substrates in G1P configuration than the GGGPs does. This finding therefore shows that the stereoselectivity of G1P in archaea is thus not solely determined by the presence of the G1P-dehydrogenases but also determined by substrate specificity of the lipid synthesis enzymes downstream of GGGPs.

A



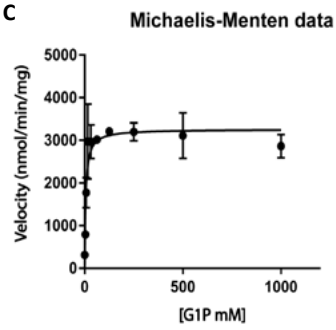
	1	2	3
GGPP	yes	yes	Yes
GGGPS_M.mar	yes	yes	Yes
G1P	no	yes	no
G3P	no	no	yes

B

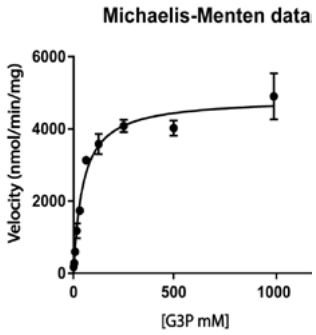


	1	2	3
Oleic acid	yes	yes	yes
CoA	yes	yes	yes
ATP	yes	yes	yes
FadD	yes	yes	yes
PlsB	yes	yes	yes
G3P	no	yes	no
G1P	no	no	yes
lipids	yes	yes	yes

C



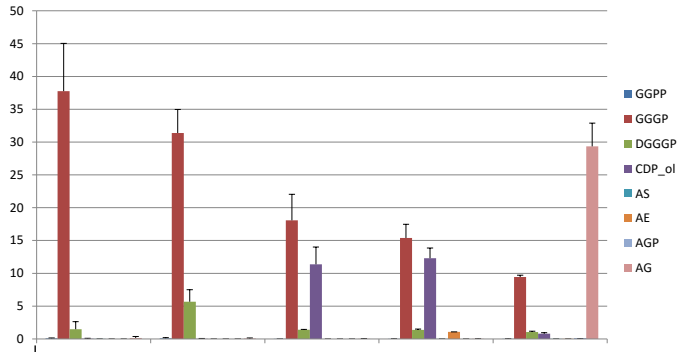
Vmax= 3259 ± 150 nmol/min/mg
Km= 5.75 ± 1.57 uM
Kcat = 1.086 ± 0.05 s^{-1}



Vmax= 4850 ± 161 nmol/min/mg
Km= 46.67 ± 6 uM
Kcat= 0.8084 ± 0.027 s^{-1}

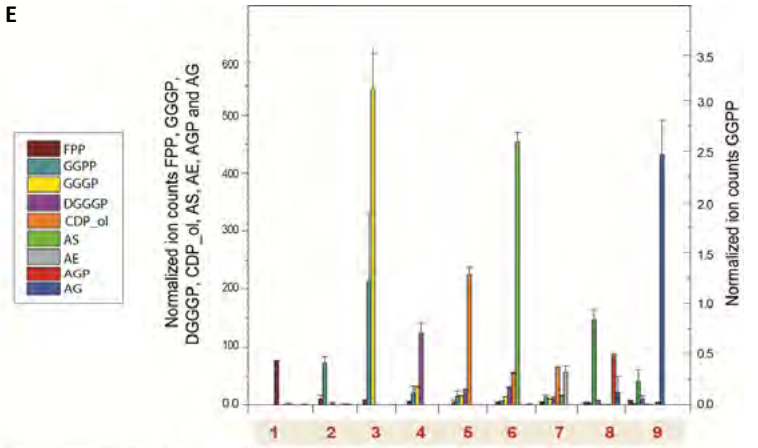
Figure 4. Specificity of archaeal lipid synthesis enzymes towards glycerophosphate substrates A. Enantiomer specificity of GGGP synthase from *M. maripaludis* to G1P and G3P **B.** PlsB acitivity towards G1P and G3P **C.** Kinetics of *M. maripaludis* GGGP synthase dependent on glycerophosphate **D.** Cascade *in vitro* reaction using G3P substrate; **E**. Cascade *in vitro* reaction using G1P substrate via AraM [348].

D



	1	2	3	4	5
GGGPs (µM)	0.26	0.26	0.26	0.26	0.26
DGGGPs (µM)	-	1	1	1	1
CarS (µM)	-	-	1	1	1
CTP (mM)	-	-	2	2	2
PssA (µM)	-	-	-	0.5	-
L-serine (mM)	-	-	-	10	-
Psd (µM)	-	-	-	0.5	-
PgsA (µM)	-	-	-	-	0.5
G3P (mM)	10	10	10	10	10
PgpA (µM)	-	-	-	-	0.5

E



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

Figure 4. Specificity of archaeal lipid synthesis enzymes towards glycerophosphate substrates. Continued

6. Confirming heterochirality of a hybrid membrane

After finding archaeal lipids in the Δ*araM* control strain and substrate promiscuity of GGGPs and DGGGPs, the chirality of the lipids was analysed by NMR. Via a saponification of the bacterial ester-lipids, the archaeal ether-lipids were purified from the lipid extracts. The purified lipids were then analysed to reveal the stereo configuration of their glycerophosphate backbone. By using lipid standards having either a G3P (**Fig. 5**, panel E) or a G1P (**Fig. 5**, panel D) configuration and a combination of both (**Fig. 5**, panel A), the chirality of the ether lipids was established. Strain **WUR182** harbouring the MEP-DOXP operon and all the ether lipid required genes, clearly showed a near 100% identity with the lipid standard in G1P configuration (**Fig. 5**, panel B). Surprisingly, also the control strain **WUR187** having the deletion in *araM* showed the presence of lipids in the G1P configuration (**Fig. 5**, panel C), with only a trace amount of G3P-lipids. Apparently, G1P is natively present in *E.coli* and the inclusion of the *araM* gene thus appeared to be redundant for the formation of the desired G1P-diether lipids. Despite this, the level of diether lipids markedly increased by the presence of intact *araM* (**Fig. 3B**, lane 3 vs. lane 6).

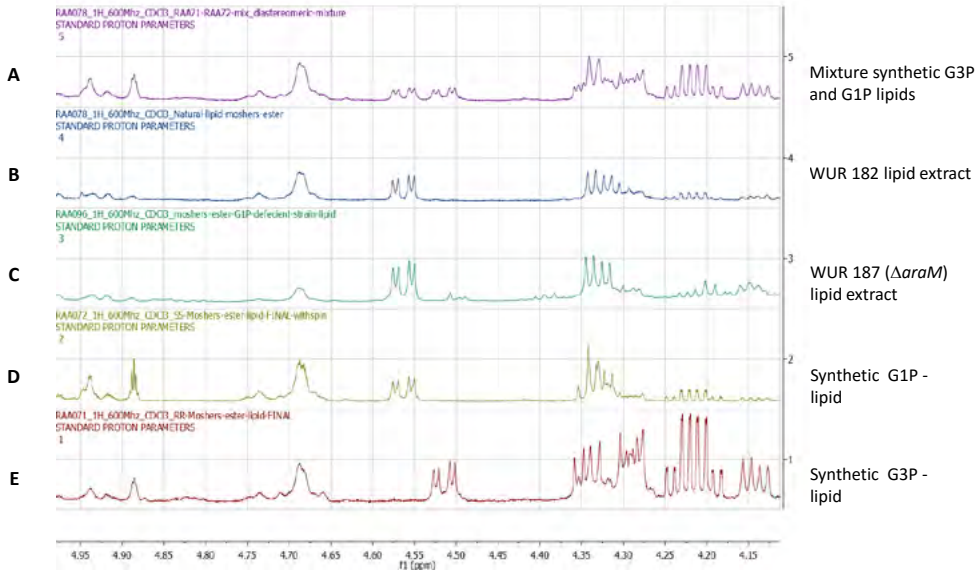


Figure 5. NMR spectra of lipid extract from WUR182 and WUR187(Δ*araM*)

Cellular characteristics and general responses to a hybrid heterochiral membrane

Biochemical adaptations to endogenous lipid synthesis

To assess phenotypical responses of *E. coli* to archaeal lipids, we characterized the endogenous lipid composition, growth performance, cell morphology and physico-chemical adaptiveness of the induced **WUR182** strain. As previously shown and confirmed here, conversion of DGGGP to AG can be catalysed by the native polar head-group machinery of *E.coli* [348]. Although introduction of *B. subtilis* PssA is necessary for AE formation, this product also relies on the additional activity of two native proteins i.e. CDP-diacylglycerol synthase (CdsA) and phosphatidylserine decarboxylase (Psd). This dependence on native *E.coli* membrane proteins [332, 352, 353], highly supports membrane incorporation and functionality of the introduced lipids. This assumed membrane integration of AG and AE might in turn trigger compensatory changes in the native lipid composition. Arguments in favour of this hypothesis are the observed sudden increase in anionic lipid content and shared substrate utilization of G-3-P for polar head groups. Secondly, the acquired lipids are also likely to cause an overall change in membrane permeability and fluidity which may elicit changes in the native lipid composition as a homeoviscous adaptation mechanism. To test this, we grouped all obtained lipidomes according to their cumulative diether lipid content (DGGGP, AG and AE). The phosphatidylglycerol (PG) pool, which typically accounts for 25% of the phospholipids [290, 353] was remarkably high in the WUR168 control strain (i.e. 39.1%) (**Fig. 6A**, lane 1).

This increase is possibly due to an excess of glycerol added to the OPT1 medium. Associated with an increase in diether lipid content we observed a gradual decrease of PG in response to the increasing AG content (**Fig. 6A**, lanes 2-11). A highest overall PG decrease of 28.1% was observed that perfectly matched with the detected amount of diether lipids (also 28.1%). This phenomenon is probably best explained by competition of CDP-DAG (CDP-diacylglycerol) and CDP-archaeol for phosphatidylglycerophosphate synthase (PgsA), the enzyme that exchanges the CDP moiety for G-3-P. Despite the trade-off between PG and AG, the overall percentage of lipids with glycerol head groups remained stable at 37.1% of the total phospholipid content. Average PE levels also remain the same regardless of the AG content.

Growth performance and genotypic changes

To evaluate potential effects of the archaeal lipids on the metabolic fitness of the **WUR182** strain during mixed membrane formation, growth was analysed. Fresh pre-cultures of **WUR182** on LB were obtained and used to inoculate the OPT1 medium without the addition of IPTG. Changing from LB to OPT1 medium first resulted in a large lag-phase that lasted for approximately 16 h, which was also observed for the control strain. We then used this starter

culture to inoculate fresh OPT1 medium supplemented with IPTG. We observed, however, an additional lag-phase for **WUR182** and **WUR187** that lasted for at least 20 h (**Fig. 6B**, triangles) compared to a non-induced culture which entered log-phase already after 7 h (**Fig. 6B**, circles). This growth retardation is best explained by some metabolic adaptation to the introduced heterologous proteins or the foreign lipid species. Alternatively, the final culture may represent a selective progeny of cells. In this case, incidental but beneficial mutations rather than a collective growth retardation of the entire population explains the growth delay.

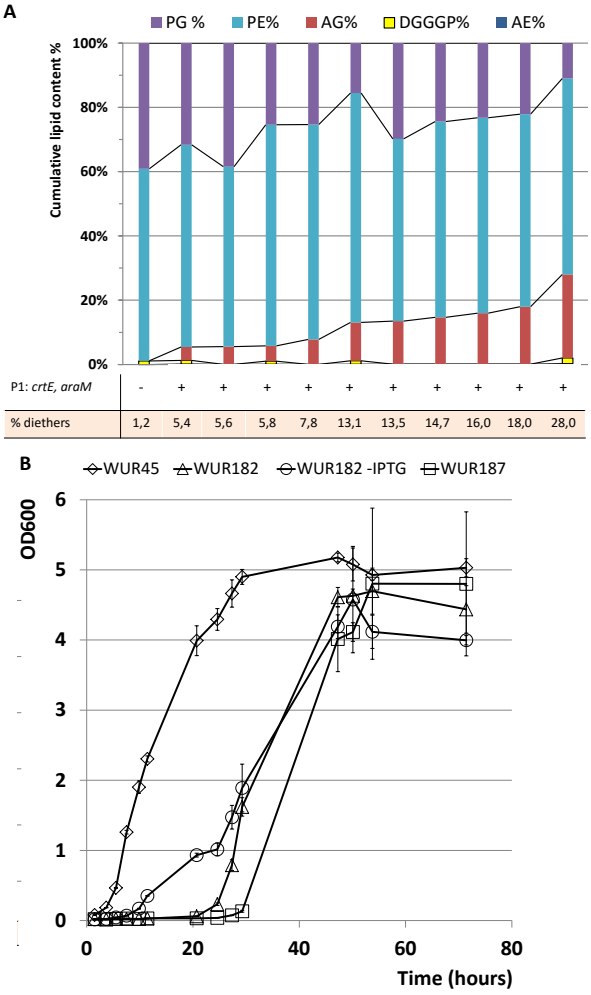


Figure 6. A. Lipidome composition in relation to the diether lipid content. Lipidomes of WUR182 were arranged according to increasing diether lipid content; the WUR168 control strain lacking pMS148 is depicted in lane 1; **B. Growth curves obtained during mixed membrane production.** Negative control WUR45 (diamonds); WUR182 with immediate induction (triangles); WUR182 with delayed induction (circles); WUR187 with *araM* deletion, with immediate induction (squares). Growth was monitored by measuring the turbidity at 600 nm (OD_{600}) using biological triplicates.

Interestingly, this latency was no longer observed after subsequent culturing of the adapted cells on fresh OPT1 with IPTG (not shown). Late inductions at OD_{600}=0.3 also caused a comparable but smaller growth retardation (**Fig. 6B**, circles). In contrast to a 20 h lag phase, growth was restored shortly after a transient plateau was reached at OD_{600}=1. An even longer lag phase (>24 h) was observed for the **WUR187** strain that lacks G1PDH (**Fig. 6B**, squares). Although onset of growth was severely affected in both **WUR182** and **WUR187**, specific growth rate of the strains did not change, as almost equal slopes were observed compared to the control. In addition to this, nearly equal cell densities were obtained, indicating that mixed membrane production does not affect the metabolic fitness of the cells.

To detect incidental mutations that could explain the initial growth retardation, we performed genome sequencing on two samples of adapted cells and compared these to a reference genome sequence. The latter was obtained by a PacBIO sequencing on strain **WUR182,** which was grown on LB without induction. For the adapted cells (sample B and C), two individual cultures were grown on OPT1 with IPTG and sampled for DNA extraction at stationary phase. These cultures showed the same growth retardation as observed before. The hybrid assemblies of the induced strains were then referenced to the *de novo* annotated genome of the non-induced cells. The comparative analysis of the "pre-induction reference genome" and the two "post-induction" genomes revealed the occurrence of 10 mutations (8 substitutions and 2 deletions) at 7 different loci (**Table 6**).

Table 6. Single nucleotide variants detected by illumina sequencing. Samples B and C were referenced against a whole genome hybrid assembly (sample A).

Ref. allele (A)	allele	sample	gene/ intergenic region	effect	Predicted function of affected gene or intergenic region
C	T	B	*mhpE – mhpT*	None	4-hydroxy-2-oxovalerate aldolase; 3-hydroxyphenylpropionic acid transporter
A	T	C	*lysZ - lysQ*	None	(tRNA-lys) ttt
G	A	C	*nu*1_2 (possible prophage)	None	possibly nohB
G	C	C	*rsxC*	E614Q	Reducer of SoxR, SoxR iron-sulfur cluster reduction factor component
G	C	C	*lptB*	V210L	LPS transport; LPS export ABC transporter ATPase
A	T	B	*yhhI*	A292A	H-repeat-associated protein, RhsE-linked, (predicted transposase)
A	G	B	*yhhI*	E293E	
G	A	B	*yhhI*	K294K	
C	*del*	C	*rop - lacI*	none	non-coding C-terminal region of *lacI*
G	*del*	C	*rop - lacI*	none	

Abbreviations: *del*: deletion

Four of these mutations occurred in intergenic regions, but not in the vicinity of genes that had any obvious relation with the observed phenotype. In addition, 3 silent mutations were found in the *yhhI* gene of sample B (predicted transposase) and 2 substitutions in *rsxC* (E614Q)(reducer of SoxR) and *lptB* (V210L)(LPS transport ABC transporter ATPase) in sample C. None of these mutations, however, was involved in central metabolism or able to directly affect phospholipid synthesis, nor did the observed mutations occur in both tested samples or in related genes. Besides the integrated operon, the recovered plasmids were also left unchanged. Therefore, we conclude that the observed lag-phase was likely due to a metabolic adaptation rather than a small selective population of mutants. We therefore also conclude that the applied membrane modifications result in a stable genotype.

Effects of the hybrid heterochiral membrane on cell morphology

WUR182 cells for which 28% diether lipid content and heterochirality was confirmed, were inspected by SEM and TEM to examine their cell morphology. SEM imaging first illustrated that the shape of the cells was radically changed (**Fig. 7A-C**). The most striking observation was the appearance of one or more lobular appendages with diameters ranging roughly between 100 and 500 nm. Occasionally, we observed a detachment or apparent budding of the appendages which resulted in separated vesicles (**Fig. 7A**, lower right panel). This phenotype appeared to affect the majority of the cells, and was only observed at a very low frequency for the late induced cells having lower amounts of archaeal lipids (i.e. 5.4%). Moreover, the cells with 28% diether lipids were generally longer (up to 12 µm in length) with irregular surfaces. TEM analysis indicated that the appendages bud from the cytoplasmic membrane with a continuous outer and inner membraneand detachment likely occurs through a vesiculation or pinching event. Furthermore, the cells formed intracellular vesicles with a high diversity in diameter, that apparently originate from the cytoplasm or the periplasmic space (**Fig. 7A**, lower left panel). Dramatic effects were also observed in cells that were induced late at early log phase (OD_{600}=0.3). Instead of forming appendages, thickening at the distal ends was observed that appeared to have its origin from within the cells (**Fig. 7B**, upper panel). The accumulation of a dense material was demonstrated within the cells, occasionally protruding the cell surface from the inside. To further analyse the nature of these dense bodies, the cells were fractionated by French-press and the resulting extract centrifuged. Here we found large cellular fragments accumulating in the insoluble fraction. SEM and TEM analysis showed the formation of extremely dense bodies covered by membrane or cell wall fragments (**Fig. 7B**, lower panel).

Physico-chemical robustness of a hybrid heterochiral membrane

To determine whether membrane robustness is affected by a hybrid and heterochiral lipid composition, cells were subjected to some physical/chemical challenges. The engineered strains were tested for temperature and butanol sensitivity. We were particularly interested

in these parameters because archaeal lipids are assumed to be responsible for cellular robustness [30, 121, 295, 354]. We included the **WUR187** strain as a primary control to assess the influence of lower quantities of archaeal lipids. As a second control strain, we included the lycopene synthesis strain (**WUR94**) to assess the effect of a high isoprenoid level in general. Cell cultures were prepared according to the IPTG optimized protocol and aliquoted. Individual samples were then exposed to the different stressors and analysed for their survival when normalized to a condition without a stressor. A heat shock survival assay was performed by exposing the cells 2 min to challenging temperatures between 42°C and 58°C and the surviving cells were normalized to control cells from each strain exposed to 37°C. We observed that the hybrid heterochiral membrane provides *E.coli* with

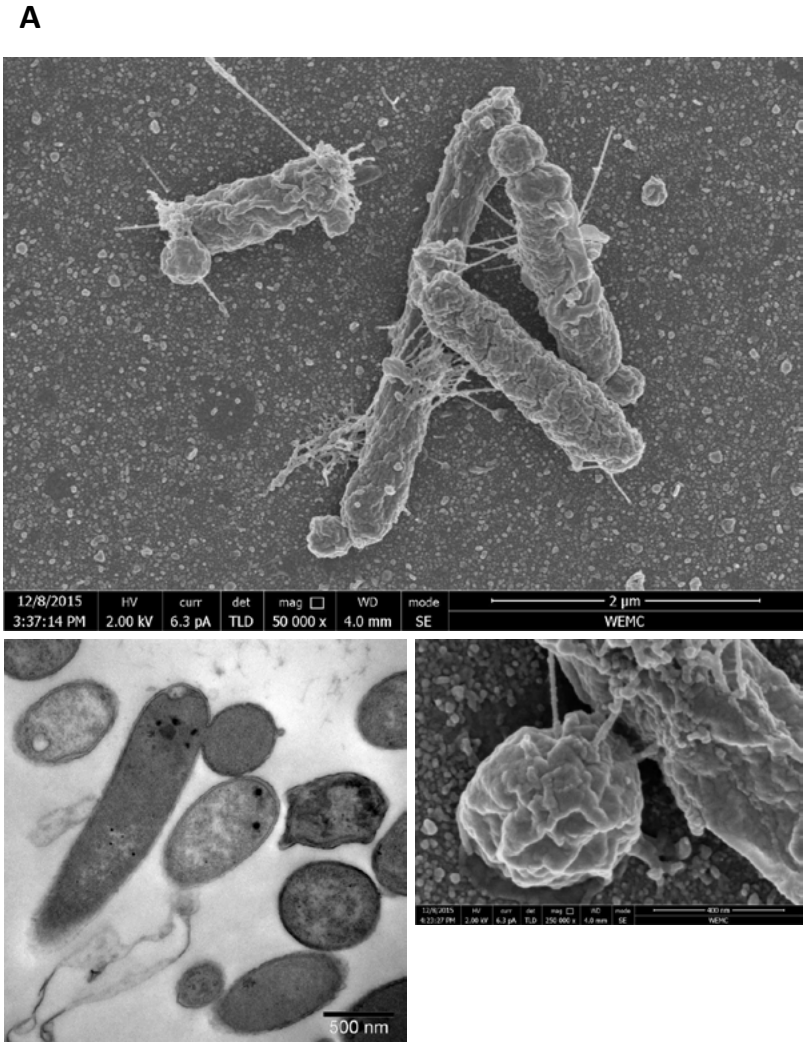


Figure 7. Electron microscopy pictures of engineered strains A. WUR182 cells containing 28% diether lipids: upper panel, SEM at 50.000 x magnification. Lower left, TEM; lower right, SEM at 250.000 x magnification.

B

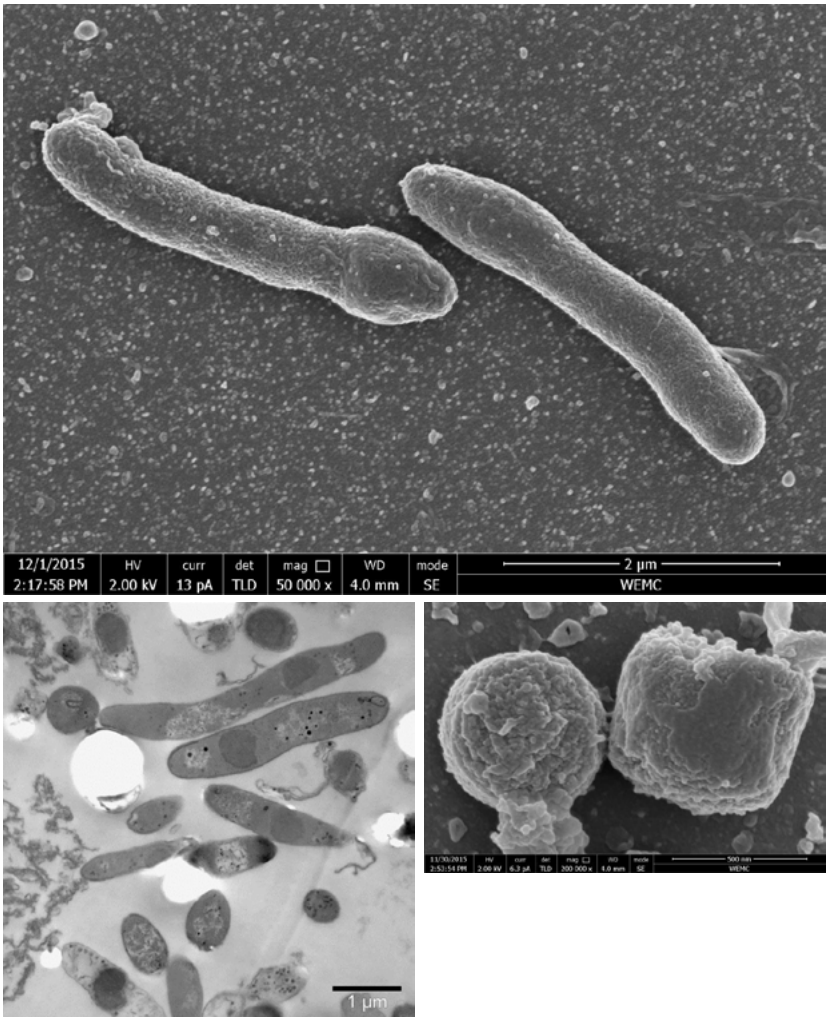


Figure 7. Electron microscopy pictures of engineered strains B. WUR182 cells containing 5.4% diether lipids: upper panel, SEM at 50.000 x magnification; Lower left, TEM; lower right, SEM at 200.000x magnification.

C

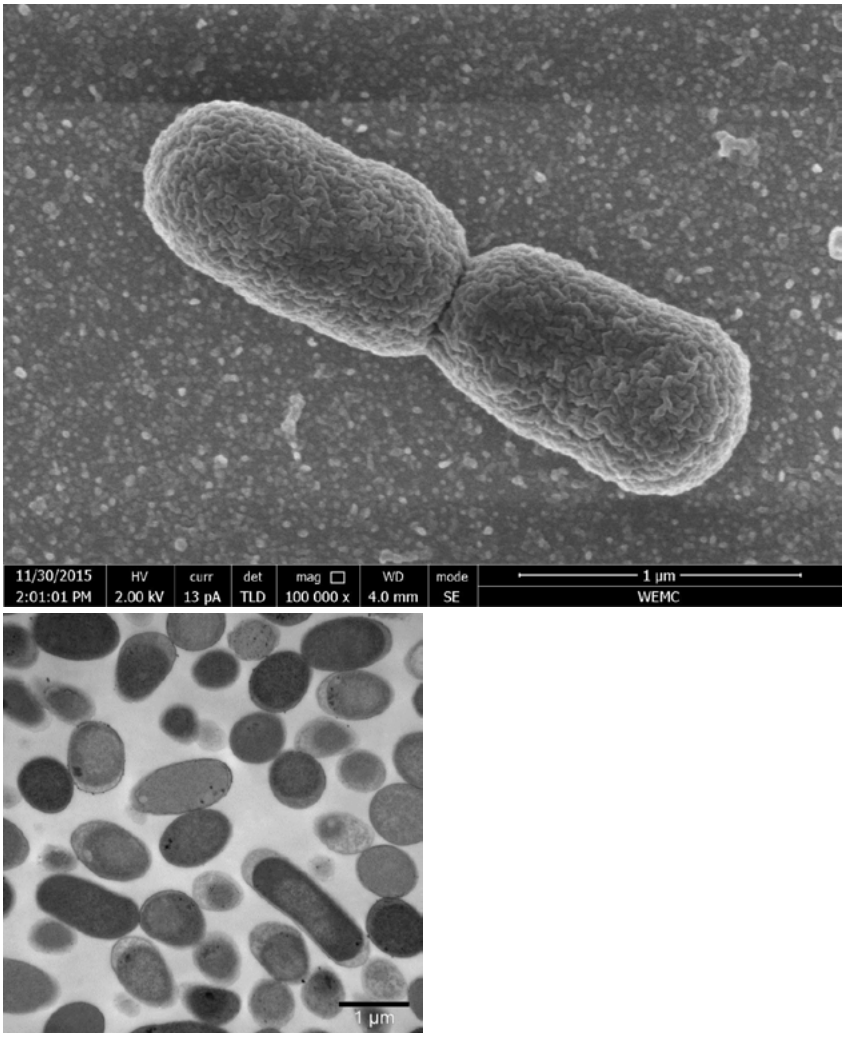


Figure 7. Electron microscopy pictures of engineered strains C. WUR182 cells containing <1% diether lipids. WUR182 cells were grown on LB without induction: upper panel, SEM at 100.000 x magnification; Lower left, TEM.

a lower heat shock survival (**Fig. 8A**, black bars). We observed a dramatic decline in the number of recovered cells compared to the control strains. Conversely, significantly more cells were recovered after cryogenic treatment at -80°C compared to wild-type cells (**Fig. 8B**, black vs. white bars). Using **WUR182**, 140% of the cells was recovered compared to only 75% for the wild-type cells. In sharp contrast, an incubation at -20°C for a comparable duration resulted in loss of viability. The presence of archaeal lipids thus seemed to have a protective effect against extreme cryogenic conditions (-80 °C). This finding suggests that the hybrid membranes are generally more robust in terms of cold stability. Interestingly,

enrichment of unsaturated diether lipids has been described as a psychrophilic adaptation in *Methanococcoides burtonii* [64] and *Halorubrum lacusprofundi* [65]. The relationship between the introduced archaeal lipids and cold stability may thus be based on an increased fluidity of the membrane. An increased fluidity, however, also implies that butanol resistance would be reduced. This is mainly because increasing membrane rigidity is shown to be effective in rationally designed tolerant bacteria [20, 24]. Since butanol concentrations higher than 1% are toxic to *E.coli*, we exposed the cells to butanol in a range of 0 to 2.5%. The 1% butanol exposure resulted in a 75% recovery of the lycopene control strain (**Fig. 8C**, squares). In contrast, both the **WUR182** and **WUR187** strains were found to survive over 100% of CFU's at a butanol concentration of 1%. The **WUR182** strain even resisted butanol levels of 1.5% where still 100% of the original amount of CFU's was recovered. Altogether, these assays suggest a higher stability of the membranes against cryogenic and butanol exposure. The lower resistance to a heat shock, however, emphasizes that the membrane modifications do not imply an overall improved robustness.

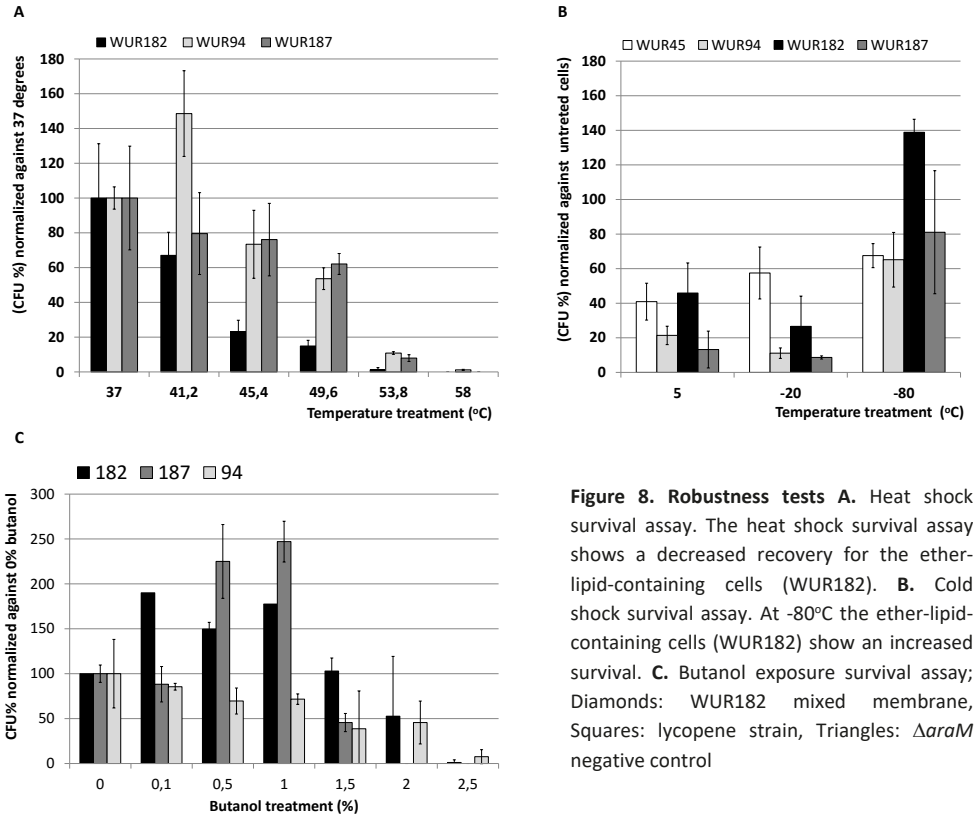


Figure 8. Robustness tests A. Heat shock survival assay. The heat shock survival assay shows a decreased recovery for the ether-lipid-containing cells (WUR182). **B.** Cold shock survival assay. At -80°C the ether-lipid-containing cells (WUR182) show an increased survival. **C.** Butanol exposure survival assay; Diamonds: WUR182 mixed membrane, Squares: lycopene strain, Triangles: Δ*araM* negative control

Discussion

Archaea and Bacteria are believed to have emerged from a last universal common ancestor [14]. The appearance at a certain point in time of a mixed membrane, composed of both ether- and ester lipids, is regarded as an inevitable transitional stage to reconcile the shared characteristics of bacteria and archaea. As this may indeed explain shared features as well as the lipid divide, a chimeric membrane should at least support a viable situation to support this event.

To simulate a mixed membrane, we combined metabolic engineering of the MEP-DOXP pathway in *E.coli* with heterologous expression of a composite archaeal biosynthesis pathway. To this end, 3 bacterial genes for the synthesis of GGPP (*crtE; P. ananatis*), G1P (*araM; B. subtilis*), and AS (*pssA; B. subtilis*) were used and 3 archaeal genes for synthesis of GGGP (*mmarc5_RS05150; M. maripaludus*), DGGGP (*mmarc5_RS08865; M. maripaludis*) and CDP-archaeol (*carS; A. fulgidus*) (**Fig. S1**). An important finding of this endeavour was the formation of a mixed membrane with reproducible archaeal lipid levels up to 28% (w/w) of the total extracted lipids. Besides the choice of genes, the MEP-DOXP engineering and tuning of the induction protocol seemed to be of high influence on the net archaeal lipid production levels. The most radical improvement, however, was observed when the engineered strain was grown on optimized medium, consisting of glycerol and citric acid as basic components [213]. We conclude that the optimized medium resulted in a substantial stimulation of the isoprenoid flux. This finding strongly suggested that during overexpression of only the 4 MEP-DOXP genes, a maximum was reached in the supply of pyruvate and glyceraldehyde-3-phosphate as substrate for DXS. On the other hand, when we searched for optimal induction parameters, we also observed that the onset of induction can radically change the outcome. Immediate induction just before inoculation led to higher levels of AG compared to delayed inductions at $OD_{600} \geq 0.3$. The highest reproducible level of archaeal lipids consisted mainly of archaetidylglycerol (92.8%), and to a lower extent of DGGGP (6.3%) and archaetidylethanolamine (0.9%). We performed chiral-selective NMR on these archaeal lipids to analyse the enantiomeric backbone of the archaeal lipids. Here we confirmed that the lipids indeed consisted exclusively of the *sn*-glycerol-1-phosphate configuration. When synthesis of G1P was disabled by disruption of the *araM* gene, the cells unexpectedly still produced a significant amount of diether lipids (≥10.6%). We therefore expected to find a high substrate promiscuity of GGGP-synthase for G-3-P which was confirmed in an *in vitro* assay. This finding could explain the unexpected lipid production and makes the postulated existence of the presently unnatural G3P-isoprenoid-ether lipid in LUCA [7] more credible. Unexpectedly, NMR analysis showed that lipids from this control strain also consisted exclusively in G-1-P configuration. This observation leads us to two more interesting conclusions. Apparently, *E.coli* can synthesize G1P endogenously under these circumstances, an activity that has never been reported to occur in this bacterium before; as no obvious G1PDH ortholog is present in *E.coli*, the responsible enzyme remains

elusive. Secondly, although *M. maripaludis* GGGP-synthase can utilize both G1P and G3P to synthesize GGGP, selectivity for G1P (and not G3P) as the only backbone molecule for archaeal lipids is brought about by the higher stereo-selective substrate specificity of the DGGGP-synthase for G1P-lipids.

Because a mixed membrane of 28%(w/w) archaeal lipids was demonstrated to sustain a viable situation, we were specifically interested in the novel phenotype of the obtained strain (WUR182). Here we focussed on establishing growth performance, endogenous lipid composition, cell morphology and physicochemical robustness. We noticed that the cells displayed a major lag phase upon induction, but eventually started to produce mixed membranes at growth rates approaching that of the parent strain. Although the occurrence of a long lag phase may have suggested a genetic adaptation, no adaptive mutations affecting central metabolism or membrane synthesis could be detected. In addition, the cells did not display another lag phase when passed to fresh OPT1 medium with IPTG again, which is indicative of a metabolic adjustment during the growth retardation. When we looked into the endogenous lipid composition we observed a significant decrease in PG that anti-correlated with the level of AG production. PG levels are primarily controlled by activity of PgsA and its substrates [355]. The involvement of an allosteric downregulation of the endogenous phosphatidylglycerol production by PgsA, however, seems unlikely. Because PgsA is also required for AG formation its activity appears not to be inhibited. The formation of PG on the other hand is also indirectly controlled by a compensatory activity of endogenous PssA that competes for the same substrate CDP-diacylglycerol [356]. Because the overall anionic lipid content (39%) was fairly high, *E.coli* PssA may have been stimulated to overproduce PS as a means to compensate with more zwitterionic lipids (PE). This mechanism seems like the most plausible explanation to the observed lower PG contents. The change in endogenous lipid composition and the membrane-associated location of the polar head-group enzymes CdsA, PgsA and PgpA, strongly supports that nascent archaeal lipids become incorporated in the membrane. Despite compensatory changes in the membrane and a fairly normal growth, SEM and TEM images revealed a high degree of morphological malformations. Most strikingly are the lobular appendages with diameters ranging between 100-500 nm. Besides extraneous appendages, the cells also disseminated these bodies by an apparent budding event and formation of vesicles, intra- and extracellularly or between the inner and outer membrane. Because of a seemingly dose-dependent response, it is quite evident that the diether lipids were directly responsible for these malformations. The exact mechanism behind the localized dissociations remains unclear, but might very well be triggered by localized accumulation of archaeal lipids in patches or the immiscibility of the archaeal lipid species with the bacterial variants. In turn this may have resulted in locally reduced membrane stabilities and the formation of vesicles to expel the foreign lipids. SEM and TEM analysis suggested the possible formation of extremely large aggregates in cells with a delayed induction (**Fig. 7B**, lower panel). These findings show that late induction at early

log phase was probably disadvantageous to proper folding of the produced proteins and therefore resulted in dense aggregates likely consisting of the heterologous gene products either with or without archaeal lipids. More importantly, it supports that the many lobular appendages and intracellular vesicles cannot be explained by the sheer presence of the heterologous expression of membrane proteins. Because inclusion bodies were detected in high abundance in both strains, the morphological changes of lobules and intracellular or extracellular vesicle formation are very likely attributed to the presence of the foreign lipids. In addition to morphological changes, we also observed a dramatic reduction in heat tolerance which may support immiscibility. An increased fluidity is likely the reason behind this reduction as we also observed a cryo-protective effect of the mixed membrane. Here it is commonly accepted that psychrophiles amend a more fluid membrane as a means to lower the phase-transition temperature of their membrane.

In conclusion, our data provides evidence that heterochiral hybrid membranes can sustain a viable situation but, at least under the tested conditions, this may result in dramatic phenotypical alterations. Our findings also support that archaeal lipids are incorporated in the membranes, but the miscibility of bacterial and archaeal lipids remains questionable which may be related to the concentration of archaeal lipids or their unsaturation. This conceivable heterogeneous distribution of lipids may in fact have been the reason for the here reported budding, and possibly also may have contributed to the lipid-divide: the divergence of the archaeal and bacterial domains.

Materials and Methods

MEP-DOXP operon construction

The *idi, ispDF,* and *dxs* genes were amplified by PCR using *E.coli* MG1655 genomic DNA as template. DNA was extracted according to the manufacturer's procedures 'Qiagen DNeasy Blood & Tissue Kit'. For gene amplifications, primers were used as indicated for the *idi, ispDF* and *dxs* genes in **Table 3**. Amplified MEP-DOXP genes were cloned in consecutive steps in pGFPuv (Clontech). Restriction digestions, electroporation and cloning procedures were carried out according to standard molecular biological techniques according to Maniatis *et al* [239].

PCR amplifications were performed according to the standard protocol of Phusion DNA polymerase (Thermo Fisher Scientific). The *idi* gene was amplified with an annealing temperature of 58°C and extension time of 80 s. The *idi* insert and pGFPuv vector were double digested for 2h with SphI and PstI (Thermo Fisher Scientific). The insert was subsequently ligated into linearized pGFPuv by T4 DNA ligase (Thermo Fisher Scientific) to yield pGFPuv-*idi* (**Table 4**, pMS03). The subsequent *ispDF* fragment was amplified at an

annealing temperature of 58°C and an extension time of 3 min. The insert was subsequently cloned into pMS03 using the XbaI and PstI restriction sites to yield pGFPuv- *idi-ispDF* (pMS08) using the XbaI and PstI restriction sites. The subsequent *dxs* fragment was amplified at an annealing temperature of 60°C and an extension time of 4 min. The insert was cloned into (pMS08) to yield the final construct pGFPuv-*idi-ispDF-dxs* (pMS11). For this final cloning the XbaI and the KpnI restriction sites were used.

Table 3: Oligonucleotide sequences. Restriction sites are underlined. For MEP-DOXP operon integrations, homologous regions are underlined.

Primer	TG	Sequence (5'→3')	RS	Primer name
Cloning primers				
BG3606	*idi*	GGCATGCCATGCAAACGGAACACG	SphI	For-a1.1 SphI_Idi
BG3599	*idi*	GCTGCAGTTATTTAAGCTGGGTAAATGCA	PstI	Rev-a1.1 idi_PstI
BG3600	*ispDF*	GCTGCAGAGGAGATATACATATGGCAACCA CTCATTTGG	PstI	For-a1.2Pst_RBS_ IspDF
BG3601	*ispDF*	GTCTAGATCATTTTGTTGCCTTAATGAGTAGCG	XbaI	Rev-a1.2 IspDF_XbaI
BG3602	*dxs*	GTCTAGAGGAGATATACTGATGAGTTTTG ATATTGCCAAATAC	XbaI	For-a1.3XbaI_rbs_dxs
BG3603	*dxs*	GCGGTACCTTATGCCAGCCAGGCC	KpnI	Rev_a1.3dxs_KpnI
BG3899	*crtE*	GAACGAATTCAGCCCGAATGACGGTCTGC	EcoRI	For_EcoRI_CrtE
BG3900	*crtE*	GAATCTTAAGGCGCGACCAGTTCCTGAG	AflII	Rev_AflII_CrtE
BG3901	*crtBI*	GCTGAGATCTGATGAAACCAACTACGGTAATTGG	BglII	For_BglII_CrtBI
BG3902	*crtBI*	CTTACTCGAGAAAGACATGGCGCTAGAG	XhoI	Rev_XhoI_CrtBI
BG4429	*BB*	GACGCGTACGGTGTCTTTTTTACCTGTTTGACC	BsiWI	For-BsiWI-PI-BB
BG4430	*BB*	GACGCTTAAGCTACCTCTGGTGAAGGAGTTGG	AflII	Rev-AflII-PII-BB
BG4476	*kanR*	TTCCGTCAGCCAGTTTAGTC	-	Rev-Kan1-BB
BG5004	*aceB*	TACATCGAAGCGTGGATCTCTGG	-	For EI–scr UP
BG5005	*aceA*	ATCTGATCGGCACGACGGAAGG	-	Rev EI–scr DOWN
Knock-in primers				
BG4885	MEP-DOXP operons	CTGACCCTGCCAGGCTACCGCCTGTTAGCGTAAACCACCA-CATAACTATG GCTACCTCTGGTGAAGGAGTTG	-	For-H1B-P2
BG4886	MEP-DOXP operons	CTGTAATTCTTCAATTTGTTGTGTACGGGTTTTCATGTGCAGAT-GCTCGGCAGTGAGCGCAACGCAATT	-	Rev-H2B-Lac O3

Abbreviations: TG: Target gene; RS: restriction enzymes; BB: backbone, represents the *lox71-kanR-lox66* knock-in cassette

A *lox71-kanR-lox66* gene cassette as selection marker was cloned to pCR2.1 topo by topo-TA cloning to yield pMS46. The purpose of the cloning was to extend the cassette with an N-terminal BsiWI and a C-terminal AflII restriction site. These sites allow the insertion of the *lox71-kanR-lox66* cassette downstream of the previously constructed MEP-DOXP

operons. The *lox71-kanR-lox66* gene template (provided by T. van Rossum) was amplified with Phusion polymerase. The PCR reaction contained: 1 µl *lox71-kanR-lox66* template, 0.2 µM forward primer (BG4429), 02 µM reverse primer (BG4430), 0.2 mM dNTPs, 1x Phusion buffer, and 1U Phusion polymerase. The amplification of *lox71-kanR-lox66* was done at an annealing temperature of 52°C and an extension time of 40s. Correct insertion was verified by sequence analysis. The main purpose of the following cloning procedure was the exchange of the *gfp* gene from the previously cloned constructs pMS03, pMS11 with the *lox71-kanR-lox66* selection marker. An Acc65I (BsiWI) and BspTI (AflII) double digestion on 2 µg pMS03 and pMS11 plasmid DNA was performed to generate a linear vector. To generate a *lox71-kanR-lox66* insert, 4 µg pMS46 was similarly digested. Digests were applied on 1% agarose gel and separated for 30 min at 100V. Correct fragments were excised from gel and purified on a DNA clean and concentrator-5 column (ZYMO Research). The *lox71-kanR-lox66* inserts were ligated with linearized vectors to generate pMS51 and pMS53 (**Table 4**).

