MICROBIOLOGY

# Disentangling the lipid divide: Identification of key enzymes for the biosynthesis of membrane-spanning and ether lipids in Bacteria

**Diana X. Sahonero-Canavesi[1]*, Melvin F. Siliakus[1], Alejandro Abdala Asbun[1], Michel Koenen[1], F. A. Bastiaan von Meijenfeldt[1], Sjef Boeren[2], Nicole J. Bale[1], Julia C. Engelman[1], Kerstin Fiege[1], Lora Strack van Schijndel[1], Jaap S. Sinninghe Damsté[1,3], Laura Villanueva[1,3]**



**Bacterial membranes are composed of fatty acids (FAs) ester-linked to glycerol-3-phosphate, while archaea have membranes made of isoprenoid chains ether-linked to glycerol-1-phosphate. Many archaeal species organize their membrane as a monolayer of membrane-spanning lipids (MSLs). Exceptions to this "lipid divide" are the production by some bacterial species of (ether-bound) MSLs, formed by tail-to-tail condensation of FAs resulting in the formation of (iso) diabolic acids (DAs), which are the likely precursors of paleoclimatological relevant branched glycerol dialkyl glycerol tetraether molecules. However, the enzymes responsible for their production are unknown. Here, we report the discovery of bacterial enzymes responsible for the condensation reaction of FAs and for ether bond formation and confirm that the building blocks of iso-DA are branched iso-FAs. Phylogenomic analyses of the key biosynthetic genes reveal a much wider diversity of potential MSL (ether)–producing bacteria than previously thought, with importantt implications for our understanding of the evolution of lipid membranes.**



## INTRODUCTION

Cells are separated from the surrounding environment by a cytoplasmic membrane composed of lipids and proteins. The typical bacterial lipid membrane consists of fatty acids (FAs) bound to a glycerol-3-phosphate (G3P) backbone via ester linkages, organized in a bilayer structure. Notably, however, some bacterial groups organize their membranes in a monolayer of membrane-spanning lipids (MSLs) formed by long-chain dicarboxylic acids that are linked to G3P through ester and, sometimes, ether bonds. Both MSL and ether bonds have been considered archetypical archaeal membrane features. Known bacterial MSLs are constituted by diabolic acids (DAs) and iso-diabolic acids (iso-DAs) (Fig. 1A and fig. S1). DAs have been encountered in members of the phyla Thermotogae and Firmicutes of the class Clostridia (*1*–*4*). iso-DAs occur in members of the genus *Thermoanaerobacter* [phylum Firmicutes, class Clostridia) (*5*), as well as in species of the subdivisions (SDs) 1, 3, 4, and 6 of the phylum Acidobacteria (*6*)]. DAs have been shown to be biosynthesized by tail-to-tail condensation of two $C_{16:0}$ FAs at the ω-1 positions (*7*). In a similar way, iso-DAs are believed to be produced by the condensation of two iso-$C_{15:0}$ FAs at the ω positions (*1*), but the enzymes responsible for the formation of these types of bacterial MSLs (i.e., MSL synthases) remain elusive. Recently, a radical *S*-adenosylmethionine (SAM) enzyme [tetraether synthase (Tes)] responsible for the archaeal tail-to-tail coupling of two ether-bound phytanyl chains, enabling the synthesis of isoprenoidal glycerol dialkyl glycerol tetraethers (GDGTs) (*8*, *9*). Because Tes homologs were detected in some bacterial genomes, they have been hypothesized to also be involved in the synthesis of bacterial MSLs (*8*).

Ether-bound membrane lipids are also a typical archaeal feature. Nevertheless, they have been found in some bacteria, sometimes together with MSLs (in Thermotogae and several Acidobacteria SDs). Non-isoprenoid alkyl glycerol ether lipids have been found in (hyper)thermophilic species of the bacterial phylum Thermotogae (*4*) in aerobic and facultative anaerobic mesophilic bacteria of the Acidobacteria SD1 and SD4 (*6*, *7*), in *Aquifex pyrophilus* (*10*), in *Ammonifex degensii* (Firmicutes, Clostridia) (*11*), in some *Planctomycetes* (*12*), and in some sulfate-reducing bacteria (*13*, *14*). In addition, alkenyl (1-alk-1′-enyl, vinyl) glycerol ether lipids, so-called plasmalogens, have been detected in nonthermophilic bacteria and suggested to play a role in cell resistance against environmental stresses (*15*, *16*). Enzymes involved in bacterial ether lipid biosynthesis have been found in select taxa. In myxobacteria, two independent pathways contributing to the biosynthesis of ether lipids have been identified: the gene product of Mxan_1676 coding for an alkylglycerone-phosphate synthase (*agps* gene) and the *elbB-elbE* gene cluster (*17*), which has also been detected in SD4 Acidobacteria (*6*). A gene encoding a plasmalogen synthase (*pls*A) has been identified in anaerobic bacteria (*18*). A modified form of *pls*A has been detected in *Thermotoga maritima* and in other bacteria producing ether-derived lipids (*19*) and proposed to be involved in the conversion of bacterial ester bonds into ether bonds generating saturated alkyl ethers.