Table 4: Plasmids and inserts used in this study

Annotation	vector	Insert 1	Insert 2	Insert 3	Insert 4
pMS03	pGFPuv	*idi*			
pMS08	pGFPuv	*idi*	*ispDF*		
pMS11	pGFPuv	*idi*	*ispDF*	*dxs*	
pMS17	pACYCduet	*crtE*	*crtBI*		
pMS46	pCR2.1 topo	*lox71-kanR-lox66*			
pMS51	pGFPuv	*idi*	*lox71-kanR-lox66*		
pMS53	pGFPuv	*idi-IspDF-dxs*	*lox71-KanR-lox66*		
pMS148	pETduet	*crtE*	*araM*		
pMS184	pETduet	*crtE*	*ΔaraM**		
pSP001	pRSF	*GGGPs M.marp*	*DGGGPs M.marp*		
pSJ139		*GGGPs M.marp*	*DGGGPs A.fulg*		
pAC027	pRSF	*GGGPs M.marp*	*DGGGPs M.marp*	*carS*	
pAC028	pRSF	*GGGPs M.marp*	*DGGGPs M.marp*	*pssA*	
pAC029	pRSF	*GGGPs M.marp*	*DGGGPs M.marp*	*carS*	*pssA*

*Δ-sign denotes deletion,

Operon integration

Homologous end extension PCR was performed by Phusion polymerase PCR (ThermoFisher Scientific) to generate integration cassettes. The following primers were used to extend the operons with homologous ends (**Table 3)**. For the EPOD-independent locus insertion, BG4885 and BG4886 primers were used and annealed at 56°C. A standard PCR protocol was performed to amplify the knock-in cassettes for the *idi*, and *idi-ispDF-dxs* operons using pMS51, and pMS53 template DNA, respectively. The resulting PCR-fragments were purified on a clean and concentrator 5 column (ZYMO research).The integration and clearing of the selection marker was performed as described previously in the gene disruption strategy

by Datsenko and Wanner [234]. Briefly, competent JM109(DE3) cells with induced pKD46 for Red-recombinase **(WUR61)** were electroporated with 375 ng purified cassette DNA at 2500V, 200Ω, 25 μF and 5 ms. After 1.5 h recovery on SOC medium at 37°C, cells were plated on LB-agar supplemented with 50 μg/ml kanamycin. Mutant colonies were screened by colony PCR using primers BG5004 and BG5005 for the EPOD-independent locus according to standard PCR conditions. Positive clones were annotated **WUR63**, and **WUR71** for the *idi-lox71-kanR-lox66* and *idi-ispDF-dxs-lox71-kanR-lox66* operons in the EPOD-independent region, respectively. Correct insertion was confirmed by sequence analysis of the amplified knock-in region with primers BG4476 and BG4886. Novel strains were transformed with pJW168 harbouring the Cre-recombinase to remove the kanamycin resistance gene leaving a *lox72* scar sequence. The cells were spread on LB-agar plates supplemented with 100 μg/ml ampicillin and 0.5 mM IPTG to induce cre-recombinase and grown at 30°C. To finalize recombination, purification of the cells and clearance of the plasmids, the colonies were re-streaked on LB-agar plates without additives and grown at 37°C. The novel strains were annotated **WUR89**, and **WUR91** for the integrated operons *idi,* and *idi-ispDF-dxs,* respectively. The mutant strains were retransformed with the lycopene reporter construct pMS17 to yield strains WUR92 and WUR94.

Table 5: Strains and respective genotypes used in this study

strain	Genome integrations	Plasmid 1	Plasmid 2
WUR45	- (Wild type JM109(DE3))	-	-
WUR61	-	pKD46	-
WUR63	*idi- lox71-kanR-lox66*	-	-
WUR71	*idi-ispDF-dxs-lox71-kanR-lox66*	-	-
WUR89	*idi*	-	-
WUR91	*idi-ispDF-dxs*	-	-
WUR92	*idi*	pMS17	-
WUR94	*idi-ispDF-dxs*	pMS17	-
WUR150	-	pMS148	pSJ139
WUR151	*idi*	pMS148	pSJ139
WUR152	*idi-ispDF-dxs*	pMS148	pSJ139
WUR153	*idi-ispDF-dxs*	pMS148	pSP001
WUR168	*idi-ispDF-dxs*	pETduet	pSJ139
WUR177	*idi-ispDF-dxs*	pMS48	pAC027
WUR178	*idi-ispDF-dxs*	pMS48	pAC028
WUR182	*idi-ispDF-dxs*	pMS48	pAC029
WUR187	*idi-ispDF-dxs*	pMS84	pAC029

Assembly of a lycopene reporter construct pMS17

The *crtE, crtB and crtI* genes were PCR-amplified using *Pantoea ananatis* genomic DNA as template DSMZ (DSM30080) Amplifications were performed at an annealing temperature of 55°C and an extension time of 35 s with the primers indicated in **Table 3**. pACYCduet vector

(Novagen) as well as the amplified *crtE* fragments were digested with EcoRI and AflII (BspTI). The vector was treated with antarctic phosphatase (NEB) to prevent self-circularization. The *crtE* insert was ligated into linearized pACYCduet by T4 ligase to obtain pACYC-*crtE*. The *crtB* and *crtI* genes constitute a bicistronic mRNA and were PCR-amplified as a single fragment. The amplified *crtBI* fragment was amplified at an annealing temperature of 58°C and an extension time of 80 s. The insert as well as the previously constructed pACYC-*crtE* plasmid were digested with BglII and XhoI. The linearized plasmid was treated with antarctic phosphatase (NEB) to prevent self-circularization. Ligation by T4 ligase yielded the completed lycopene reporter pACYC-*crtEBI*. From this point on this construct was referred to as pMS17 (**Table 4**).

Lycopene quantification

An aliquot (10 ml) from each culture was centrifuged at 4,700 x*g* for 10 min. The cell pellet was resuspended in 1 ml MQ, and transferred to a clean 2-ml screw-cap vial and centrifuged (12,000 x*g*, 5 min). The residual cell pellet was suspended in 500 µl acetone and incubated 30 min at 55°C to promote extraction and centrifuged again (20 min, 16,000 x*g*, 4°C). The supernatant containing lycopene was transferred to a clean screw-cap vial. The extracts were kept in the dark to prevent light induced break-down of lycopene. This extraction procedure was repeated two times and the three lycopene extracts pooled to a final volume of 1.5 ml. The extract was centrifuged additionally at 16,000 x*g* for 2 min to pellet incidental cell-debris. A 250-µl sample was diluted with 750 µl acetone and the absorbance was measured at 472 nm in a chemical-resistant cuvette (Brand GmbH). The lycopene concentration was determined by interpolation of the measured absorbance to a standard calibration curve of lycopene (Sigma-Aldrich). The calibration curve of lycopene dissolved in acetone resulted in the following equation (y=0.2254x-0.0077) established at an R^2 =0.9998. The total lycopene concentration was normalized to the dry cell weights (DCW) per 10 ml induced culture. For DCW determinations, a 10-ml culture aliquot was centrifuged (4,700 x*g*, 2 min), and the pellet was resuspended in 1 ml MQ and transferred to a pre-weighed aluminium dish. The cells were dried for 24 h at 105°C and weighed. The lycopene production was expressed as "µg lycopene/mg DCW". All obtained values were derived from biological triplicates.

Plasmid 1, pMS148 and pMS184 construction

Plasmid pSJ130 (pETduet-*araM*) harbouring the *Bacillus subtilis* G1P-dehydrogenase was used as vector for cloning of pMS148. The *crtE* gene of *Pantoea ananatis* was excised from the cloning vehicle pMS14 (pCR2.1-*crtE*) which was already used to construct pMS17. The *crtE* gene and pSJ130 were both digested with EcoRI and ligated to obtain pMS148. A negative control plasmid with a deleted section of the *araM* gene was made by EcoRV and BmgBI double digestion of pMS148 and a subsequent blunt ligation from the linearized plasmid to obtain pMS184.

Plasmid 2, pSP001 and pAC series construction

The *mmarc7_RS05150* gene encoding the *Methanococcus maripaludis* geranylgeranyl-glyceryl-phosphate synthase was cloned to the EcoRI and HindIII sites of the pRSF-duet vector. The obtained plasmid was complemented with the *mmarc5_RS08865* gene encoding the *Methanococcus maripaludis* 2,3-di-*O*-geranylgeranylglycerol phosphate synthase at the NdeI and XhoI site to yield pSP001. To construct pAC027, the codon optimized *AF1749* gene encoding the CarS enzyme from *Archaeoglobus fulgidus* was cloned to the NcoI and HindIII sites of pSP001 vector. To obtain pAC028, the *pssA* gene encoding the phosphatidylserine synthase enzyme from *Bacillus subtilis* was cloned to the NotI and AflII sites of pSP001 vector. Finally, to construct pAC029, the *pssA* gene encoding the phosphatidylserine synthase enzyme from *Bacillus subtilis* was cloned to the NotI and AflII sites of pAC027 vector.

Growth conditions and lipid analysis

Newly obtained *E.coli* strains were grown aerobically on LB at 37°C, and where appropriate supplemented with 100 µg/ml ampicillin (Amp^{100}) and 50 µg/ml kanamycin (Kan^{50}) to obtain starter cultures. To obtain secondary starter cultures, 25 ml OPT1 medium without IPTG was inoculated at a 1/100 volume and grown overnight at 37°C. OPT1 medium contained (Glycerol 1% (v/v), KH_2PO_4 2.4% (w/v), $(NH_4)_2HPO_4$ 0.4% (w/v), citric acid 0.17% (w/v) as adapted from the previous formulation [213]. After autoclaving the OPT1 medium, filter-sterilized $NiCl_2$ 1 mM, $MgSO_4$ 0.012 mM and 1x MEM vitamin solution (Sigma-Aldrich 58970C) were added and where appropriate IPTG added at 0.1 mM.

For lipid production assays, 220 ml LB or OPT1 medium was inoculated in triplicates using the primary or secondary starter cultures resp. from a 1/100 volume inoculum and grown aerobically overnight at 37°C. The obtained cells were cooled on ice and pelleted by centrifugation (15 min at 4,700 x*g,* 4°C). The supernatant was removed, and the cells resuspended in 10 ml lysis buffer (50 mM Tris-HCL pH=8, 1 mM EDTA, 150 mM NaCL, 1 mM DTT, 0.1 mM PMSF). The cell suspensions were passed through a French press twice at 16,000 psi to lyse the cells. The collected crude extracts were then centrifuged (10 min at 10,000 x*g*, 4°C) to remove insoluble debris. To collect the membrane fractions, the cleared lysate was ultracentrifuged (80 min at 120,000 x*g*, 4°C). The resulting pellets were resuspended in 500-1000 µl lysis buffer by pipetting. To obtain lipid extracts, the rapid extraction and purification of lipids method by Bligh and Dyer was performed [341]. The membrane fractions were quantified and normalized for their protein content (Roti Nanoquant method, ROTH) and diluted to 1 ml with lysis buffer. Prior to dissolving the membranes in 3.75 ml chloroform:methanol (1:2), 2 µl of 20 µM eicosane was added as internal standard. After addition of 1 ml MQ, the samples were vortexed and 1.25 ml chloroform added. The samples were again vortexed and 1.25 ml MQ added. After vortexing and centrifugation (15 min at 2,000 rpm), the chloroform phases were transferred to a clean glass tube. To purify the extracts, 2.25 ml authentic upper phase was added and vigorously vortexed. The chloroform phase was again separated after centrifugation (15 min at 2,000 rpm). The chloroform was evaporated under a stream of

nitrogen gas. The resulting lipid pellets were dissolved in 300 µl chloroform:methanol (1:2), again pelleted under a stream of nitrogen and dissolved in 100 µl methanol. Finally, the lipid extracts were briefly centrifuged to pellet insoluble particles.

LC-MS analysis

The samples obtained by lipid extraction and *in vitro* reactions were analysed using an Accela1250 high-performance liquid chromatography system coupled with an electrospray ionization mass spectrometry (ESI-MS) Orbitrap Exactive (Thermo Fisher Scientific). A volume of 5 µl per sample was used for the analysis. The LC-MS method parameters used in this study were the same as described previously [22]. The capillary and the tube lens voltage were set to -75 V and -190 V, respectively. For the samples from *in vivo* experiments, a sample volume of 5 µl was injected into a Shim-pack XR-ODS/C8/Phenyl column with dimension 3.0 mm I.D. x 75 mm (Shimadzu) operating at 55°C with a flow rate of 400 µl/min. Mobile phase A (10 mM ammonium formate with 0.1% formic acid in water:acetonitrile 40:60 (v/v) and mobile phase B (10 mM ammonium formate with 0.1% formic acid in acetonitrile:isopropanol 10:90 (v/v) were used as follows: initial condition started with 40% B, a linear gradient was started in 2 min from 40% to 43% B, gradient from 43% to 50% B in 0.1 min, 54% B isocratic for the next 9.9 min, linear gradient from 54% to 70% B in 0.1 min, 99% B isocratic for the following 5.9 min, gradient from 99% to 40% B in 0.1 min and 40% B isocratic for the last 1.9 min[43]. The MS settings used for this analysis were the same as described above. The Thermo XCalibur processing software was used for data analysis and the Genesis algorithm for automated peak detection and integration was applied to this analysis.

PacBIo and Illumina whole genome sequencing

One ml **WUR182** cell suspensions were obtained from cultures grown on LB (control) or OPT1 medium as described in the section 'growth conditions and lipid analysis'. Chromosomal DNA was isolated according to the gram positive DNA purification protocol of the EPICENTRE masterpure DNA purification kit (MGPO4100). Quality and quantity of DNA isolates were checked by QUBIT quantification (Thermo Fisher Scientific). Next Generation Sequencing was executed by Baseclear B.V. Leiden. A *de novo* hybrid assembly was done by means of PacBio sequencing on the unaltered isolate of WUR182 cells (sample A), and on two independent cultures after induction on OPT1 (samples B and C) by Illumina sequencing. The *de novo* hybrid assembled genome was annotated using the Baseclear annotation pipeline based on Prokka version 1.6 Prokaryotic Genome Annotation System. Illumina generated sequence reads were aligned against the assembled hybrid genome using the 'Map reads to reference' option of the CLC Genomics Workbench version 8.5.1. Output of the variant detection was given in single nucleotide variants, multiple nucleotide variants and short InDels. Identified mutations were interactively analysed using the BaseClear online annotation browser. Intergenic regions were inspected for possible alterations of promoter regions and transcription factor sequences by means of the online transcriptional regulatory network database (http://regulondb.ccg.unam.mx/).

Stereo-selective NMR

The NMR spectra were recorded on a Varian Unity Inova 600 MHz spectrometer using deuterated chloroform. The signals are given as in parts per million (δ, ppm) and were determined relative to the proton resonance of chloroform at 7.26 ppm as the internal standard. All Mosher's ester derivatives (synthetic and naturally derived) were isolated as mixtures (*sn*-3 synthetic, *sn*-1 synthetic, mixture of *sn*-1 and *sn*-3 synthetic, *E. coli* extracted and *E. coli* G1P deficient strain extracted lipids). The spectrum of the synthetic (*R,R*)-lipid (*sn*-3 stereochemistry) shows a pair of sharp double doublets at 4.525 and 4.505 ppm whereas in the spectrum of the synthetic (*S,R*)-lipid we observe similar double doublets at 4.575 and 4.555 ppm, for all naturally derived Mosher's esters double doublets at 4.575 and 4.555 ppm were observed.

In vitro activity assays

GGGP synthase activity

In vitro reactions were performed in 100 µl of assay buffer containing a final concentration of 50 mM Tris-HCl pH7.5, 10 mM $MgCl_2$, 60 mM NaCl, 100 mM imidazole, 0.08% DDM and 4% glycerol. Where specified, 100 µM GGPP, 10 mM G3P, 10 mM G1P and the indicated amount of purified enzymes were added to the reaction mixture. The reactions were incubated at 37°C for 1 h. The kinetics assays were performed using the same mixture but the reactions were incubated at 37 °C for 2 h. For determining the *in vitro* bacterial PlsB activity, 100 µl of reaction volume was used containing the following assay buffer: 50 mM Tris-HCl pH 8, 10 mM $MgCl_2$, 100 mM KCl, 20% glycerol, 2mM DTT, 2.8 mM lipids (DOPG:DOPE, 2:1). Where specified, 200 µM oleic acid, 10 mM G-3-P, 10 mM G-1-P, 1mM ATP, 40 µM CoA and the indicated amount of purified enzymes were added to the reaction mixture. Reactions were incubated at 37°C for 3 h and the products were extracted two times with 0.3 ml of *n*-butanol. Extracted lipids were evaporated under a stream of nitrogen gas and resuspended in 50 µl of methanol for the LC-MS analysis.

In vitro reconstruction of AE and AG

For the *in vitro* reconstitution of AE and AG, 100 µl of reaction volume was used containing the following assay buffer: 50 mM Tris-HCl pH 7.5, 10 mM $MgCl_2$, 15 mM NaCl, 25 mM, Imidazole, 0.02% DDM, 1% glycerol, 100 ul GGPP and 10 mM G3P. The indicated amount of purified enzymes was added to the reaction mixture. Reactions were incubated at 37 °C for 1 h as described previously, and the products were extracted twice with 0.3 ml of *n*-butanol. Extracted lipids were evaporated under a stream of nitrogen gas and resuspended in 50 µl of methanol for the LC-MS analysis.

Scanning electron microscopy

Droplets (150 µl) of culture medium containing bacteria or isolated beads from bacteria were put on poly-L-lysine coated cover slips (Corning art. 354085) for 1 h to immobilize the samples on the surface. The glasses with the attached cells were put in 2.5% glutaraldehyde in 0.05 M sodium cacodylate buffer pH 7.2 at room temperature for one h, rinsed 3 times in the same buffer and post fixed for 1 h in 1% OsO4 (w/v) in the same buffer. Hereafter they were washed with water twice, dehydrated in a graded ethanol series (10, 30, 50, 70, 90, 100, 100%) and were critical point dried with carbon dioxide (Leica EM CPD 300). The glasses were attached on a sample holder by carbon adhesive tabs (EMS Washington USA), sputter coated with tungsten (Leica EM SCD 500) and analysed and digitally imaged with a field emission scanning electron microscope (FEI Magellan 400). Sample preparation, imaging and measurements were performed by the Wageningen Electron Microscopy Centre (WEMC) facility.

Transmission electron microscopy of thin sections

E.coli cells in stationary phase were fixed by incubation in 2.5% glutaraldehyde in 0.05 M sodium cacodylate buffer pH 7.2 for 1.5 h at room temperature. After three washes in the same buffer a post fixation was done in 1% (w/v) OsO_4 and 1.5% (w/v) potassium ferricyanide in the same buffer for 1 h at room temperature and subsequently rinsed three times in water. One volume of the sample suspensions were mixed with equal volumes of 3% low melting point agarose. Between each step the samples were centrifuged for 5 min at 8,000 rpm and the pellets were suspended again in the next liquid. After centrifugation the agarose was solidified at 0°C and the parts containing the pelleted sample were cut in very small pieces. The sample pieces were dehydrated for one h in 1% uranylacetate (w/v) in ethanol 50%, infiltrated with modified Spurr resin mixture (Serva) and polymerised for 8 h at 70°C. The samples were ultra-thin (70-80 nm) sectioned with an ultra-microtome (Reichert, Ultracut S), post-stained with 5 times diluted Reynold's lead citrate for 5 min and analysed with a transmission electron microscope (JEOL, JEM 1011). Micrographs were made with a digital camera (Olympus, Veleta) and measurements were done with iTEM software. Sample preparation, imaging and measurements were performed by the Wageningen Electron Microscopy Centre (WEMC) facility.

Physicochemical survival assays

WUR182, WUR187, WUR45 and WUR94 cells were grown on OPT1 medium to stationary phase as described in the section 'growth conditions and lipid analysis'. Cultures were cooled on ice for 15 min. Cell turbidity of the cultures was measured spectrophotometrically at 600 nm as a reference to subsequent dilutions. An aliquote of cells was diluted to an OD_{600} = 1. A subsequent aliquot was again diluted 1000x in 50 ml pre-cooled OPT1 medium. At least 25 samples were aliquoted at 20 µl per ep and snap-frozen on dry ice and stored at -80°C. As a reference, 20 µl cells were plated on LB-agar supplemented with appropriate antibiotics in triplicate in a dilution series of 10, 100 and 1000x. The cells were grown overnight at 37°C

and the resulting colonies counted to estimate the exact amount of cells per aliquote. For cold treatments, cells which were not snap-frozen were stored at -20°C or 5°C. To estimate the survival rate after cold treatments, the cells stored at -20°C, 5°C and -80°C were thawed on ice for 10 min. The cells were then diluted with pre-cooled OPT1 medium and plated on LB-agar at appropriate volumes. The amount of raised colonies were referenced to the earlier determined cell concentration without any treatment and expressed as a percentage of CFU recovery. For heat shock treatments, cells stored at -80°C were thawed on ice for 10 min. The samples were exposed to a heat treatment for exactly 120 seconds at a calibrated 41.2, 45.34, 49.6, 53.8, and 58°C in triplicate. Immediately after the heat shock, the cells were incubated on ice to allow recovery for at least 5 min. After recovery, the cells were suspended in 980 µl SOC medium and incubated at 37°C for 1 h. Appropriate volumes of cells were plated on LB-agar to quantify the percentage of CFU recovery. For butanol exposure assays, the heat shock treatment protocol was applied, with the omission of the heat shock. Here, the SOC medium was supplemented with the following concentrations at 0, 0.5, 1.0, 1.5, 2.0 and 2.5% butanol (v/v). To avoid evaporation, the cells were transferred to screw-cap tubes during incubations.

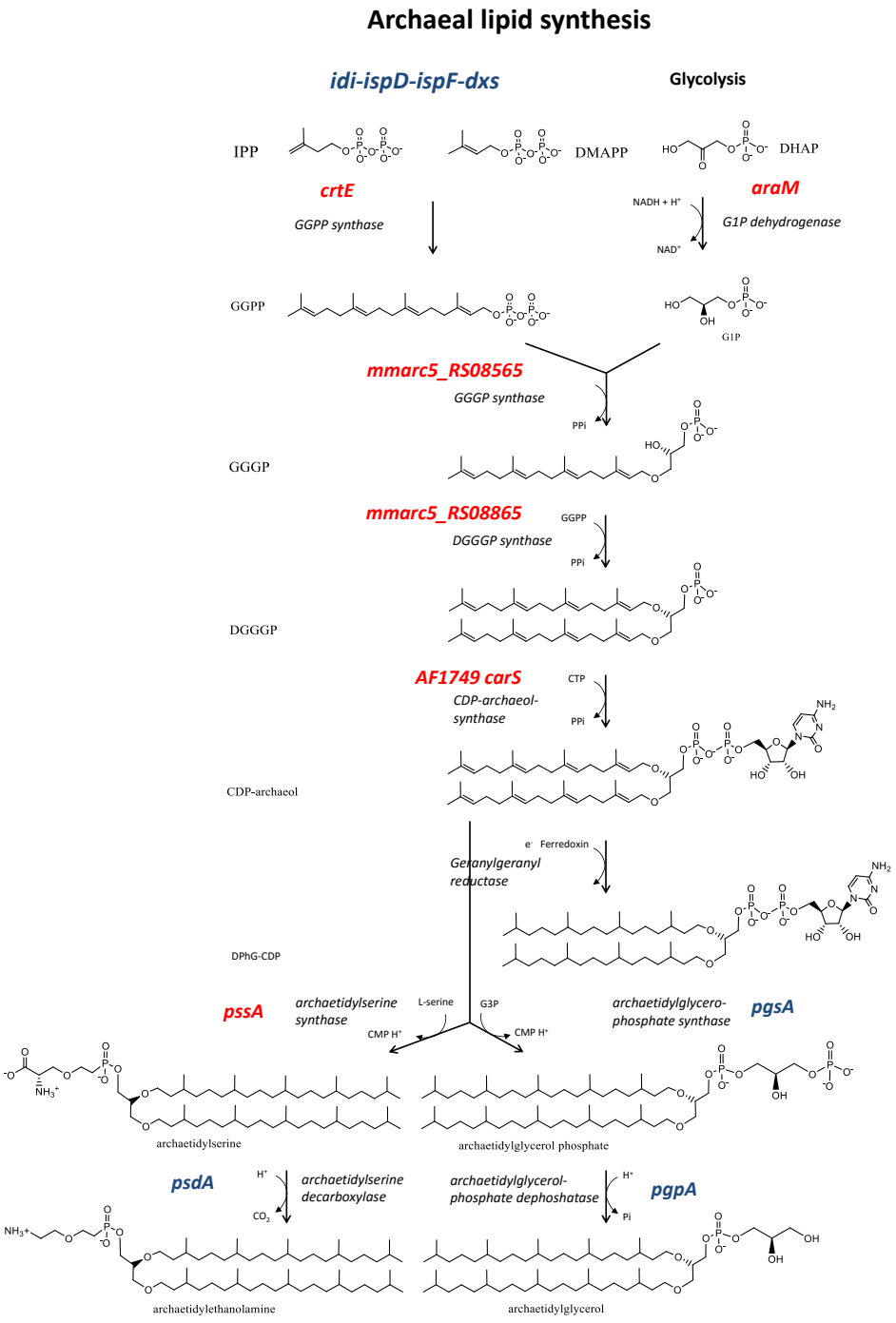
Acknowledgements

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

Authors contributions

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

Supplementary data



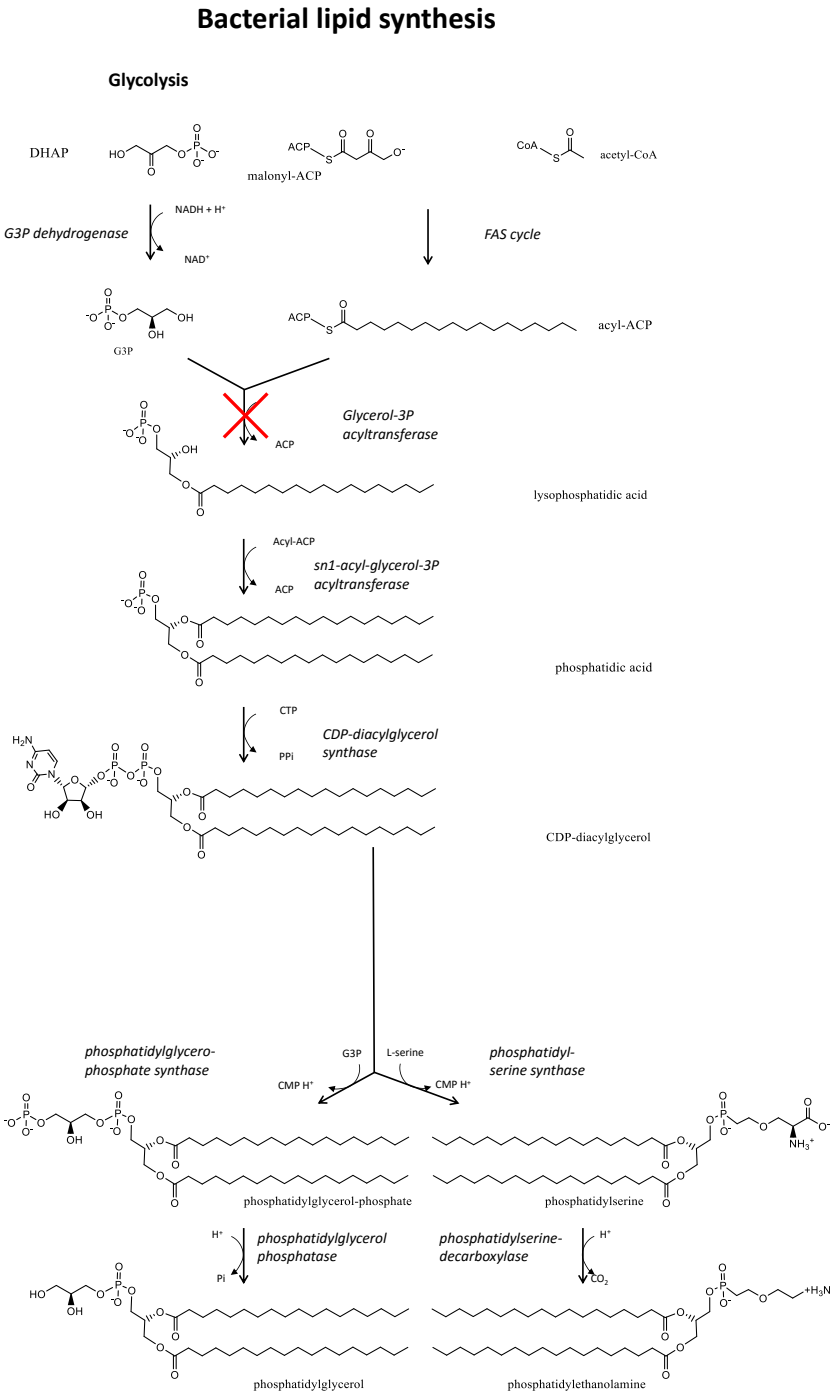


Figure S1. Biosynthesis routes of archaeal and bacterial phospholipids. In red, the introduced genes for archaeal core lipid synthesis. In blue, *E.coli* genes responsible for promiscuous activity in the synthesis of archaeal lipids.

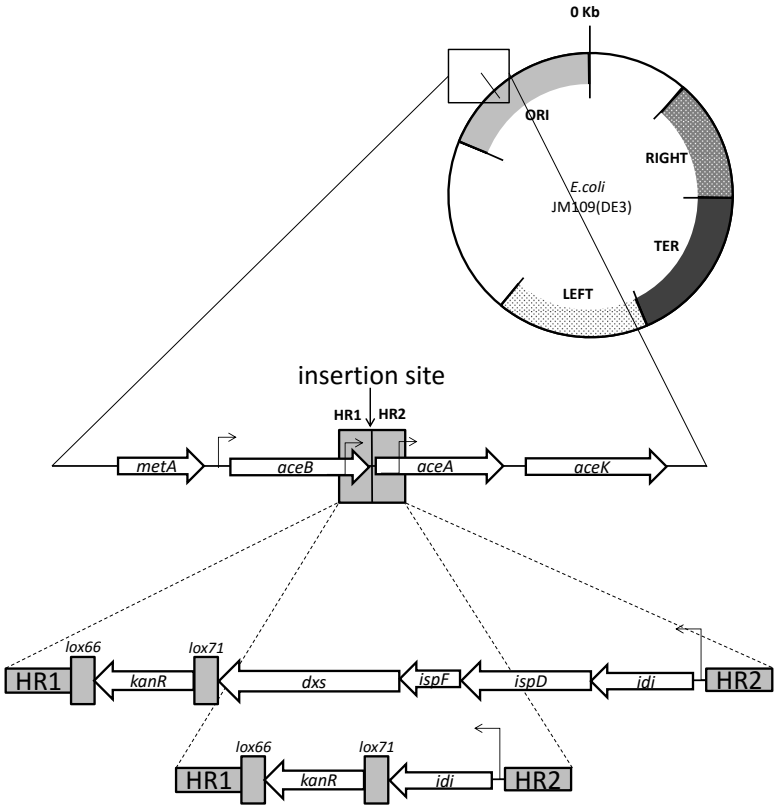


Figure S2. Integration locus and operon constructs. For chromosomal integration of synthetic MEP-DOXP operons, a locus between the *aceB* and *aceA* genes was selected encoding the non-essential: malate synthase A and isocitrate lyase respectively. Abbreviations: HR, homologous region (adapted from Bryant *et al.* [223]).

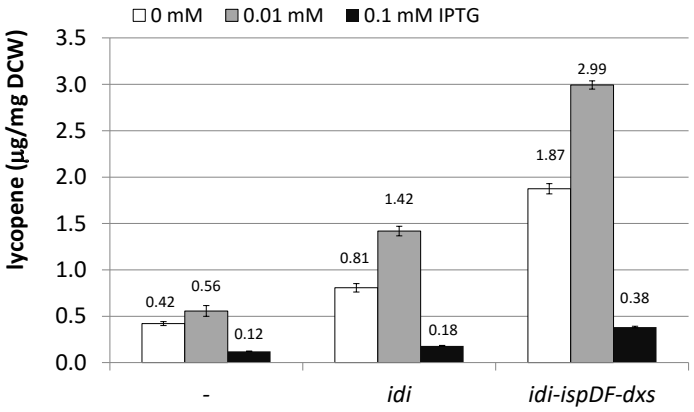


Figure S3. Lycopene production. The integration of *idi* only or *idi-spDF-dxs* operon is compared with a wild-type control for production of lycopene to estimate MEP-DOXP upregulation.

Chapter 7

Identification of *Thermococcus kodakarensis* genes involved in GDGT synthesis

Melvin F. Siliakus[1†], Sebastiaan K. Spaans[1†], Kira S. Makarova[3], Rie Matsumi[4], Antonella Caforio[2], Arnold J.M. Driessen[2], Willem M.de Vos[1], Servé W.M. Kengen[1], John van der Oost[1]

[1]Laboratory of Microbiology, Wageningen University and Research Centre, Stippeneng 4, 6708 WE Wageningen, The Netherlands
[2]Department of Molecular Microbiology, Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, 9747 AG Groningen, The Netherlands
[3]National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, 8600 Rockville Pike, Bethesda, MD, 20894, USA
[4]Department of Microbiology, The Ohio State University, 105 Biological Sciences Building, 484 West 12th Avenue, Columbus, OH 43210, USA

[†]Authors contributed equally

Abstract

A typical feature of archaea is that their membranes are composed of ether lipids, unlike the ester lipid membranes of bacteria and eukarya. Apart from a diether state, also condensation to membrane spanning tetra-ether lipid occurs at least in some archaea. The enzyme that catalyzes this condensation reaction has not yet been identified. In this study, we set out to identify the enzyme(s) involved in the synthesis of tetraether lipids in *Thermococcus kodakarensis*, more specifically the one(s) responsible for the condensation of two C_{20} isoprenyl chains, i.e. the putative tetraether synthase(s). To do this, a list of genes that are potentially involved in tetraether lipid synthesis was composed and prioritized based on a couple of characteristics assumed to be important for a tetraether synthase. This resulted in the selection of 12 genes, which were targeted for overexpression and disruption. Subsequent diether/ tetraether lipid analysis of the mutant strains showed that at least three of these genes, *TK1548*, *TK1836* and *TK1957* play a role in controlling the amount of glycerol dialkylglycerol tetraethers (GDGTs) in the archaeal membrane. Whereas *TK1548*, annotated as serine-pyruvate aminotransferase, has no obvious link with lipid biosynthesis, both *TK1836* and *TK1957* have putative roles as geranylgeranyl hydrogenase and polyprenyltransferase, respectively. Their potential role in GDGT synthesis is discussed. Although the key enzyme responsible for tetraether synthase remains to be identified, our analyzed lipid contents strongly suggest that unsaturated diether lipids serve as substrate for the missing link in tetraether synthesis.

Keywords: archaea; ether lipid biosynthesis; membrane lipids, GDGT, tetraether lipids, mass spectrometry (MS), *Thermococcus kodakarensis*

Introduction

Micro-organisms are able to colonize and survive an extraordinary diversity of environments. Most strikingly, some not only survive, but even thrive under extreme conditions such as very high temperatures (thermophiles), salinities (halophiles), or extreme pH levels (acido- and alkaliphiles). Though the majority of the first identified extremophiles belong to the archaea, they are also abundant amongst the bacterial domain. Since their discovery, extremophiles have attracted much attention because of both their biotechnological potential and their remarkable physiological capabilities. Obviously, the main focus of how these microbes withstand the perturbations of cellular integrity has been on the cytoplasmic membrane, as the membrane shields the inner homeostatic conditions of the cell from the hostile environment. Before the discovery of extremophiles, the membrane composition and adaptation strategies of mesophilic bacteria were well described [246, 289, 357, 358].

The cytoplasmic membrane generally consists primarily of a phospholipid bilayer. In bacteria, the framework molecule of these phospholipids is phosphatidic acid. Phosphatidic acid is composed of two acyl-chains which are esterified to the backbone *sn*-glycerol-3-phosphate. Phosphatidic acid serves as a basis for a high variety of phospholipids, which are adjusted to the requirements of the membrane. The main membrane parameters that are subject to change in response to varying physicochemical conditions are fluidity and permeability [44, 359]. In response to physicochemical stressors, bacteria regulate the saturation and length of the acyl chains to keep the membrane parameters within acceptable boundaries. Archaea have a different lipid architecture and thus also differ in their homeoviscous adaptation strategies [71]. Instead of phosphatidic acid, archaea typically use archaetidic acid as the framework molecule of their phospholipids. Archaetidic acid, which is also known as (S)-2,3-di-*O*-geranylgeranyl-*sn*-1-glyceryl phosphate (DGGGP), differs from phosphatidic acid in several key aspects. For example, the hydrophobic moiety is composed of highly branched C_{20}-isoprenoid chains, rather than the bacterial acyl chains. Moreover, these chains are bound by ether bonds (as opposed to the bacterial ester bonds) to *sn*-glycerol-1-phosphate, the stereochemical opposite of the bacterial *sn*-glycerol-3-phosphate [292, 295]. These key differences are unique to both domains and have therefore been termed “the lipid divide” [7]. Hence, phosphatidic acid and archaetidic acid are regarded as marker molecules between the prokaryotic lineages. Archaetidic acid is considered as the framework molecule of a broad range of archaeal membrane lipids [31]. Nevertheless, archaeal membranes typically consist of a variation of two main core structures: *sn*-2,3-diphytanyl glycerol diether, which in this study will simply be referred to as diether (and is also known as archaeol), and *sn*-2,3-dibiphytanyl diglycerol tetraether, which in this study will be referred to as tetraether (and is also known as glycerol-dialkyl-glycerol tetraether (GDGT), or caldarchaeol). The diether lipids are generated from archaetidic acid by reducing the unsaturated isoprenoid chains, forming C_{20}-phytanyl chains, and by subsequent addition of a polar headgroup to the glycerol moiety.

Tetraether lipids are bipolar, consisting of saturated C_{40}-biphytanyl chains (a molecule that is composed of two condensed C_{20}-phytanyl units) and a polar head group at each end of that chain. Because of their bipolarity, tetraether lipids form a monolayer membrane, whereas diether lipids form a bilayer structure that closely resembles the traditional double leaflet bilayer of bacteria [9, 292, 295].

The variation observed in tetraether lipids involves the incorporation of one or more pentacycles, the incorporation of a single hexamer, the addition of methyl groups, or the incorporation of bridges between the isoprenoids [68, 71, 360, 361]. Tetraether lipids are believed to contribute considerably to an increased stability, permeability and resistance of the membrane [121, 359]. Moreover, the diether:tetraether ratio appears to vary substantially between archaea at different conditions. And for some hyperthermophilic archaea the amount of tetraethers has been shown to increase gradually as the growth temperature rises [68, 153, 362]. This suggests that the diether:tetraether ratio might be regulated to adapt membrane fluidity and permeability. Additionally, for some hyperthermophilic archaea it has been shown that in response to increasing temperatures, the number of incorporated pentacycles also increases [70, 153, 359]. However, as thermal adaptation has frequently been studied in archaea, it soon became clear that not all archaea form tetraethers and that different archaeal species respond to temperature fluctuations in different ways.

In general, tetraethers have been shown to be the dominant lipid species in Crenarchaeota and Thaumarchaeota, whereas members of the Euryarchaeota synthesize diethers, tetraethers, or both [9]. However, the exact route leading to tetraether synthesis is still unknown. Nevertheless, two different pathways have been postulated (**Fig. 1**) [9, 292, 295]. In the first pathway (which we will refer to as the traditional route), DGGGP is first converted to saturated diether lipids, after which two of these molecules are coupled via the yet unknown route by a tail-to-tail condensation of the isoprenoid chains to form one tetraether molecule [10, 292, 295]. Evidence for the traditional route comes from pulse chase and labelling experiments using cell extracts of the euryarchaeon *Thermoplasma acidophilum* incubated with ^{14}C-mevalonate, which is a precursor for the synthesis of the isoprenoid building blocks IPP and DMAPP [10, 292]. These experiments showed that the label first incorporates into diethers until saturation and only then into tetraethers. Moreover, when an inhibitor of squalene epoxidase (known as terbinafine) was used, an inhibition of tetraether formation was observed. In addition, pulse labelling in the presence of terbinafine showed an accumulation of radioactivity in the diether fraction and a decrease in radioactivity in the tetraether fraction, which was reversed by removal of the inhibitor [10, 363]. These observations are interesting as, in the traditional route, the putative tetraether synthase is predicted to be an analog of squalene or phytoene epoxidase [10].

However, no incorporation of radiolabelled diether lipids (i.e. archaeol) into tetraether lipids (i.e. GDGT-0) was observed *in vivo* in the euryarchaeon *Methanospirillum hungatei* [364].

Moreover, from a chemical point of view, the required carbon-carbon bond formation of tetraethers between diethers (i.e. the condensation of two saturated phytanyl chains of archaeol) is considered quite unusual and some studies indeed suggest that the presence of unsaturations is a prerequisite for the formation of tetraethers [9, 92, 365]. These observations are contradicting the ones described above and suggest that saturated diether lipids are not the precursor of tetraether lipids. This discrepancy might, for example, be explained by the possibility that a diether accumulation upon terbinafine exposure is the result of growth retention due to toxicity rather than the inhibition of a specific enzyme. This alternative might very well be true as it has been demonstrated several times that the ratio of diether-to-tetraether is highly dependent on growth phase as well [31, 68, 360, 366].

Recently, an alternative route for the formation of tetraether lipids has been proposed in which the tail-to-tail condensation is thought to take place directly after the synthesis of various forms of isoprenoid chains (i.e. with or without pentacycle moieties) [9]. According to this model, the tail-to-tail condensation of two C_{20} isoprenyl molecules is thus suggested to occur before attachment of the hydrocarbon chains to glycerol-1-phosphate and the subsequent reduction of the chains (**Fig. 1**). The consequence of this proposal is that it eliminates the requirement of the unusual condensation of two saturated phytanyl chains present in diether lipids (i.e. archaeol). The alternative route implies that diether lipids are not the precursor of tetraether lipids and that the ether-bond forming enzymes, GGGP synthase and DGGGP synthase (responsible for the coupling of the isoprenyl chains to the glycerol unit), can accomodate not one but various prenyl species as substrate. These substrates include both C_{20} and C_{40} isoprenoid chains (for the formation of diether lipids and tetraether lipids, respectively) with or without ring moieties. Thus, in this alternative pathway, tetraether lipids are formed by creating C_{40} isoprenoid chains first, which are subsequently used as a substrate for the GGGP synthase and DGGGP synthase to attach both the first and second glycerol units. However, like the traditional route, the enzyme and the hypothetical intermediates responsible for the condensation of the C_{20} isoprenyl chains is still unknown.

Despite the high interest in the biochemical mechanism, no information is available about the enzymes involved in the tail-to-tail condensation. This lack of knowledge might partly be due to the long-term absence of archaeal genetic toolboxes. Here, the anaerobic, sulphur reducing, hyperthermophilic Euryarchaeon *Thermococcus kodakarensis* was used as a model organism that meets a couple of important criteria that enable such a study: (i) *T. kodakarensis* has previously been confirmed to have both diether lipids and tetraether lipids in its cytoplasmic membrane [367],(ii) *T. kodakarensis* has a genetic toolbox available[368-371], (iii) It has a relatively small and fully sequenced genome of an estimated 2306 coding DNA sequences [372], many of which have been functionally characterized already, and (iv) *T. kodakarensis* is fairly easy to grow in a laboratory setting. By performing comparative bioinformatics analyses, we selected several candidate genes that might encode for enzyme(s) involved in the tail-to-tail condensation reaction.

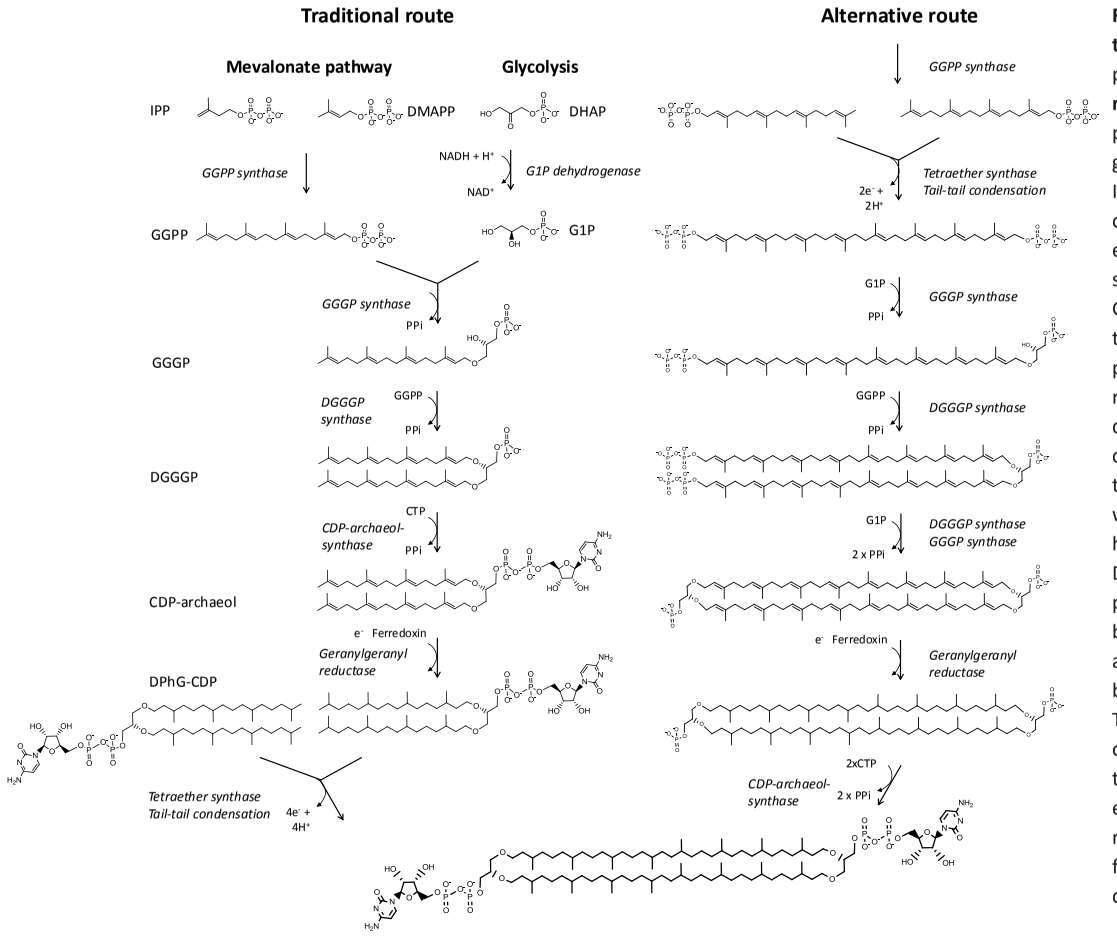


Figure 1. Two proposed pathways for tetraether synthesis. The first proposed pathway is represented by the "**Traditional route**". This pathway proposes a substrate product relationship between archaeol and glycerol-dialkyl-glycerol-tetraethers (GDGTs). Initial steps are completed by full synthesis of DGGGP. From this stage, archaeol can either bifurcate towards full synthesis of saturated diethers or towards synthesis of GDGT's by the successive condensation of the two tail ends of two perpendicularly positioned archaeols. The condensation reaction is here presented by two saturated diphytanylglycerol-CDP (DPhG-CDP) and the concomitant release of two electrons and two protons. Although the saturated diether was the first proposed substrate (as depicted here), some studies suggest the unsaturated DGGGP is a more likely substrate. The second pathway that has been proposed is presented by the "**Alternative route**". In this pathway a C_{40} hydrocarbon chain is synthesized first by condensation of two molecules of GGPP. The GDGT is finalized by the attachment of the glycerol moiety to the tail ends of the C_{40} chains by ether bond formations by either GGGP or DGGGP synthesis. These two reactions are considered equivalent to the formation of GGGP and DGGGP, except for the different chain length of the substrates.

Moreover, we composed a list of characteristics that were expected to be possessed by a tetraether synthase. By scoring the candidate genes according to these characteristics, we ultimately selected 12 candidate genes that were targeted for overexpression and deletion. The generated mutant strains were analysed for their diether/tetraether lipid composition.

Results

Selection of candidate genes

To select gene(s) that are potentially involved in tetraether synthesis, two independent bioinformatics approaches were performed. For both analyses, we assumed a direct correlation between the presence of putative genes (genotype) and the presence of tetraether lipids (phenotype). In addition to archaea producing both diether and tetraether lipids, there are also a number of archaea reported within the phylum of Euryarchaeota that completely lack tetraethers [71]. In the first analysis, we made use of these phenotypes by selecting a set of reference species that are phylogenetically closely related to *T. kodakarensis* and that have a confirmed presence or absence of tetraethers in their membranes. The selected reference genomes belonging to the first group (i.e. confirmed presence of tetraethers) include the hyperthermophiles *Pyrococcus furiosus* and *Thermoplasma acidophilum* [10, 373-375] and those of the second group (i.e. confirmed absence) include the haloalkaliphilic *Natronomonas pharaonis* and the hyperthermophilic *Methanopyrus kandleri* [299, 376]. The genome of *T. kodakarensis* was screened for homologs present in the first group, but absent in the second group, by performing a phylogenetic profiler single gene search (using the IMG system of JGI). This resulted in the identification of 57 *T. kodakarensis* genes that may potentially be involved in tetraether synthesis (**Table. S1A**).

In the second comparative genome analysis we attempted to correlate the presence of specific arCOGs (archaeal Clusters of Orthologous Genes [377] in 149 sequenced archaeal genomes to species with a confirmed tetraether lipid membrane. The arCOGs included in the analysis (**Table. S2**) were related to the following categories: phospholipid biosynthesis, isoprenoid mevalonate biosynthesis, uncharacterized proteins associated with mevalonate biosynthesis, terpene cyclases, fatty acid synthesis, and phytoene/squalene synthesis or saturases. However, none of the arCOGs analysed showed a strong correlation with species that have a confirmed tetraether lipid membrane. Assuming correct membrane profiling, this analysis suggests that tetraether synthesis is either catalysed by different enzymes over different archaeal lineages, or that the responsible enzyme is present in all archaea whose synthesis or activity is tightly regulated causing different phenotypes (e.g. by post-translational modification, temperature activation, etc.). Alternatively, one cannot rule out that the reported lipid analyses (tetraether/diether presence) are correct.

Since no correlation was observed, it was decided to select all uncharacterized *T. kodakarensis* genes that are present in the mentioned arCOGs categories instead. This led to the selection of 32 additional genes. To exclude other members of the archaeal lipid biosynthetic pathway (e.g. IPP synthase, GGPP synthase, GGGP synthase, etc.) we compared the selected genes with experimentally characterized genes. Here we made use of homologs from various archaea that have a confirmed enzymatic activity in reactions leading to isoprenoid and CDP-archaeol synthesis. To do this, characterized genes from each individual step in lipid synthesis [34, 298, 304, 306, 348, 351, 378] were analysed by BLAST search against the genome of *T. kodakarensis*. *T. kodakarensis* genes that showed a high sequence similarity to these characterized genes were excluded from this list (**Fig. S1**). However, the candidate genes *TK1324* and *TK1957* both showed sequence similarity to known DGGGP synthases, making it, based on this information only, impossible to decide if both or only one (and if the latter, which one) encode for a DGGGP synthase. Moreover, both are annotated as 4-hydroxybenzoate polyprenyltransferase belonging to the UbiA-like prenyltransferase COG0382, and it was therefore decided to retain and prioritize both of them. This ultimately led to the selection of 27 additional genes that may be involved in tetraether lipid biosynthesis (**Table S1B**).

To reach a final selection, it was decided to prioritize the resulting 84 genes based on characteristics assumed to be important for a putative tetraether synthase. These characteristics were (in a non-specific order): (i) predicted prenyl binding. (ii) the presence of an FAD or Flavin mononucleotide (FMN) binding domain, which is potentially important for tetraether synthases because carbon-carbon bond formations are often aided by FAD cofactors as dual electron acceptor [379]. (iii) the presence of aspartic acid rich DDxxD motifs. These motifs contribute to the substrate binding of prenyltransferases, which are responsible for the transfer of prenyl groups and chain elongation of isoprenoids [337, 380, 381]. (iv) Homology to squalene epoxidase. This has been suggested by Kon *et al*., who, upon addition of a squalene epoxidase inhibiter (i.e. terbinafine) to the growth medium of *Thermoplasma acidophilum*, observed an accumulation of diether lipids at the expense of tetraether lipid biosynthesis [10]. (v) Transcriptional regulation by the *Thermococcales* heat shock protein Phr. Phr is a transcriptional regulator that represses transcription under normal growth conditions by binding to the promoters of certain genes and that is released from them upon heat shock [382-384]. Since previous studies showed an increase in tetraether lipid content upon thermal treatment, transcription of the putative tetraether synthase may also be Phr-dependent [68, 69, 153]. Altogether, this scoring resulted in a ranking of the candidate genes (**Table S1**). In case the candidate genes had an equal score, priority was given to the presence of prenyl binding or -transferase domains. By doing this, we selected 10 priority genes (in addition to *TK1324 & TK1957*) that were targeted for overexpression and disruption (**Table 1**). These genes were: *TK1088, TK0929, TK2159, TK1836, TK0509, TK0513, TK1173, TK1477, TK1548* and *TK1732*.

Moreover, an additional column was included in **Table 1** showing the presence or absence of transmembrane spanning helices. This information may become valuable as it can demonstrate the validity of either the classical or alternative synthesis pathway. This is because in the traditional biosynthesis pathway, the tetraether lipid biosynthesis is proposed to occur by condensing two diether lipids, the synthesis of which already takes place in the membrane[295]. In contrast, however, in the alternative pathway, the condensation reaction leading to the C_{40} isoprenoid chains takes place in the early steps of the ether lipid biosynthesis and is therefore more likely to be catalyzed by a cytosolic enzyme [9].

Construction of mutant strains

To screen for enzymes involved in tetraether lipid formation, 12 genes of interest (GOIs) were selected. For each of these genes, we attempted to create a disruption strain and a strain overexpressing the GOI. *T. kodakarensis* strain TS559 was chosen as a parent strain to utilize the previously developed selection and counter-selection procedure in rich medium (**Fig. S2A**) [368, 369, 385]. Disruption strains were created as described previously [368]. Transformation of the disruption vectors (**Table 2**) resulted in agmatine prototrophs for 11 out of the 12 GOIs. *TK2159* was the only GOI for which no transformants were obtained, not even after 3 independent transformation attempts. Transformants of the remaining 11 GOIs were confirmed to be 6-methyl purine sensitive ($6MP_S$), a phenotype conferred by the expression of *TK0664*. Next, disruption of the GOIs as well as precise and markerless deletion of the selection and counter-selection cassette was stimulated by plating the transformants on plates containing 6MP and agmatine. For each of the remaining GOIs, genomic DNA of 30 spontaneous 6MP resistant ($6MP_R$) colonies was analyzed by diagnostic PCR. In this analysis, we screened for the deletion of both the GOIs and of the selection and counter-selection cassette. Ultimately, 6 of the 12 selected GOIs were successfully deleted. For the remaining 5 (*TK2159* did not result in transformants), only revertants (i.e. parental) were obtained. For the 6 successful GOI disruptions, representative clones having the desired final genome configurations were designated *T. kodakarensis* strain Δ*TK0513*, Δ*TK1324*, Δ*TK1477*, Δ*TK1548*, Δ*TK1732*, and Δ*TK1836*, in which the Δ denotes disruption and the number indicates which TK-gene has been deleted (**Table 2**).

Table 1. Top-10 ranking candidate genes

Gene locus	Gene annotation	Predicted Prenyl binding	FAD/FMN binding domain	DDxxD motif	SQE homology	Phr-regulation	TMH	Total score
TK1088	Geranygeranyl hydrogenase	+	+	+	+	+	0	**5**
TK0929	3-octaprenyl-4-hydroxybenzoate decarboxylase	+	+	+		+	0	**4**
TK2159	Predicted oxidoreductase		+		+	+	1	**3**
TK1836	Geranygeranyl hydrogenase	+	+	+			0	**3**
TK0509	3-octaprenyl-4-hydroxybenzoate decarboxylase UbiX	+	+				0	**2**
TK0513	Hypothetical protein	+				+	0	**2**
TK1173	Undecaprenyl diphosphate synthase	+				+	0	**2**
TK1477	Hypothetical protein			+		+	0	**2**
TK1548	Probable serine—glyoxylate aminotransferase			+		+	0	**2**
TK1732	Dolichol-phosphate mannosyltransferase	+		+			4	2

Abbreviations: b/t, binding or transferase; FAD, Flavin adenine dinucleotide; FMN, Flavin mononucleotide; SQE, squalene epoxidase; Phr, the Phr-transcriptional regulator; TMH, transmembrane spanning helices

Table 2. *T. kodakarensis* strains and plasmids used in this study

Strain or plasmid	Relevant characteristics	Source or reference
Strains		
TS559	Δ*pyrF*; Δ*trpE*::*pyrF*; Δ*TK0664*; Δ*TK0149*	[385]
Disruption strains		
Δ*TK0513*	TS559 derivative, Δ*TK0513*	This study
Δ*TK1324*	TS559 derivative, Δ*TK1324*	This study
Δ*TK1477*	TS559 derivative, Δ*TK1477*	This study
Δ*TK1548*	TS559 derivative, Δ*TK1548*	This study
Δ*TK1732*	TS559 derivative, Δ*TK1732*	This study
Δ*TK1836*	TS559 derivative, Δ*TK1836*	This study
Overexpression strains		
OE0509	TS559 derivative, Δ*TK1765*::*TK0509, TK0149*	This study
OE0513	TS559 derivative, Δ*TK1765*::*TK0513, TK0149*	This study
OE0929	TS559 derivative, Δ*TK1765*::*TK0929, TK0149*	This study
OE1088	TS559 derivative, Δ*TK1765*::*TK1088, TK0149*	This study
OE1173	TS559 derivative, Δ*TK1765*::*TK1173, TK0149*	This study
OE1324	TS559 derivative, Δ*TK1765*::*TK1324, TK0149*	This study
OE1477	TS559 derivative, Δ*TK1765*::*TK1477, TK0149*	This study
OE1548	TS559 derivative, Δ*TK1765*::*TK1548, TK0149*	This study
OE1732	TS559 derivative, Δ*TK1765*::*TK1732, TK0149*	This study
OE1836	TS559 derivative, Δ*TK1765*::*TK1836, TK0149*	This study
OE1957	TS559 derivative, Δ*TK1765*::*TK1957, TK0149*	This study
OE2159	TS559 derivative, Δ*TK1765*::TK2159, TK0149	This study
Plasmids		
pTS700	Basis for the creation of "A"-plasmids, Harboring *TK0664; TK0149*	[368]
pSKS022	Derivative of pSecPyrF [386], enabling integration at *TK1765 locus*. Basis for the creation of vector backbone of OE plasmids, harboring *TK0149*	[387] and [386]
Disruption vectors*		
pΔ"GOI" (12x)	*TK0664*; *TK0149*; Δ*"GOI"*	This study
Integration vectors*		
pOE"GOI" (12x)	*TK0149*; Δ*TK1765*::*"GOI"*	This study

* = for each gene of interest (GOI) a disruption (Δ) vector and integration (or overexpression; OE) vector was created. The GOIs were *TK0509*, *TK0513*, *TK0929*, *TK1088*, *TK1173*, *TK1324*, *TK1477*, *TK1548*, *TK1732*, *TK1836*, *TK1957*, and *TK2159*.

Overexpression strains were obtained by the following procedure. First, a universal vector backbone was created to accommodate the integration of an additional copy of the GOIs into the chromosome of *T. kodakarensis*. Next, the 12 selected GOIs were individually cloned into this vector backbone, placing them under control of the strong constitutive *hmtB*-promoter [385], resulting in 12 unique overexpression vectors (**Table 2**). Subsequent transformation of these vectors in *T. kodakarensis* TS559 ultimately resulted in agmatine prototrophs for all GOIs. For each GOI, genomic DNA of 5 agmatine prototrophic clones was analyzed by diagnostic PCR to confirm integration of the various overexpression vectors. Moreover, the

diagnostic PCR was also used to discriminate between transformants that were the result of a single homologous recombination (resulting in the integration of the entire donor plasmid DNA integrated into the genome of *T. kodakarensis*) and transformants that were the result of two crossover events (resulting in the integration of only the GOI and selection marker (*TK0149*)). This discrimination was done since only the transformants resulting from two crossover events can be considered stable (and are thus preferred), while the transformants resulting from a single homologous recombination may ultimately revert to the original genome. Nevertheless, are both variants considered sufficient for this study as long as the agmatine prototrophy selection pressure is maintained. The diagnostic PCRs showed that preferred double crossover mutants were obtained for *TK1324, TK1477, TK1548, TK1732, TK1088, TK1957* and *TK2159,* but not for *TK0509, TK0513, TK1836, TK1173 and TK0929*. Next, the entire integration locus was PCR amplified and subsequent sequencing confirmed that the integrated additional copy of the GOIs did not contain any undesired mutations. Representative clones, containing correct sequences of the additional copies of the GOIs were designated *T. kodakarensis* strains: OE0509, OE0513, OE0929, OE1088, OE1173, OE1324, OE1468, OE1477, OE1548, OE1732, OE1836, OE1957 and OE2159, in which the OE denotes 'overexpression' and the number specifies the introduced TK gene (**Table 2**).

Finally, we aimed to confirm enhanced transcription or loss of transcription of the overexpressed and knock-out strains respectively. To this end, we isolated total RNA from all OE- and Δ-strains to use in a semi-quantitative Reverse Transcriptase-PCR (RT-PCR) to detect the level of transcription of the respective GOI's. The primers used for the diagnostic PCRs were designed such that amplification of a particular GOI resulted in a PCR product of ~250 bp in case of expression. The results shown in **Fig. 2** confirm increased transcription levels of the GOIs in their corresponding OE-strain and the absence of transcription in their corresponding Δ-strains. The housekeeping gene *puf* (*TK0308*) [388], encoding the α-subunit of the elongation factor-1 complex, was used for normalization and as positive control in the RT-PCRs for the OE- and Δ-strains, respectively.

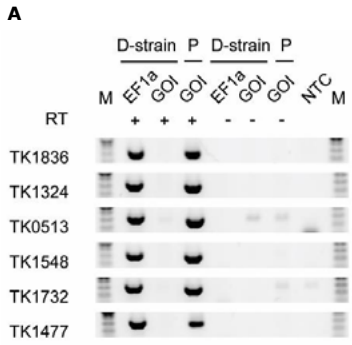
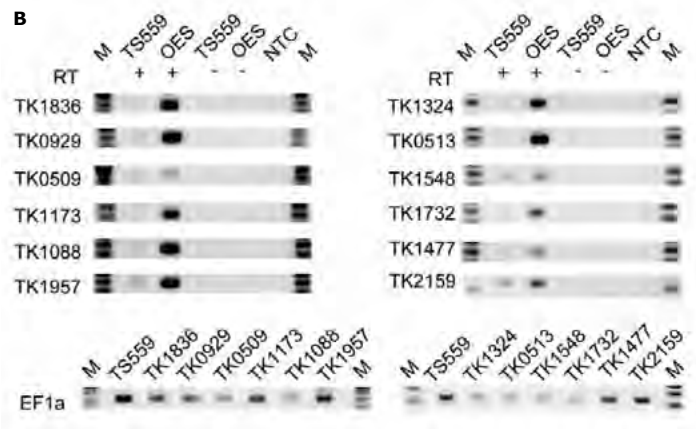


Figure 2. Semi quantitative Reverse Transcriptase-PCR (RT-PCR) analysis Total RNA was isolated from the parental and mutant strains and subjected to a DNase I treatment to degrade traces of isolated chromosomal DNA. cDNA was created via a reverse transcriptase reaction and subsequently used as template in a diagnostic PCR to confirm (**A**) loss of expression or (**B**) overexpression of the genes of interest in the disruption and overexpression strains, respectively. All RT-PCR products had the appropriate size on SYBR safe stained agarose gels. The housekeeping gene elongation factor 1-α was used for normalization. Abbreviations: D-strain, disruption strain; P, parent strain TS559; M, marker; EF1a, elongation factor 1-α; GOI, gene of interest; NTC, no template control; RT, Reverse transcriptase; TS559, parent strain; OES, overexpression strain; TK#, gene locus of GOI. "+", RT present; "-", RT absent



Phenotypical changes elicited by the OE- and Δ-strains

To assess the involvement of the corresponding GOIs in tetraether lipid biosynthesis, the total diether and tetraether lipid content of the OE- and Δ-strains was determined and compared with the lipid content of the parental strain TS559. To do this, we prepared overnight cultures (>16 h) for each strain in triplicate and harvested the cells. The membrane lipids were then hydrolysed to cleave off polar head groups and subjected to a total lipid extraction to isolate core lipids. These isolates were subsequently characterized for their diether and tetraether lipid content by UHPLC-API/MS by monitoring their ions at (m/z=653.6 $[M-H]^+$) and (m/z=1302.3 $[M-H]^+$), respectively. Here, the total amounts of diether and tetraether lipids (in µg) turned out to be highly variable between the different strains, even after normalization per mg dry cell weight (DCW) (**Fig. 3**, in (w/w)%). For example, the total amount of diether and tetraether lipids obtained for strain OE0929 was nearly twice as high as that obtained for the parental strain TS559, while hardly any diether and tetraether lipids were obtained from strain OE1324.

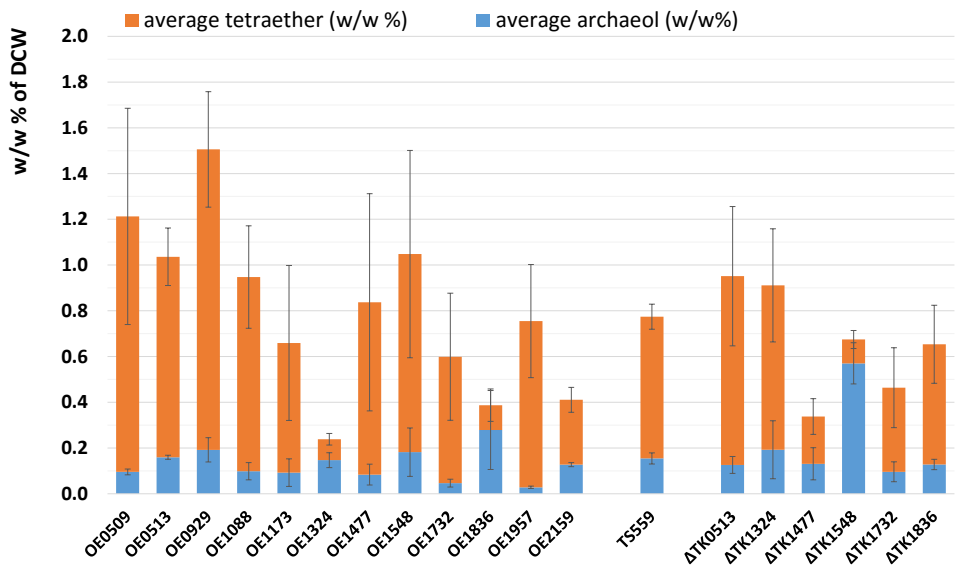


Figure 3. Average diether and tetraether lipid content per strain. orange: average tetraether lipid content (i.e. GDGT-0), blue: average diether lipid content (i.e. archaeol).

To allow a more accurate and realistic comparison of the membrane composition of the different strains, the average diether and tetraether lipid contents per strain was expressed as a percentage of their cumulative amount (**Fig. 4**). For this analysis we also included the detection of unsaturated diether lipids which were missing in the preliminary analysis, by performing HPLC/ESI-MS to detect intact polar lipids (IPLs). This showed that the parental strain TS559 has a tetraether lipid content of 80% ± 3.9, which roughly agrees with the results obtained by Matsuno *et al*. [68]. Concerning the possible involvement of the GOIs in tetraether lipid biosynthesis, the most striking effects were observed for the strains OE1957 and Δ*TK1548*, for which a tetraether lipid content of 96.1% ± 1.1 and 15.1% ± 2.8 (w/w) was detected, respectively. Moreover, an increased but insignificant (p=0.06) tetraether lipid content was also observed for the OE-strain of *TK0509,* having a tetraether lipid content of 91.7% ± 2.3 (w/w). In addition, a significant decrease in the tetraether lipid content of the OE-strain of *TK1836* and *TK1324* was shown. Interestingly, though disruption of *TK1548* led to a dramatic decrease in the tetraether lipid content, no notable increase was observed upon its overexpression. With the exception of the disruption of *TK1548,* no significant changes in the tetraether lipid content (TE:DE ratio) of the other disruption strains were observed. In previous studies, it was established that the tetraether lipid content may vary markedly dependent on the growth phase [68, 366]. As this may have contributed to the different levels of diether and tetraether lipids as well, we determined the dependence of the lipid composition on turbidity in the parental strain TS559. Here we observed that the tetraether content reaches a near stable plateau

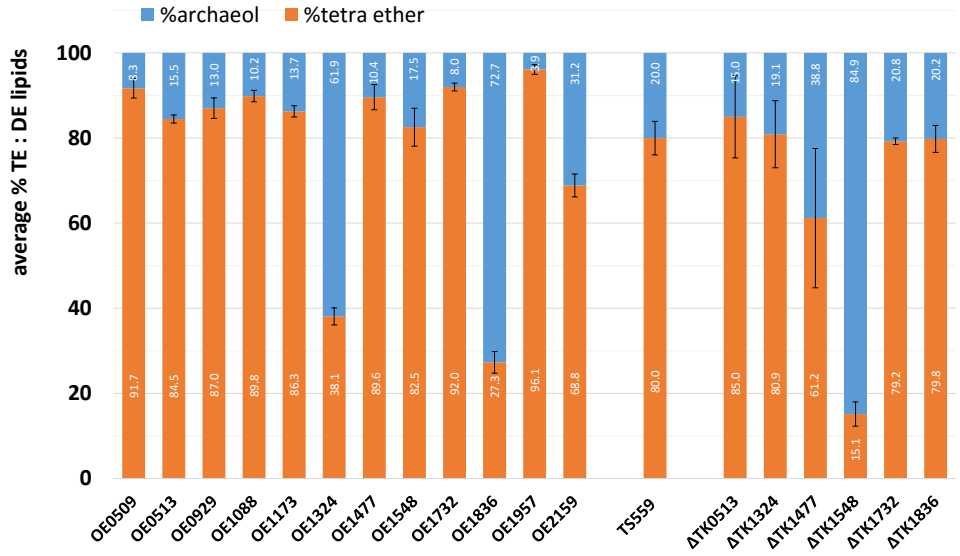


Figure 4. Average % of tetraether and diether lipids per strain The tetraether lipid content (in orange) and diether lipid content (in blue) of three biological replicates was normalized to the total amount of lipids measured per strain (i.e. to the sum of GDGT-0 and archaeol). The vertical numbers (in white) show the average fraction (in %) of tetraether lipids per strain.

after 11 h of growth (stationary phase) (**Fig. S3).** This finding reinforces our postulation that the *TK1548, TK1836, TK1324* and *TK1957* genes are involved in determining the tetraether content. To obtain more insight in the mechanism that causes these shifts in diether:tetraether lipid (DE:TE) ratios, we performed a total ion monitoring on the lipid extracts (without acid hydrolysis), to selectively detect both saturated and unsaturated intact polar lipids species (**Table S4**, IPLs with head groups, hexose, phosphohexose and aminohexose phosphohexose). Interestingly, intact tetraether polar lipids were not detected, probably due to insufficient solubilisation during the extraction procedure. When the diether IPLs were analysed instead, we noticed a marked variation in the level of unsaturated diether lipids (3 unsaturations per lipid) between the mutant strains (**Fig. 5**).

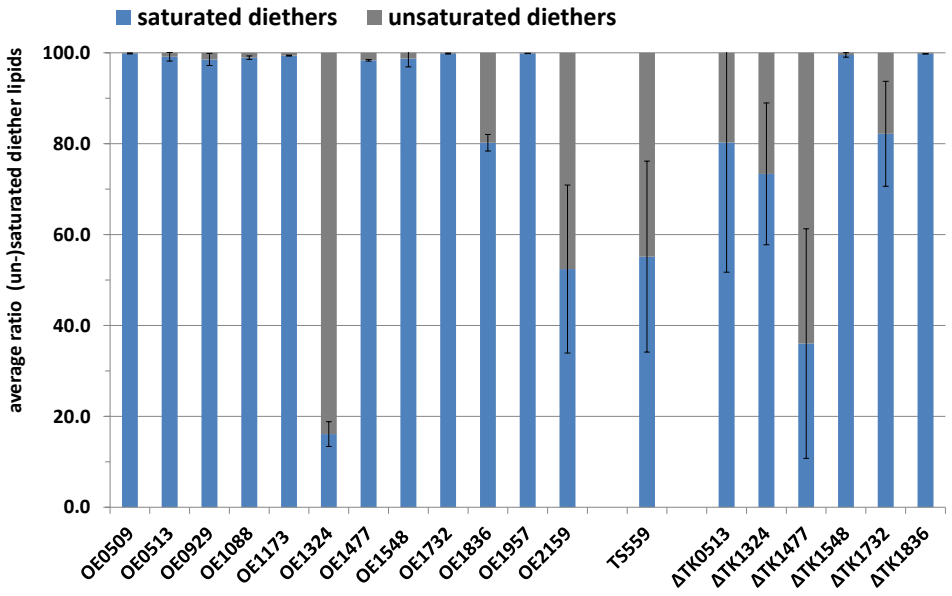


Figure 5. Representation of the cumulative saturated and unsaturated diether lipids per strains The saturated diether levels (in blue) and unsaturated diether lipids (in grey) of three biological replicates were normalized to the total amounts of intact polar diether lipids (IPLs) measured per strain (i.e. the sum of all intact polar lipids of archaeols). As such, the blue bars are representative of the total diether lipid peaks depicted in **Fig. 3**.

Moreover, the previous variations in tetraether content specifically correlates with either a decrease or increase of unsaturated or saturated diether levels. More specifically, mutants with both a low tetraether levels (OE1836 and Δ*TK1548*) and a high tetraether levels (as observed for OE1957) show a marked decrease in unsaturated diether lipids (summarized in **Table S5**, marked). This finding highly suggests that unsaturated diethers may serve as shared substrate for the formation of both tetraether lipids and fully saturated diether lipids. Additionally, the unsaturated lipids (which were not included in the preliminary test) explain the overall low amounts of lipids obtained for OE1324 and ΔTK1477 (**Fig. 3**, lanes 6 & 16).

To investigate whether the altered membrane composition (diether:tetraether lipid ratio) has a detrimental effect on the morphology of *T. kodakarensis*, samples of the parental strain and mutant strains showing the biggest extremes in their membrane lipid composition (strain OE1957 and Δ*TK1548*) were studied by transmission electron microscopy (TEM). In addition, samples of the mutant strain OE1324 were included to check for its cellular integrity as this strain showed low overall lipid content consisting of high levels of unsaturation (**Fig. 3**, lane 6). The obtained images showed neither a difference on the morphology of the cells, nor on the physical appearance of the membranes of *T. kodakarensis* (**Fig. 6**, upper and lower right panel).

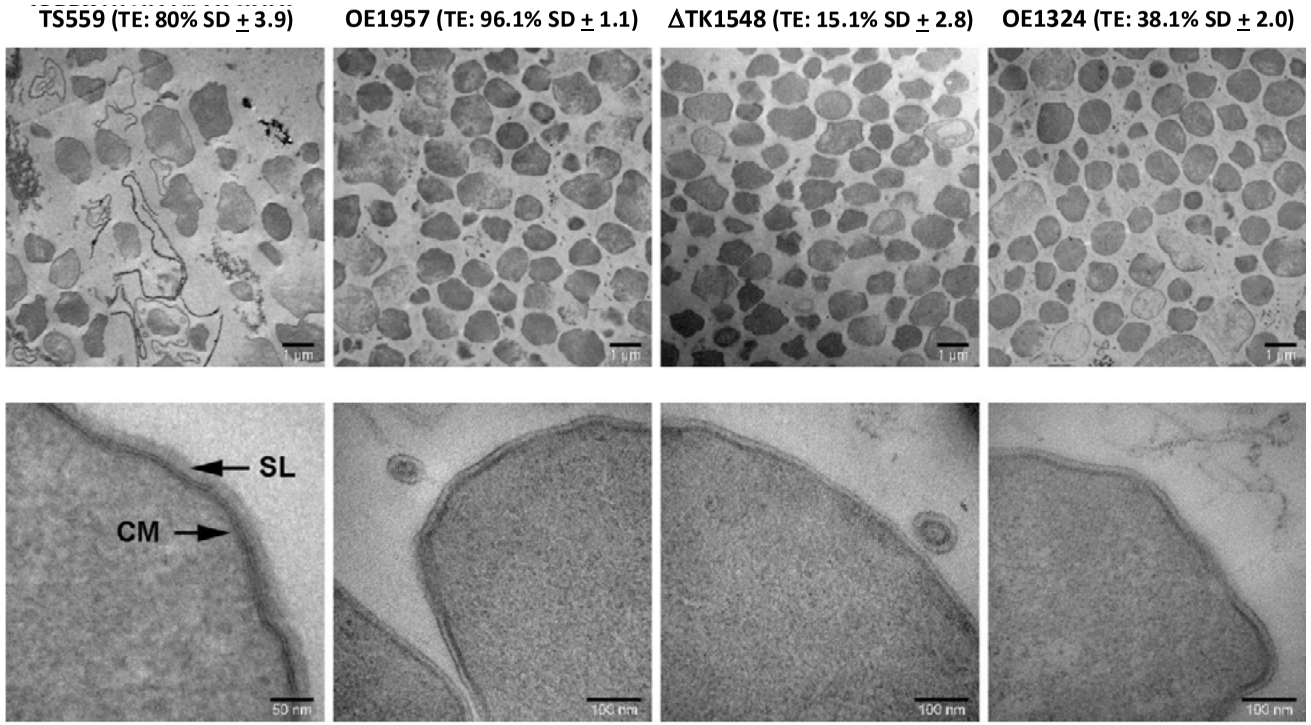


Figure 6. Thin cross-sections of *T. kodakarensis* strains TS559, OE1957, Δ*TK1548* & OE1324 analysed by Transmission Electron Microscopy (TEM). Strain OE1957 and Δ*TK1548* were selected for TEM as these strains showed the highest and lowest tetraether lipid content (TE), respectively. Strain OE1324 was included to check its cellular integrity as hardly any diether or tetraether lipids were obtained for this strain in the membrane characterization experiment. The upper micrographs show multiple cells of each strain (bars 1 µm). The lower micrographs show the cytoplasmic membrane (CM) and S-layer (SL) of individual cells (bar TS559 50 nm, bars OE1957, Δ*TK1548* & OE1324 100 nm). The tetraether lipid content (TE) of each strain is shown between brackets.

Discussion

In this study, we set out to identify genes involved in the synthesis of tetraether lipids. More particularly, we focused on the enzyme(s) responsible for the condensation of two C_{20} isoprenyl chains. To do this, we composed a list of genes that are potentially involved in tetraether synthesis, and prioritized these based on characteristics assumed to be important for a tetraether synthase. This resulted in the selection of 12 genes, which were targeted for overexpression and disruption. A subsequent lipid analyses on the mutant strains, strongly suggests that at least four of these genes (*TK1548, TK1836, TK1324* and *TK1957)* indeed play an intimate role in the synthesis of tetraether lipids. In addition, the genes that appear to be non-viable upon disruption may be interesting in this respect as well. Especially, the candidates that did not yield a disruption strain but instead showed a slight increase in the tetraether lipid content upon its overexpression, such as *TK0509* and *TK1088* may be of significance. This is because the membrane of *T. kodakarensis* exists primarily of tetraether lipids (i.e. 80.0%). Losing the ability to synthesize tetraether lipids, may possibly result in a lethal phenotype.

One of the genes that seems to affect the synthesis of tetraether lipids is *TK1548*. Interestingly, this gene is annotated to encode a serine-glyoxylate aminotransferase. This gene was selected from the list of candidate genes as a priority gene since it encompasses a degenerate DDxxD motif, is associated with substrate binding of prenyltransferases, and it has been shown to be slightly upregulated during heat shock (**Table 1**). Disruption of *TK1548* resulted in a strain (Δ*TK1548)* with significantly lower tetraether lipid contents (**Fig. 3** and **4**, lanes 17), which is a strong indication that the enzyme encoded by this gene plays an important role in tetraether synthesis. However, despite the disruption of *TK1548,* a considerable amount of tetraethers could still be detected (15.1% ± 2.8). This either suggests there is not just one encoding gene responsible for tetraether synthase and hence there may be multiple isozymes involved as well, or that *TK1548* encodes for an enzyme that indirectly influences tetraether synthesis. Moreover, since this gene has been annotated as an aminotransferase it is also possible that the annotation about amino substrate specificity is wrong, meaning that it could be involved in lipid synthesis and that the observed effect may be elicited directly. Interestingly, overexpression of *TK1548* did not result in a significant effect (**Fig. 4**, lane 8), which would agree with the conclusion that it is not a tetraether synthase, but might also simply imply it is not a rate controlling step in the pathway. Furthermore the IPL analyses showed that Δ*TK1548* results in membranes with very low levels of unsaturated diether lipids. This suggests that unsaturated diether lipids may act as a substrate for tetraether synthesis and hence correlates with the shortage of tetraether lipids. A second gene that has an influence on the tetraether lipid content is *TK1836*. This gene has been annotated as a geranylgeranyl hydrogenase and was selected as a priority gene for its predicted prenyl binding, and the presence of a FAD/FMN binding

domain and DDxxD motif (**Table 1**). Though disruption of *TK1836* does not seem to have a detrimental effect on the tetraether lipid content of the membrane, overexpression does (**Fig. 4**, lanes 19 and 10 resp.). These results imply that *TK1836* is involved in lipid biosynthesis, but make it very unlikely to be a tetraether synthase. Instead, phylogenetic analysis suggests that *TK1836* probably belongs to a dehydrogenase family that has been shown to be a homolog of prenyl reductases [389, 390]. In case of a true reductase, the results obtained for this mutant again suggest that the presence of unsaturations in the precursor molecule of tetraether lipids is a prerequisite to the formation of tetraether lipids. This would be an interesting finding, as the absence or presence of double bonds in the C_{20} isoprenyl chains (i.e. before the condensation reaction) is a major difference between the traditional and alternative tetraether biosynthesis hypothesis, respectively [9]. Correspondingly, the diether IPL analysis also showed a significant decrease in the amount of unsaturated diethers which agrees with the suggestion that *TK1836* encodes for a reductase. Equivalent to *TK1548*, a low level of unsaturated diether may imply a shortage of substrate for tetraether synthesis that caused the low tetraether levels.

A third gene that seems to be involved in tetraether lipid synthesis is *TK1957*, which has been annotated as a 4-hydroxybenzoate polyprenyltransferase. Though this gene shows sequence similarity to experimentally characterized DGGGP synthases [304, 348] (which should have led to the exclusion of this gene from the list of candidate genes), it was decided to retain it and to select it as a priority gene for further analysis. The reason for this was that another candidate gene, also annotated as a 4-hydroxybenzoate polyprenyltransferase, i.e. *TK1324*, showed sequence similarity to DGGGP synthases as well. This fact makes it difficult to decide which of these two genes encodes a DGGGP synthase or which may in potential be a tetraether synthase. It was therefore decided to retain both and to select them as priority genes for further analysis. Surprisingly, overexpression of *TK1957* resulted in a strain with a membrane that almost exclusively consists of tetraether lipids (96.1% SD ± 1.1; **Fig. 4**, lane 11). This is a strong indication that *TK1957* is involved in the tetraether lipid synthesis. Moreover, the observation that disruption of *TK1957* appears to be non-viable is quite interesting in this respect as well. However, though these results imply that *TK1957* may be a tetraether synthase, a similar reasoning as described above would also apply to a DGGGP-synthase. This is because, in both the traditional and alternative biosynthesis hypothesis, DGGGP synthases are key enzymes in the formation of archaeal membrane lipids (**Fig. 1**). However, from the small amount of detected diether lipids, the main share consisted of saturated IPLs only (99.9%). This, however, cannot be explained by an excess of DGGGP-synthase activity and strongly argues against the possibility of a DGGGP-synthase. Instead of a postulated shortage of unsaturated diethers (Δ*TK1548* and OE1836), an enhanced conversion to tetraether lipids may explain the observed shortage of unsaturated diethers. Contrary to *TK1957*, overexpression of *TK1324* (the other gene annotated as a DGGGP synthase) resulted in a significant decrease in the tetraether/diether lipid ratio (**Fig. 4**). This implies that *TK1324* is most likely neither a DGGGP synthase nor a

tetraether synthase. However, two important observations suggest that *TK1324* is involved in the biosynthesis of lipids as well. Firstly, we obtained a significant decrease in the amount of total lipids (di- and tetraethers) upon its overexpression (**Fig. 3**), and that the unsaturated diether lipid level expanded dramatically. This strain however, also shows that a decrease in tetraether levels can also correlate with a high level of unsaturated diether lipids.

In conclusion, we provide evidence that *TK1548, TK1836, TK1324* and *TK1957* seem to be interesting candidates to further study tetraether synthesis. In addition, our data strongly suggests that unsaturated diether lipids are substrate to the tail-to-tail condensation in tetraether synthesis and compete with geranylgeranyl-reductases that form the saturated diether lipids. Alltogether, these analyses could provide a better understanding of the tetraether lipid biosynthesis pathway and may ultimately lead to the identification of the tetraether synthase(s).

Materials and Methods

Selection of candidate genes

Two independent bioinformatics approaches were performed. In the first analysis, a set of reference species were selected that are phylogenetically closely related to *T. kodakarensis*, and that have a confirmed presence (*Pyrococcus furiosus* and *Thermoplasma acidophilum*) or absence (*Natronomonas pharaonis* and *Methanopyrus kandleri*) of tetraethers in their membranes [10, 299, 374-376]. The genome of *T. kodakarensis* was screened for homologs present in the first two, but which are absent in the latter two by performing a phylogenetic profiler single gene search using the IMG system of JGI (https://img.jgi.doe.gov/cgi-bin/m/main.cgi). In the second analysis, it was attempted to correlate the presence of specific arCOGs (archaeal Clusters of Orthologous Genes [377] in 149 sequenced archaeal genomes to species with a confirmed tetraether lipid membrane.

The list of genes resulting from two bioinformatics approaches was prioritized based on the following characteristics: (i) predicted prenyl binding, (ii) the presence of a FAD or Flavin mononucleotide (FMN) binding domain or Rosmann Fold [391], (iii) the presence of aspartic acid rich DDxxD motifs, (iv) homology to squalene epoxidase, and (v) transcriptional regulation by the *Thermococcales* heat shock protein Phr. Screening for (i) and (ii) was done by scanning the protein sequences against the InterPro collection of protein signature databases (http://www.ebi.ac.uk/Tools/pfa/iprscan5/). The screening for (iii) was done by manually checking the protein sequences for presence of a DDxxD motif. To identify possible homologs of squalene epoxidase (iv), the selected genes were included in a BLAST search using the *Saccharomyces cerevisiae* ERG1 squalene epoxidase sequence (EC1.14.99.7) as a query. The *S. cerevisiae* squalene epoxidase was selected, as a recently

created three-dimensional model of this protein allowed for the identification of crucial residues [392]. To determine if the transcription of the candidate genes may be influenced by Phr (v), we made use of the results of a previously published microarray experiment that was performed to identify the Phr-dependent heat shock regulon in *T. kodakarensis* [382]. The presence or absence of transmembrane spanning helices was determined by constructing hydrophobicity plots using the online Transmembrane Hidden Markov Model (TMHMM http://www.cbs.dtu.dk/services/TMHMM-2.0/) tool to predict probability of transmembrane spanning segments. From this analysis it was shown that the following gene products (*TK1324*, *TK1732*, *TK1957*) have high probability of being integral membrane proteins or membrane associated (*TK2159*).

Microorganisms and culture conditions

E. coli DH5α was used for general DNA manipulation, and *E. coli* strains were cultivated at 37°C in LB medium supplemented with 100 μg/ml ampicillin. *T. kodakarensis* strains and plasmids used in this study are listed in **Table 2**. *T. kodakarensis* strains were cultivated anaerobically in stoppered serum bottles at 85°C in a nutrient-rich medium (ASW-YT) [370], composed of 0.8x artificial seawater (0.8x ASW), 5.0 g/L yeast extract, 5.0 g/L tryptone, 1 ml/L KOD vitamin mixture (1000x) and 0.8 mg/L resazurine. The headspace of all cultures was exchanged for N_2. Before inoculation, either 2.0 g/L sulfur (ASW-YT-S^0) or 5.0 g/L pyruvate (ASW-YT-Pyr) was added to the medium, as well as $Na_2S.9H_2O$ to reduce the dissolved oxygen, until the resazurin became colorless. When required, the media were supplemented with 500 μM agmatine and/or 75 μM 6-methyl purine (6MP) [368, 369]. Cultures were continuously shaken (50 rpm) during growth. In the case of plate culture, 1% (wt/vol) Gelrite was added to solidify the medium together with 2 ml/L polysulfide solution [370] instead of S^0. Cells inoculated on plates were incubated at 85°C for at least 48 h in an air-tight anaerobic jar containing an Anaerocult A mini (Merck Millipore, Darmstadt, Germany) to maintain anaerobic conditions.

Construction of plasmids and mutant strains

Plasmids incapable of autonomous replication in *T. kodakarensis* were constructed and amplified in *Escherichia coli*, carrying the genes of interest (GOIs) or knockout constructs adjacent to the previous developed selection and, in the case of the disruption vectors, counter-selection cassette (*TK0149; TK0664*) [368, 369]. The GOIs and selection cassette (in case of the OE vectors) and the selection/counter-selection cassette (in case of the disruption vectors) were flanked by a ~500-1000 bp region homologous to *T. kodakarensis* genome that are natively located immediately up- and downstream of the target locus (**Fig. S2**). Standard molecular biology procedures were used for PCR amplifications, plasmid construction, *E. coli* DH5α transformation, transformants selection, plasmid isolation, and sequence verification. The sequences of oligonucleotide primers used in this study are listed in **Table S3**.

The vectors created to disrupt or overexpress the GOIs are shown in **Table 2**. The disruption vectors were constructed as described before [368]; first, so-called "A"-plasmids were created by amplifying each GOI along with ~500-700 bp of flanking DNA and cloning the corresponding amplicon into pTS700 in a ligation-independent mechanism. Next, the "A"-plasmids served as a foundation plasmid from which so-called "B"-plasmids, or disruption vectors, were generated. This was done by a single-step PCR-based procedure wherein the original GOI was deleted from the plasmid, while leaving the flanking DNA that will target integration of the entire B-plasmid to the TS559 genome. Two separate initial integration events are possible for incorporation of the entire B-plasmid to the genome, and each generates an intermediate strain that, when 6MP-based counter-selective pressure is applied, can undergo an internal recombination event to either restore the original genome or produce a strain containing a genome with the desired, targeted deletion (**Fig. S2A**).

The overexpression vectors were constructed as follows. First an inverse PCR using the primer-pair BG4881-BG4882 and the previously constructed integration vector pSKS022 [387] as template was performed. pSKS022 is a derivative of pSecPyrF [386] and was chosen as template, as integration in the genome at the *ChiA* locus targeted by this plasmid has shown to be efficient and non-disturbing. The PCR resulted in the amplification of a vector backbone for the overexpression vectors that contained the flanking regions needed for integration as well as the strong constitutive *hmtB*-promoter. The PCR thus removed the *csg*-promoter, the *TM*-gene expression cassette, the synthetic transcription terminator (T_8), the 500-bp pop-out homology, and the counter-selection marker *TK0664* from pSKS022. The linear vector backbone created by this PCR was extended with a unique AgeI restriction site, which was incorporated in the primer BG4882. Next, the GOIs were amplified by Phusion polymerase PCR (Thermo Fisher Scientific) using the corresponding primers listed in **Table S3**. The reverse primers were extended with a unique AgeI restriction site, allowing for easy and directional (blunt-sticky) ligation in the amplified vector backbone, creating the desired vectors as shown in **Figure S2b**. Transcription of the GOIs was designed to be controlled by the strong constitutive *hmtB*-promoter [385].

The *T. kodakarensis* Δ- and OE-strains shown in **Table 2** were generated by transforming the corresponding vectors into *T. kodakarensis* TS559 (Δ*pyrF*; Δ*trpE*::*pyrF*; Δ*TK0664*; Δ*TK0149*) [368, 369, 385]. Disruption strains were created as described previously [368]. Overexpression strains were created by a slightly adapted approach in which the counter-selection marker (*TK0664*) was not involved and desired mutants were thus simply selected by growing colonies on medium lacking agmatine. An important difference between these two methods is that, as a result of excluding *TK0664*, the selection marker (*TK0149*) cannot be removed from the genome after obtaining initial transformants. This means that, in contrast to the disruption strains, the created overexpression strains do not allow another round of the procedure to make additional genome modifications. Diagnostic PCRs were performed to select and confirm genotypes of desired strains. To do this, 6MP-resistant

colonies growing on ASW-YT plates supplemented with agmatine (in the case of the disruption strains) or agmatine-prototrophic colonies growing on ASW-YT plates without agmatine (in the case of the overexpression strains) were picked and used to inoculate liquid ASW-YT-S^0 cultures (with and without agmatine, respectively), which were incubated o/n at 85°C. Next, 1.5-ml culture was sampled using a syringe and genomic DNA was isolated according to the protocol for Gram-positive bacteria of the GeneJET genomic DNA purification kit (Thermo Scientific). Purified DNA was then used as template in three diagnostic PCRs. In the first PCR, the primers used were complementary to the sequence upstream of the desired integration locus as well as to the C-terminus of the integrated GOI (illustrated by primers A and B in **Fig. S2**). Similarly, a second PCR was included using primers that were complementary to the sequence downstream of the desired integration locus as well as to the N-terminus of the integrated GOI (illustrated by primers C and D). A third PCR was performed to identify double crossover mutants, using primers that were complementary to the sequence downstream and upstream of the integration locus (illustrated by primers A and D). A generated PCR product having the correct size for all three reactions suggested the desired double crossover genotype. In addition, fragments containing the additional copy of the GOI were PCR amplified from the genome of the desired clones and sent for sequencing to confirm the absence of undesired mutations. Sequences of the primers are listed in **Table S3**.

RNA isolation and RT-PCR

To confirm loss of transcription or increased mRNA levels of the GOIs in the disruption strains or overexpression strains, respectively, reverse transcriptase PCR (RT-PCR) was performed. To do this, the parental strain TS559 and the various mutant strains created in this study were cultivated in 120-ml serum bottles containing 40 ml ASW-YT-S^0 medium, which in the case of the disruption strains were also supplemented with 500 µM agmatine (Sigma-aldrich) and 75 µM 6MP. Samples of 10 ml were removed and cells were harvested by centrifugation (4700x *g*, 30 min, 4°C). Total RNA was isolated by the following method. Cell pellets were resuspended in 1 ml Trizol and 200µl chloroform was added. The mixtures were shaken by hand for 15 seconds, followed by incubation at RT for 3 min. Next, the samples were centrifuged (13,000x *g*, 15 min, 4°C) and the supernatants were transferred to fresh Eppendorf tubes. 5 µg glycogen and 500 µl isopropanol were added and the mixtures were incubated at RT for 10 min. The samples were centrifuged (13,000x *g*, 15 min, 4°C) and the pellets were washed in 75% ethanol, air-dried for 10 min and resuspended in 20 µl MQ water. RNA concentrations were determined with a Nanodrop ND-1000 and the total RNA samples were stored at -20°C until further processing.

Before RT-PCR, the total RNA samples (1 µg each) were treated with DNaseI, to remove residual genomic DNA. cDNA was prepared by following the protocol for first strand cDNA synthesis of the Thermo Scientific RevertAid RT kit (20 µl reactions), using random hexamers

as primers. Gene-specific cDNA fragments were amplified by conventional PCR using the freshly prepared cDNA as template and primer pairs developed to specifically target the genes of interest (**Table S3**). The housekeeping gene *TK0308 (puf)*, encoding the α-subunit of the elongation factor-1 complex (EF-1-α), was included for normalization [388]. All primer pairs were designed to produce a specific product of ~250 bp. All reactions were run for the amount of cycles needed to make the respective product in the reactions containing cDNA of the parental strain TS559 just visible. The only exception to this are the reactions used for normalization (*TK0308)*, which were run for 18 cycles. PCR amplification products were analyzed by agarose gel electrophoresis. Control experiments were included by following the same procedure, but without addition of RevertAid Reverse Transcriptase enzyme to the RT reaction mixture.

Core lipid extraction and analysis by UHP-APCI/MS

Triplicate cultures of the various mutant strains and parental strain TS559 were cultured on ASW-YT-Pyr. When the cultures reached the stationary phase, cells were harvested by 10 min centrifugation at 4700x *g* and lyophilized overnight. To extract the core lipids of the various strains, lyophilized cells were hydrolysed with 2ml 1.5N HCl in methanol by reflux at 130°C for 2 h. The hydrolysate was adjusted to pH 7 with 2N KOH-methanol (MeOH) (1:1, vol/vol) and, after addition of water to a final 1:1 ratio of H_2O-MeOH, extracted three times with dichloromethane (DCM). The DCM fractions were collected and dried over sodium sulfate. The resulting extract was dried under nitrogen and dissolved in hexane–2-propanol (99:1, vol/vol), filtered over a 0.45-µm polytetrafluoroethylene filter, and analysed by high-performance liquid chromatography (HPLC)–atmospheric pressure chemical ionization (APCI) mass spectrometry (MS) for archaeol and GDGTs. The UHPLC-APCI/MS analysis was done according to Hopmans *et al*. [393], with some modifications. Briefly, analysis was performed on an Agilent 1260 UHPLC coupled to a 6130 quadrupole MSD in selected ion monitoring (SIM) mode. Separation was achieved on two UHPLC silica columns (BEH HILIC columns, 2.1 x 150 mm, 1.7 µm; Waters) in series, fitted with a 2.1 x 5 mm pre-column of the same material (Waters) and maintained at 30˚C. Archaeol and GDGTs were eluted isocratically for 10 min with 10% B, followed by a linear gradient to 18% B in 20 min, then a linear gradient to 100% B in 20 min, where A is hexane and B is hexane: isopropanol (9:1). Flow rate was 0.2 ml/min. Total run time is 60 min with a 20 min re-equilibration. Source settings were identical to Schouten *et al*. (2007)[394]. Typical injection volume was 5 µl of a 1 mg/ml solution. *M/z* of the protonated molecules of archaeol and isoprenoid GDGTs were monitored. Archaeol and GDGTs were quantified by comparing the response in the appropriate mass chromatogram to the response of known amounts of standards (archaeol obtained from Avanti Polar Lipids, Inc. Alabaster, Alabama, USA; GDGT-0 prepared according to Sinninghe Damste *et al*., 2002 [395]).

Intact polar lipid analysis by HPLC/ ESI-MS

For analysis of the intact polar lipids, total lipids of T. kodakarensis were extracted using a modified Bligh–Dyer procedure (Bligh and Dyer 1959). 1-3 mg of freeze-dried cell material was used and dissolved in a solvent mixture of phosphate-buffer (0.05 M, pH 7.4)/MeOH/ DCM 0.8/2/1 (v/v). A 1 ml aliquot of the solvent mixture was sonicated for 10 min. The sample was centrifuged (2,500 x*g*, 1 min), and the supernatant removed. The sonication procedure was repeated twice after which DCM and phosphate buffer were added to a volume ratio of 0.9/1/1. After centrifuging (2,500 x*g*, 2 min) the DCM layer was collected. The residue was re-extracted twice following the same procedure. The combined DCM layers were concentrated by evaporating solvent under a stream of N2 gas. Lipid extracts were re-dissolved in hexane:2-propanol:water (72:27:1) at a concentration of 1 mg mL-1 and filtered through a 0.45 µm regenerated cellulose (RC) filter (Alltech Associates Inc., Deerfield, IL) prior to injection.

Intact polar lipids (IPLs) were analysed according to Sturt *et al.* (2004) with some modifications. An Ultimate 3000 RS UHPLC, equipped with thermostatted auto-injector and column oven, coupled to a Q Exactive Orbitrap MS with Ion Max source with heated electrospray ionization (HESI) probe (Thermo Fisher Scientific, Waltham, MA), was used. Separation was achieved on a YMC-Triart Diol-HILIC column (250 x 2.0 mm, 1.9 µm particles, pore size 12 nm; YMC Co., Ltd, Kyoto, Japan) maintained at 30 °C. The following elution program was used with a flow rate of 0.2 mL min^{-1}: 100% A for 5 min, followed by a linear gradient to 66% A: 34% B in 20 min, maintained for 15 min, followed by a linear gradient to 40% A: 60% B in 15 min, followed by a linear gradient to 30%A:70%B in 10 min, where A = hexane/2-propanol/formic acid/14.8 M NH_{3aq} (79:20:0.12:0.04 [volume in volume in volume in volume, v/v/v/v]) and B = 2-propanol/water/formic acid/ 14.8 M NH_{3aq} (88:10:0.12:0.04 [v/v/v/v]). Total run time was 70 min with a re-equilibration period of 20 min in between runs. HESI settings were as follows: sheath gas (N_2) pressure 35 (arbitrary units), auxiliary gas (N_2) pressure 10 (arbitrary units), auxiliary gas (N2) T 50 ˚C, sweep gas (N_2) pressure 10 (arbitrary units), spray voltage 4.0 kV (positive ion ESI), capillary temperature 275 °C, S-Lens 70 V. Lipids were analyzed with a mass range of *m/z* 375 to 2000 (resolution 70,000), followed by data dependent MS^2 (resolution 17,500), in which the ten most abundant masses in the mass spectrum(with the exclusion of isotope peaks) were fragmented successively (stepped normalized collision energy 15, 22.5, 30; isolation window 1.0 *m/z*). A dynamic exclusion window, with a mass tolerance of 3 ppm, of 6 sec was applied. In addition an inclusion list was used with a mass tolerance of 3 ppm, targeting the following compounds: **Table S4.** The Q Exactive was calibrated within a mass accuracy range of 1 ppm using the Thermo Scientific Pierce LTQ Velos ESI Positive Ion Calibration Solution (containing a mixture of caffeine, MRFA, Ultramark 1621, and *N*-butylamine in an acetonitrilemethanol-acetic solution).

Transmission electron microscopy (TEM) of thin sections

Cells of *T. kodakarensis* strains TS559, OE1324, OE1957 and Δ*TK1548* cultured in ASW-YT-Pyr and harvested in the stationary phase were fixed by incubation in 3% glutaraldehyde in 0.1 M sodium cacodylate buffer pH 7.2 for 1.5 h at room temperature. After three washes in the same buffer, a post fixation was done in 1% (w/v) OsO_4 and 1.5% (w/v) potassium ferricyanide in the same buffer for 1 h at room temperature and subsequently rinsed three times in water. One volume of the sample suspensions were mixed with equal volumes of 3% low melting point agarose. Between each step the samples were centrifuged for 5 min at 8,000 rpm (5,220 x*g*) and the pellets were suspended again in the next liquid. After centrifugation the agarose was solidified at 0°C and the parts containing the pelleted sample were cut in very small pieces. The sample pieces were dehydrated in a graded ethanol series (10, 30, 50, 70, 90, 100, 100%), infiltrated with modified Spurr resin mixture (Serva) and polymerized for 8 h at 70°C. The samples were ultra-thin (70-80 nm) sectioned with an ultra-microtome (Reichert, Ultracut S), post-stained with 5 times diluted Reynold's lead citrate for 5 min and analyzed with a transmission electron microscope (JEOL, JEM 1011). Micrographs were made with a digital camera (Olympus, Veleta). Sample preparation and imaging were performed by the Wageningen Electron Microscopy Centre (WEMC) facility.

Acknowledgements

This work was carried out within the BE-Basic R&D Program, which was granted a FES (Fonds Economische Structuurversterking) subsidy from the Dutch Ministry of Economic affairs, agriculture and innovation (EL&I). We thank Tiny Franssen Verheijen from the Wageningen Electron Microscopy Centre (WEMC) for her excellent work on providing electron micrographs. Furthermore, we thank Jaap Sinninghe Damsté, Laura Villanueva, Ellen C. Hopmans for their fruitfull discussions, and Michel Koenen from NIOZ for his excellent technical assistance.

Supplemental Information

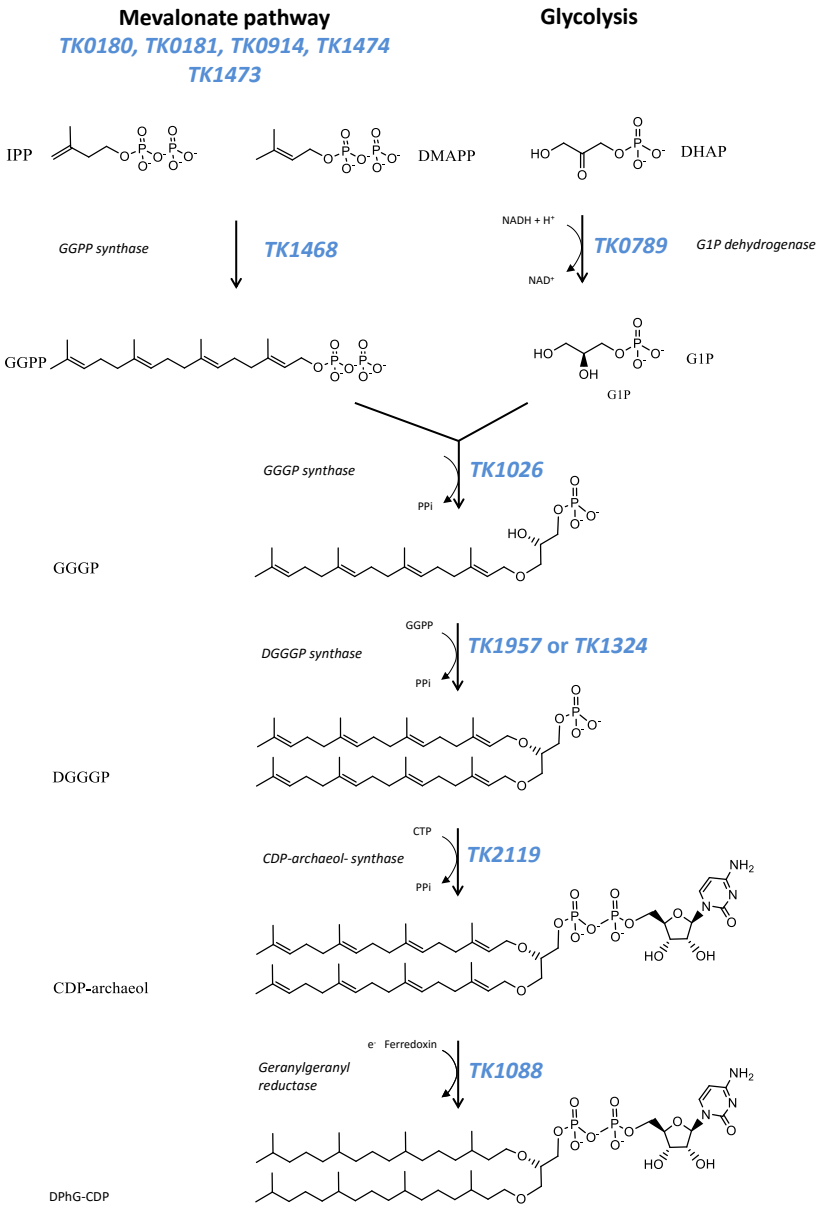


Figure S1. Hypothetical pathway responsible for the biosynthesis of CDP-archaeol in *T. kodakarensis* CDP-archaeol is the CDP-activated precursor of diether polar lipids. The pathway was reconstructed by aligning sequences of experimentally characterized genes, and shown to be responsible for the reactions leading to isoprenoid and CDP-archaeol synthesis in various archaea, against the genome of *T. kodakarensis* [34, 298, 304, 306, 348, 351, 378]. Analysis of *TK1957* and *TK1324* overexpression in this study showed that both genes are unlikely responsible for DGGGP-synthesis.

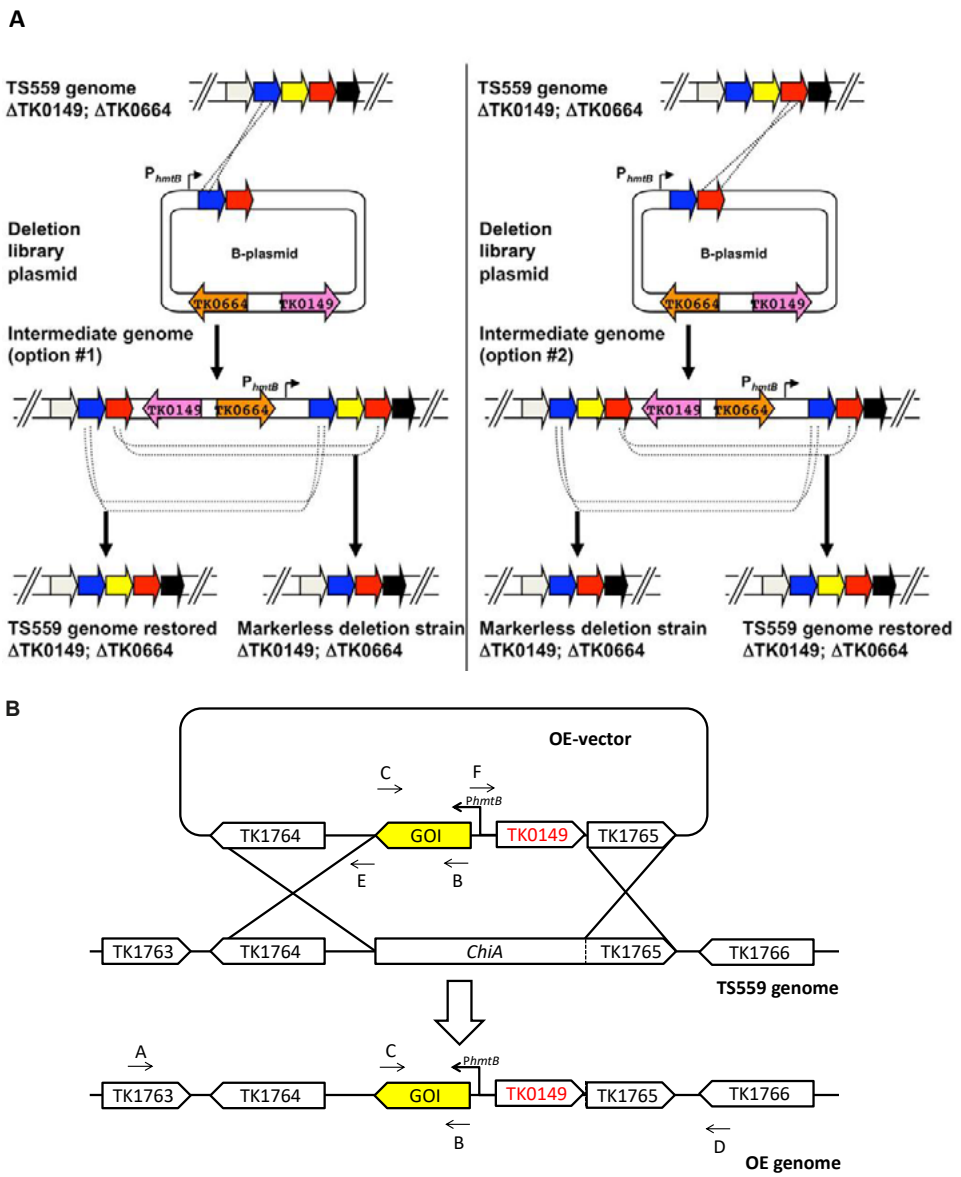


Figure S2A. Schematic representation of the disruption and overexpression procedure in *T. kodakarensis* Use of disruption vectors (B-plasmids) to generate markerless knockouts in the genome of *T. kodakarensis* TS559. The image has been reused from Hileman, T.H., and Santangelo, T.J. (2012) [368]. A hypothetical region of the genome of *T. kodakarensis* strain TS559 is shown at top, with the left and right panels depicting the two possible integration events yielding agmatine-prototrophic intermediate strains from the diagramed B-plasmid. Both intermediate strains #1 and #2 contain direct repeats flanking the target locus, and dependent on recombination responsible for excision upon counter-selection with 6MP, either the original TS559 genome can be restored or the desired deletion genome generated. **Fig. S2B** Use of overexpression vectors (OE-vector) to generate *T. kodakarensis* strains overexpressing the gene of interest (GOI). The GOI is integrated at the *ChiA*-locus of *T. kodakarensis* strain

TS559 via a double homologous recombination event yielding agmatine-prototrophic OE-strains. The letters A-D represent the binding sites of the primers used in the diagnostic PCRs. The letters E & F represent the binding sites of the primers used to amplify the vector backbone using pSKS022 as a template. Abbreviations: *TK0149*, agmatine selection marker; *TK0664*, counter-selection marker; P_{hmtb}, strong constitutive promoter.

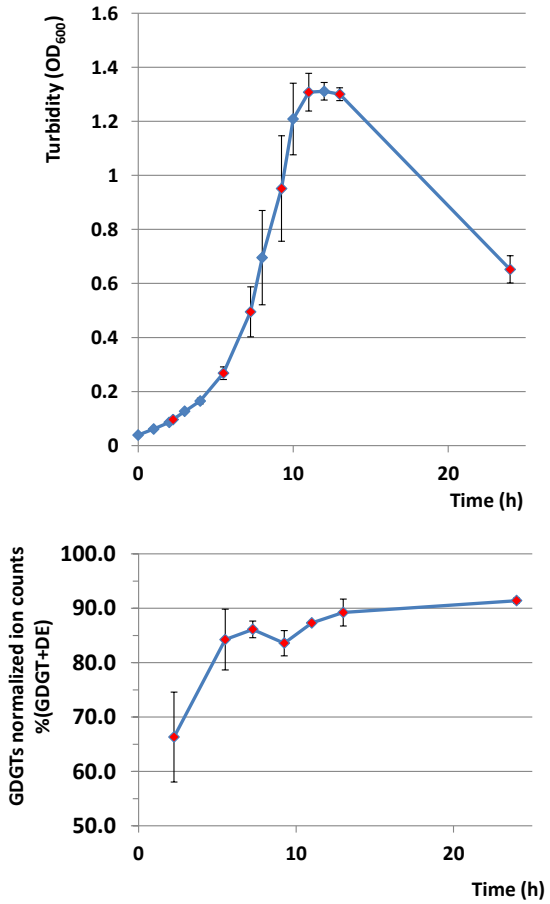


Figure S3. Growth phase dependent tetraether levels Upper panel, *T. kodakarensis* T5559 cells grown on ASW-YT pyr were sampled at different ODs (red dots); Lower panel, the OD dependent tetraether levels (%) were determined and expressed as % of the total detected amount of lipids showing near stability after 11 h of growth at stationary phase.

Table S1. genes that are potentially involved in tetraether synthesis

A Results of phylogenetic profiler single gene search

Gene locus	Gene annotation	prenyl binding	FAD binding	DDxxD	SQE hom	Phr	TMH	+ score	selected
TK0038	archaeal flagellin B1 precursor			+		+	1	2	
TK0039	archaeal flagellin B2 precursor			+		+	1	2	
TK0040	archaeal flagellin B3 precursor					+	1	1	
TK0041	archaeal flagellin B4 precursor					+	1	1	
TK0042	archaeal flagellin B5 precursor					+	1	1	
TK0049	archaeal flagella-related membrane protein J						9	0	
TK0135	indolepyruvate: ferredoxin oxidoreductase, beta subunit					+	1	1	
TK0136	indolepyruvate: ferredoxin oxidoreductase, alpha subunit					+	0	1	
TK0139	acetyl-CoA synthetase II (NDP forming), alpha subunit			+		+	0	2	
TK0147	spermidine synthase					+	1	1	
TK0172	hypothetical membrane protein			+			6	1	
TK0182	hypothetical protein						0	0	
TK0188	hypothetical protein						0	0	
TK0306	DEAD/DEAH box RNA helicase					+	0	1	
TK0314	hydrolase, metallo-beta-lactamase superfamily					+	0	1	
TK0428	riboflavin biosynthesis protein RibA		+				0	1	
TK0537	peroxiredoxin, AhpC/TSA family					+	0	1	
TK0624	multisubunit sodium/hydrogen antiporter, MnhC subunit						3	0	
TK0711	hypothetical protein						0	0	
TK0810	oligosaccharyl transferase, STT3 subunit					+	13	1	
TK0817	2-oxoacid:ferredoxin oxidoreductase, alpha subunit						0	0	
TK0834	transcription regulator, Lrp/AsnC family			+			0	1	
TK0882	Agmatinase						0	0	
TK0893	pyruvate fromate-lyase activating enzyme-related protein					+	0	1	
TK0929	3-octaprenyl-4-hydroxybenzoate decarboxylyase	+	+	+		+	0	4	x
TK0950	asparagine synthase (glutamine-hydrolyzing)						0	0	
TK0951	XPA-binding protein 1 homolog						0	0	
TK0971	hypothetical membrane protein						4	0	

Gene locus	Gene annotation	prenyl binding	FAD binding	DDxxD	SQE hom	Phr	TMH	+ score	selected
TK1008	Fe-containing alcohol dehydrogenase						0	0	
TK1045	cobalamin adenosyltransferase					+	0	1	
TK1057	iron(II) uptake regulation protein						0	0	
TK1219	membrane bound hydrogenase, MbxH' subunit					+	14	1	
TK1220	membrane bound hydrogenase, MbxH subunit					+	13	1	
TK1227	predicted regulator of amino acid metabolism, containing ACT domain					+	0	1	
TK1258	hypothetical membrane protein					+	7	1	
TK1285	transcription regulator, Lrp/AsnC family					+	0	1	
TK1296	hypothetical protein						0	0	
TK1386	metallophosphoesterase, calcineurin superfamily					+	0	1	
TK1393	anaerobic glycerol 3-phosphate dehydrogenase		+				0	1	
TK1396	glycerol kinase			+		+	0	2	
TK1561	hypothetical membrane protein					+	5	1	
TK1592	S-adenosylmethionine decarboxylase proenzyme					+	0	1	
TK1659	small neutral amino acid transporter						5	0	
TK1718	oligosaccharyl transferase, STT3 subunit					+	11	1	
TK1732	dolichol-phosphate mannosyltransferase	+		+			4	2	x
TK1806	probable dihydroorotate dehydrogenase, electron transfer subunit		+				0	0	
TK1812	hypothetical membrane protein					+	2	1	
TK1820	membrane-associated metalloprotease					+	6	1	
TK1851	hypothetical membrane protein, conserved					+	5	1	
TK1852	hypothetical membrane protein, conserved					+	4	1	
TK1890	histone macroH2A-related protein (C-terminus)						0	0	
TK2052	predicted transcription regulator, Lrp/AsnC family						0	0	
TK2087	membrane bound hydrogenase, MbhH subunit					+	13	1	
TK2112	hypothetical protein					+	0	1	
TK2257	deoxycytidylate deaminase					+	0	1	
TK2275	RNA-binding protein, containing PUA domain					+	0	1	
TK2296	hypothetical protein, conserved, radical SAM superfamily						0	0	

B Uncharacterized *T. kodakarensis* genes present in the selected arCOGs

Gene locus	COG	COG annotation	Gene annotation	prenyl binding	FAD FMN	DDxxD	SQE hom	Phr	TMH	score	chosen
TK0059	COG0558	Phosphatidylglycerophosphate synthase	CDP-alcohol phosphatidyltransferase					+	3	1	
TK0179	COG1545	Predicted nucleic-acid-binding protein containing a Zn-ribbon	Predicted nucleic-acid-binding protein containing a Zn-ribbon					+	0	1	
TK0180	COG0183	Acetyl-CoA acetyltransferase	acetyl-CoA acetyltransferase			+		+	0	2	
TK0217	COG0214	Pyridoxine biosynthesis enzyme	Pyridoxine biosynthesis enzyme					+	0	1	
TK0283	COG0119	Isopropylmalate/homocitrate/citramalate synthases	homocitrate synthase						0	0	
TK0509	COG0163	3-polyprenyl-4-hydroxybenzoate decarboxylase	3-octaprenyl-4-hydroxybenzoate decarboxylase UbiX	+	+				0	2	x
TK0513	COG1689	Uncharacterized protein conserved in archaea	hypothetical protein	+				+	0	2	x
TK0551	COG1052	Lactate dehydrogenase and related dehydrogenases	D-isomer specific 2-hydroxyacid dehydrogenase					+	0	1	
TK0683	COG1052	Lactate dehydrogenase and related dehydrogenases	glyoxylate reductase						0	0	
TK0789	COG0371	Glycerol dehydrogenase and related enzymes	glycerol-1-phosphate dehydrogenase						0	0	
TK0858	COG1267	Phosphatidylglycerophosphatase A and related proteins	hypothetical protein						0	0	
TK0929	COG0043	3-polyprenyl-4-hydroxybenzoate decarboxylase and related decarboxylases	3-octaprenyl-4-hydroxybenzoate decarboxylyase	+	+	+		+	0	4	x
TK1088	COG0644	Dehydrogenases (flavoproteins)	geranylgeranyl hydrogenase	+	+	+	+	+	0	5	x
TK1173	COG0020	Undecaprenyl pyrophosphate synthase	undecaprenyl diphosphate synthase	+				+	0	2	x
TK1324	COG0382	4-hydroxybenzoate polyprenyltransferase and related prenyltransferases	4-hydroxybenzoate octaprenyltransferase	+					6	1	x
TK1444	COG0083	Homoserine kinase	Homoserine kinase						0	0	
TK1473	COG1608	Predicted archaeal kinase	amino acid kinase						0	0	
TK1477	COG1355	Predicted dioxygenase	hypothetical protein			+		+	0	2	x
TK1548	COG0075	Serine-pyruvate aminotransferase/archaeal aspartate aminotransferase	probable serine--glyoxylate aminotransferase			+		+	0	2	x
TK1806	COG0167	Dihydroorotate dehydrogenase	probable dihydroorotate dehydrogenase		+				0	1	
TK1831	COG0153	Galactokinase	Galactokinase						0	0	
TK1836	COG0644		geranylgeranyl hydrogenase	+	+	+			0	3	x
TK1849	COG4072	Uncharacterized protein conserved in archaea	hypothetical protein					+	0	1	
TK1957	COG0382	4-hydroxybenzoate polyprenyltransferase and related prenyltransferases	4-hydroxybenzoate octaprenyltransferase	+					8	1	x
TK2119	COG0575	CDP-diglyceride synthetase	hypothetical protein						1	0	
TK2159	COG1233		predicted oxidoreductase		+		+	+	1	3	x
TK2279	COG0558	Phosphatidylglycerophosphate synthase	bifunctional sugar nucleotidyltransferase/CDP-alcohol phosphatidyltransferase synthase					+	3	1	

Table S2. arCOGs included in the second comparative genome analysis.

COG0020	Undecaprenyl pyrophosphate synthase
COG0039	Malate/lactate dehydrogenases
COG0043	3-polyprenyl-4-hydroxybenzoate decarboxylase and related decarboxylases
COG0075	Serine-pyruvate aminotransferase/archaeal aspartate aminotransferase
COG0083	Homoserine kinase
COG0119	Isopropylmalate/homocitrate/citramalate synthases
COG0142	Geranylgeranyl pyrophosphate synthase
COG0153	Galactokinase
COG0163	3-polyprenyl-4-hydroxybenzoate decarboxylase
COG0167	Dihydroorotate dehydrogenase
COG0183	Acetyl-CoA acetyltransferase
COG0210	Superfamily I DNA and RNA helicases
COG0214	Pyridoxine biosynthesis enzyme
COG0240	Glycerol-3-phosphate dehydrogenase
COG0280	Phosphotransacetylase
COG0282	Acetate kinase
COG0318	Acyl-CoA synthetases (AMP-forming)/AMP-acid ligases II
COG0331	(acyl-carrier-protein) S-malonyltransferase
COG0371	Glycerol dehydrogenase and related enzymes
COG0382	4-hydroxybenzoate polyprenyltransferase and related prenyltransferases
COG0407	Uroporphyrinogen-III decarboxylase
COG0417	DNA polymerase elongation subunit (family B)
COG0439	Biotin carboxylase
COG0469	Pyruvate kinase
COG0558	Phosphatidylglycerophosphate synthase
COG0575	CDP-diglyceride synthetase
COG0578	Glycerol-3-phosphate dehydrogenase
COG0644	Dehydrogenases (flavoproteins)
COG0688	Phosphatidylserine decarboxylase
COG0761	Penicillin tolerance protein
COG1022	Long-chain acyl-CoA synthetases (AMP-forming)
COG1024	Enoyl-CoA hydratase/carnithine racemase
COG1052	Lactate dehdrogenase and related dehydrogenases
COG1154	Deoxyxylulose-5-phosphate synthase
COG1183	Phosphatidylserine synthase
COG1233	Phytoene dehydrogenase-related protein
COG1239	Mg-chelatase subunit ChlI
COG1250	3-hydroxyacyl-CoA dehydrogenase
COG1257	Hydroxymethylglutaryl-CoA reductase
COG1267	Phosphatidylglycerophosphatase A and related proteins
COG1304	L-lactate dehydrogenase (FMN-dependent) and related alpha-hydroxy acid dehydrogenases
COG1331	Highly conserved protein containing a thioredoxin domain
COG1348	Nitrogenase subunit NifH (ATPase)
COG1355	Predicted dioxygenase
COG1429	Cobalamin biosynthesis protein CobN and related Mg-chelatases
COG1443	Isopentenyldiphsophate isomerase
COG1502	Phosphatidylserine/phosphatidylglycerophosphate/cardiolipin synthases and related enzymes

COG1545	Predicted nucleic -acid-binding protein containing a Zn-ribbon
COG1562	Phytoene/squalene synthetase
COG1577	Mevalonate kinase
COG1608	Predicted archaeal kinase
COG1646	Predicted phosphate-binding enzymes, TIM-barrel fold
COG1657	Squalene cyclase
COG1689	Uncharacterized protein conserved in archaea
COG1960	Acyl-CoA dehydrogenases
COG2025	Electron transfer flavoprotein, alpha subunit
COG2057	Acyl CoA-acetate/3-ketoacid CoA transferase, beta subunit
COG2086	Electron transfer flavoprotein, alpha subunit
COG2098	Uncharacterized protein conserved in archaea
COG2227	2-polyprenyl-3-methyl-5-hdyroxy-6-metoxy-1,4-benzoquinol methylase
COG2324	Predicted membrane protein
COG2440	Ferredoxin-like protein
COG2710	Nitrogenase molybdenum-iron protein, alpha and beta chains
COG3407	Mevalonate pyrophsophate decarboxylase
COG3425	3-hydroxy-3-methylglutaryl CoA synthase
COG3890	Phosphomevalonate kinase
COG4072	Uncharacterized protein conserved in archaea
COG4851	Protein involved in sex pheromone biosynthesis
COG5190	TFIIF-interacting CTD phosphatases, including NLI-interacting factor
COG5243	HRD ubiquitin ligase complex, ER membrane component
COG5272	Ubiquitin

Table S3. Sequences of oligonucleotide primers used in this study

Oligonucleotide	Sequence (5'→3')	Application
Disruption vectors		
0010509	GGAGGTGAATTTCGTATGGACTTATACGTCAGAAAAGC	*TK0509* initial amplification
0020509	GGTGAAGGATTTCCTCCTTTATTATTTCCTTGCTCTCC	*TK0509* initial amplification
0160509	AGCCCTATCAAGGTCGTCCAGCTGTACCATCCAAT CACCTAACGTTCGCTAGGGACGCTT	*TK0509* inverse PCR to construct B plasmids
0170509	AAGCGTCCCTAGCGAACGTTAGGTGATTGGATGGTA CAGCTGGACGACCTTGATAGGGCT	*TK0509* inverse PCR to construct B plasmids
0010513	GGAGGTGAATTTCAGGGGATTATCGAGTATCTTCAGG	*TK0513* initial amplification
0020513	GGTGAAGGATTTCCTTCTCTATCTTATGCCCCCTATCC	*TK0513* initial amplification
0160513	CGCCTGTTCTACTCCATAGGAGGCGGTAGAATCTT TTAAACAATCCATCCTTTCGAAAGC	*TK0513* inverse PCR to construct B plasmids
0170513	GCTTTCGAAAGGATGGATTGTTTAAAAGATTCTACC GCCTCCTATGGAGTAGAACAGGCG	*TK0513* inverse PCR to construct B plasmids
0010929	GGAGGTGAATTTCGGCATTCATCGAAGAGGACTATCG	*TK0929* initial amplification
0020929	GGTGAAGGATTTCCATGTCGTTTAAAGAAAGTGTGACC	*TK0929* initial amplification
0160929	AGACCTTTTGGAAGTTGGTATGGTGATAGCT CACTCCCCCATTTCCCTAATCTTCCTGAC	*TK0929* inverse PCR to construct B plasmids
0170929	GTCAGGAAGATTAGGGAAATGGGGGAGTGAGCTA TCACCATACCAACTTCCAAAAGGTCT	*TK0929* inverse PCR to construct B plasmids
0011088	GGAGGTGAATTTCCAGCGAGGAAGTACTATAACATCG	*TK1088* initial amplification
0021088	GGTGAAGGATTTCACCCTGATCTTGATCTTCAGTTAGG	*TK1088* initial amplification
0161088	AGCGCTAAACCCCTGGAGGTGAGCGAATGAT CAAGGAGCTTCTTTAACCAGTCTGAACCC	*TK1088* inverse PCR to construct B plasmids
0171088	GGGTTCAGACTGGTTAAAGAAGCTCCTTGATCATTC GCTCACCTCCAGGGGTTTAGCGCT	*TK1088* inverse PCR to construct B plasmids
0011173	GGAGGTGAATTTCAGCTCAAAGACCTCATAGTGTTCC	*TK1173* initial amplification
0021173	GGTGAAGGATTTCAGGACTGCGAATCCTTTACTTTCG	*TK1173* initial amplification
0161173	TTAATAAGCTTTAGCGGAAAAATTTAAAAATTGCCCT CACTCTATCTTCTTCCTCTTGAG	*TK1173* inverse PCR to construct B plasmids
0171173	CTCAAGAGGAAGAAGATAGAGTGAGGGCAATTTTTA AATTTTTCCGCTAAAGCTTATTAA	*TK1173* inverse PCR to construct B plasmids
0011324	GGAGGTGAATTTCCTCACCAGTTCTCTCTCTTTATTGC	*TK1324* initial amplification
0021324	GGTGAAGGATTTCAATACGACACACCGGTTTACGTTGG	*TK1324* initial amplification
0161324	CTATTTTCATTTTCCTGCACCCCTAAGACAGGTATC ACTCAATAGCTCCCAAAATAGGAC	*TK1324* inverse PCR to construct B plasmids
0171324	GTCCTATTTTGGGAGCTATTGAGTGATACCTGTCT TAGGGGTGCAGGAAAATGAAAATAG	*TK1324* inverse PCR to construct B plasmids
0011477	GGAGGTGAATTTCCCTCTAACAAGTCCTTCAAATACG	*TK1477* initial amplification
0021477	GGTGAAGGATTTCGTTCACAAAGAGGCTATAGAGTTCG	*TK1477* initial amplification
0161477	AAATATGGAAAAGGGCTTATCACACTTTCGCATCT CCATCACCCCCAATGGATTTTAGCT	*TK1477* inverse PCR to construct B plasmids

Oligonucleotide	Sequence (5'→3')	Application
0171477	AGCTAAAATCCATTGGGGGTGATGGAGATGCGAAAG TGTGATAAGCCCTTTTCCATATTT	*TK1477* inverse PCR to construct B plasmids
0011548	GGAGGTGAATTTCGAAAGTCCCAAATCTATTGTGTAGG	*TK1548* initial amplification
0021548	GGTGAAGGATTTCTATCATGGGAGTCTACATCTTTACC	*TK1548* initial amplification
0161548	CCTTGAGCAACTAAGGGCGAGGTGATAACTTTC CAATTCTCCCTCTTTTCCCCTTCTTAT	*TK1548* inverse PCR to construct B plasmids
0171548	ATAAGAAGGGGAAAAGAGGGAGAATTGGAAAGTTA TCACCTCGCCCTTAGTTGCTCAAGG	*TK1548* inverse PCR to construct B plasmids
0011732	GGAGGTGAATTTCAAAGAGAACAGTATTCCCATCACC	*TK1732* initial amplification
0021732	GGTGAAGGATTTCGGAAGGAGAGTTGAGAGACTGACC	*TK1732* initial amplification
0161732	AAAATAGGGAGCAACGAGCAGAACCCTCATTCAACT GTCATCTGCGACCCTCTCGTAGAA	*TK1732* inverse PCR to construct B plasmids
0171732	TTCTACGAGAGGGTCGCAGATGACAGTTGAAT GAGGGTTCTGCTCGTTGCTCCCTATTTT	*TK1732* inverse PCR to construct B plasmids
0011836	GGAGGTGAATTTCCAACGAGCTTTAATAACCTTCTCG	*TK1836* initial amplification
0021836	GGTGAAGGATTTCGCCAAGTAAAAGCTCTCGTTGTCC	*TK1836* initial amplification
0161836	TTTTAAAACCACCCCATCCGAGGGGTGAAGGGTGAG AAGATGGCCAGATGGAAGATGGGC	*TK1836* inverse PCR to construct B plasmids
0171836	GCCCATCTTCCATCTGGCCATCTTCT CACCCTTCACCCCTCGGATGGGGTGGTTTTAAAA	*TK1836* inverse PCR to construct B plasmids
0011957	GGAGGTGAATTTCAGCTGGTAAAACCTGACATTATTGC	*TK1957* initial amplification
0021957	GGTGAAGGATTTCAAGAAGAGGAGGAGAAAACTGAGG	*TK1957* initial amplification
0161957	CTTCCGACAGAATTTAGTTAGGTGCCCCTATGGATA TGGATTTTGAGAAAGAACTGATTA	*TK1957* inverse PCR to construct B plasmids
0171957	TAATCAGTTCTTTCTCAAAATCCATATC CATAGGGGCACCTAACTAAATTCTGTCGGAAG	*TK1957* inverse PCR to construct B plasmids
0012159	GGAGGTGAATTTCGTCAACTATGACCGGAGCCTCTACG	*TK2159* initial amplification
0022159	GGTGAAGGATTTCGAAGGTATCAAAAATCCCAAAGTGG	*TK2159* initial amplification
0162159	CCACTCTACAAATGTAAAGAAAAGTGCCGCTCT ACCACCTGAAATGCCAAGGGTGAGGAA	*TK2159* inverse PCR to construct B plasmids
0172159	TTCCTCACCCTTGGCATTTCAGGTGGTAG AGCGGCACTTTTCTTTACATTTGTAGAGTGG	*TK2159* inverse PCR to construct B plasmids
Overexpression vectors		
BG4881	CACCTCCGTGATATTATCTATTAC	vbb amplification (pSKS022)
BG4882	CGCACCGGTCCGGACAACACCCCTTGAGCTTTG	vbb amplification (pSKS022)
BG4934	ATGCTATACAGATTGGTCTCCCACATCCC	*TK1173* amplification
BG4935	CGCACCGGTCTACCTTCCAAAGCGCCGCTGTCTC	*TK1173* amplification
BG4936	ATGCTGAGGGAGATACTCTCAC	*TK0929* amplification
BG4937	CGCACCGGTTCACACCCTCGCCCTCTCAAAC	*TK0929* amplification
BG4938	ATGAAGTACGATGTCCTCATAATCGGCGGC	*TK1836* amplification
BG4939	GCGACCGGTTACTCCTCCTGGAGGCCCATCTTCCAT	*TK1836* amplification
BG4940	ATGAAAGTCATTGTCGCAATAACGGGGGC	*TK0509* amplification
BG4941	GCGACCGGTTCAAGGTCGTCCAGCTGTAC	*TK0509* amplification

Oligonucleotide	Sequence (5'→3')	Application
BG4942	ATGGCATCTGTTTCAGCTGTTATAAGG	*TK1324* amplification
BG4943	GCGACCGGTCGCTGACGAGGGCTATTTTCAT	*TK1324* amplification
BG5196	ATGACCTGGAAGTACGACGTCGTTG	*TK1088* amplification
BG5197	GCGACCGGTGAAGAAGGGGTTCAGACTGGTT	*TK1088* amplification
BG5200	ATGGAATTCAAGGCATTCATCGAG	*TK1957* amplification
BG5201	GCGACCGGTGGCGTCTTGAACGTTATGACAG	*TK1957* amplification
BG5202	ATGGGCTCGAAGCTTGATGAG	*TK0513* amplification
BG5203	GCGACCGGTTCCCGCACGTTAGTCCTACAC	*TK0513* amplification
BG5204	ATGGAGCTCCACTTTGACATGACC	*TK1548* amplification
BG5205	GCGACCGGTGGGGAAAAGAGGGAGAATTG	*TK1548* amplification
BG5584	ATGACAGTTGAAGTCTCCGTG	*TK1732* amplification
BG5585	GCGACCGGTAACGAGCAGAACCCTCATTC	*TK1732* amplification
BG5586	ATGGAGATGGTGGTAAGGTATCCC	*TK1477* amplification
BG5587	GCGACCGGTCTAGCGCCTGAAGACTATG	*TK1477* amplification
BG5588	ATGAGTAAAGGAGTTGTACATCC	*TK2159* amplification
BG5589	GCGACCGGTGCCGCTCAGAGGTACCATTCGG	*TK2159* amplification
Diagnostic primers		
BG4871	CATCCTCACAGGGACGTTTG	OExxxx diagnosis & sequencing (primer A in Figure S2b*)*
BG4873	ACAGACTTGAGGTGGAAAGG	OExxxx diagnosis & sequencing (primer D in Figure S2b*)*
BG4662	GCTTGGGTCATGGACAGGCTC	RT-PCR: *TK0308* (F)
BG4663	TCTTAACCTGCTCGGCGACC	RT-PCR: *TK0308* (R)
BG5570	TAGAAGGATAACCGCAGGTG	RT-PCR: *TK1477* (F)
BG5571	CGATGGAGTGCTCGTATTTG	RT-PCR: *TK1477* (R)
BG5572	TGGAAACTACCTGGCGTCTC	RT-PCR: *TK1836* (F)
BG5573	AGTCGGCTCCCTTCTTTACC	RT-PCR: *TK1836* (R)
BG5574	CTCACTGAACCCAGATAAGG	RT-PCR: *TK1324* (F)
BG5575	GATTCTCCCGTCCATTGAAG	RT-PCR: *TK1324* (R)
BG5576	AAGACGGCGGCTACTGCTTC	RT-PCR: *TK0513* (F)
BG5577	CGCCAAGACCAACGGCAAAC	RT-PCR: *TK0513* (R)
BG5578	TGGTGAAGCCGAAGTACAAG	RT-PCR: *TK1548* (F)
BG5579	AACGCTCCAATAACCGTGAC	RT-PCR: *TK1548* (R)
BG5580	GGTTGGGCGTTTCCTATGAG	RT-PCR: *TK1732* (F)
BG5581	TGCACCGGGTTCGATCATTC	RT-PCR: *TK1732* (R)
BG5723	GCTATCTCCAGCTCGCTAAC	RT-PCR: *TK0929* (F)
BG5724	CTCAAGAAGCTCCCGATTCC	RT-PCR: *TK0929* (R)
BG5725	GGATGCGGGCATAATCACAG	RT-PCR: *TK0509* (F)
BG5726	GCTTGGGCATGAGGTTGTTG	RT-PCR: *TK0509* (R)
BG5727	CGCTGTAGGCGATCTGATAG	RT-PCR: *TK1173* (F)
BG5728	AACATAGCCCTCGCATACGG	RT-PCR: *TK1173* (R)
BG5729	GCGGAGCGTAGGTGTTTATG	RT-PCR: *TK1088* (F)
BG5730	ACAAGGAAGTCAGCGGAGTG	RT-PCR: *TK1088* (R)
BG5733	ATGAGGAACGCAGAGTACAG	RT-PCR: *TK1957* (F)
BG5734	ATAGCCGTTGGTGAGATTGG	RT-PCR: *TK1957* (R)
BG5737	TCTCCTTGATGCCGATGTTC	RT-PCR: *TK2159* (F)
BG5738	AGATAGGCGACAACGAGTTC	RT-PCR: *TK2159* (R)

Table S4. Detected IPL-adduct masses in the HPLC / ESI-MS

	m/z		
Saturated	H+	NH4+	Na+
PH	895,700	912,726	917,682
HPH	1057,753	1074,779	1079,735
H(NH2)PH	1056,769		1078,751
Unsaturated	H+	NH4+	Na+
PH	883,606	900,632	
HPH	1045,659	1062,685	1067,641
H(NH2)PH	1044,675		1066,657

Fragmentation indicated the presence of unsaturated diether lipids with 3 unsaturations per lipid on one chain only. Abbreviations: PH: phosphohexose; HPH: hexose-phophohexose; H(NH2)PH: aminohexose phosphohexose,

Table S5. Summarized Tetraether levels and diether compositions for knockout strains (Δ) and overexpression strains (OE) with altered tetraether levels

Gene	Δ (TE:DE)%	DE (sat)	DE (unsat)	OE (TE:DE)%	DE (sat)	DE (unsat)
TK1548	15.1	99.6	0.4	82.5	91.8	1.2
TK1836	79.8	99.8	0.2	27.3	52.4	19.8
TK1957	†	†	†	96.1	99.9	0.1
TK1324	80.9	73.4	26.6	38.1	16.1	83.9
		wildtype			**wildtype**	
T.k. TS559	80.0	55.2	44.8	80	55.2	44.8

The data indicate that both low obtained tetraether levels (Δ*TK1548* and OE1836) and high tetraether levels (*OE1957*) correlate with decreased unsaturated diether levels compared to the wild type TS559 strain. †*: lethality*; (TE:DE%): percentage of tetraethers over total amount of core lipids (tetraethers and diethers)

Chapter 8

Thesis summary and general discussion

As a means to improve *Escherichia coli* robustness, one of the pillars of synthetic biological approaches is dedicated to membrane engineering. As described in **Chapter 1**, rational design of membranes already proved to be a promising technique with few applications already being implemented. Alternatively, much can be learned from nature. As archaea are labelled as naturally robust microbes, the membrane properties of these organisms may in fact be the key to many of the biological constraints encountered in industry. This thesis describes the metabolic engineering of a hybrid membrane of archaeal and bacterial lipids in *E.coli,* and the characterization of the accompanying properties of this engineered strain. The main objective of this venture was to investigate whether mixed membranes can support a biologically viable cell, and to determine what phenotypical changes are elicited. **Chapter 2** provides an overview of membrane adaptations to temperature, pH and pressure for both bacterial and archaeal extremophiles. The main focus of this study lies on how nature deals with extreme habitats and thus what to expect from a mixed membrane. Here we made an inventory of the encountered membrane lipid compositions near the survival boundaries at physico-chemical constraints. To do this, the most robust extremophiles known to date were first subjected to a comparative analysis. We then looked into the encountered lipid compositions and searched for correlations within the respective survival limits. This overview showed that archaea are generally dominating eco-systems with high temperatures and low pH. On the other hand, bacteria seem to dominate habitats with low temperature and slightly but insignificantly also to high pH. As the adaptation mechanisms to low and high temperature and low and high pH are fundamentally different, we conclude that bacterial 'phosphatidic acid' based lipids are better customized to fluidizing modifications of the membrane. Archaetidic acid based lipids on the other hand seem to be better equipped for rigidifying modifications of the membrane most likely due to the methyl branches of the isoprenoids. On the basis of these observations, it is therefore tempting to suggest that a mixed membrane will gain heat and pH dependent stability. Here we stress that these rigidifying adaptations are probably the most useful for industry. In **Chapter 3**, we embarked on the construction of a platform strain of *E.coli* with a stable but inducible upregulation of the MEP-DOXP pathway. A successful strategy was shown to be the directed integration of an artificial MEP-DOXP operon composed of four rate controlling steps. Although flux was never actually measured, the accumulation of a carotenoid end-product 'lycopene' showed that a ~5-fold flux increase was accomplished. Here, a maximal amount of roughly 3 µg/mg dry cell weight (DCW) was measured. This finding however also implied it would not suffice to generate a mixed membrane with an appreciable amount of archaeal lipids. The average amount of lipids in *E.coli* is estimated to be roughly 10% of the DCW. This implies that a maximum of 3% archaeal lipids of the total membrane content could be reached. This finding underlined the need for further MEP-DOXP engineering or a downregulation of the endogenous lipid synthesis pathway. A down-regulation of bacterial lipids would in the first place increase the ratio to archaeal lipids, but perhaps also

redistribute carbon flux to the introduced archaeal lipid pathway. In **Chapter 4** we generated a conditional knock-down of lysophosphatidic acid biosynthesis. To accomplish this, we generated a *plsX* knock-out strain and provided the alternative *plsB* gene with a repressible and inducible promoter. Here we observed that despite a successful engineering, the system fails to repress the synthesis of phospholipids. Because of this insufficiency, we complemented the system with an overproduction of apo-ACP, which is a potent competitive inhibitor of PlsB. This complementation demonstrated that cell growth was severely impaired, suggesting that the desired silenced membrane synthesis was obtained. To understand the true nature of the retarded growth, we performed a proteomics analysis on the mutant cells during induction and without induction. A reversal of expression in enzymes involved in fatty acid metabolism was shown from anabolic to catabolic. Instead of a compensatory expression of the fatty acid synthesis pathway, a relatively high expression of the fatty acid degradation pathway was observed. These findings showed that this system is suitable to combine with the archaeal biosynthesis pathway. In **Chapter 5**, the *in vitro* and *in vivo* assembly of the archaeal lipids archaetidylglycerol and archaetidylethanolamine was tested. Here, the aim of the study was to identify candidate genes from bacterial origin suitable for mixed membrane synthesis. This led to the successful reconstruction of archaetidylglycerol and archaetidyl-ethanolamine both *in vitro* as *in vivo*. Besides a selection of archaeal enzymes that show heterologous activity in *Escherichia coli*, a promiscuous activity of endogenous lipid synthesis enzymes PgsA, PgpA and Psd and the *Bacillus subtilis* PssA was shown towards the archaeal lipid precursor CDP-archaeol. With the combined products and knowledge gained from the MEP-DOXP upregulation, phosphatidic acid downregulation and archaeal lipid reconstruction, we embarked on the assembly of a mixed membrane as described in **Chapter 6**. From preliminary assays of combining the three systems it soon became clear that neither the combination of the MEP-DOXP upregulation and lipid synthesis nor a joint system including a downregulation of phosphatidic acid synthesis would lead to the desired level of archaeal lipids. We attempted to boost these levels by tuning the down-regulatory system or alternatively by providing cerulenin, a fatty acid synthesis inhibitor, without significant improvement. A successful strategy however proved to be the usage of a statistically optimized medium (OPT1). A combination of MEP-DOXP engineering and OPT1 proved to be a vital combination to boost archaeal lipid levels to a reproducible average of 28%. It was shown here that these levels of archaeal lipid production do not lead to loss of viability or affect growth of *E.coli* besides an extensive delay upon induction. A genome analysis comparing induced and uninduced cells showed the occurrence of only few insignificant mutations which cannot account for the growth retardation. Scanning electron microscopy revealed that the cells are prone to dramatic morphological aberrations. Here we observed a high number of lobular appendages. High resolution TEM images indicated that these appendages are released from the membrane, which suggests an expelling mechanism possibly to liberate the cells from a disadvantageous membrane composition. Furthermore, in this chapter it is demonstrated that although GGGP-

and DGGGP-synthases from *Methanococcus maripaludis* can synthesize DGGGP in G3P configuration, this production is absent in *E.coli* or only present in trace amounts. Remarkably, the deletion of a functional *araM* gene for G1P synthesis was shown to be redundant because *E.coli* can generate G1P naturally, or as a side-effect of cultivation on OPT1 medium. When testing the physicochemical robustness we observed a beneficial effect under cryogenic conditions and butanol exposures but not to heat shock, as anticipated. As discussed in Chapter 2, the main characteristic that contribute to thermal robustness of archaea is the ability to form membrane spanning bipolar lipids. Because the enzyme(s) responsible for the putative tail-to-tail condensation are currently unknown we set out to identify this system in *Thermococcus kodakarensis* in **Chapter 7**. Here we constructed a series of strains with deletions or overexpression from a selection of high-probability candidate genes. Although the biochemical nature and the responsible gene was not included, we identified four genes that highly influenced the ultimate level of tetraether lipids. By total ion monitoring MS of the isolated lipids we found that the level of tetraethers highly correlates with the amount of unsaturated diether lipids. This makes a substrate-product relationship between unsaturated diether lipids and tetraethers very likely, and ultimately increases the odds of encountering the true responsible gene.

General Discussion

1. Introduction

The global demand for chemicals and energy is increasing steadily and unfortunately accompanied by a detrimental impact on the environment. These developments, but also the depletion of the natural resources necessitates a transition from fossil fuels to biomass. Clean technologies are thus required for industry to comply with the ever growing demand for energy and chemicals. One of the promising approaches that attracted widespread attention for quite some time is microbial green chemical productions by engineered micro-organisms through 'synthetic biology'. This scientific field combines different disciplines and aims to create, control and programme cellular behaviour, leading to improved so-called microbial cell factories. Here, the focus lies mainly on the design and optimization of biochemical pathways in microbes to enable or increase the production of interesting compounds. One of the limitations however, is that these metabolic modifications may cause an enormous stress on the micro-organisms. Because of this, applied research also focusses on developing new methods for survival or alleviation of the demanding production conditions. One example is to enhance product export by the engineering of transporters that specifically transport a product of interest like amino acids and alkanes [396-399]. The most interesting compounds contributing to a green economy, however, are alcohols and fatty acids. This is mainly because of their high energy storage, easy production and potential to cut CO_2 emissions [400]. For

microbial production of these compounds, however, it has long been recognized that the toxicodynamic site of interaction is usually focussed on the membrane [401-403]. Especially long chain alcohols are known to fluidize the membrane by intercalating between the phospholipids. This ultimately leads to lysis of the production host. To deal with this problem, complementing traits are generally focussed on stabilizing the cytoplasmic membranes to increase solvent tolerance. Such fortification strategies through membrane engineering are usually done by computational or rational design. One can model or predict the outcome of membrane dynamics *in silico* and apply this knowledge to the host. Rational design on the other hand can be done by reproducing promising homeoviscous adaptation mechanisms to the host. Beneficial membrane modifications can be identified by exposure of compounds to multiple micro-organisms or copying membrane modifications from naturally adapted micro-organisms. In this study, we intended to produce a membrane composed partially of archaeal lipids as a fortification strategy. This endeavour is based on the observation that some archaea can withstand high levels of solvents, low pH and high temperatures.

Are archaeal membranes truly more robust?

Because many archaea thrive in extreme habitats, it is generally believed that their membranes, composed of diether or tetrether lipids, are more stable than their bacterial counterparts. However, are archaeal membranes truly more robust? To address this question, a couple of biophysical studies on liposomes have been done to compare archaeal lipids with bacterial lipids [28, 156, 404, 405]. These studies consistently concluded that archaeal lipids have a higher thermal stability and lower permeability to protons and carboxyfluorescein. With the exception of one study [121], these reports did not involve lipids from extreme hyperthermophilic bacteria, nor was the question ever addressed *in vivo*. In **Chapter 2**, we therefore listed the growth boundaries of archaea and bacteria and the corresponding membrane adaptation mechanisms. To come to an unbiased answer to the question on archaeal membrane stability, an unambiguous perspective of robustness is required. To address this, we listed all published data on the growth range boundaries, instead of just the growth optimum. A comparison of archaeal and bacterial robustness however also requires a precise specification of one particular physicochemical challenge. This is mainly because adaptability of the barrier function is found to be flexible but not universal. Different physicochemical perturbations elicit different adaptation mechanisms in the cytoplasmic membrane which is especially illustrated in the following example. It is generally known that some archaea, like methanogens, can tolerate high levels of organic solvents in deep sea oil reservoirs at high temperatures [406]. However, when comparing the hyperthermophilic archaeon *Thermococcus kodakarensis* (also tetraether lipids) with a mesophilic bacterium like *Escherichia coli*, they seem to have nearly comparable NOAEL (no observed adverse effect level) and LD50 (median lethal dose) to butanol exposure (**Fig. S1**). This specific comparison shows that robustness cannot be generalized nor attributed to membrane composition alone. Neither *T. kodakarensis*, nor *E.*

coli are commonly exposed to butanol. Consequently, one cannot exclude the possibility that e.g. maintaining the functionality of membrane proteins poses a bigger challenge than the membrane itself. To come to a justified conclusion about archaeal and bacterial membrane robustness, we thus only compared the membrane adaptations of various extremophiles to heat, cold, acidity, alkalinity and pressure in **Chapter 2**. In this case, variables other than membrane robustness are assumed to be equal and eliminated from the comparison. Besides the (pseudo)peptidoglycan cell wall, the phospholipid composition is only one of few true chemical differences between archaea and bacteria. The comparative analysis from **Chapter 2** thus provides the most legitimate answer to the question whether the membrane is accountable for a difference between survival boundaries. Our findings confirmed that the archaeal domain is relatively abundant in environments with the highest temperatures and lowest pH. Bacterial lipids were not observed above 100°C or below pH 1 in contrast to archaeal lipids. The bacteria, on the other hand, seem to dominate the coldest environments and to a lower extent also to alkalinity. Archaea were not observed growing below -2°C or above pH 11 in contrast to bacteria. On the basis of the encountered lipid compositions we postulate that phosphatidic acid-based lipids are superior in inflicting membrane fluidization, whereas archaetidic acid-based lipids are generally better accommodated for rigidifying modifications. A mechanistic foundation that supports this is the fact that fluidization is generally aimed at providing more space between hydrocarbon chains at the hydrophobic centre of the bilayer. This adjustment is most efficiently done by introduction of double bonds in fatty acids to generate cone-shaped lipids. Conversely, this may be less efficiently accomplished in archaeal lipids which retain a cylinder shaped lipid despite the presence of double bonds. On the other hand, rigidifying modifications are generally aimed at providing less space and motion at the interface of the bilayer, with membrane spanning lipids forming a monolayer as extreme example. A consequence of this is a denser packing which culminates in an extremely low permeability to solutes in archaeal membranes. One minor flaw of the extremophile comparison is the fact that we cannot exclude that measurement errors are included in the analysis by selection of studies. It e.g. cannot be ruled out that in the future, bacteria with higher growth temperatures or archaea with much lower growth temperatures will be isolated. However, due to the fact that classical sampling and cultivation methods do not favour archaea over bacteria, or vice versa, it seems unlikely that the observed boundaries have been prone to a major selection bias. Support for this is found in a recent meta-genomics and transcriptomics sequence analysis study of the subglacial lake Vostok (Antarctica)[407] which confirms one of the conclusions drawn. Here, the metagenome showed a distribution of 94% bacterial sequences, 6% eukaryal and only a trace amount of archaeal sequences. This study on a highly confined environment supports the bacterial dominance of cryogenic habitats as was also concluded from our comparative analysis.

2. Contemplations on the mixed membrane hypothesis

In **Chapter 6**, we introduced an entire archaeal lipid synthesis pathway in *E.coli* and attempted to simultaneously down-regulate the endogenous bacterial lipid pathway (**S1**). The main goal of the construction of a mixed membrane was to generate a robust membrane and meanwhile provide the possibility to study the proposed hypothesis about the archaeal and bacterial divergence. Wächtershauser postulated his mixed membrane hypothesis in 2003 [14] to account for some seemingly incompatible facts. Archaea and bacteria share a high number of orthologous genes, mainly for transcription and translation, and are therefore most certainly descendants from a primeval common ancestor cell [408]. A multi-component genetic machinery must therefore have existed and due to its heritable character also been confined within a cellular lipid membrane. Extant archaea and bacteria, however, have highly distinct membrane lipids, which raises questions about the composition of the pre-cell membrane. Wächtershäuser, proposed that this might have been a mixed heterochiral membrane. Through selective events, a less stable unfavourable mixed membrane segregated into membranes with a predominance of fatty-acyl-ester-G3P lipids in a bacterial ancestor, and isoprenoid-ether-G1P lipids dominating in an archaeal ancestor. Alternatively, a homochiral bacterial or archaeal membrane evolved first in pre-cells. In this scenario, a subsequent membrane transformation in either one of the domains followed through an inevitable intermediary heterochiral membrane. This last event, however, is regarded as counter-selective and its occurrence thus highly questionable. In case it would not have been counter-selective, we would come across hybrid heterochiral membranes to this day as well. These deductions thus attribute the first scenario as the most plausible one, and inevitably leads to the question why a heterochiral membrane in pre-cells evolved in the first place. The answer to this equally important question should be sought in identifying the conditions that support stability of mixed membranes instead, rather than testing its stability at mesophilic circumstances. Despite the fact that some studies on heterochirality are available [17, 409-411], consensus about the stability of heterochiral membranes has not yet been reached. To our humble opinion, this is because it is dependent on what is considered to be a stable membrane. Whereas some authors put emphasis on low permeability, however, more importantly is probably only the ability of lipids to self-assemble into a closed bilayer system. Both the level of permeability and fluidity to modern-times mesophiles are subordinate because the conditions where pre-cells evolved are largely unknown, but was likely heterogeneous in all aspects. The conditions where initial life evolved are hypothesized to be from deep sea hydrothermal vents, mineral cavities or at the interface of natural proton and hydroxide ion laminar flows [13, 347]. At this stage in the evolution of LUCA, it has also been speculated that the proto-membrane self-assembled from lipids from non-enzymatic origin [14, 412]. Peretó *et al.* and Lombard *et al.*, however both point at the phylogenomic observation that biosynthesis of fatty acids and isoprenoids and a member of the two superfamilies of G1P

and G3P synthesizing dehydrogenases were likely present in LUCA already. If true, this does not discard the 'abiotic source of lipids' hypothesis, but is difficult to reconcile with a LUCA having all pathways available at certain point in time.

The reason why mixed membranes evolved in LUCA, may possibly not be due to selection pressures that favoured the characteristics of a mixed membrane. Instead, the solution may be that before emergence of fully developed LUCA cells, individual constituents relied on exogenous abiotic sources that were irregular and unreliable in their supply. The consequence may have been that mixing hydrocarbon chains provided higher chance of survival. Another consequence of this is that enzymatic pathways for the synthesis of hydrocarbon chains appeared in a later stage than the lipid assembly enzymes and as a result required broad substrate promiscuity. Until now, for many segments of the lipid assembly machinery, rudimentary substrate promiscuities indeed have been reported. In the proposed scenario: (i) early stage pre-LUCA-cells may have had non-stereoselective glycerol-phosphate dehydrogenases [413]. Our findings of substrate promiscuity of both GGGP-synthase and DGGGP-synthase from *M. maripaludis* towards G3P and the unusual bacterial G1P production in *E.coli* (**Chapter 6**) supports the possible coexistence of G1P and G3P in this pre-LUCA cell. In this stage fully developed enzymes for polar head group attachment were also already present. Koga specifically points at the fact that these enzymes have very broad specificity even to unnatural lipids like isoprenoid-esters [7]. In this stage, these cells had access to free fatty acids and perhaps also to isoprenoids from non-enzymatic origin. The abiotic source of fatty acids could have been, as indicated by Lombart *et al.*[13], through Fisher-Tropsch synthesis involving CO and H_2 reactions in the presence of iron catalysts under the conditions of hydrothermal vents. Additionally, the organic carbon content in meteorites contains fatty acids and glycerol monoacyl derivates which are shown to readily self-assemble to primitive membranes [354]. The fact that both archaeal and bacterial species are capable of generating the uncommon fatty acyl ether lipids [5, 6] supports this hypothesis to some extent. (ii) In a subsequent stage, a too limited supply of fatty acids may have provoked the invention of the fatty acid synthesis pathway (FAS) and prenyl-synthases. This is further supported by the fact that both enzyme families are present in contemporary archaea and bacteria which were ultimately recruited for different purposes [413]. (iii) In the final stage before the divergence of the archaeal and bacterial domain, stereo-selective glycerol-phosphate dehydrogenases evolved which triggered the separation into two domains.

3. MEP-DOXP up-regulation and subsequent improvements on flux

The success of engineering a mixed membrane in *E.coli* is for a great deal dependent on an enriched pool of the isoprenoid building blocks isopentenyl-diphosphate (IPP) and dimethylallyl-diphosphate (DMAPP). For this reason we first embarked on the engineering of a platform strain which could rely on two completely different strategies. The first option

was to introduce a complete heterologous HMG-CoA reductase (mevalonate) pathway. On this occasion, the desirable absence of innate flux regulation provides a major benefit as already indicated by Martin V.J.J. *et al.* [237]. The main challenge of this approach would have been the introduction of an imperatively high number of heterologous genes (8 in total) and the subsequent refactoring of that pathway. Furthermore, this strategy would have been burdened with difficulties when attempting to obtain stable integrations for such a high number of genes. The alternative was to up-regulate the innate MEP-DOXP pathway as described in **Chapter 3,** (**Fig. S2A**). The main advantage of the latter strategy is that only endogenous enzymes need to be manipulated and heterologous gene expression is not required. To our experience, however, the up-regulation of the MEP-DOXP pathway by overexpression of the *idi*, *ispDF* and *dxs* genes alone did not lead to very high yields. The main question that arose was whether innate regulatory inhibitions were counteracting the effect of overexpression. We thus had the choice to further elucidate the encountered flux limitation or systematically trying additional over-expressions on other MEP-DOXP genes. A few studies are available demonstrating that higher isoprenoid yields can be achieved by other strategies that mitigate subsequent limitations [210-212]. One example is by targeting the pool of (methylerythritol cyclodiphosphate) MEC that leaks into the extracellular environment. Zhou *et al.* show that this loss can be confined by overexpression of *ispG* and the knock-out of potential MEC efflux pumps [210]. Another option is given by Bitok *et al.* who demonstrated a feedback and feedforward inhibitory mechanism on IspF *in vitro* [211]. To study the role for the latter mechanism during up-regulations we tested the overexpression of *idi* and *dxs* with either *ispD or ispD* and *ispF*. These overexpressions were compared with a construct without the latter genes (*idi-dxs*). This *idi-dxs* construct could in potential have demonstrated a substantial isoprenoid decrease in case of such feedback mechanism was involved during *in vivo* up-regulation of the MEP-DOXP pathway. One could argue however that this setup did not allow us to draw firm conclusions about any regulatory event *in vivo,* given the fact that the inhibitory complex wasn't measured nor the carbon flux. Secondly, the accumulation of farnesyldiphosphate (FPP) to complex with IspF and 2C-methyl-D-erythritol 4-phosphate (MEP) was rather disputable. On the other hand, this setup did allow the determination of the individual constraints of IspD and IspF on the up-regulation, and as such could have motivated us to perform metabolomics analysis to study the involvement of a potential feedback inhibition. In this set-up, however, we showed that IspF does not provide a major contribution nor limitation during MEP-DOXP up-regulations and that feedback inhibition on IspF is therefore unlikely to occur. When Zhang *et al.* reported the implications of genotype improvement and statistical medium optimization [213], their findings validated that physiological control elements at the level of the MEP-DOXP enzymes are unlikely to play a major role.

On the basis of our archaeal lipid production assays (**Chapter 6**) we concluded that two other constraints may have limited further MEP-DOXP up-regulations as determined in **Chapter 3**. Firstly, we demonstrated that DGGGP-synthase from *Methanococcus maripaludis*, instead

of *Archaeoglobus fulgidus*, for a great deal improved the lipid (isoprenoid) production. This particular observation showed that isoprenoid flux may possibly be further restricted by the downstream isoprenoid pathway. It is conceivable that other lipid intermediates (GGPP or GGGP) were accumulating in the strain with the *A. fulgidus* gene besides the measured end-products. This finding also pointed at the risk of relying flux increase strategies on the detection of only one metabolite. Metabolomics analysis would have been desirable to adjust or pinpoint individual rate limitations. Retrospectively, the measured lycopene levels were therefore a bad representation of flux increase in **Chapter 3** as well. In the case of DGGGP-synthases, the *A. fulgidus* gene appeared to be a major limitation to the production of the measured end-products archaetidylglycerol (AG) and DGGGP and supports the conclusions being drawn from (**Chapter 3**) that production levels can benefit from optimizing the expression of the downstream pathway of the isoprenoid variants. Secondly, another even higher constraint appeared when a substantial improvement was observed by changing LB medium to OPT1 defined medium [213]. Although the reported OPT1 medium was not provided with a rationale behind stimulating isoprenoid synthesis [213], the OPT1 recipe provides some insight. The most contributing constituents are glycerol and citric acid. The glycerol is a readily used carbon source that upon internalization becomes phosphorylated. The resulting glycerol-3-phosphate is than converted to dihydroxyacetone-phosphate which upon conversion to glyceraldehyde-3-phosphate is readily used as substrate for the MEP-DOXP pathway. The citric acid on the other hand, diminishes conversion of pyruvate to oxaloacetate for the TCA cycle. A consequence of citric acid uptake is thus a redistribution of pyruvate as second required substrate for the MEP-DOXP pathway. The effect of this medium, however, could possibly also be reached by a metabolic engineering approach targeting the pyruvate branch-point. A proposed strategy would be to co-express gluconeogenic phosphoenolpyruvate carboxykinase (*pckA*) which converts oxaloacetate back to phosphoenol-pyruvate (**Fig. 1**). This overexpression can be complemented with an overexpression of glycerol-phosphate synthase (*gapB*) that converts glycerate to glyceraldehyde-3-phosphate which increases the second required substrate of the MEP-DOXP pathway. This approach may possibly result in an increased MEP-DOXP flux that is comparable to the use of the OPT1 medium.

Initially we intended to compare the transcription levels of integrated operons between a set of strategically chosen loci dependent on extensive protein occupancy domains (EPODs). At the chosen *he*EPOD *(highly expressed)* position we expected to find higher transcription levels as compared to the EPOD-independent locus and the *ts*EPOD (*transcriptionally silent*) locus. The supposedly varyingtranscription initiation levels were also hypothesized to result in different levels of MEP-DOXP up-regulation as a consequence of varying degrees of expressions. Retrospectively, a successful integration at the initially favoured *he*EPOD was unlikely to significantly improve the MEP-DOXP up-regulation. When we tested the integrated operons, we observed that a lower expression level of *idi*, *ispDF* and *dxs* was

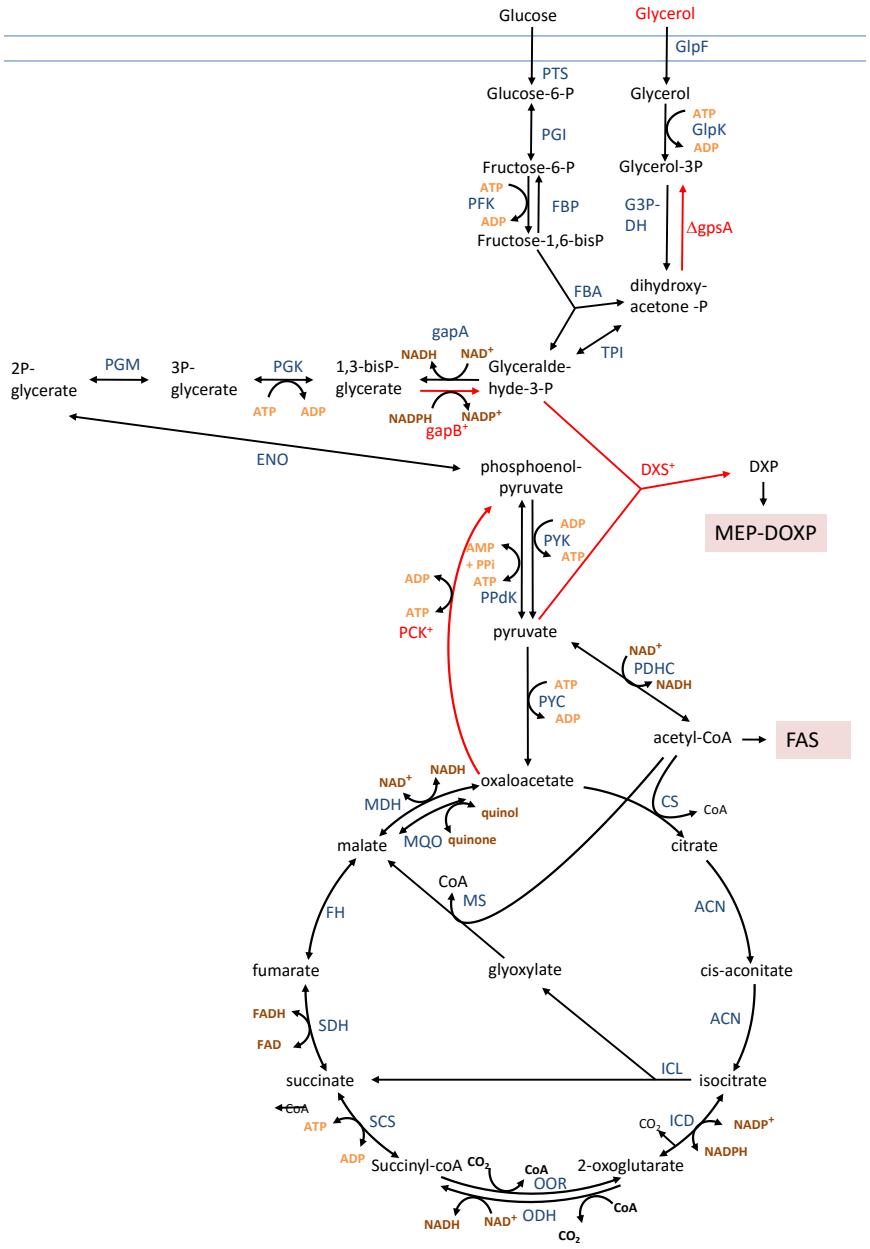


Figure 1. Proposed strategy to increase MEP-DOXP flux and a complementing approach to obtain an auxotrophic conditional knock-down of phosphatidic acid synthesis Red arrows and markings indicate targets for overexpression or knock-out. Overproduction of GapB: glyceraldehyde-3-phosphate dehydrogenase; DXS: 1-deoxy-D-xylulose-5-phosphate synthase; Pck: phosphoenol pyruvate carboxykinase are proposed to increase MEP-DOXP carbon flux. A conditional down-regulation of phosphatidic acid synthesis can be obtained by generating a *gpsA* (glycerol-3-phosphate dehydrogenase) knock-out strain and providing inadequate glycerol levels.

causing a significantly increased production of lycopene compared to the plasmid-encoded variants. A higher transcription from the *he*EPOD would therefore not have contributed to higher isoprenoid yields as well.

4. Down-regulation of the endogenous lipid synthesis,

To down-regulate *E.coli's* endogenous lipid synthesis, we focussed on the first committed step of phospholipid synthesis, i.e. glycerol-3-phosphate acyl-transferase. This conversion requires two substrates, which are acyl-ACP and glycerol-3-phosphate. As the fatty acid component is also required for purposes other than phospholipids, targeting the fatty acid synthesis (FAS) cycle would have been a too risky endeavour. Especially, the necessity for many membrane proteins to become post-translationally modified by acylation and the precursor role in cofactors like biotin would have required an extensive tuning when targeting fatty acid synthesis. As an alternative to the FAS cycle, we attempted to silence lysophosphatidic acid synthesis being the first committed step of phospholipid synthesis. In bacteria, this catalysis is performed by the *plsX-plsY* pathway. In *E.coli*, this is by exception and for yet unknown reasons also performed by the alternative *plsB* pathway [270]. To conditionally knock-down lysophosphatidic acid synthesis we therefore exchanged the *plsB* promoter region with a tunable rhamnose promoter in a *plsX* knock-out strain (**Fig. S2A** and **S2B**). Because of insufficient repression of the rhamnose promoter, however, we had to complement the strain with an overexpression of *acpP* (acyl carrier protein). This overexpression leads to an excess of the inactive apo-ACP isoform, which is known to competitively inhibit the redundant PlsB [277]. Here we observed dramatic growth retardation when inducing both PlsB and ACP overexpression, which suggested that membrane synthesis was silenced. We looked into the mechanism of action by performing a proteome analysis of the induced strain compared to a wild-type *plsB* strain. This led us to conclude that expression of fatty acid synthesis was down-regulated and fatty acid degradation was up-regulated in the conditional knock-down strain. By exogenously adding fatty acids and lysophophatidic acids we concluded that a shortage of available acyl-ACP was likely the cause of the growth defect. Although all data supported this conclusion, a major limitation of the experimental setup was that the concentration of key metabolites was never measured directly (i.e. acyl-ACP and fluxes) but solely deduced from the growth performances, proteome data and fatty acid and LPA rescue assay. Secondly, one can argue that an overproduced PlsB and apo-ACP pool cannot represent a realistic physiological control mechanism to down-regulate membrane synthesis. Nevertheless, it is hard to believe that a competitive inhibition by a cofactor (apo-ACP) has no biological reason or significance. In addition, *E.coli* is also equipped with an ACP-phosphodiesterase (*acpH)* that converts holo-ACP back to inactive apo-ACP, for which the physiological role is largely unknown [414]. Taken also these factors into account, it is conceivable that the proposed silencing mechanism is implemented to abruptly silence lipid synthesis when needed.

As a follow up to this study we therefore suggest to repeat the proteomics analysis to statistically validate the observed effects and to complement this data with metabolomics data. Furthermore, this study can be extended with transcriptomics specifically targeting ACP-synthase (*acpS*) and ACP-phosphodiesterase (*acpH*), to determine the conditions or growth phases at which apo-ACP allows detection of physiological levels of apo-ACP.

When we implemented the PlsB-ACP inhibitory system in combination with the archaeal lipid synthesis, we did not observe a significant decrease in the amount of endogenous lipids (not shown). One of the main determining factors was the fact that production of archaeal lipids was dependent on the presence of glycerol whereas the down-regulation of bacterial lipids was dependent on rhamnose. Because these inducers are both carbon sources and had to be combined with IPTG at the same time, it is unclear whether a combination of glycerol and rhamnose could elicit both desired effects, or catabolite repression was frustrating either one of the systems. In addition to this, we had to introduce the required *acpP* gene into the pMS148 plasmid which may have resulted in different expression levels as established during the proteome analysis (**Fig. S2B**). These factors, along with an incomplete characterization of induction of the down-regulatory system may have caused the inadequacy of our attempts to lower the fraction of bacterial lipids. An alternative that was more likely to cause significantly lower levels of endogenous lipids was the supplementation of glucose to elicit catabolite repression via the CAP domain on the rhamnose promoter of *PlsB*. However, also the archaeal lipid synthesis might be affected, as both the pMS148 and pAC029 plasmids and integrated MEP-DOXP operon are also controlled by a CRP or LacI repression. As an alternative to the down-regulation of endogenous lipid synthesis via *plsB*, we also attempted to generate a disruption of glycerol-3-phosphate dehydrogenase (*gpsA*). Despite having access to a G3P transporter (GlpT) [415], this mutant could not be isolated as glycerol-3P auxotroph when attempting to disrupt this essential gene (*gpsA*). As an alternative, one could try to obtain a Δ*gpsA* mutant by providing glycerol instead. *E.coli* is able to take up exogenous glycerol into the cytoplasm via the GlpF transporter and GlpK kinase system. The glycerol channel (GlpF) facilitates the uptake of glycerol upon which the kinase (GlpK) traps the glycerol inside the cell as glycerol-3-phosphate. A depletion of glycerol-3-phosphate in a *gpsA*–knockout strain can subsequently be obtained by providing insufficient or no glycerol as a means to conditionally knock-down phosphatidic acid synthesis (**Fig. 1**). This approach however, is presently hampered by the committed use of OPT1 medium. This is because glycerol is the foremost constituent of OPT1 because of its crucial role in MEP-DOXP upregulation in the current strain (WUR182). This detail can be solved by implementation of the previously discussed overexpression of *pck* and *gapB* to circumvent the as yet compulsory usage of the OPT1 medium to further increase isoprenoid flux.

5. A hybrid heterochiral membrane and its implications

One of the prominent achievements of this study was the production of a membrane consisting of 28% archaeal lipids. Basically two observations support that the foreign lipids were incorporated in the membrane. Firstly, the archaeal head group attachments were almost completely dependent on enzymes endogenous to *E. coli.* As all of these enzymes are membrane embedded, the archaeal lipids must have been present at that location as well. Secondly, we observed that the endogenous lipid pool of phosphatidylglycerol (PG) decreased proportionally to the increase of archaetidylglycerol (AG). As discussed in **Chapter 6** we postulate that this seemingly mutual exclusiveness was a result of the balanced regulation between anionic and zwitterionic lipids, and not by a lack of polar head group substrate glycerol-3-phosphate. The mechanism that keeps anionic and zwitterionic phospholipids in equilibrium is regulated via the acidic phospholipid content in the membrane. These levels elicit a compensatory synthesis of zwitterionic phosphatidyl-ethanolamine (PE) lipids by activation of phosphatidyl serine synthase (PssA). It has been shown that an ionic interaction of acidic phospholipids (PG) with PssA activates the enzyme from its latent cytosolic form [416]. Because the PG levels that normally stimulate PssA were relatively low, this regulation thus must have occurred via anionic AG as well. Accordingly, this conclusion supports the incorporation of AG in the membrane. Despite a sustained balance between the anionic and zwitterionic lipids and a fairly normal growth behaviour, the EM pictures showed dramatic morphological malformations. Here we observed a high number of extraneous appendages and subsequent release of vesicles. Although the exact cause is not determined, this finding suggests that the *E.coli* cells attempt to expel the foreign lipids, instead of maintaining lipid homogeneity. Because of this, it may indicate that the mixed membrane situation is not sustainable *in vivo*, and that a homogeneously mixed lipid membrane was never obtained. Interestingly, the *in vitro* study of Shimada and Yamagishi disclosed a comparable mol/mol ratio in their mixed lipid systems [17]. In this study, the obtained liposomes were shown to increase in thermal robustness compared to homochiral bacterial liposomes. This can only be reached when stable homogeneous structures are being formed. Shortly after this study, Balleza *et al.* also confirmed the stability of mixed lipid vesicles using synthetic lipids [417]. In this study, archaeal lipids diphytanyl phosphatidylcholine (DPhPC) were mixed at a 1:1 ratio with bacterial dipalmitoyl-phosphatidylcholine (DPPC) or dihexadecylphosphatidylcholine (DHPC). These mixtures led to stable vesicles with solid or finely reticulated domains (**Fig. 2A** and **B**) depending on the presence of ester or ether bonds in the archaeal species respectively. Over time however, they also observed the formation of solid coalescent domains from the reticulated domains indicating low miscibility of the archaeal and bacterial lipids used (**Fig. 2C**). In our strain, however, the synthesized archaeal lipids were chemically different from the lipids used in these *in vitro* studies (unsaturated diethers versus saturated diether or bipolar tetraether lipids by Shimada and Yamagishi [17] or saturated diethers lipids using homochiral G3P

backbones with choline headgroups only by Balezza *et al.* [417]. Despite this, the mixed vesicles from Balleza *et al.* are supportive of low miscibility in the hybrid membrane constructed *in vivo*.

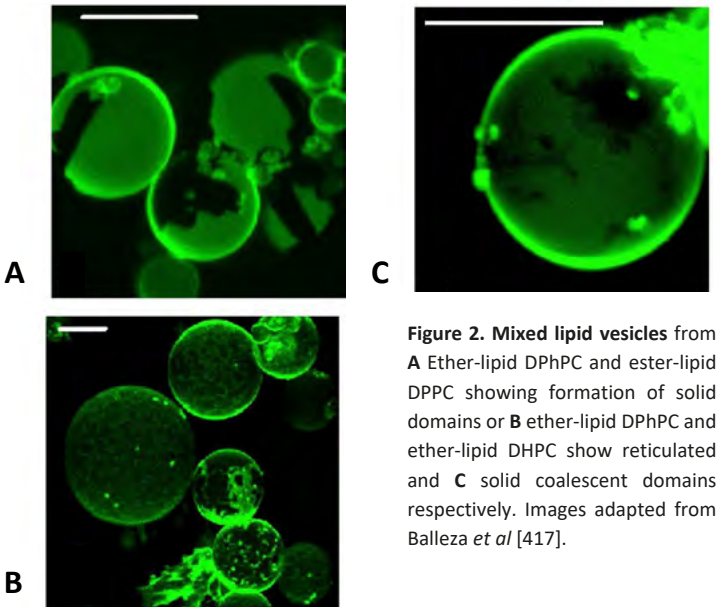


Figure 2. Mixed lipid vesicles from **A** Ether-lipid DPhPC and ester-lipid DPPC showing formation of solid domains or **B** ether-lipid DPhPC and ether-lipid DHPC show reticulated and **C** solid coalescent domains respectively. Images adapted from Balleza *et al* [417].

There are however also *in vivo* findings that support miscibility of mixed lipids. Unsaturated isoprenoids like undecaprenyl-phosphate (UndP) are well known to function as sugar carriers in bacterial membranes. These compounds are known to readily incorporate between the bacterial fatty-acid lipids [418], but their quantities are very low. A few natural mixed membranes, which are likely homochiral, are shown *in vivo*. Gattinger *et al.* reported the existence of a membrane composed of 83.4% ether-linked, and 5.6% ester-linked fatty acid hydrocarbon chains mixed with isoprenoid-ether lipids in the euryarchaeon *Methanothermus fervidus* [6]. Despite the indicated uncertainties concerning chemical disparity, our study confirms that heterochirality of the incorporated glycerol-phosphates lead to a viable situation which does not affect growth. We postulate that the unsaturated nature or the concentration of the foreign lipids may have had a significant influence on the observed phenotypical changes for which the outcome may completely change once the archaeal lipids become saturated or when the archaeal lipid content increases.

6. Hybrid membranes: a solution to solvents?

The most obvious reason why archaea are labelled as being naturally robust, is because most harsh conditions are counteracted by the efficient rigidifying ability of their lipids. Typical examples of membrane stability that may become useful in industry are those against high temperature, low pH, but also stability against high concentrations of solvents.

Because of this, we tested the WUR182 strain for its tolerance to heat, cold and butanol exposure (**Chapter 6**). With the exception of cold exposure, firm conclusions cannot be drawn yet because of the absence of a proper control strain. We deliberately neglected to include a wild-type strain. The inclusion of a strain with a comparable genotype but different phenotype seemed like a more suitable control. Therefore, we included a control strain that was expected not to contain archaeal lipids due to a dysfunctional *araM* gene for glycerol-1-phosphate synthesis. However, from the LC-MS analysis it appeared that this control strain also had archaeal lipids albeit at lower quantities (10% vs. 28%). Alternatively, the mixed membrane strains are referenced to the included lycopene strain. Using this strain as a reference, however, should be approached with caution. This is because lycopene is known to intercalate in archaeal membranes. In bacteria, polar carotenoids are described to buffer the fluidizing effect of *cis*-unsaturated fatty acids, but the effect of the non-polar carotenoid lycopene on bacterial membranes is unknown. In the robustness assays, we were expecting to find a higher survival after high temperature exposures because of the rigidifying effect by the methyl branches of the archaeal lipids. The results, however, indicated a lower tolerance to heat shocks compared to the strain with low levels of ether-lipid and the lycopene control. Interestingly, we also observed an increased tolerance to butanol exposure and also to cryogenic conditions. The cryoprotectant effect is best explained by a reduced rupturing of membranes by ice crystals.

To explain the observation of a reduced heat shock survival, two possibilities come to mind. Firstly, the produced diether lipids are fully unsaturated. It is generally accepted that the unsaturated intermediate DGGGP, displays a level of stability to acidity and heat that is similar to ester-lipids because of the allylic ether bonds [419]. The consequence of unsaturated archaeal lipids on membrane fluidity however is ambiguous. The record-holder of hyperthermophilic growth, *Methanopyrus kandleri* contains high levels of unsaturated lipids including DGGGP and AG [66, 420], and were also detected in *Thermococcus kodakarensis* in **Chapter 7**. The psychrophilic archaeon *Methanococcoides burtonii* however is also highly enriched with unsaturated lipids [64]. One factor that substantially changes the stability of unsaturated lipids is the position of the double bonds. An ether-lipid is more labile when the unsaturation is positioned adjacent to the ether bond by forming an allylic ether [419]. This lability however is completely abolished in case the unsaturations are not positioned at the C1 or C2 relative to the ether bond. This situation leads to maximally 6 unsaturations per diether lipid and possibly also occurs to some extent in both *M. kandleri* and *M. burtonii*. A comparison with the unsaturated lipids encountered in the hyperthermophile *M. kandleri* and psychrophiles *M. burtonii* may thus not be completely justified. On the other hand, as hydrolysis of the allylic ether bond was established complete at 80°C during HCL-methanolysis [419], the allylic lability may not be involved in the decreased heat shock survival which was performed below 58°C under non-acidified conditions for only 2 min. Instead, the membranes ability to maintain a closed and intact bilayer (stability) when mixed with ester-lipids may explain the observation. In this case

the trend that cells with 28% archaeal lipids are less tolerant to heat shocks than cells with 10% archaeal lipids is an interesting observation. It suggests that higher archaeal lipid ratios lead to a lower ability to maintain temperature dependent stability. Conversely, a mixed membrane may not benefit from an archaeal lipid percentage of only 28%. Although, the 28% ratio is the highest accomplished level of mixed lipids *in vivo* thus far, it may simply not be high enough to obtain beneficial effects on thermal stability. Despite this possibility, the trend we observed in the heat shock assay is not indicative of an improved survival when the ratio increases. However, it cannot be ruled out.

The set-up of the butanol exposure assay also did not allow us to draw firm conclusions about solvent robustness. Interestingly, a 1.5% butanol exposure led to a relatively higher level of recovered colonies compared to the lycopene control strain. This finding is indicative of a higher robustness. A likely explanation for this increased robustness might be the steric hindrance of methyl-groups preventing intercalation of the alcohols between the membrane lipids. One could argue that a 1-h exposure was simply not long enough to claim an increased butanol survival; however a 2.5% butanol exposure was high enough to lyse nearly 100% of the cells regardless of the strain being used. At a concentration of 1.5% butanol, 100% of the cells with a mixed membrane survived the 1-h butanol exposure. Applying the mixed membrane may therefore seem like a promising strategy to process or produce alcohols. More research however, is required with the inclusion of a proper control strain to determine whether growth is also tolerated in presence of alcohols rather than measuring survival after a transient exposure.

7. Perspectives

Future research should first focus on the issues of miscibility and saturation. We would like to address the miscibility issue by approaching it from two different angles. Determining miscibility can first be done by isolating the observed budding appendages or released vesicles from the strain with the hybrid membrane. This can be done either with or without a brief sonication treatment to accelerate the detachment of the appendages. The resulting vesicles can be analysed by LC-MS to determine whether enrichment of archaeal lipids in the vesicles has taken place. In case no enrichment took place, the archaeal lipids may become homogeneously divided over the plasma-membrane or perhaps aggregated in patches/rafts dispersed throughout the membrane. As a secondary approach to assess the miscibility, one could develop a fluorescent probe that specifically binds to ether-bonds, or use an unspecific probe like 1-(4-trimethylammoniumphenyl)-6-phenyl 1,3,5-hexatriene *p*-toluene-sulfonate (TMA-DPH)[417, 421] to visualize the dispersion of archaeal lipids. Ideally, these experiments are also done with cells which are equipped with a functional geranylgeranyl reductase to compare saturated with unsaturated archaeal lipids. As already emphasized by Koga and Morii [419], unsaturated core lipids, DGGGP and thus also the obtained unsaturated AG, may contain relatively labile allyl ether bonds. These bonds have

an equal stability in terms of heat and acid tolerance as bacterial acyl-ester lipids. Besides this difference by the level of saturation, a difference in miscibility between unsaturated and saturated diether lipids is also likely to occur. Unfortunately, packing efficiency or dynamics of both saturated and unsaturated archaeal lipids have never been modelled in molecular dynamics studies. Therefore it is thus not possible to predict their different behaviour. In contrast, for bipolar tetraether lipids with its molecular variants several studies are available [123, 422, 423].

This brings us back to the discussion on the mixed membrane hypothesis of the common ancestor of archaea and bacteria. What are the implications of our findings for this hypothesis? Two main findings described in **Chapter 6** support the primordial existence of a mixed membrane in LUCA: (i) the viability of a membrane containing 28% archaeal lipids and (ii) the promiscuous specificity of bacterial enzymes for archaeal lipid substrates and *vice versa.* Several key questions, however, remain unanswered. In this respect, the question whether archaeal and bacterial lipids result in homogeneous mixtures has currently the most important implication for the mixed membrane hypothesis. In case the archaeal lipids are not miscible but expelled trough budding, it is interesting to study whether this is an active budding process like formation of outer membrane vesicles (OMV's)[424], or accelerated by spontaneous lipid segregation. To come to a proper biophysical characterization of a mixed membrane, the viscosity of the membrane can be measured by performing fluorescence anisotropy. A temperature dependent proton permeability of the mixed membrane can be determined by preparing liposomes loaded with a pH buffer. Proton permeability can be assessed by monitoring the pH from a fluorescent pH indicator dye (pyranine) and pulsing the liposome mixtures with acid according to Nichols and Deamer [425].

Besides the key questions concerning miscibility, viscosity and permeability of the membrane, many more minor issues remain unsolved. Firstly, the question remains whether higher levels of archaeal lipids are also tolerated and whether these lipids are sustainable over an extended period of time or ultimately counter-selected. This issue can be solved once the proposed engineering has been accomplished to make the use of OPT1 medium redundant. Growing a strain with the proposed modifications (**Fig. 1**) and provided with limiting glycerol concentrations might then result in even higher archaeal lipid levels than detected thus far. Another issue that is interesting to study is the composition of the peptidoglycan cell wall. Here one might expect to observe some dramatic effects. The endogenous pool of farnesyl-diphosphate is natively converted to undecaprenyl-diphopshate (UndP) which functions as a carrier to transport LPS constituents from the inner membrane to the periplasmic space via the lpt translocase system. Due to an up-regulated synthesis of GGGP, the level of available UndP might become limiting with dramatic consequences to the peptidoglycan layer. Another interesting issue is the location of the archaeal lipids, i.e. do they end up in the inner or outer membrane, or both? Lipids from the outer-membrane are usually transported by flippases and therefore the novel archaeal lipids might only accumulate

in the inner membrane. This question can be addressed by separating the inner from the outer membrane by making use of a sucrose density gradient ultra-centrifugation [426]. As a final prospect to an application of the mixed membrane, the strain is a highly interesting tool for the identification of the missing link in glycerol dialkyl glycerol tetraether (GDGT) lipid synthesis. Here the envisioned enzyme system attaches two diether lipids to form a single membrane spanning tetraether lipid via a tail-to-tail condensation reaction [9]. Both the enzyme(s) and the involved substrates remain unidentified to this day. The mixed membrane strain can provide a powerful tool for the identification of the responsible enzyme, because all of the obviously anticipated substrates are present within the strain. The genes identified in **Chapter 7**, or a series of heterologously expressed genes from a mesophilic archaeon like *Methanospirillum hungatei* (T_{opt}=37°C, 59%TE's) with confirmed tetraether lipids can be used in the mixed membrane model to screen for GDGT synthase activity.

Supplemental Information

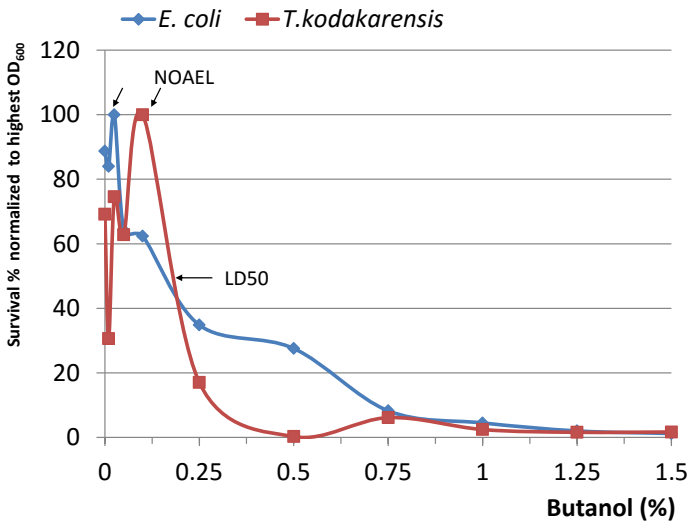


Figure S1. Survival ratio's after butanol exposure in *E. coli* and *T. kodakarensis*. *E. coli* has a NOAEL (No observed adverse effect level) of 0.025% butanol and an LD50 at around 1.9 %. *T. kodakarensis* is determined to have a NOAEL level of 0.025% and a comparable LD50 around 0.19% butanol.

A

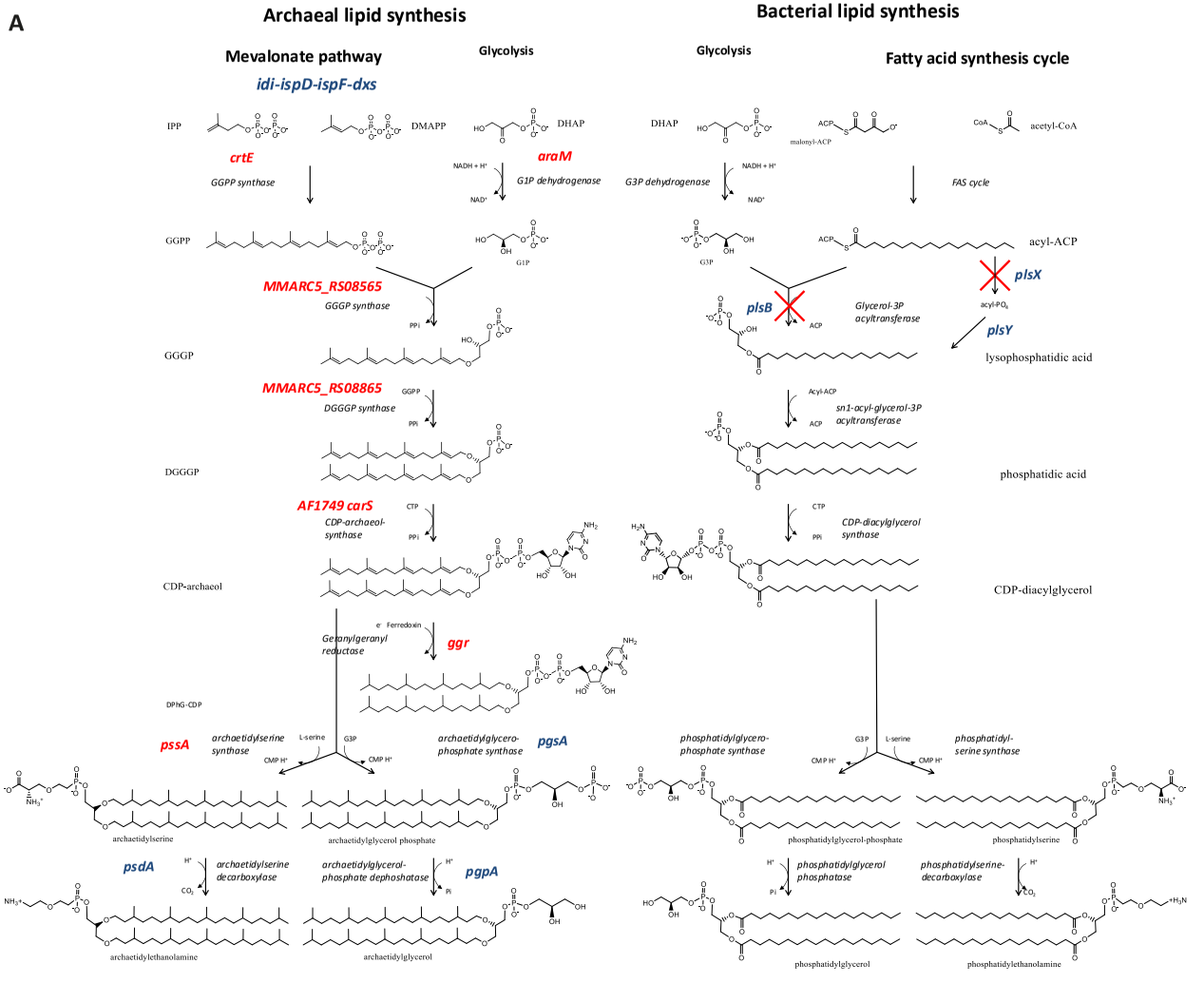

Archaeal lipid synthesis
Bacterial lipid synthesis
Mevalonate pathway
idi-ispD-ispF-dxs
Glycolysis
Glycolysis
Fatty acid synthesis cycle
IPP
DMAPP
DHAP
DHAP
malonyl-ACP
acetyl-CoA
crtE
GGPP synthase
araM
G1P dehydrogenase
G3P dehydrogenase
FAS cycle
GGPP
G1P
G3P
acyl-ACP
MMARC5_RS08565
GGGP synthase
plsB
Glycerol-3P acyltransferase
plsX
plsY
acyl-PO4
GGGP
lysophosphatidic acid
MMARC5_RS08865
DGGGP synthase
Acyl-ACP
sn1-acyl-glycerol-3P acyltransferase
DGGGP
phosphatidic acid
AF1749 carS
CDP-archaeol-synthase
CDP-diacylglycerol synthase
CDP-archaeol
CDP-diacylglycerol
Geranylgeranyl reductase
ggr
DPhG-CDP
pssA
archaetidylserine synthase
archaetidylglycero-phosphate synthase
pgsA
phosphatidylglycero-phosphate synthase
phosphatidyl-serine synthase
archaetidylserine
archaetidylglycerol phosphate
phosphatidylglycerol-phosphate
phosphatidylserine
psdA
archaetidylserine decarboxylase
archaetidylglycerol-phosphate dephoshatase
pgpA
phosphatidylglycerol phosphatase
phosphatidylserine-decarboxylase
archaetidylethanolamine
archaetidylglycerol
phosphatidylglycerol
phosphatidylethanolamine

Figure S2A.

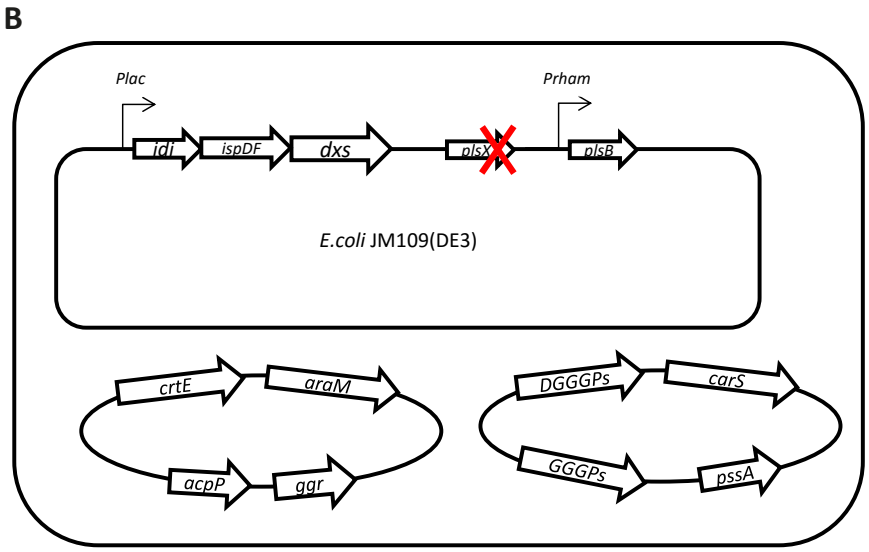


Figure S2A. Schematic representation of the initial setup of the mixed membrane engineering. With the exception of the geranylgeranyl reductase gene (*ggr*), all genes and modifications were implemented as indicated in **Fig. S2B** Abbreviations: *acpP*: acyl carrier protein; *araM*: glycerol-1-phosphate dehydrogenase; *carS*: CDP-archaeol synthase; *crtE*: geranylgeranyl-diphosphate synthase; *DGGGPs*: di-O-geranylgeranyl-glycerylphosphate synthase; *dxs*: 1-deoxy-D-xylulose-5-phosphate synthase; *GGGPs*: geranylgeranylglyceryl-phosphate synthase; *ggr*: geranylgeranyl reductase; *idi*: isopentenyl-diphosphate isomerase; *ispDF*: 2C-methyl-D-erythritol-4-phosphate cytidylyltransferase and 2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase resp.; *pgpA*: phosphatidylglycerol phosphatase; *pgsA*: phosphatidylglycero-phosphatesynthase; *plsB*: glycerol-3-phosphate acyltransferase; *plsX*: phosphate-acyltransferase; *plsY*: glycerol-3-phosphate acyltransferase; *psdA*: phosphatidylserine decarboxylase*;pssA*: phosphatidylserine synthase;

References

1. Woese, C.R. and G.E. Fox, *Phylogenetic structure of the prokaryotic domain: the primary kingdoms.* Proc Natl Acad Sci U S A, 1977. **74**(11): p. 5088-90.
2. Woese, C.R., O. Kandler, and M.L. Wheelis, *Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya.* Proc Natl Acad Sci U S A, 1990. **87**(12): p. 4576-9.
3. Ouverney, C.C. and J.A. Fuhrman, *Marine planktonic archaea take up amino acids.* Appl Environ Microbiol, 2000. **66**(11): p. 4829-33.
4. Damsté, J.S.S., et al., *Structural characterization of diabolic acid-based tetraester, tetraether and mixed ether/ester, membrane-spanning lipids of bacteria from the order Thermotogales.* Archives of Microbiology, 2007. **188**(6): p. 629-641.
5. Huber, R., et al., *Aquifex pyrophilus gen. nov. sp. nov., Represents a Novel Group of Marine Hyperthermophilic Hydrogen-Oxidizing Bacteria.* Systematic and Applied Microbiology, 1992. **15**(3): p. 340-351.
6. Gattinger, A., M. Schloter, and J.C. Munch, *Phospholipid etherlipid and phospholipid fatty acid fingerprints in selected euryarchaeotal monocultures for taxonomic profiling.* FEMS Microbiology Letters, 2002. **213**(1): p. 133-139.
7. Koga, Y., *Early evolution of membrane lipids: how did the lipid divide occur?* J Mol Evol, 2011. **72**(3): p. 274-82.
8. Langworthy, T.A., *Lipids of Bacteria living in extreme environments*, in *Membrane lipids of prokaryotes*, A.K. Felix Bronner, Editor. 1982.
9. Villanueva, L., J.S. Damste, and S. Schouten, *A re-evaluation of the archaeal membrane lipid biosynthetic pathway.* Nat Rev Microbiol, 2014. **12**(6): p. 438-48.
10. Kon, T., et al., *Effects of a squalene epoxidase inhibitor, terbinafine, on ether lipid biosyntheses in a thermoacidophilic archaeon, Thermoplasma acidophilum.* J Bacteriol, 2002. **184**(5): p. 1395-401.
11. Chong, P.L., et al., *On physical properties of tetraether lipid membranes: effects of cyclopentane rings.* Archaea, 2012. **2012**: p. 138439.
12. Hopmans, E.C., et al., *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, 2000. **14**(7): p. 585-9.
13. Lombard, J., P. López-García, and D. Moreira, *The early evolution of lipid membranes and the three domains of life.* Nat Rev Micro, 2012. **10**(7): p. 507-515.
14. Wachtershauser, G., *From pre-cells to Eukarya--a tale of two lipids.* Mol Microbiol, 2003. **47**(1): p. 13-22.
15. Kandler, O., *Cell-Wall Biochemistry in Archaea and Its Phylogenetic Implications.* Journal of Biological Physics, 1994. **20**(1-4): p. 165-169.
16. Morigaki, K., et al., *Autopoietic Self-Reproduction of Chiral Fatty Acid Vesicles.* Journal of the American Chemical Society, 1997. **119**(2): p. 292-301.
17. Shimada, H. and A. Yamagishi, *Stability of Heterochiral Hybrid Membrane Made of Bacterial sn-G3P Lipids and Archaeal sn-G1P Lipids.* Biochemistry, 2011. **50**(19): p. 4114-4120.
18. Sinensky, M., *Homeoviscous adaptation--a homeostatic process that regulates the viscosity of membrane lipids in Escherichia coli.* Proc Natl Acad Sci U S A, 1974. **71**(2): p. 522-5.
19. Marr, A.G. and J.L. Ingraham, *Effect of Temperature on the Composition of Fatty Acids in Escherichia Coli.* J Bacteriol, 1962. **84**(6): p. 1260-7.
20. Kanno, M., et al., *Isolation of butanol- and isobutanol-tolerant bacteria and physiological characterization of their butanol tolerance.* Appl Environ Microbiol, 2013. **79**(22): p. 6998-7005.
21. Chang, Y.Y. and J.E. Cronan, *Membrane cyclopropane fatty acid content is a major factor in acid resistance of Escherichia coli.* Molecular microbiology, 1999. **33**(2): p. 249-259.
22. Suwannakham, S. and S.T. Yang, *Enhanced propionic acid fermentation by Propionibacterium acidipropionici mutant obtained by adaptation in a fibrous-bed bioreactor.* Biotechnol Bioeng, 2005. **91**(3): p. 325-37.
23. Wada, H., Z. Combos, and N. Murata, *Enhancement of chilling tolerance of a cyanobacterium by genetic manipulation of fatty acid desaturation.* 1990. **347**(6289): p. 200-203.
24. Tan, Z., et al., *Membrane engineering via trans unsaturated fatty acids production improves Escherichia coli robustness and production of biorenewables.* Metabolic Engineering, 2016. **35**: p. 105-113.
25. Zhao, Y., et al., *Expression of a cloned cyclopropane fatty acid synthase gene reduces solvent formation in Clostridium acetobutylicum ATCC 824.* Applied and environmental microbiology, 2003. **69**(5): p. 2831-2841.
26. Brown, D.A., et al., *Bipolar tetraether archaeosomes exhibit unusual stability against autoclaving as studied by dynamic light scattering and electron microscopy.* Chemistry and Physics of Lipids, 2009. **159**(2): p. 95-103.
27. Chong, P.L.-G., *Archaebacterial bipolar tetraether lipids: Physico-chemical and membrane properties.* Chemistry and Physics of Lipids, 2010. **163**(3): p. 253-265.

28. Komatsu, H. and P.L.-G. Chong, *Low permeability of liposomal membranes composed of bipolar tetraether lipids from thermoacidophilic archaebacterium Sulfolobus acidocaldarius.* Biochemistry, 1998. **37**(1): p. 107-115.
29. Benvegnu, T., L. Lemiegre, and S. Cammas-Marion, *New generation of liposomes called archaeosomes based on natural or synthetic archaeal lipids as innovative formulations for drug delivery.* Recent Pat Drug Deliv Formul, 2009. **3**(3): p. 206-20.
30. Koga, Y., *Thermal adaptation of the archaeal and bacterial lipid membranes.* Archaea, 2012. **2012**: p. 789652.
31. Koga, Y. and H. Morii, *Recent advances in structural research on ether lipids from archaea including comparative and physiological aspects.* Biosci Biotechnol Biochem, 2005. **69**(11): p. 2019-34.
32. Yokoi, T., et al., *Archaeal phospholipid biosynthetic pathway reconstructed in Escherichia coli.* Archaea, 2012. **2012**: p. 438931.
33. Lai, D., et al., *Reconstruction of the archaeal isoprenoid ether lipid biosynthesis pathway in Escherichia coli through digeranylgeranylglyceryl phosphate.* Metab Eng, 2009. **11**(3): p. 184-91.
34. Isobe, K., et al., *Geranylgeranyl reductase and ferredoxin from Methanosarcina acetivorans are required for the synthesis of fully reduced archaeal membrane lipid in Escherichia coli cells.* J Bacteriol, 2014. **196**(2): p. 417-23.
35. Eze, M.O., *Phase transitions in phospholipid bilayers: Lateral phase separations play vital roles in biomembranes.* Biochemical Education, 1991. **19**(4): p. 204-208.
36. Russell, N.J., *Cold adaptation of microorganisms.* Philos Trans R Soc Lond B Biol Sci, 1990. **326**(1237): p. 595-608, discussion 608-11.
37. Mykytczuk, N.C., et al., *Fluorescence polarization in studies of bacterial cytoplasmic membrane fluidity under environmental stress.* Prog Biophys Mol Biol, 2007. **95**(1-3): p. 60-82.
38. Konings, W.N., et al., *The cell membrane plays a crucial role in survival of bacteria and archaea in extreme environments.* Antonie Van Leeuwenhoek, 2002. **81**(1-4): p. 61-72.
39. Shivaji, S. and J.S. Prakash, *How do bacteria sense and respond to low temperature?* Arch Microbiol, 2010. **192**(2): p. 85-95.
40. Chintalapati, S., M.D. Kiran, and S. Shivaji, *Role of membrane lipid fatty acids in cold adaptation.* Cell Mol Biol (Noisy-le-grand), 2004. **50**(5): p. 631-42.
41. Chattopadhyay, M.K., *Mechanism of bacterial adaptation to low temperature.* J Biosci, 2006. **31**(1): p. 157-65.
42. Russell, N.J., *Psychrophilic bacteria--molecular adaptations of membrane lipids.* Comp Biochem Physiol A Physiol, 1997. **118**(3): p. 489-93.
43. Suutari, M. and S. Laakso, *Microbial fatty acids and thermal adaptation.* Crit Rev Microbiol, 1994. **20**(4): p. 285-328.
44. Denich, T.J., et al., *Effect of selected environmental and physico-chemical factors on bacterial cytoplasmic membranes.* J Microbiol Methods, 2003. **52**(2): p. 149-82.
45. Gruner, S.M., et al., *Lipid polymorphism: the molecular basis of nonbilayer phases.* Annu Rev Biophys Biophys Chem, 1985. **14**: p. 211-38.
46. Kiran, M.D., et al., *Psychrophilic Pseudomonas syringae requires trans-monounsaturated fatty acid for growth at higher temperature.* Extremophiles, 2004. **8**(5): p. 401-10.
47. Zsiros, O., et al., *Induction of polyunsaturated fatty-acid synthesis enhances tolerance of a cyanobacterium, Cylindrospermopsis raciborskii, to low-temperature photoinhibition.* Indian J Biochem Biophys, 2000. **37**(6): p. 470-6.
48. Oshima, M. and A. Miyagawa, *Comparative studies on the fatty acid composition of moderately and extremely thermophilic bacteria.* Lipids, 1974. **9**(7): p. 476-80.
49. Shivaji, S., et al., *Sphingobacterium antarcticus sp. nov., a Psychrotrophic Bacterium from the Soils of Schirmacher Oasis, Antarctica.* International Journal of Systematic and Evolutionary Microbiology, 1992. **42**(1): p. 102-106.
50. Jagannadham, M.V., et al., *Carotenoids of an Antarctic psychrotolerant bacterium, Sphingobacterium antarcticus, and a mesophilic bacterium, Sphingobacterium multivorum.* Arch Microbiol, 2000. **173**(5-6): p. 418-24.
51. Chattopadhyay, M., et al., *Carotenoid pigments of an antarctic psychrotrophic bacteriumMicrococcus roseus: temperature dependent biosynthesis, structure, and interaction with synthetic membranes.* Biochemical and biophysical research communications, 1997. **239**(1): p. 85-90.
52. Gruszecki, W.I. and K. Strzałka, *Carotenoids as modulators of lipid membrane physical properties.* Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease, 2005. **1740**(2): p. 108-115.
53. Chattopadhyay, M.K., et al., *Carotenoid pigments of an antarctic psychrotrophic bacterium Micrococcus roseus: temperature dependent biosynthesis, structure, and interaction with synthetic membranes.* Biochem Biophys Res Commun, 1997. **239**(1): p. 85-90.
54. Guan, Z., et al., *The polar lipids of Clostridium psychrophilum, an anaerobic psychrophile.* Biochim Biophys Acta, 2013. **1831**(6): p. 1108-12.
55. Spring, S., et al., *Characterization of novel psychrophilic clostridia from an Antarctic microbial mat: description of Clostridium frigoris sp. nov., Clostridium lacusfryxellense sp. nov., Clostridium bowmanii sp. nov. and Clostridium psychrophilum sp. nov. and reclassification of Clostridium laramiense as Clostridium estertheticum subsp. laramiense subsp. nov.* Int J Syst Evol Microbiol, 2003. **53**(Pt 4): p. 1019-29.

56. Huston, A.L., B. Methe, and J.W. Deming, *Purification, characterization, and sequencing of an extracellular cold-active aminopeptidase produced by marine psychrophile Colwellia psychrerythraea strain 34H.* Appl Environ Microbiol, 2004. **70**(6): p. 3321-8.
57. Wan, X., et al., *Effect of cerulenin on fatty acid composition and gene expression pattern of DHA-producing strain Colwellia psychrerythraea strain 34H.* Microb Cell Fact, 2016. **15**(1): p. 30.
58. Auman, A.J., et al., *Psychromonas ingrahamii sp. nov., a novel gas vacuolate, psychrophilic bacterium isolated from Arctic polar sea ice.* Int J Syst Evol Microbiol, 2006. **56**(Pt 5): p. 1001-7.
59. Breezee, J., N. Cady, and J.T. Staley, *Subfreezing growth of the sea ice bacterium "Psychromonas ingrahamii".* Microb Ecol, 2004. **47**(3): p. 300-4.
60. Mykytczuk, N.C.S., et al., *Bacterial growth at −15 °C; molecular insights from the permafrost bacterium Planococcus halocryophilus Or1.* The ISME Journal, 2013. **7**(6): p. 1211-1226.
61. D'Amico, S., et al., *Psychrophilic microorganisms: challenges for life.* EMBO Reports, 2006. **7**(4): p. 385-389.
62. Preston, C.M., et al., *A psychrophilic crenarchaeon inhabits a marine sponge: Cenarchaeum symbiosum gen. nov., sp. nov.* Proceedings of the National Academy of Sciences, 1996. **93**(13): p. 6241-6246.
63. Franzmann, P.D., et al., *Methanogenium frigidum sp. nov., a psychrophilic, H2-using methanogen from Ace Lake, Antarctica.* International Journal of Systematic and Evolutionary Microbiology, 1997. **47**(4): p. 1068-1072.
64. Nichols, D.S., et al., *Cold adaptation in the Antarctic Archaeon Methanococcoides burtonii involves membrane lipid unsaturation.* J Bacteriol, 2004. **186**(24): p. 8508-15.
65. Gibson, J.A., et al., *Unsaturated diether lipids in the psychrotrophic archaeon Halorubrum lacusprofundi.* Syst Appl Microbiol, 2005. **28**(1): p. 19-26.
66. Sprott, G.D., B.J. Agnew, and G.B. Patel, *Structural features of ether lipids in the archaeobacterial thermophiles Pyrococcus furiosus, Methanopyrus kandleri, Methanothermus fervidus, and Sulfolobus acidocaldarius.* Canadian Journal of Microbiology, 1997. **43**(5): p. 467-476.
67. Gilmore, S.F., et al., *Role of squalene in the organization of monolayers derived from lipid extracts of Halobacterium salinarum.* Langmuir, 2013. **29**(25): p. 7922-30.
68. Matsuno, Y., et al., *Effect of growth temperature and growth phase on the lipid composition of the archaeal membrane from Thermococcus kodakaraensis.* Biosci Biotechnol Biochem, 2009. **73**(1): p. 104-8.
69. Lai, D., J.R. Springstead, and H.G. Monbouquette, *Effect of growth temperature on ether lipid biochemistry in Archaeoglobus fulgidus.* Extremophiles, 2008. **12**(2): p. 271-8.
70. Uda, I., et al., *Variation in molecular species of polar lipids from thermoplasma acidophilum depends on growth temperature.* Lipids, 2001. **36**(1): p. 103-5.
71. Oger, P.M. and A. Cario, *Adaptation of the membrane in Archaea.* Biophys Chem, 2013. **183**: p. 42-56.
72. Franzmann, P.D., et al., *Methanogenium frigidum sp. nov., a psychrophilic, H2-using methanogen from Ace Lake, Antarctica.* Int J Syst Bacteriol, 1997. **47**(4): p. 1068-72.
73. Blocher, D., et al., *Physicochemical characterization of tetraether lipids from Thermoplasma acidophilum. V. Evidence for the existence of a metastable state in lipids with acyclic hydrocarbon chains.* Biochim Biophys Acta, 1990. **1024**(1): p. 54-60.
74. Escriba, P.V., *Membrane-lipid therapy: a new approach in molecular medicine.* Trends Mol Med, 2006. **12**(1): p. 34-43.
75. Sinensky, M., *Temperature control of phospholipid biosynthesis in Escherichia coli.* J Bacteriol, 1971. **106**(2): p. 449-55.
76. Patel, B.K.C., J.H. Skerratt, and P.D. Nichols, *The Phospholipid Ester-linked Fatty Acid Composition of Thermophilic Bacteria.* Systematic and Applied Microbiology, 1991. **14**(4): p. 311-316.
77. Ray, P.H., D.C. White, and T.D. Brock, *Effect of growth temperature on the lipid composition of Thermus aquaticus.* J Bacteriol, 1971. **108**(1): p. 227-35.
78. Yokoyama, A., et al., *Thermocryptoxanthins: novel intermediates in the carotenoid biosynthetic pathway of Thermus thermophilus.* Arch Microbiol, 1996. **165**(5): p. 342-5.
79. Langworthy, T.A., et al., *Iso- and Anteiso-Branched Glycerol Diethers of the Thermophilic Anaerobe Thermodesulfotobacterium commune.* Syst Appl Microbiol, 1983. **4**(1): p. 1-17.
80. Jahnke, L.L., et al., *Signature lipids and stable carbon isotope analyses of Octopus Spring hyperthermophilic communities compared with those of Aquificales representatives.* Appl Environ Microbiol, 2001. **67**(11): p. 5179-89.
81. Jung, S., J.G. Zeikus, and R.I. Hollingsworth, *A new family of very long chain alpha,omega-dicarboxylic acids is a major structural fatty acyl component of the membrane lipids of Thermoanaerobacter ethanolicus 39E.* Journal of Lipid Research, 1994. **35**(6): p. 1057-65.
82. Sanghoo Lee, S.K., Jai Neung Kim, and Seunho Jung, *Structural analyses of the novel phosphoglycolipids containing the unusual very long bifunctional acyl chain, α,ω-13-16-dimethyloctacosanedioate in Thermoanaerobacter ethanolicus.* Bull. Korean Chem. Soc. , 2002. **23**(12): p. 1778-1784.

83. Russell, N.J. and N. Fukunaga, *A comparison of thermal adaptation of membrane lipids in psychrophilic and thermophilic bacteria.* FEMS Microbiology Letters, 1990. **75**(2): p. 171-182.
84. Huber, R., et al., *Thermosipho africanus gen. nov., Represents a New Genus of Thermophilic Eubacteria within the "Thermotogales".* Systematic and Applied Microbiology, 1989. **12**(1): p. 32-37.
85. Kashefi, K., et al., *Use of Fe (III) as an electron acceptor to recover previously uncultured hyperthermophiles: isolation and characterization of Geothermobacterium ferrireducens gen. nov., sp. nov.* Applied and Environmental Microbiology, 2002. **68**(4): p. 1735-1742.
86. Takai, K., et al., *Cell proliferation at 122 C and isotopically heavy CH4 production by a hyperthermophilic methanogen under high-pressure cultivation.* Proceedings of the National Academy of Sciences, 2008. **105**(31): p. 10949-10954.
87. Cario, A., et al., *Membrane homeoviscous adaptation in the piezo-hyperthermophilic archaeon Thermococcus barophilus.* Front Microbiol, 2015. **6**: p. 1152.
88. De Rosa, M., et al., *Lipids of Thermococcus celer, a sulfur-reducing archaebacterium: Structure and biosynthesis.* Systematic and Applied Microbiology, 1987. **9**(1–2): p. 1-5.
89. Hafenbradl, D., M. Keller, and K.O. Stetter, *Lipid analysis of Methanopyrus kandleri.* FEMS Microbiology Letters, 1996. **136**(2): p. 199-202.
90. Ulrih, N.P., D. Gmajner, and P. Raspor, *Structural and physicochemical properties of polar lipids from thermophilic archaea.* Appl Microbiol Biotechnol, 2009. **84**(2): p. 249-60.
91. De Rosa, M., et al., *Effects of temperature on ether lipid composition of Caldariella acidophila.* Phytochemistry, 1980. **19**(5): p. 827-831.
92. Dannenmuller, O., et al., *Membrane properties of archaeal macrocyclic diether phospholipids.* Chemistry, 2000. **6**(4): p. 645-54.
93. Mathai, J.C., G.D. Sprott, and M.L. Zeidel, *Molecular mechanisms of water and solute transport across archaebacterial lipid membranes.* J Biol Chem, 2001. **276**(29): p. 27266-71.
94. Oger, P.M. and M. Jebbar, *The many ways of coping with pressure.* Res Microbiol, 2010. **161**(10): p. 799-809.
95. Casadei, M.A., et al., *Role of membrane fluidity in pressure resistance of Escherichia coli NCTC 8164.* Appl Environ Microbiol, 2002. **68**(12): p. 5965-72.
96. Zeng, X., et al., *Pyrococcus CH1, an obligate piezophilic hyperthermophile: extending the upper pressure-temperature limits for life.* ISME J, 2009. **3**(7): p. 873-876.
97. Yayanos, A.A., *Evolutional and ecological implications of the properties of deep-sea barophilic bacteria.* Proceedings of the National Academy of Sciences, 1986. **83**(24): p. 9542-9546.
98. Yayanos, A.A., A.S. Dietz, and R. Van Boxtel, *Obligately barophilic bacterium from the Mariana Trench.* Proceedings of the National Academy of Sciences, 1981. **78**(8): p. 5212-5215.
99. Zeng, X., et al., *Pyrococcus CH1, an obligate piezophilic hyperthermophile: extending the upper pressure-temperature limits for life.* The ISME journal, 2009. **3**(7): p. 873-876.
100. Bartlett, D.H., *Microbial Life at High-Pressures.* Science Progress, 1992. **76**(301-02): p. 479-496.
101. Carl O. Wirsen, H.W.J., Stuart G. Wakeham, and Elizabeth A. Canuel, *Membrane lipids of a psychrophilic and barophilic deep-sea bacterium.* Current Microbiology, 1987. **14**(6): p. 319-322.
102. Allen, E.E., D. Facciotti, and D.H. Bartlett, *Monounsaturated but Not Polyunsaturated Fatty Acids Are Required for Growth of the Deep-Sea Bacterium Photobacterium profundum SS9 at High Pressure and Low Temperature.* Applied and Environmental Microbiology, 1999. **65**(4): p. 1710-1720.
103. Yano, Y., et al., *Adaptive changes in membrane lipids of barophilic bacteria in response to changes in growth pressure.* Appl Environ Microbiol, 1998. **64**(2): p. 479-85.
104. Mangelsdorf, K., et al., *A quantitative assessment of pressure dependent adaptive changes in the membrane lipids of a piezosensitive deep sub-seafloor bacterium.* Organic Geochemistry, 2005. **36**(11): p. 1459-1479.
105. Jebbar, M., et al., *Microbial diversity and adaptation to high hydrostatic pressure in deep-sea hydrothermal vents prokaryotes.* Extremophiles, 2015. **19**(4): p. 721-40.
106. Kaneshiro, S.M. and D.S. Clark, *Pressure effects on the composition and thermal behavior of lipids from the deep-sea thermophile Methanococcus jannaschii.* J Bacteriol, 1995. **177**(13): p. 3668-72.
107. Stephanopoulos, G., M.E. Garefalaki, and K. Lyroudia, *Genes and related proteins involved in amelogenesis imperfecta.* J Dent Res, 2005. **84**(12): p. 1117-26.
108. Yuk, H.G. and D.L. Marshall, *Adaptation of Escherichia coli O157:H7 to pH alters membrane lipid composition, verotoxin secretion, and resistance to simulated gastric fluid acid.* Appl Environ Microbiol, 2004. **70**(6): p. 3500-5.
109. Fozo, E.M. and R.G. Quivey, Jr., *Shifts in the membrane fatty acid profile of Streptococcus mutans enhance survival in acidic environments.* Appl Environ Microbiol, 2004. **70**(2): p. 929-36.
110. Kim, B.H., et al., *The formation of cyclopropane fatty acids in Salmonella enterica serovar Typhimurium.* Microbiology, 2005. **151**(Pt 1): p. 209-18.

111. Giotis, E.S., et al., *Role of branched-chain fatty acids in pH stress tolerance in Listeria monocytogenes.* Appl Environ Microbiol, 2007. **73**(3): p. 997-1001.
112. Baker-Austin, C. and M. Dopson, *Life in acid: pH homeostasis in acidophiles.* Trends Microbiol, 2007. **15**(4): p. 165-71.
113. De Rosa, M., A. Gambacorta, and J.D. Bu'lock, *Effects of pH and temperature on the fatty acid composition of bacillus acidocaldarius.* J Bacteriol, 1974. **117**(1): p. 212-4.
114. Wichlacz, P.L., R.F. Unz, and T.A. Langworthy, *Acidiphilium angustum sp. nov., Acidiphilium facilis sp. nov., and Acidiphilium rubrum sp. nov.: Acidophilic Heterotrophic Bacteria Isolated from Acidic Coal Mine Drainage.* International Journal of Systematic and Evolutionary Microbiology, 1986. **36**(2): p. 197-201.
115. Wakao, N., et al., *ACIDIPHILIUM MULTIVORUM SP. NOV., AN ACIDOPHILIC CHEMOORGANOTROPHIC BACTERIUM FROM PYRITIC ACID MINE DRAINAGE.* The Journal of General and Applied Microbiology, 1994. **40**(2): p. 143-159.
116. Matsubara, H., et al., *Alicyclobacillus acidiphilus sp. nov., a novel thermo-acidophilic, omega-alicyclic fatty acid-containing bacterium isolated from acidic beverages.* Int J Syst Evol Microbiol, 2002. **52**(Pt 5): p. 1681-5.
117. Mykytczuk, N.C., et al., *Cytoplasmic membrane fluidity and fatty acid composition of Acidithiobacillus ferrooxidans in response to pH stress.* Extremophiles, 2010. **14**(5): p. 427-41.
118. Weijers, J.W., et al., *Membrane lipids of mesophilic anaerobic bacteria thriving in peats have typical archaeal traits.* Environ Microbiol, 2006. **8**(4): p. 648-57.
119. Damste, J.S., et al., *13,16-Dimethyl octacosanedioic acid (iso-diabolic acid), a common membrane-spanning lipid of Acidobacteria subdivisions 1 and 3.* Appl Environ Microbiol, 2011. **77**(12): p. 4147-54.
120. Macalady, J.L., et al., *Tetraether-linked membrane monolayers in Ferroplasma spp: a key to survival in acid.* Extremophiles, 2004. **8**(5): p. 411-9.
121. Elferink, M.G., et al., *Stability and proton-permeability of liposomes composed of archaeal tetraether lipids.* Biochim Biophys Acta, 1994. **1193**(2): p. 247-54.
122. Schleper, C., et al., *Picrophilus gen. nov., fam. nov.: a novel aerobic, heterotrophic, thermoacidophilic genus and family comprising archaea capable of growth around pH 0.* J Bacteriol, 1995. **177**(24): p. 7050-9.
123. Gabriel, J.L. and P.L. Chong, *Molecular modeling of archaebacterial bipolar tetraether lipid membranes.* Chem Phys Lipids, 2000. **105**(2): p. 193-200.
124. Shimada, H., et al., *Effects of pH and temperature on the composition of polar lipids in Thermoplasma acidophilum HO-62.* J Bacteriol, 2008. **190**(15): p. 5404-11.
125. Boyd, E.S., et al., *Temperature and pH controls on glycerol dibiphytanyl glycerol tetraether lipid composition in the hyperthermophilic crenarchaeon Acidilobus sulfurireducens.* Extremophiles, 2011. **15**(1): p. 59-65.
126. Krulwich, T.A., *Alkaliphilic Prokaryotes.* The Prokaryotes, 2006(2): p. 283-308.
127. Ye, Q., et al., *Alkaline anaerobic respiration: isolation and characterization of a novel alkaliphilic and metal-reducing bacterium.* Appl Environ Microbiol, 2004. **70**(9): p. 5595-602.
128. Rainey, F.A., et al., *The genus Nocardiopsis represents a phylogenetically coherent taxon and a distinct actinomycete lineage: proposal of Nocardiopsaceae fam. nov.* Int J Syst Bacteriol, 1996. **46**(4): p. 1088-92.
129. Li, Y., et al., *Clostridium thermoalcaliphilum sp. nov., an anaerobic and thermotolerant facultative alkaliphile.* Int J Syst Bacteriol, 1994. **44**(1): p. 111-8.
130. Clejan, S., et al., *Membrane lipid composition of obligately and facultatively alkalophilic strains of Bacillus spp.* J Bacteriol, 1986. **168**(1): p. 334-40.
131. Janto, B., et al., *Genome of alkaliphilic Bacillus pseudofirmus OF4 reveals adaptations that support the ability to grow in an external pH range from 7.5 to 11.4.* Environmental microbiology, 2011. **13**(12): p. 3289-3309.
132. Clejan, S., et al., *Membrane lipid composition of obligately and facultatively alkalophilic strains of Bacillus spp.* Journal of bacteriology, 1986. **168**(1): p. 334-340.
133. Hauß, T., et al., *Squalane is in the midplane of the lipid bilayer: implications for its function as a proton permeability barrier.* Biochimica et Biophysica Acta (BBA) - Bioenergetics, 2002. **1556**(2–3): p. 149-154.
134. Nogi, Y. and C. Kato, *Taxonomic studies of extremely barophilic bacteria isolated from the Mariana Trench and description of Moritella yayanosii sp. nov., a new barophilic bacterial isolate.* Extremophiles. **3**(1): p. 71-77.
135. Knoblauch, C., K. Sahm, and B.B. Jorgensen, *Psychrophilic sulfate-reducing bacteria isolated from permanently cold arctic marine sediments: description of Desulfofrigus oceanense gen. nov., sp. nov., Desulfofrigus fragile sp. nov., Desulfofaba gelida gen. nov., sp. nov., Desulfotalea psychrophila gen. nov., sp. nov. and Desulfotalea arctica sp. nov.* Int J Syst Bacteriol, 1999. **49 Pt 4**: p. 1631-43.
136. Heinen, W., H.P. Klein, and C.M. Volkmann, *Fatty acid composition of Thermus aquaticus at different growth temperatures.* Arch Mikrobiol, 1970. **72**(2): p. 199-202.
137. LI, Y., et al., *Clostridium thermoalcaliphilum sp. nov., an Anaerobic and Thermotolerant Facultative Alkaliphile.* International Journal of Systematic and Evolutionary Microbiology, 1994. **44**(1): p. 111-118.

138. Prowe, S.G. and G. Antranikian, *Anaerobranca gottschalkii sp. nov., a novel thermoalkaliphilic bacterium that grows anaerobically at high pH and temperature.* International Journal of Systematic and Evolutionary Microbiology, 2001. **51**(2): p. 457-465.
139. Clejan, S., et al., *Membrane lipid composition of obligately and facultatively alkalophilic strains of Bacillus spp.* Journal of Bacteriology, 1986. **168**(1): p. 334-340.
140. Balk, M., et al., *Isolation and characterization of a new CO-utilizing strain, Thermoanaerobacter thermohydrosulfuricus subsp. carboxydovorans, isolated from a geothermal spring in Turkey.* Extremophiles, 2009. **13**(6): p. 885-894.
141. Lee, S., et al., *Structural analyses of the novel phosphoglycolipids containing the unusual very long bifunctional acyl chain, alpha,omega-13,16-dimethyloctacosanedioate in Thermoanaerobacter ethanolicus.* 2002.
142. Namwong, S., et al., *Halococcus thailandensis sp. nov., from fish sauce in Thailand.* International Journal of Systematic and Evolutionary Microbiology, 2007. **57**(10): p. 2199-2203.
143. Feng, J., et al., *Halorubrum alkaliphilum sp. nov., a novel haloalkaliphile isolated from a soda lake in Xinjiang, China.* International Journal of Systematic and Evolutionary Microbiology, 2005. **55**(1): p. 149-152.
144. Castillo, A.M., et al., *Halorubrum orientale sp. nov., a halophilic archaeon isolated from Lake Ejinor, Inner Mongolia, China.* International Journal of Systematic and Evolutionary Microbiology, 2006. **56**(11): p. 2559-2563.
145. Xu, Y., et al., *Natrialba hulunbeirensis sp. nov. and Natrialba chahannaoensis sp. nov., novel haloalkaliphilic archaea from soda lakes in Inner Mongolia Autonomous Region, China.* International Journal of Systematic and Evolutionary Microbiology, 2001. **51**(5): p. 1693-1698.
146. Xu, Y., P. Zhou, and X. Tian, *Characterization of two novel haloalkaliphilic archaea Natronorubrum bangense gen. nov., sp. nov. and Natronorubrum tibetense gen. nov., sp. nov.* International Journal of Systematic and Evolutionary Microbiology, 1999. **49**(1): p. 261-266.
147. Lanzotti, V., et al., *A complex lipid with a cyclic phosphate from the archaebacterium Natronococcus occultus.* Biochimica et Biophysica Acta (BBA) - Lipids and Lipid Metabolism, 1989. **1001**(1): p. 31-34.
148. Tindall, B.J., H.N.M. Ross, and W.D. Grant, *Natronobacterium gen. nov. and Natronococcus gen. nov., Two New Genera of Haloalkaliphilic Archaebacteria.* Systematic and Applied Microbiology, 1984. **5**(1): p. 41-57.
149. Grant, W.D., et al., *Polar Lipids in Methanogen Taxonomy.* Microbiology, 1985. **131**(12): p. 3277-3286.
150. Marteinsson, V.T., et al., *Thermococcus barophilus sp. nov., a new barophilic and hyperthermophilic archaeon isolated under high hydrostatic pressure from a deep-sea hydrothermal vent.* International Journal of Systematic and Evolutionary Microbiology, 1999. **49**(2): p. 351-359.
151. Takai, K., et al., *Palaeococcus ferrophilus gen. nov., sp. nov., a barophilic, hyperthermophilic archaeon from a deep-sea hydrothermal vent chimney.* International Journal of Systematic and Evolutionary Microbiology, 2000. **50**(2): p. 489-500.
152. Sako, Y., et al., *Aeropyrum pernix gen. nov., sp. nov., a novel aerobic hyperthermophilic archaeon growing at temperatures up to 100 degrees C.* Int J Syst Bacteriol, 1996. **46**(4): p. 1070-7.
153. Sprott, G.D., M. Meloche, and J.C. Richards, *Proportions of diether, macrocyclic diether, and tetraether lipids in Methanococcus jannaschii grown at different temperatures.* J Bacteriol, 1991. **173**(12): p. 3907-10.
154. Hu, L., et al., *Halorubrum luteum sp. nov., isolated from Lake Chagannor, Inner Mongolia, China.* International Journal of Systematic and Evolutionary Microbiology, 2008. **58**(7): p. 1705-1708.
155. Bowers, K.J. and J. Wiegel, *Temperature and pH optima of extremely halophilic archaea: a mini-review.* Extremophiles, 2011. **15**(2): p. 119-128.
156. Yamauchi, K., et al., *Archaebacterial lipids: highly proton-impermeable membranes from 1,2-diphytanyl-sn-glycero-3-phosphocholine.* Biochim Biophys Acta, 1993. **1146**(2): p. 178-82.
157. van de Vossenberg, M.J.L.C., M.A.J. Driessen, and N.W. Konings, *The essence of being extremophilic: the role of the unique archaeal membrane lipids.* Extremophiles, 1998. **2**(3): p. 163-170.
158. BOWMAN, J.P., et al., *Novel Psychrobacter Species from Antarctic Ornithogenic Soils.* International Journal of Systematic and Evolutionary Microbiology, 1996. **46**(4): p. 841-848.
159. Wan, X., et al., *Effect of cerulenin on fatty acid composition and gene expression pattern of DHA-producing strain Colwellia psychrerythraea strain 34H.* Microbial Cell Factories, 2016. **15**(1): p. 1-13.
160. Auman, A.J., et al., *Psychromonas ingrahamii sp. nov., a novel gas vacuolate, psychrophilic bacterium isolated from Arctic polar sea ice.* International Journal of Systematic and Evolutionary Microbiology, 2006. **56**(5): p. 1001-1007.
161. Breezee, J., N. Cady, and T.J. Staley, *Subfreezing Growth of the Sea Ice Bacterium "Psychromonas ingrahamii".* Microbial Ecology, 2004. **47**(3): p. 300-304.
162. Girard, A.E., *A comparative study of the fatty acids of some micrococci.* Canadian Journal of Microbiology, 1971. **17**(12): p. 1503-1508.
163. Shivaji, S., et al., *Isolation and identification ofMicrococcus roseus andPlanococcus sp. from schirmacher oasis, antarctica.* Journal of Biosciences. **13**(4): p. 409-414.

164. Jagannadham, V.M., et al., *Carotenoids of an Antarctic psychrotolerant bacterium, Sphingobacterium antarcticus, and a mesophilic bacterium, Sphingobacterium multivorum.* Archives of Microbiology, 2000. **173**(5): p. 418-424.
165. Drucker, D.B., *Chemotaxonomic fatty-acid fingerprints of some streptococci with subsequent statistical analysis.* Canadian Journal of Microbiology, 1974. **20**(12): p. 1723-1728.
166. Schleifer, K.H. and R. Kilpper-Bälz, *Transfer of Streptococcus faecalis and Streptococcus faecium to the Genus Enterococcus nom. rev. as Enterococcus faecalis comb. nov. and Enterococcus faecium comb. nov.* International Journal of Systematic and Evolutionary Microbiology, 1984. **34**(1): p. 31-34.
167. ELSDEN, S.R., et al., *The Lipid Fatty Acids of Proteolytic Clostridia.* Microbiology, 1980. **118**(1): p. 115-123.
168. Jackson, S., et al., *Analysis of proline reduction in the nosocomial pathogen Clostridium difficile.* J Bacteriol, 2006. **188**(24): p. 8487-95.
169. Moule, A.L. and S.G. Wilkinson, *Polar Lipids, Fatty Acids, and Isoprenoid Quinones of Alteromonas putrefaciens (Shewanella putrefaciens).* Systematic and Applied Microbiology, 1987. **9**(3): p. 192-198.
170. Pivnick, H., *PSEUDOMONAS RUBESCENS, A NEW SPECIES FROM SOLUBLE OIL EMULSIONS.* Journal of Bacteriology, 1955. **70**(1): p. 1-6.
171. Dees, S.B., et al., *Chemical and Phenotypic Characteristics of Flavobacterium thalpophilum Compared with Those of Other Flavobacterium and Sphingobacterium Species.* International Journal of Systematic and Evolutionary Microbiology, 1985. **35**(1): p. 16-22.
172. HOLMES, B., et al., *Flavobacterium thalpophilum, a New Species Recovered from Human Clinical Material.* International Journal of Systematic and Evolutionary Microbiology, 1983. **33**(4): p. 677-682.
173. Kaneda, T., *Iso- and anteiso-fatty acids in bacteria: biosynthesis, function, and taxonomic significance.* Microbiol Rev, 1991. **55**(2): p. 288-302.
174. Nonomura, H. and Y. Ohara, *Distribution of the actinomycetes in soil. IV. The isolation and classification of the genus Streptosporangium.* J Ferment Technol, 1960. **38**: p. 405-409.
175. Ueki, A. and T. Suto, *CELLULAR FATTY ACID COMPOSITION OF SULFATE-REDUCING BACTERIA.* The Journal of General and Applied Microbiology, 1979. **25**(3): p. 185-196.
176. Campbell, L.L. and J.R. Postgate, *Classification of the spore-forming sulfate-reducing bacteria.* Bacteriological Reviews, 1965. **29**(3): p. 359-363.
177. Roger, P., et al., *Characterization of Streptococcus salivarius growth and maintenance in artificial saliva.* J Appl Microbiol, 2011. **111**(3): p. 631-41.
178. McElhaney, R.N. and K.A. Souza, *The relationship between environmental temperature, cell growth and the fluidity and physical state of the membrane lipids in Bacillus stearothermophilus.* Biochimica et Biophysica Acta (BBA) - Nucleic Acids and Protein Synthesis, 1976. **443**(3): p. 348-359.
179. apos, et al., *Numerical Analysis of Fatty Acid Profiles in the Identification of Staphylococci.* Microbiology, 1985. **131**(8): p. 2023-2033.
180. Kloos, W.E. and K.H. Schleifer, *Isolation and Characterization of Staphylococci from Human Skin II. Descriptions of Four New Species: Staphylococcus warneri, Staphylococcus capitis, Staphylococcus hominis, and Staphylococcus simulans1.* International Journal of Systematic and Evolutionary Microbiology, 1975. **25**(1): p. 62-79.
181. Tsu, I.-H., et al., *Isolation and characterization of Desulfovibrio senezii sp. nov., a halotolerant sulfate reducer from a solar saltern and phylogenetic confirmation of Desulfovibrio fructosovorans as a new species.* Archives of Microbiology, 1998. **170**(4): p. 313-317.
182. Moss, C.W., et al., *Cultural Characteristics and Fatty Acid Composition of Propionibacteria.* Journal of Bacteriology, 1969. **97**(2): p. 561-570.
183. DARLAND, G. and T.D. BROCK, *Bacillus acidocaldarius sp.nov., an Acidophilic Thermophilic Spore-forming Bacterium.* Microbiology, 1971. **67**(1): p. 9-15.
184. Loginova, L.G., et al., *Thermus ruber sp. nov., nom. rev.* International Journal of Systematic and Evolutionary Microbiology, 1984. **34**(4): p. 498-499.
185. Henssen, A. and E. Schnepf, *Studies on thermophilic actinomycetes.* Archiv für Mikrobiologie, 1967. **57**(3): p. 214-231.
186. Thies, E., T. Jenkins, and F. Stutzenberger, *Effects of the detergent Tween 80 on Thermomonospora curvata.* World Journal of Microbiology and Biotechnology, 1994. **10**(6): p. 657-663.
187. Kaneda, T., *Fatty acids of the genus Bacillus: an example of branched-chain preference.* Bacteriological Reviews, 1977. **41**(2): p. 391-418.
188. Herrero, A.A., R.F. Gomez, and M.F. Roberts, *Ethanol-induced changes in the membrane lipid composition of Clostridium thermocellum.* Biochimica et Biophysica Acta (BBA) - Biomembranes, 1982. **693**(1): p. 195-204.
189. Freier, D., C.P. Mothershed, and J. Wiegel, *Characterization of Clostridium thermocellum JW20.* Appl Environ Microbiol, 1988. **54**(1): p. 204-211.

190. Taylor, J. and R.J. Parkes, *The Cellular Fatty Acids of the Sulphate-reducing Bacteria, Desulfobacter sp., Desulfobulbus sp. and Desulfovibrio desulfuricans.* Microbiology, 1983. **129**(11): p. 3303-3309.
191. Kaneda, T., *Fatty acids in the genus Bacillus. II. Similarity in the fatty acid compositions of Bacillus thuringiensis, Bacillus anthracis, and Bacillus cereus.* J Bacteriol, 1968. **95**(6): p. 2210-6.
192. Koehler, T.M., *Bacillus anthracis physiology and genetics.* Molecular Aspects of Medicine, 2009. **30**(6): p. 386-396.
193. Suzuki, K.-i., et al., *Chemotaxonomic characterization of a radiotolerant bacterium, Arthrobacter radiotolerans: Description of Rubrobacter radiotolerans gen. nov., comb. nov.* FEMS Microbiology Letters, 1988. **52**(1-2): p. 33-39.
194. Yoshinaka, T., K. Yano, and H. Yamaguchi, *Isolation of Highly Radioresistant Bacterium, Arthrobacter radiotolerans nov. sp.* Agricultural and Biological Chemistry, 1973. **37**(10): p. 2269-2275.
195. Langworthy, T.A., et al., *Iso- and Anteiso-Branched Glycerol Diethers of the Thermophilic Anaerobe Thermodesulfotobacterium commune.* Systematic and Applied Microbiology, 1983. **4**(1): p. 1-17.
196. Jantzen, E., T. Bergan, and K. Bøvre, *GAS CHROMATOGRAPHY OF BACTERIAL WHOLE CELL METHANOLYSATES.* Acta Pathologica Microbiologica Scandinavica Section B Microbiology and Immunology, 1974. **82**(6): p. 785-798.
197. Chan, K. and O. Leung, *Nutrition and growth of the moderately halophilic bacteria Micrococcus morrhuae K-17 and Micrococcus luteus K-15.* Microbios, 1978. **25**(100): p. 71-84.
198. Su, Y., et al., *Physiological and fermentation properties of Bacillus coagulans and a mutant lacking fermentative lactate dehydrogenase activity.* Journal of industrial microbiology & biotechnology, 2011. **38**(3): p. 441-450.
199. Jackson, T.J., R.F. Ramaley, and W.G. Meinschein, *Fatty acids of a non-pigmented, thermophilic bacterium similar to Thermus aquaticus.* Archiv für Mikrobiologie, 1972. **88**(2): p. 127-133.
200. Brock, T.D. and H. Freeze, *Thermus aquaticus gen. n. and sp. n., a nonsporulating extreme thermophile.* Journal of bacteriology, 1969. **98**(1): p. 289-297.
201. Kaneda, T., E.J. Smith, and D.N. Naik, *Fatty acid composition and primer specificity of de novo fatty acid synthetase in Bacillus globispores, Bacillus insolitus, and Bacillus psychrophilus.* Canadian journal of microbiology, 1983. **29**(12): p. 1634-1641.
202. Nakamura, L., *Bacillus psychrophilus sp. nov., nom. rev.* International Journal of Systematic and Evolutionary Microbiology, 1984. **34**(2): p. 121-123.
203. Balk, M., et al., *Isolation and characterization of a new CO-utilizing strain, Thermoanaerobacter thermohydrosulfuricus subsp. carboxydovorans, isolated from a geothermal spring in Turkey.* Extremophiles, 2009. **13**(6): p. 885-894.
204. Oshima, M. and A. Miyagawa, *Comparative studies on the fatty acid composition of moderately and extremely thermophilic bacteria.* Lipids, 1974. **9**(7): p. 476-480.
205. Oshima, T. and K. IMAHORI, *Description of Thermus thermophilus (Yoshida and Oshima) comb. nov., a nonsporulating thermophilic bacterium from a Japanese thermal spa.* International Journal of Systematic and Evolutionary Microbiology, 1974. **24**(1): p. 102-112.
206. Ro, D.-K., et al., *Production of the antimalarial drug precursor artemisinic acid in engineered yeast.* Nature, 2006. **440**(7086): p. 940-943.
207. Ajikumar, P.K., et al., *Isoprenoid pathway optimization for Taxol precursor overproduction in Escherichia coli.* Science, 2010. **330**(6000): p. 70-4.
208. Reiling, K.K., et al., *Mono and diterpene production in Escherichia coli.* Biotechnology and bioengineering, 2004. **87**(2): p. 200-212.
209. Yuan, L.Z., et al., *Chromosomal promoter replacement of the isoprenoid pathway for enhancing carotenoid production in E. coli.* Metab Eng, 2006. **8**(1): p. 79-90.
210. Zhou, K., et al., *Metabolite profiling identified methylerythritol cyclodiphosphate efflux as a limiting step in microbial isoprenoid production.* PLoS One, 2012. **7**(11): p. e47513.
211. Bitok, J.K. and C.F. Meyers, *2C-Methyl-d-erythritol 4-phosphate enhances and sustains cyclodiphosphate synthase IspF activity.* ACS Chem Biol, 2012. **7**(10): p. 1702-10.
212. Kim, S.W. and J.D. Keasling, *Metabolic engineering of the nonmevalonate isopentenyl diphosphate synthesis pathway in Escherichia coli enhances lycopene production.* Biotechnol Bioeng, 2001. **72**(4): p. 408-15.
213. Zhang, C., et al., *Combining genotype improvement and statistical media optimization for isoprenoid production in E. coli.* PLoS One, 2013. **8**(10): p. e75164.
214. Immethun, C.M., et al., *Microbial production of isoprenoids enabled by synthetic biology.* Front Microbiol, 2013. **4**: p. 75.
215. Matsumi, R., et al., *Isoprenoid biosynthesis in Archaea–biochemical and evolutionary implications.* Research in Microbiology, 2011. **162**(1): p. 39-52.
216. Rohmer, M., et al., *Isoprenoid biosynthesis in bacteria: a novel pathway for the early steps leading to isopentenyl diphosphate.* Biochem J, 1993. **295 (Pt 2)**: p. 517-24.
217. Hunter, W.N., *The non-mevalonate pathway of isoprenoid precursor biosynthesis.* J Biol Chem, 2007. **282**(30): p. 21573-7.

218. Fujisaki, S., et al., *Disruption of the structural gene for farnesyl diphosphate synthase in Escherichia coli.* J Biochem, 2005. **137**(3): p. 395-400.
219. Bouhss, A., et al., *The biosynthesis of peptidoglycan lipid-linked intermediates.* FEMS Microbiol Rev, 2008. **32**(2): p. 208-33.
220. Samuel, G. and P. Reeves, *Biosynthesis of O-antigens: genes and pathways involved in nucleotide sugar precursor synthesis and O-antigen assembly.* Carbohydr Res, 2003. **338**(23): p. 2503-19.
221. Willenbrock, H. and D.W. Ussery, *Chromatin architecture and gene expression in Escherichia coli.* Genome Biol, 2004. **5**(12): p. 252.
222. Martinez-Antonio, A., A. Medina-Rivera, and J. Collado-Vides, *Structural and functional map of a bacterial nucleoid.* Genome Biol, 2009. **10**(12): p. 247.
223. Bryant, J.A., et al., *Chromosome position effects on gene expression in Escherichia coli K-12.* Nucleic Acids Res, 2014. **42**(18): p. 11383-92.
224. Dame, R.T., O.J. Kalmykowa, and D.C. Grainger, *Chromosomal macrodomains and associated proteins: implications for DNA organization and replication in gram negative bacteria.* PLoS Genet, 2011. **7**(6): p. e1002123.
225. Sousa, C., V. de Lorenzo, and A. Cebolla, *Modulation of gene expression through chromosomal positioning in Escherichia coli.* Microbiology, 1997. **143**(6): p. 2071-2078.
226. Postow, L., et al., *Topological domain structure of the Escherichia coli chromosome.* Genes Dev, 2004. **18**(14): p. 1766-79.
227. Fritsche, M., et al., *A model for Escherichia coli chromosome packaging supports transcription factor-induced DNA domain formation.* Nucleic Acids Res, 2012. **40**(3): p. 972-80.
228. Vora, T., A.K. Hottes, and S. Tavazoie, *Protein occupancy landscape of a bacterial genome.* Mol Cell, 2009. **35**(2): p. 247-53.
229. Kim, S.-W., et al., *High-level production of lycopene in metabolically engineered E. coli.* Process Biochemistry, 2009. **44**(8): p. 899-905.
230. Schmidt-Dannert, C., D. Umeno, and F.H. Arnold, *Molecular breeding of carotenoid biosynthetic pathways.* Nat Biotechnol, 2000. **18**(7): p. 750-3.
231. Vadali, R.V., et al., *Enhanced lycopene productivity by manipulation of carbon flow to isopentenyl diphosphate in Escherichia coli.* Biotechnol Prog, 2005. **21**(5): p. 1558-61.
232. Albrecht, M., N. Misawa, and G. Sandmann, *Metabolic engineering of the terpenoid biosynthetic pathway of Escherichia coli for production of the carotenoids β-carotene and zeaxanthin.* Biotechnology Letters, 1999. **21**(9): p. 791-795.
233. Vondenhoff, G.H., et al., *Characterization of peptide chain length and constituency requirements for YejABEF-mediated uptake of microcin C analogues.* J Bacteriol, 2011. **193**(14): p. 3618-23.
234. Datsenko, K.A. and B.L. Wanner, *One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products.* Proc Natl Acad Sci U S A, 2000. **97**(12): p. 6640-5.
235. Baba, T., et al., *Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection.* Mol Syst Biol, 2006. **2**: p. 2006 0008.
236. Yang, J. and L. Guo, *Biosynthesis of β-carotene in engineered E. coli using the MEP and MVA pathways.* Microbial cell factories, 2014. **13**(1): p. 1.
237. Martin, V.J., et al., *Engineering a mevalonate pathway in Escherichia coli for production of terpenoids.* Nat Biotechnol, 2003. **21**(7): p. 796-802.
238. Wang, C.W., M.K. Oh, and J.C. Liao, *Engineered isoprenoid pathway enhances astaxanthin production in Escherichia coli.* Biotechnol Bioeng, 1999. **62**(2): p. 235-41.
239. Maniatis, T., E.F. Fritsch, and J. Sambrook, *Molecular cloning: a laboratory manual*. Vol. 545. 1982: Cold Spring Harbor Laboratory Cold Spring Harbor, NY.
240. Satterwhite, D.M., *Isopentenyldiphosphate delta-isomerase.* Methods Enzymol, 1985. **110**: p. 92-9.
241. Lang, B., et al., *High-affinity DNA binding sites for H-NS provide a molecular basis for selective silencing within proteobacterial genomes.* Nucleic Acids Res, 2007. **35**(18): p. 6330-7.
242. Pougach, K., et al., *Transcription, processing and function of CRISPR cassettes in Escherichia coli.* Mol Microbiol, 2010. **77**(6): p. 1367-79.
243. Hengen, P.N., et al., *Information analysis of Fis binding sites.* Nucleic Acids Res, 1997. **25**(24): p. 4994-5002.
244. Goodman, S.D., et al., *In vitro selection of integration host factor binding sites.* J Bacteriol, 1999. **181**(10): p. 3246-55.
245. Hales, L.M., R.I. Gumport, and J.F. Gardner, *Determining the DNA sequence elements required for binding integration host factor to two different target sites.* J Bacteriol, 1994. **176**(10): p. 2999-3006.
246. Zhang, Y.M. and C.O. Rock, *Membrane lipid homeostasis in bacteria.* Nat Rev Microbiol, 2008. **6**(3): p. 222-33.
247. Janssen, H.J. and A. Steinbuchel, *Fatty acid synthesis in Escherichia coli and its applications towards the production of fatty acid based biofuels.* Biotechnol Biofuels, 2014. **7**(1): p. 7.
248. Nguyen, C., et al., *Trapping the dynamic acyl carrier protein in fatty acid biosynthesis.* 2014. **505**(7483): p. 427-431.

249. Issartel, J.P., V. Koronakis, and C. Hughes, *Activation of Escherichia coli prohaemolysin to the mature toxin by acyl carrier protein-dependent fatty acylation.* Nature, 1991. **351**(6329): p. 759-61.
250. Agarwal, V., et al., *Structure of the enzyme-acyl carrier protein (ACP) substrate gatekeeper complex required for biotin synthesis.* Proc Natl Acad Sci U S A, 2012. **109**(43): p. 17406-11.
251. Anderson, M.S., C.E. Bulawa, and C.R. Raetz, *The biosynthesis of gram-negative endotoxin. Formation of lipid A precursors from UDP-GlcNAc in extracts of Escherichia coli.* J Biol Chem, 1985. **260**(29): p. 15536-41.
252. Jordan, S.W. and J.E. Cronan, Jr., *A new metabolic link. The acyl carrier protein of lipid synthesis donates lipoic acid to the pyruvate dehydrogenase complex in Escherichia coli and mitochondria.* J Biol Chem, 1997. **272**(29): p. 17903-6.
253. Lu, Y.J., et al., *Acyl-phosphates initiate membrane phospholipid synthesis in Gram-positive pathogens.* Mol Cell, 2006. **23**(5): p. 765-72.
254. Nakayama, H., K. Kurokawa, and B.L. Lee, *Lipoproteins in bacteria: structures and biosynthetic pathways.* FEBS Journal, 2012. **279**(23): p. 4247-4268.
255. Stanley, P., V. Koronakis, and C. Hughes, *Acylation of Escherichia coli Hemolysin: A Unique Protein Lipidation Mechanism Underlying Toxin Function.* Microbiology and Molecular Biology Reviews, 1998. **62**(2): p. 309-333.
256. Morris, T.W., K.E. Reed, and J.E. Cronan, Jr., *Lipoic acid metabolism in Escherichia coli: the lplA and lipB genes define redundant pathways for ligation of lipoyl groups to apoprotein.* J Bacteriol, 1995. **177**(1): p. 1-10.
257. Lin, S. and J.E. Cronan, *Closing in on complete pathways of biotin biosynthesis.* Mol Biosyst, 2011. **7**(6): p. 1811-21.
258. Cronan, J.E., Jr. and G.L. Waldrop, *Multi-subunit acetyl-CoA carboxylases.* Prog Lipid Res, 2002. **41**(5): p. 407-35.
259. Joshi, V.C. and S.J. Wakil, *Studies on the mechanism of fatty acid synthesis. XXVI. Purification and properties of malonyl-coenzyme A--acyl carrier protein transacylase of Escherichia coli.* Arch Biochem Biophys, 1971. **143**(2): p. 493-505.
260. Tsay, J.T., et al., *Isolation and characterization of the beta-ketoacyl-acyl carrier protein synthase III gene (fabH) from Escherichia coli K-12.* J Biol Chem, 1992. **267**(10): p. 6807-14.
261. Heath, R.J. and C.O. Rock, *Inhibition of beta-ketoacyl-acyl carrier protein synthase III (FabH) by acyl-acyl carrier protein in Escherichia coli.* J Biol Chem, 1996. **271**(18): p. 10996-1000.
262. Choi, K.H., R.J. Heath, and C.O. Rock, *beta-ketoacyl-acyl carrier protein synthase III (FabH) is a determining factor in branched-chain fatty acid biosynthesis.* J Bacteriol, 2000. **182**(2): p. 365-70.
263. Edwards, P., et al., *Cloning of the fabF gene in an expression vector and in vitro characterization of recombinant fabF and fabB encoded enzymes from Escherichia coli.* FEBS Lett, 1997. **402**(1): p. 62-6.
264. Toomey, R.E. and S.J. Wakil, *Studies on the mechanism of fatty acid synthesis. XV. Preparation and general properties of beta-ketoacyl acyl carrier protein reductase from Escherichia coli.* Biochim Biophys Acta, 1966. **116**(2): p. 189-97.
265. Kass, L.R. and K. Bloch, *On the enzymatic synthesis of unsaturated fatty acids in Escherichia coli.* Proc Natl Acad Sci U S A, 1967. **58**(3): p. 1168-73.
266. Clark, D.P., et al., *Beta-hydroxydecanoyl thio ester dehydrase does not catalyze a rate-limiting step in Escherichia coli unsaturated fatty acid synthesis.* Biochemistry, 1983. **22**(25): p. 5897-902.
267. Mohan, S., et al., *An Escherichia coli gene (FabZ) encoding (3R)-hydroxymyristoyl acyl carrier protein dehydrase. Relation to fabA and suppression of mutations in lipid A biosynthesis.* J Biol Chem, 1994. **269**(52): p. 32896-903.
268. Heath, R.J. and C.O. Rock, *Enoyl-acyl carrier protein reductase (fabI) plays a determinant role in completing cycles of fatty acid elongation in Escherichia coli.* J Biol Chem, 1995. **270**(44): p. 26538-42.
269. Bergler, H., et al., *The enoyl-[acyl-carrier-protein] reductase (FabI) of Escherichia coli, which catalyzes a key regulatory step in fatty acid biosynthesis, accepts NADH and NADPH as cofactors and is inhibited by palmitoyl-CoA.* Eur J Biochem, 1996. **242**(3): p. 689-94.
270. Yoshimura, M., T. Oshima, and N. Ogasawara, *Involvement of the YneS/YgiH and PlsX proteins in phospholipid biosynthesis in both Bacillus subtilis and Escherichia coli.* BMC Microbiol, 2007. **7**: p. 69.
271. Ray, T.K. and J.E. Cronan, Jr., *Acylation of sn-glycerol 3-phosphate in Escherichia coli. Study of reaction with native palmitoyl-acyl carrier protein.* J Biol Chem, 1975. **250**(21): p. 8422-7.
272. Davis, M.S. and J.E. Cronan, Jr., *Inhibition of Escherichia coli acetyl coenzyme A carboxylase by acyl-acyl carrier protein.* J Bacteriol, 2001. **183**(4): p. 1499-503.
273. Heath, R.J. and C.O. Rock, *Regulation of fatty acid elongation and initiation by acyl-acyl carrier protein in Escherichia coli.* J Biol Chem, 1996. **271**(4): p. 1833-6.
274. Heath, R.J., S. Jackowski, and C.O. Rock, *Guanosine tetraphosphate inhibition of fatty acid and phospholipid synthesis in Escherichia coli is relieved by overexpression of glycerol-3-phosphate acyltransferase (plsB).* J Biol Chem, 1994. **269**(42): p. 26584-90.
275. Flugel, R.S., et al., *Holo-(acyl carrier protein) synthase and phosphopantetheinyl transfer in Escherichia coli.* J Biol Chem, 2000. **275**(2): p. 959-68.
276. Elovson, J. and P.R. Vagelos, *Acyl carrier protein. X. Acyl carrier protein synthetase.* J Biol Chem, 1968. **243**(13): p. 3603-11.

277. Keating, D.H., M.R. Carey, and J.E. Cronan, Jr., *The unmodified (apo) form of Escherichia coli acyl carrier protein is a potent inhibitor of cell growth.* J Biol Chem, 1995. **270**(38): p. 22229-35.
278. Kitagawa, M., et al., *Complete set of ORF clones of Escherichia coli ASKA library (a complete set of E. coli K-12 ORF archive): unique resources for biological research.* DNA Res, 2005. **12**(5): p. 291-9.
279. My, L., et al., *Transcription of the Escherichia coli Fatty Acid Synthesis Operon fabHDG Is Directly Activated by FadR and Inhibited by ppGpp.* Journal of Bacteriology, 2013. **195**(16): p. 3784-3795.
280. Henry, M.F. and J.E. Cronan Jr, *Escherichia coli transcription factor that both activates fatty acid synthesis and represses fatty acid degradation.* Journal of Molecular Biology, 1991. **222**(4): p. 843-849.
281. Jiang, P. and J.E. Cronan, Jr., *Inhibition of fatty acid synthesis in Escherichia coli in the absence of phospholipid synthesis and release of inhibition by thioesterase action.* J Bacteriol, 1994. **176**(10): p. 2814-21.
282. Cho, H. and J.E. Cronan, Jr., *Defective export of a periplasmic enzyme disrupts regulation of fatty acid synthesis.* J Biol Chem, 1995. **270**(9): p. 4216-9.
283. Voelker, T.A. and H.M. Davies, *Alteration of the specificity and regulation of fatty acid synthesis of Escherichia coli by expression of a plant medium-chain acyl-acyl carrier protein thioesterase.* J Bacteriol, 1994. **176**(23): p. 7320-7.
284. Westra, E.R., et al., *H-NS-mediated repression of CRISPR-based immunity in Escherichia coli K12 can be relieved by the transcription activator LeuO.* Molecular Microbiology, 2010. **77**(6): p. 1380-1393.
285. Albert, H., et al., *Site-specific integration of DNA into wild-type and mutant lox sites placed in the plant genome.* Plant J, 1995. **7**(4): p. 649-59.
286. Palmeros, B., et al., *A family of removable cassettes designed to obtain antibiotic-resistance-free genomic modifications of Escherichia coli and other bacteria.* Gene, 2000. **247**(1-2): p. 255-64.
287. Baba, T., et al., *The applications of systematic in-frame, single-gene knockout mutant collection of Escherichia coli K-12.* Methods Mol Biol, 2008. **416**: p. 183-94.
288. Kazakov, T., et al., *Escherichia coli peptidase A, B, or N can process translation inhibitor microcin C.* J Bacteriol, 2008. **190**(7): p. 2607-10.
289. Cronan, J.E., *Bacterial membrane lipids: where do we stand?* Annual reviews in microbiology, 2003. **57**(1): p. 203-224.
290. Dowhan, W., *Molecular basis for membrane phospholipid diversity: why are there so many lipids?* Annual review of biochemistry, 1997. **66**(1): p. 199-232.
291. Martin, W. and M.J. Russell, *On the origins of cells: a hypothesis for the evolutionary transitions from abiotic geochemistry to chemoautotrophic prokaryotes, and from prokaryotes to nucleated cells.* Philosophical Transactions of the Royal Society of London B: Biological Sciences, 2003. **358**(1429): p. 59-85.
292. Jain, S., A. Caforio, and A.J. Driessen, *Biosynthesis of archaeal membrane ether lipids.* Front Microbiol, 2014. **5**: p. 641.
293. Kamekura, M. and M. Kates, *Structural diversity of membrane lipids in members of Halobacteriaceae.* Bioscience, biotechnology, and biochemistry, 1999. **63**(6): p. 969-972.
294. Kates, M. *Archaebacterial lipids: structure, biosynthesis and function*. in *Biochemical Society Symposium*. 1991.
295. Koga, Y. and H. Morii, *Biosynthesis of ether-type polar lipids in archaea and evolutionary considerations.* Microbiol Mol Biol Rev, 2007. **71**(1): p. 97-120.
296. Zhang, D. and C.D. Poulter, *Biosynthesis of archaebacterial ether lipids. Formation of ether linkages by prenyltransferases.* Journal of the American Chemical Society, 1993. **115**(4): p. 1270-1277.
297. Boucher, Y., M. Kamekura, and W.F. Doolittle, *Origins and evolution of isoprenoid lipid biosynthesis in archaea.* Molecular microbiology, 2004. **52**(2): p. 515-527.
298. Tachibana, A., et al., *Potassium-stimulating mechanism of geranylgeranyl diphosphate synthase of Methanobacterium thermoformicicum SF-4.* Journal of biochemistry, 1993. **114**(3): p. 389-392.
299. Tachibana, A., *A novel prenyltransferase, farnesylgeranyl diphosphate synthase, from the haloalkaliphilic archaeon, Natronobacterium pharaonis.* FEBS Lett, 1994. **341**(2-3): p. 291-4.
300. Han, J.S., et al., *Kinetic study of sn-glycerol-1-phosphate dehydrogenase from the aerobic hyperthermophilic archaeon, Aeropyrum pernix K1.* European Journal of Biochemistry, 2002. **269**(3): p. 969-976.
301. Nishihara, M. and Y. Koga, *Purification and properties of sn-glycerol-1-phosphate dehydrogenase from Methanobacterium thermoautotrophicum: characterization of the biosynthetic enzyme for the enantiomeric glycerophosphate backbone of ether polar lipids of Archaea.* Journal of biochemistry, 1997. **122**(3): p. 572-576.
302. Payandeh, J., et al., *The crystal structure of (S)-3-O-geranylgeranylglyceryl phosphate synthase reveals an ancient fold for an ancient enzyme.* Journal of Biological Chemistry, 2006. **281**(9): p. 6070-6078.
303. Peterhoff, D., et al., *A comprehensive analysis of the geranylgeranylglyceryl phosphate synthase enzyme family identifies novel members and reveals mechanisms of substrate specificity and quaternary structure organization.* Molecular microbiology, 2014. **92**(4): p. 885-899.

304. Hemmi, H., et al., *(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 membrane lipids.* J Biol Chem, 2004. **279**(48): p. 50197-203.
305. Zhang, H., et al., *Total synthesis of geranylgeranylglyceryl phosphate enantiomers: substrates for characterization of 2, 3-O-digeranylgeranylglyceryl phosphate synthase.* Organic letters, 2006. **8**(5): p. 943-946.
306. Jain, S., et al., *Identification of CDP-archaeol synthase, a missing link of ether lipid biosynthesis in Archaea.* Chem Biol, 2014. **21**(10): p. 1392-401.
307. Warner, T.G. and E.A. Dennis, *Phosphatidylserine decarboxylase: analysis of its action towards unsaturated and saturated phosphatidylserine and the effect of Triton X-100 on activity.* Archives of biochemistry and biophysics, 1975. **167**(2): p. 761-768.
308. Kent, C., *Eukaryotic phospholipid biosynthesis.* Annual review of biochemistry, 1995. **64**(1): p. 315-343.
309. Raetz, C.R. and W. Dowhan, *Biosynthesis and Function of Phospholipids in Escherichia coli.* J. Biol. Chem., 1990. **265**: p. 1235-8.
310. Yung, B. and A. Kornberg, *Membrane attachment activates dnaA protein, the initiation protein of chromosome replication in Escherichia coli.* Proceedings of the National Academy of Sciences, 1988. **85**(19): p. 7202-7205.
311. Sekimizu, K., B.Y.-M. Yung, and A. Kornberg, *The dnaA protein of Escherichia coli. Abundance, improved purification, and membrane binding.* Journal of Biological Chemistry, 1988. **263**(15): p. 7136-7140.
312. Xia, W. and W. DowHAN, *In vivo evidence for the involvement of anionic phospholipids in initiation of DNA replication in Escherichia coli.* Proceedings of the National Academy of Sciences, 1995. **92**(3): p. 783-787.
313. De Vrije, T., et al., *Phosphatidylglycerol is involved in protein translocation across Escherichia coli inner membranes.* Nature: international weekly journal of science, 1988. **334**(6178): p. 173-175.
314. Lill, R., W. Dowhan, and W. Wickner, *The ATPase activity of SecA is regulated by acidic phospholipids, SecY, and the leader and mature domains of precursor proteins.* Cell, 1990. **60**(2): p. 271-280.
315. Phoenix, D.A., et al., *OmpF-Lpp signal sequence mutants with varying charge hydrophobicity ratios provide evidence for a phosphatidylglycerol-signal sequence interaction during protein translocation across the Escherichia coli inner membrane.* Journal of Biological Chemistry, 1993. **268**(23): p. 17069-17073.
316. Dowhan, W., *CDP-diacylglycerol synthase of microorganisms.* Biochimica et Biophysica Acta (BBA)-Lipids and Lipid Metabolism, 1997. **1348**(1): p. 157-165.
317. Morii, H. and Y. Koga, *CDP-2,3-Di-O-geranylgeranyl-sn-glycerol:L-serine O-archaetidyltransferase (archaetidylserine synthase) in the methanogenic archaeon Methanothermobacter thermautotrophicus.* J Bacteriol, 2003. **185**(4): p. 1181-9.
318. Daiyasu, H., et al., *A study of archaeal enzymes involved in polar lipid synthesis linking amino acid sequence information, genomic contexts and lipid composition.* Archaea, 2005. **1**(6): p. 399-410.
319. Koga, Y., *From promiscuity to the lipid divide: on the evolution of distinct membranes in Archaea and Bacteria.* Journal of molecular evolution, 2014. **78**(3-4): p. 234-242.
320. Lombard, J., P. López-García, and D. Moreira, *Phylogenomic investigation of phospholipid synthesis in archaea.* Archaea, 2012. **2012**.
321. Nishijima, M. and C. Raetz, *Membrane lipid biogenesis in Escherichia coli: identification of genetic loci for phosphatidylglycerophosphate synthetase and construction of mutants lacking phosphatidylglycerol.* Journal of Biological Chemistry, 1979. **254**(16): p. 7837-7844.
322. Heacock, P.N. and W. Dowhan, *Construction of a lethal mutation in the synthesis of the major acidic phospholipids of Escherichia coli.* Journal of Biological Chemistry, 1987. **262**(27): p. 13044-13049.
323. Gopalakrishnan, A.S., et al., *Structure and expression of the gene locus encoding the phosphatidylglycerophosphate synthase of Escherichia coli.* Journal of Biological Chemistry, 1986. **261**(3): p. 1329-1338.
324. Dowhan, W., *A retrospective: use of Escherichia coli as a vehicle to study phospholipid synthesis and function.* Biochimica et Biophysica Acta (BBA)-Molecular and Cell Biology of Lipids, 2013. **1831**(3): p. 471-494.
325. Lu, Y.-H., et al., *Three phosphatidylglycerol-phosphate phosphatases in the inner membrane of Escherichia coli.* Journal of Biological Chemistry, 2011. **286**(7): p. 5506-5518.
326. Icho, T. and C. Raetz, *Multiple genes for membrane-bound phosphatases in Escherichia coli and their action on phospholipid precursors.* Journal of bacteriology, 1983. **153**(2): p. 722-730.
327. Dillon, D.A., et al., *The Escherichia coli pgpB gene encodes for a diacylglycerol pyrophosphate phosphatase activity.* Journal of Biological Chemistry, 1996. **271**(48): p. 30548-30553.
328. Icho, T., *Membrane-bound phosphatases in Escherichia coli: sequence of the pgpB gene and dual subcellular localization of the pgpB product.* Journal of bacteriology, 1988. **170**(11): p. 5117-5124.
329. Funk, C., L. Zimniak, and W. Dowhan, *The pgpA and pgpB genes of Escherichia coli are not essential: evidence for a third phosphatidylglycerophosphate phosphatase.* Journal of bacteriology, 1992. **174**(1): p. 205-213.

330. Kanfer, J.N. and E.P. Kennedy, *Synthesis of phosphatidylserine by Escherichia coli.* Journal of Biological Chemistry, 1962. **237**(1): p. PC270-PC271.
331. Bae-Lee, M.S. and G.M. Carman, *Phosphatidylserine synthesis in Saccharomyces cerevisiae. Purification and characterization of membrane-associated phosphatidylserine synthase.* Journal of Biological Chemistry, 1984. **259**(17): p. 10857-10862.
332. Matsumoto, K., *Phosphatidylserine synthase from bacteria.* Biochimica et Biophysica Acta (BBA) - Lipids and Lipid Metabolism, 1997. **1348**(1–2): p. 214-227.
333. Sciara, G., et al., *Structural basis for catalysis in a CDP-alcohol phosphotransferase.* Nature communications, 2014. **5**.
334. Li, Q.-X. and W. Dowhan, *Studies on the mechanism of formation of the pyruvate prosthetic group of phosphatidylserine decarboxylase from Escherichia coli.* Journal of Biological Chemistry, 1990. **265**(7): p. 4111-4115.
335. Li, Q.-X. and W. Dowhan, *Structural characterization of Escherichia coli phosphatidylserine decarboxylase.* Journal of Biological Chemistry, 1988. **263**(23): p. 11516-11522.
336. Schuiki, I. and G. Daum, *Phosphatidylserine decarboxylases, key enzymes of lipid metabolism.* IUBMB life, 2009. **61**(2): p. 151-162.
337. Wang, K. and S.-i. Ohnuma, *Chain-length determination mechanism of isoprenyl diphosphate synthases and implications for molecular evolution.* Trends in biochemical sciences, 1999. **24**(11): p. 445-451.
338. Lee, P.C., et al., *Directed evolution of Escherichia coli farnesyl diphosphate synthase (IspA) reveals novel structural determinants of chain length specificity.* Metabolic engineering, 2005. **7**(1): p. 18-26.
339. Guldan, H., et al., *Functional Assignment of an Enzyme that Catalyzes the Synthesis of an Archaea-Type Ether Lipid in Bacteria.* Angewandte Chemie International Edition, 2011. **50**(35): p. 8188-8191.
340. Guldan, H., R. Sterner, and P. Babinger, *Identification and Characterization of a Bacterial Glycerol-1-phosphate Dehydrogenase: Ni2+-Dependent AraM from Bacillus subtilis†.* Biochemistry, 2008. **47**(28): p. 7376-7384.
341. Bligh, E.G. and W.J. Dyer, *A rapid method of total lipid extraction and purification.* Can J Biochem Physiol, 1959. **37**(8): p. 911-7.
342. Farmer, W.R. and J.C. Liao, *Improving lycopene production in Escherichia coli by engineering metabolic control.* Nat Biotechnol, 2000. **18**(5): p. 533-7.
343. Morii, H. and Y. Koga, *CDP-2, 3-di-O-geranylgeranyl-sn-glycerol: L-serine O-archaetidyltransferase (archaetidylserine synthase) in the methanogenic archaeon Methanothermobacter thermautotrophicus.* Journal of bacteriology, 2003. **185**(4): p. 1181-1189.
344. Wagner, S., et al., *Tuning Escherichia coli for membrane protein overexpression.* Proceedings of the National Academy of Sciences, 2008. **105**(38): p. 14371-14376.
345. Kaufmann, A., et al., *Cysteine-directed cross-linking demonstrates that helix 3 of SecE is close to helix 2 of SecY and helix 3 of a neighboring SecE.* Biochemistry, 1999. **38**(28): p. 9115-9125.
346. Damen, C.W., et al., *Enhanced lipid isomer separation in human plasma using reversed-phase UPLC with ion-mobility/ high-resolution MS detection.* Journal of lipid research, 2014. **55**(8): p. 1772-1783.
347. Sojo, V., A. Pomiankowski, and N. Lane, *A bioenergetic basis for membrane divergence in archaea and bacteria.* PLoS Biol, 2014. **12**(8): p. e1001926.
348. Caforio, A., et al., *Formation of the ether lipids archaetidylglycerol and archaetidylethanolamine in Escherichia coli.* Biochem J, 2015. **470**(3): p. 343-55.
349. van den Brink-van der Laan, E., J.A. Killian, and B. de Kruijff, *Nonbilayer lipids affect peripheral and integral membrane proteins via changes in the lateral pressure profile.* Biochim Biophys Acta, 2004. **1666**(1-2): p. 275-88.
350. Zhang, Z., et al., *High-level production of membrane proteins in E. coli BL21 (DE3) by omitting the inducer IPTG.* Microbial cell factories, 2015. **14**(1): p. 1.
351. Nemoto, N., T. Oshima, and A. Yamagishi, *Purification and characterization of geranylgeranylglyceryl phosphate synthase from a thermoacidophilic archaeon, Thermoplasma acidophilum.* Journal of biochemistry, 2003. **133**(5): p. 651-657.
352. Sparrow, C.P. and C.R. Raetz, *Purification and properties of the membrane-bound CDP-diglyceride synthetase from Escherichia coli.* Journal of Biological Chemistry, 1985. **260**(22): p. 12084-91.
353. Lu, Y.-H., et al., *Three Phosphatidylglycerol-phosphate Phosphatases in the Inner Membrane of Escherichia coli.* Journal of Biological Chemistry, 2011. **286**(7): p. 5506-5518.
354. Maurer, S.E., et al., *Chemical evolution of amphiphiles: glycerol monoacyl derivatives stabilize plausible prebiotic membranes.* Astrobiology, 2009. **9**(10): p. 979-87.
355. Ohta, A., et al., *Cloning of genes involved in membrane lipid synthesis: effects of amplification of phosphatidylglycerophosphate synthase in Escherichia coli.* J Bacteriol, 1981. **147**(2): p. 552-62.
356. Saha, S.K., et al., *A regulatory mechanism for the balanced synthesis of membrane phospholipid species in Escherichia coli.* Biosci Biotechnol Biochem, 1996. **60**(1): p. 111-6.

357. Parsons, J.B. and C.O. Rock, *Bacterial Lipids: Metabolism and Membrane Homeostasis.* Progress in lipid research, 2013. **52**(3): p. 249-276.
358. Zhang, Y.-M. and C.O. Rock, *Transcriptional regulation in bacterial membrane lipid synthesis.* Journal of lipid research, 2009. **50**(Supplement): p. S115-S119.
359. Albers, S.V., et al., *Adaptations of the archaeal cell membrane to heat stress.* Front Biosci, 2000. **5**: p. D813-20.
360. Meador, T.B., et al., *Thermococcus kodakarensis modulates its polar membrane lipids and elemental composition according to growth stage and phosphate availability.* Front Microbiol, 2014. **5**: p. 10.
361. Morii, H., et al., *A novel ether core lipid with H-shaped C 80-isoprenoid hydrocarbon chain from the hyperthermophilic methanogen Methanothermus fervidus.* Biochimica et Biophysica Acta (BBA)-Lipids and Lipid Metabolism, 1998. **1390**(3): p. 339-345.
362. Butland, G., et al., *eSGA: E. coli synthetic genetic array analysis.* Nat Methods, 2008. **5**(9): p. 789-95.
363. Nemoto, N., et al., *Characterization of the precursor of tetraether lipid biosynthesis in the thermoacidophilic archaeon Thermoplasma acidophilum.* Extremophiles, 2003. **7**(3): p. 235-243.
364. Poulter, C.D., T. Aoki, and L. Daniels, *Biosynthesis of isoprenoid membranes in the methanogenic archaebacterium Methanospirillum hungatei.* Journal of the American Chemical Society, 1988. **110**(8): p. 2620-2624.
365. Eguchi, T., Y. Nishimura, and K. Kakinuma, *Importance of the isopropylidene terminal of geranylgeranyl group for the formation of tetraether lipid in methanogenic archaea.* Tetrahedron Letters, 2003. **44**(16): p. 3275-3279.
366. Kramer, J.K.G. and F.D. Sauer, *Changes in the diether-to-tetraether-lipid ratio during cell growth in Methanobacterium thermoautotrophicum.* FEMS Microbiology Letters, 1991. **83**(1): p. 45-50.
367. Morikawa, M., et al., *Purification and characterization of a thermostable thiol protease from a newly isolated hyperthermophilic Pyrococcus sp.* Appl Environ Microbiol, 1994. **60**(12): p. 4559-66.
368. Hileman, T.H. and T.J. Santangelo, *Genetics techniques for Thermococcus kodakarensis.* Recent advances in genomic and genetic studies in the Archaea, 2012: p. 19.
369. Santangelo, T.J. and J.N. Reeve, *Genetic tools and manipulations of the hyperthermophilic heterotrophic archaeon Thermococcus kodakarensis*, in *Extremophiles Handbook*. 2011, Springer. p. 567-582.
370. Sato, T., et al., *Targeted gene disruption by homologous recombination in the hyperthermophilic archaeon Thermococcus kodakaraensis KOD1.* Journal of bacteriology, 2003. **185**(1): p. 210-220.
371. Sato, T., et al., *Improved and versatile transformation system allowing multiple genetic manipulations of the hyperthermophilic archaeon Thermococcus kodakaraensis.* Applied and environmental microbiology, 2005. **71**(7): p. 3889-3899.
372. Fukui, T., et al., *Complete genome sequence of the hyperthermophilic archaeon Thermococcus kodakaraensis KOD1 and comparison with Pyrococcus genomes.* Genome Res, 2005. **15**(3): p. 352-63.
373. Lobasso, S., et al., *Coupled TLC and MALDI-TOF/MS analyses of the lipid extract of the hyperthermophilic archaeon Pyrococcus furiosus.* Archaea, 2012. **2012**: p. 957852.
374. Sugai, A., et al., *The Core Lipid Composition of the 17 Strains of Hyperthermophilic Archaea, Thermococcales.* Journal of Oleo Science, 2004. **53**(1): p. 41-44.
375. Lobasso, S., et al., *Coupled TLC and MALDI-TOF/MS Analyses of the Lipid Extract of the Hyperthermophilic Archaeon Pyrococcus furiosus.* Archaea, 2012. **2012**: p. 10.
376. Nishihara, M., et al., *Structural analysis by reductive cleavage with LiAlH4 of an allyl ether choline-phospholipid, archaetidylcholine, from the hyperthermophilic methanoarchaeon Methanopyrus kandleri.* Archaea, 2002. **1**(2): p. 123-31.
377. Makarova, K.S., Y.I. Wolf, and E.V. Koonin, *Archaeal clusters of orthologous genes (arCOGs): an update and application for analysis of shared features between thermococcales, methanococcales, and methanobacteriales.* Life, 2015. **5**(1): p. 818-840.
378. Soderberg, T., A. Chen, and C.D. Poulter, *Geranylgeranylglyceryl phosphate synthase. Characterization of the recombinant enzyme from Methanobacterium thermoautotrophicum.* Biochemistry, 2001. **40**(49): p. 14847-14854.
379. Leferink, N.G., et al., *The growing VAO flavoprotein family.* Archives of biochemistry and biophysics, 2008. **474**(2): p. 292-301.
380. Liang, P.-H., T.-P. Ko, and A.H.J. Wang, *Structure, mechanism and function of prenyltransferases.* European Journal of Biochemistry, 2002. **269**(14): p. 3339-3354.
381. Ogura, K. and T. Koyama, *Enzymatic Aspects of Isoprenoid Chain Elongation.* Chemical Reviews, 1998. **98**(4): p. 1263-1276.
382. Kanai, T., et al., *Identification of the Phr-dependent heat shock regulon in the hyperthermophilic archaeon, Thermococcus kodakaraensis.* J Biochem, 2010. **147**(3): p. 361-70.
383. Liu, W., et al., *Crystal structure of the archaeal heat shock regulator from Pyrococcus furiosus: a molecular chimera representing eukaryal and bacterial features.* Journal of molecular biology, 2007. **369**(2): p. 474-488.

384. Vierke, G., et al., *A novel archaeal transcriptional regulator of heat shock response.* Journal of Biological Chemistry, 2003. **278**(1): p. 18-26.
385. Santangelo, T.J. and J.N. Reeve, *Thermococcus kodakarensis genetics: TK1827-encoded β-glycosidase, new positive-selection protocol, and targeted and repetitive deletion technology.* Applied and environmental microbiology, 2010. **76**(4): p. 1044-1052.
386. Takemasa, R., et al., *Thermococcus kodakarensis as a host for gene expression and protein secretion.* Applied and environmental microbiology, 2011. **77**(7): p. 2392-2398.
387. Spaans, S.K., *Thermococcus kodakarensis: the key to affordable hydrogen production*. 2016, Wageningen University: Wageningen.
388. Borges, N., et al., *Thermococcus kodakarensis mutants deficient in di-myo-inositol phosphate use aspartate to cope with heat stress.* Journal of bacteriology, 2010. **192**(1): p. 191-197.
389. Hemmi, H., et al., *Menaquinone-specific prenyl reductase from the hyperthermophilic archaeon Archaeoglobus fulgidus.* Journal of bacteriology, 2005. **187**(6): p. 1937-1944.
390. Naparstek, S., Z. Guan, and J. Eichler, *A predicted geranylgeranyl reductase reduces the ω-position isoprene of dolichol phosphate in the halophilic archaeon, Haloferax volcanii.* Biochimica et Biophysica Acta (BBA)-Molecular and Cell Biology of Lipids, 2012. **1821**(6): p. 923-933.
391. Huijbers, M.M., et al., *Flavin dependent monooxygenases.* Archives of biochemistry and biophysics, 2014. **544**: p. 2-17.
392. Nowosielski, M., et al., *Detailed mechanism of squalene epoxidase inhibition by terbinafine.* J Chem Inf Model, 2011. **51**(2): p. 455-62.
393. Hopmans, E.C., S. Schouten, and J.S.S. Damsté, *The effect of improved chromatography on GDGT-based palaeoproxies.* Organic Geochemistry, 2016. **93**: p. 1-6.
394. Schouten, S., et al., *Archaeal and bacterial glycerol dialkyl glycerol tetraether lipids in hot springs of Yellowstone National Park.* Applied and Environmental Microbiology, 2007. **73**(19): p. 6181-6191.
395. Damsté, J.S.S., et al., *Crenarchaeol the characteristic core glycerol dibiphytanyl glycerol tetraether membrane lipid of cosmopolitan pelagic crenarchaeota.* Journal of Lipid Research, 2002. **43**(10): p. 1641-1651.
396. Kell, D.B., et al., *Membrane transporter engineering in industrial biotechnology and whole cell biocatalysis.* Trends in biotechnology, 2015. **33**(4): p. 237-246.
397. Beauprez, J.J., et al., *Influence of C4-dicarboxylic acid transporters on succinate production.* Green Chemistry, 2011. **13**(8): p. 2179-2186.
398. Becker, M., et al., *Glutamate efflux mediated by Corynebacterium glutamicum MscCG, Escherichia coli MscS, and their derivatives.* Biochimica et Biophysica Acta (BBA)-Biomembranes, 2013. **1828**(4): p. 1230-1240.
399. Foo, J.L. and S.S.J. Leong, *Directed evolution of an E. coli inner membrane transporter for improved efflux of biofuel molecules.* Biotechnology for biofuels, 2013. **6**(1): p. 1.
400. Naik, S.N., et al., *Production of first and second generation biofuels: a comprehensive review.* Renewable and Sustainable Energy Reviews, 2010. **14**(2): p. 578-597.
401. Ly, H.V. and M.L. Longo, *The influence of short-chain alcohols on interfacial tension, mechanical properties, area/molecule, and permeability of fluid lipid bilayers.* Biophysical Journal, 2004. **87**(2): p. 1013-1033.
402. Huffer, S., et al., *Role of alcohols in growth, lipid composition, and membrane fluidity of yeasts, bacteria, and archaea.* Applied and environmental microbiology, 2011. **77**(18): p. 6400-6408.
403. Ingram, L.O., *Microbial tolerance to alcohols: role of the cell membrane.* Trends in Biotechnology, 1986. **4**(2): p. 40-44.
404. Yamauchi, K., et al., *Archaebacterial lipid models. Highly thermostable membranes from 1,1'-(1,32-dotriacontamethylene)-bis(2-phytanyl-sn-glycero-3-phosphocholine).* Journal of the American Chemical Society, 1990. **112**(8): p. 3188-3191.
405. Chang, E.L., *Unusual Thermal Stability of Liposomes Made from Bipolar Tetraether Lipids.* Biochemical and Biophysical Research Communications, 1994. **202**(2): p. 673-679.
406. Stetter, K., et al., *Hyperthermophilic archaea are thriving in deep North Sea and Alaskan oil reservoirs.* Nature, 1993. **365**(6448): p. 743-745.
407. Shtarkman, Y.M., et al., *Subglacial Lake Vostok (Antarctica) accretion ice contains a diverse set of sequences from aquatic, marine and sediment-inhabiting bacteria and eukarya.* PLoS One, 2013. **8**(7): p. e67221.
408. Kandler, O., *The early diversification of life and the origin of the three domains: a proposal*, in *Thermophiles: The keys to molecular evolution and the origin of life?* , A.W.W.M. Juergen Wiegel, Editor. 1998, Taylor and Francis: Athens, USA. p. 19-28.
409. Sprott, G., et al., *Stability of liposomes prepared from archaeobacterial lipids and phosphatidylcholine mixtures.* Cells and Materials, 1996. **6**: p. 143-155.
410. Longo, G.S., D.H. Thompson, and I. Szleifer, *Stability and phase separation in mixed monopolar lipid/bolalipid layers.* Biophysical journal, 2007. **93**(8): p. 2609-2621.

411. Fan, Q., et al., *Stability against temperature and external agents of vesicles composed of archaeal bolaform lipids and egg PC.* Biochimica et Biophysica Acta (BBA)-Biomembranes, 1995. **1240**(1): p. 83-88.
412. Monnard, P.A. and D.W. Deamer, *Membrane self-assembly processes: Steps toward the first cellular life.* The Anatomical Record, 2002. **268**(3): p. 196-207.
413. Peretó, J., P. López-García, and D. Moreira, *Ancestral lipid biosynthesis and early membrane evolution.* Trends in biochemical sciences, 2004. **29**(9): p. 469-477.
414. Thomas, J. and J.E. Cronan, *The enigmatic acyl carrier protein phosphodiesterase of Escherichia coli: genetic and enzymological characterization.* J Biol Chem, 2005. **280**(41): p. 34675-83.
415. Lemieux, M.J., Y. Huang, and D.N. Wang, *Glycerol-3-phosphate transporter of Escherichia coli: structure, function and regulation.* Res Microbiol, 2004. **155**(8): p. 623-9.
416. Shibuya, I., *Metabolic regulation and biological functions of phospholipids in Escherichia coli.* Progress in lipid research, 1992. **31**(3): p. 245-299.
417. Balleza, D., et al., *Ether-versus ester-linked phospholipid bilayers containing either linear or branched apolar chains.* Biophysical journal, 2014. **107**(6): p. 1364-1374.
418. Tatar, L.D., et al., *An Escherichia coli undecaprenyl-pyrophosphate phosphatase implicated in undecaprenyl phosphate recycling.* Microbiology, 2007. **153**(8): p. 2518-2529.
419. Koga, Y. and H. Morii, *Special methods for the analysis of ether lipid structure and metabolism in archaea.* Anal Biochem, 2006. **348**(1): p. 1-14.
420. Hafenbradl, D., et al., *A novel unsaturated archaeal ether core lipid from the hyperthermophile Methanopyrus kandleri.* Systematic and applied microbiology, 1993. **16**(2): p. 165-169.
421. Kuhry, J.-G., et al., *TMA-DPH: a suitable fluorescence polarization probe for specific plasma membrane fluidity studies in intact living cells.* Cell biophysics, 1983. **5**(2): p. 129-140.
422. Bulacu, M., X. Périole, and S.J. Marrink, *In silico design of robust bolalipid membranes.* Biomacromolecules, 2011. **13**(1): p. 196-205.
423. Chugunov, A.O., et al., *Liquid but durable: molecular dynamics simulations explain the unique properties of archaeal-like membranes.* Scientific reports, 2014. **4**.
424. Kulp, A. and M.J. Kuehn, *Biological functions and biogenesis of secreted bacterial outer membrane vesicles.* Annu Rev Microbiol, 2010. **64**: p. 163-84.
425. Nichols, J.W. and D.W. Deamer, *Net proton-hydroxyl permeability of large unilamellar liposomes measured by an acid-base titration technique.* Proc Natl Acad Sci U S A, 1980. **77**(4): p. 2038-42.
426. Hancock, R. and H. Nikaido, *Outer membranes of gram-negative bacteria. XIX. Isolation from Pseudomonas aeruginosa PAO1 and use in reconstitution and definition of the permeability barrier.* Journal of Bacteriology, 1978. **136**(1): p. 381-390.

Acknowledgements

Many of the sayings that you get to hear prior to a PhD life, are true. It's a trajectory of personal development, a struggle and truly a life changing experience. During my research I also endured the euphoria of doing three steps forward, and the frustrations of two steps (or more) backwards. One of the clichés that intrigued me the most is the path that describes your research scope during these four years. In the first year one starts with a very narrow scope on the topic. You'll start to fence off your research questions and explore the topic using the knowledge you are equipped with during your education. Sooner or later you will notice that your comfort zone of knowledge does no longer suffice and without expanding your horizon your chosen path will lead you to dead ends. This is the point at which multiple paths are explored and your scope is at its widest. Ultimately, by the end of the contract, where stress-levels reach the sky, the scope becomes narrow, equivalent to the first year. You can imagine my sense of relieve when I recognized a return to a straight trajectory at the end of my PhD. Obviously, I could not have done this without the many inspiring people by my side showing me the right way.

John, dank voor het vertrouwen en de vele inspirerende gesprekken. Je bent werkelijk een onuitputtelijke bron van ideëen. Ik zal nog met regelmaat denken aan de momenten dat ik door de bomen het bos niet meer zag. Al snel bleek dat het de beste remedie is om een uur door te brengen in jouw kantoor. Hoewel ik altijd naar buiten stapte met nog meer bomen, kon ik daarna altijd de hoofdzaken weer onderscheiden en met frisse moed weer verder gaan.

Servé, voor mij ben je de ideale begeleider geweest. Ik heb enorm veel bewondering voor je wetenschappelijke inzichten, je analytisch vermogen en dat in combinatie met een flinke dosis geduld. Jouw bereidheid om over de kleinste details met mij in gesprek te gaan hebben geleid tot mijn persoonlijke ontwikkeling en waren cruciaal voor het slagen van dit project.

In het bijzonder wil ik ook mijn paranimfen en kantoorgenoten bedanken. **Bas**, hartelijk bedankt voor de leuke tijd. Allereerst vond ik het een eer om jouw paranimf te mogen zijn, en ben ik blij dat je dat voor mij ook wilde doen. Ik zal nog lang denken aan de vele momenten van lachen aan de labtafel. Aan jou heb ik ook mijn interesse in anaerobe sulfaatreducerende hyperthermofiele archaea te danken. **Franklin**, I'm happy to have shared the final two years being your colleague. In this short time period, we became close friends and started a few collaborations which gave me insight in the beautifull world of phages. Also to you I think it is an honor having you as my paranimf. **Nico**, bedankt voor alle fijne gesprekken. Ik heb vanaf het begin bewondering gehad je vlotte babbel en je inspirerende vermogen om samenwerkingen te starten. Ik heb veel van je mogen geleerd. Ook voor jou hoop ik dat onze samenwerking uiteindelijk wordt beloond. **Teunke,** Ik heb enorm genoten van onze tijd als kantoorgenoten en heb veel van je geleerd. We hebben

veel steun aan elkaar gehad in moeizame tijden. De vele gesprekken die ik met jou, maar ook samen met **John R., Sjoerd** heb gehad over onze behaalde resultaten waren enorm leerzaam.

To all the colleagues of the Bacterial Genetics group, **Mark L, Vincent, Bram, Amos, Tom, Elleke, Tessa, Slavtscho, Brenda, Tijn, Hanne,** all the **CRISPR people: Stan, Edze, Matthijs, Daan, Jorrit, Marnix, Yifan, Tim, Patrick, Prarthana, Raymond, Jochem** and also to **Rebecca, Sjon.** Thank you all for the friendly and professional atmosphere making the entire 4 years a wonderfull time.

A special thank you to all my students: **Roy, Eva, Menno, Dennis, Jessica, Frits, Miquel, Boris** for their dedication and contribution to this project. Each of your projects raised essential information that contributed to the success of this project. I enjoyed mentoring each one of you and wish you all the best in your further carreers.

Bijzonder veel dank voor al het personeel van het laboratorium van Microbiologie. **Willem,** dank voor het beschikbaar stellen van een bijzondere wetenschappelijke omgeving. Ik kijk enorm naar je op, en zal nog lang met plezier terugdenken aan de PhD-PostDoc meetings. Ook veel dank aan de bijzonder getalenteerde technicians binnen Microbiologie, **Wim**, **Anja**, **Tom vG**, **Tom S**, **Monica**, **Phillippe**, Ik heb jullie bereidheid om te helpen zeer gewaardeerd.

Furthermore, I like to express my gratitude to our collaborators of the **Molecular Microbiology** lab from the Rijksuniversiteit Groningen. **Arnold Driessen,** zonder jou visie was dit project nooit van de grond gekomen. Dank voor het delen van je kennis over lipiden en membranen. Ik heb met veel plezier de vele bezoekjes doorgebracht in je afdeling. **Antonella**,thank you very much for the collaboration. I enjoyed our brainstorm sessions and will never forget the excitement when we analysed the breakthrough data. **Samta**, thanks for being an inspirational post-doc. Your work was indispensable in the beginning of this project.

Beste **Tiny** van het Wageningen WEMC facility, Dank voor de schitterende electronen miscroscopie fotos, en je bereidheid om de samples te optimaliseren. Zonder jou hadden de cellen geen gezicht.

Alle leden van **NIOZ, Jaap, Laura, Ellen** en **Michel** hartelijk dank voor jullie fantastische samenwerking.

Lieve familie **Papa**, **Mama**, en **Rebekka**, **Alex, Damian, Marcus en Gwyn.** Heel erg veel dank voor jullie steun en liefde. Ik heb jullie interesse in mijn werk altijd enorm gewaardeerd. Nu dat ik deze drukke periode in mijn leven kan afsluiten, ben ik blij met het vooruitzicht weer meer tijd met jullie te kunnen doorbrengen. Bedankt ook voor jullie vele ondersteuningen

waarmee jullie de tijd hebben gecreëerd om het werken mogelijk te maken. Ook aan mijn schoonfamilie, **Pa, Ma** en **Richard, Chantal, Elynn en Jaylen**. Dank voor jullie hulp en interesse. Lieve **Dick, Heleen, Nick, Brenda, Megghane, Sander,** en familieleden **tante Priscilla** en **oom Tino** ook dank voor jullie steun en enthousiasme.

Aan mijn lieve kinderen, **Renée** en **Robin,** Schatjes; heel veel dank voor jullie liefde en steun. Jullie zijn mijn bron van inspiratie en ben er bijzonder trots op jullie papa te mogen zijn. Op moeilijke momenten hoef ik alleen maar aan mijn meisjes te denken om kracht uit te putten.

Lieve, lieve **Sandra**, de laatste woorden die vaak het eerst worden gelezen zijn aan jou gericht. Wat ben ik trots dat ik bij jou hoor. Woorden schieten te kort om jou te bedanken voor je steun en liefde. Het feit dat je al die tijd in mij bent blijven geloven zijn het allerbelangrijkste geweest voor het voltooien van dit werk. Zonder jou was dit boek nooit tot stand gekomen en daarvoor ben ik je eeuwig dankbaar. Lieve schat, ik hou ontzettend veel van je.

About the author

Melvin Franciscus Siliakus was born on April 8, 1983, in Arnhem, the Netherlands. After completing secondary school (HAVO) at the Merlet College in Cuyk the Netherlands in 2001, he started his Bsc in Biological and Medical laboratory research at the HAN university (Hogeschool Arnhem, Nijmegen), with a major in biochemistry. His bachelor thesis was performed in the Frans Kuypers laboratory of the Children's Hospital Oakland Research Institute, Oakland, USA under supervision of Dr. Eric Soupene, where he studied the activity, oligomerization and subcellular localization of a human acyl-CoA synthetase which resulted in a co-authorship of a publication on this topic. After obtaining the Bachelor of applied sciences in 2006, he continued with his Msc in Biomolecular Sciences at the Radboud University of Nijmegen, the Netherlands, with a dual major in pathobiology and toxicology. His first Master thesis was performed under the supervision of Dr. William Leenders at the department of pathology in the Radboud University Medical Center where he studied the membrane localization of native and heterologously expressed dynactin p150^{glued} in epithelial brain tumor cells. His second Master thesis for the toxicology major was performed under supervision of Dr. Dinant Kroese at the department Quality of Life R&D of TNO in Zeist, where he studied the relation of repeated dose and reproductive toxicity in a comparative analysis study. After obtaining his MSc degree in 2009, he continued with his PhD at the laboratory of Microbiology at Wageningen University in the department of bacterial genetics under supervision of his promoter prof. dr. John van der Oost and his co-promoter dr. Servé W.M. Kengen. During this PhD thesis, he worked on the engineering and characterization of a hybrid heterochiral membrane in *Escherichia coli*, presented here in this thesis. During this research, he established a collaboration within the department of marine microbiology and biogeochemistry (MMB) of the NIOZ Royal Netherlands Institute for Sea Research headed by prof. dr. Jaap Sinninghe Damsté. This research was focussed on the identification of a tetraether synthase in the hyperthermophilic archaeon *Thermococcus Kodakarensis* which culminated in a post-doctoral position at NIOZ where he is currently still working.

List of publications

Siliakus M.F., van der Oost J., Kengen S.W.M., (2016) Adaptations of archaeal and bacterial membranes to variations in temperature, pH and pressure. *Submitted in* Extremophiles

Siliakus M.F., Caforio A., Jain S., Kengen S.W.M., Driessen A., van der Oost J., (2016) Generation of a hybrid heterochiral membrane in *Escherichia coli*. *Manuscript in preparation*

Siliakus M.F., Spaans S.K., Makarova K., Matsumi R., Caforio A., Driessen A.J.M., de Vos W.M., Kengen S.W.M., van der Oost J., (2016) Identification of *Thermococcus kodakarensis* genes involved in GDGT synthesis. *Manuscript in preparation*

Siliakus M.F., Rozas M., van Rossum T., van der Oost J., Kengen S.W.M., (2016) Modulated gene expression of fatty acid and phospholipid metabolism by PlsB and apo-ACP. *Manuscript in preparation*

Nobrega F.L., Costa A.R., Santos J.F., **Siliakus M.F.**, van Lent J.W.M., Kengen S.W.M., Azaredo J., Kluskens L.D., (2016) Genetically manipulated phages with improved pH resistance for oral administration in veterinary medicine. *Submitted* in Scientific Reports

Caforio A., Jain S., Fodran P., **Siliakus M.**, Minnaard A.J., van der Oost J., Driessen A.J., (2015) Formation of the ether lipids archaetidylglycerol and archaetidylethanolamine in Escherichia coli. Biochem J. 470:343-355 doi: 10.1042/BJ20150626

Claassens N.J., **Siliakus M.F.,** Creutzburg S., Nijsse B., Spaans S.K., Quax T.E.F., Schaap P., van der Oost J., (2016) Codon usage algorithms and transcriptional tuning improve heterologous production of membrane proteins in *Escherichia coli*. *Manuscript in preparation*

Co-author affiliations

Laboratory of Microbiology, Department of Agrotechnology and Food Sciences, Wageningen University, Stippeneng 4, 6708 WE, Wageningen, the Netherlands.
John van der Oost, Servé W.M. Kengen, Sebastiaan K. Spaans[a], Willem de Vos, Miquel Rozas, Teunke van Rossum, Franklin L. Nobrega, Nico J. Claassens, Sjoerd Creutzburg, Tessa Quax[c]

Department of Molecular Microbiology, Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, 9747 AG Groningen, The Netherlands; The Zernike Institute for Advanced Materials, University of Groningen, 9747 AG Groningen, The Netherlands
Arnold J. Driessen, Antonella Caforio, Samta Jain[b], Marten Exterkate

Stratingh Institute for Chemistry, University of Groningen, Nijenborgh 7, 9747 AG Groningen, The Netherlands
Adriaan J. Minnaard, Peter Fodran, Ruben L.H. Adringa, Varsha R. Jumde

National Center for Biotechnology Information, National Library of Medicine, National Institutes of Health, 8600 Rockville Pike, Bethesda, MD, 20894, USA
Kira S. Makarova

Department of Microbiology, The Ohio State University, 105 Biological Sciences Building, 484 West 12th Avenue, Columbus, OH 43210, USA
Rie Matsumi

Centre of Biological Engineering, University of Minho, Campus de Gualtar, 4710-057 Braga, Portugal
Ana Rita Costa, José F. Santos, Joana Azeredo, Leon D. Kluskens

Laboratory of Virology, Department of Plant Sciences, Wageningen University, Droevendaalsesteeg 1, 6708 PB Wageningen, The Netherlands
Jan W.M. van Lent

Laboratory of Systems and Synthetic Biology, Wageningen University, Stippeneng 4, 6708 WE, Wageningen, the Netherlands.
Bart Nijsse, Peter Schaap

Current address:

[a]The Janssen Pharmaceutical Companies, Johnson and Johnson, Archimedesweg 4, Leiden, The Netherlands

[b]Department of Medicine, Section of Infectious Diseases, Boston University School of Medicine, 02118 Boston, Massachusetts, USA

[c]Molecular Biology of Archaea, Institute für Biologie II, Albert-Ludwigs-Universität Freiburg, Schänzlestr. 1, 79104 Freiburg, Germany

Overview of completed training activities

Name of the course	Graduate school/institute	Year
a. **Discipline specific activities**		
Metabolomics for microbial systems biology	TU Delft BSDL	2011
BE Basic Symposium.I pitch/poster	BE-BASIC	2012
BE Basic Symposium.II presentation	BE-BASIC	2013
NWO meeting / poster	NWO	2012
ALW mol.gen poster	ALW	2012
ALW mol.gen	ALW	2011
ALW mol.gen presentation	ALW	2013
BE-Basic symp.III pitch/poster	BE-BASIC	2014
b. **General courses**		
Scientific writing	WGS	2013
Career perspectives CCP	WGS	2014
VLAG PhD intro week	VLAG	2012
Presentation skills (PS)	WGS	2014
Project & time management	VLAG	2012
c. **Optionals**		
Preparation research proposal	Bacterial genetics	constant
BacGen group meeting	Bacterial genetics	weekly
PhD-Post-Doc meeting	Microbiology lab	bi-monthly
PhD study trip	Microbiology lab	2013

The research described in this thesis was carried out within the BE-Basic R&D Program, which was granted a FES (Fonds Economische Structuurversterking) subsidy from the Dutch Ministry of Economic affairs, agriculture and innovation (EL&I)

Cover design: Iliana Boshoven-Gkini, ©AgileColor | www.AgileColor.com
Thesis lay-out: Iliana Boshoven-Gkini, ©AgileColor | www.AgileColor.com
Printed: GVO drukkers & vormgevers B.V., Ede | www.gvo.nl