The reason why bacteria synthesize membrane-spanning and ether lipids and how these features were evolutionarily acquired remain poorly understood, but it has been speculated that both the presence of ether bonds and a membrane organization based on a monolayer of MSLs confer membrane stability (*20*, *21*) as shown for archaeal isoprenoidal GDGTs (*22*). Determination of how bacterial ether lipids and MSLs are synthesized is important

[1]Department of Marine Microbiology and Biogeochemistry (MMB), NIOZ Royal Netherlands Institute for Sea Research, PO Box 59, Den Burg 1790 AB, Netherlands. [2]Laboratory of Biochemistry, Wageningen University and Research, Stippeneng 4, Wageningen 6708 WE, Netherlands. [3]Utrecht University, Faculty of Geosciences, Department of Earth Sciences, PO Box 80.021, Utrecht 3508 TA, Netherlands.
*Corresponding author. Email: diana.sahonero@nioz.nl

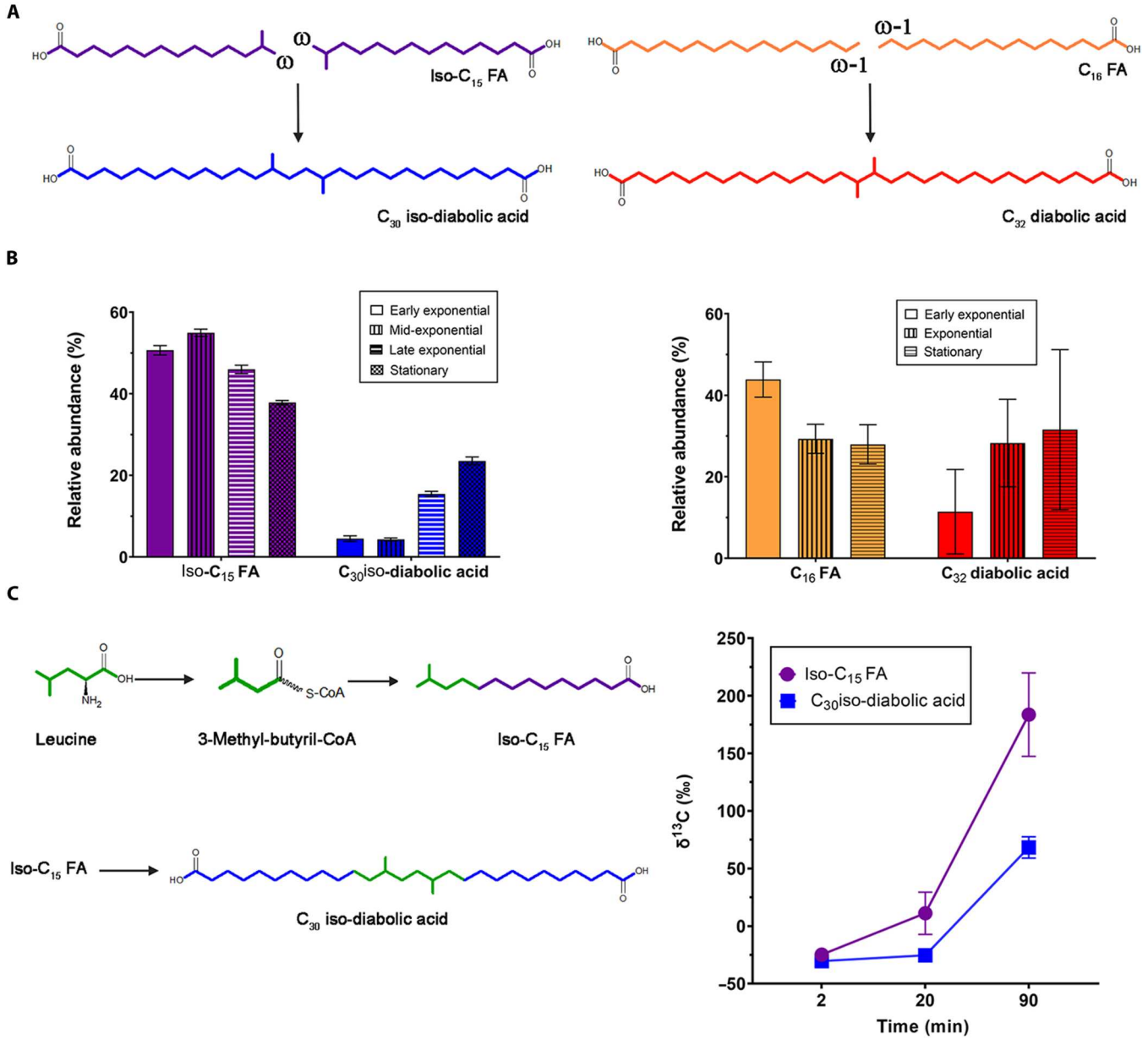


**Fig. 1. MSLs are produced via condensation of FAs.** (**A**) The $C_{30}$ iso-DA is thought to be formed by coupling two iso-$C_{15}$ FAs at their ultimate (ω) carbon atoms. The synthesis of $C_{32}$ DA proceeds through coupling of two *n*-$C_{16}$ FAs at the penultimate (ω-1) carbon atoms. (**B**) Relative abundance (%) of core lipids of iso-$C_{15}$ and iso-DA $C_{30}$ in cultures of *T. ethanolicus* grown at 60°C across growth phases (see table S1 and Supplementary Text) and of $C_{16}$ and DA $C_{32}$ in cultures of *T. maritima* grown at 80°C across growth phases (see fig. S2, table S3, and Supplementary Text for further details). In both experiments, the relative abundance of MSLs increases with growth. All experiments were performed in triplicate, and the error bars indicate ± SD. (**C**) Labeling experiment using $^{13}$C-labeled leucine added to cultures of *T. ethanolicus* led to the formation of labeled iso-$C_{15}$, and subsequently to labeled $C_{30}$ iso-DA in a time-course experiment (see table S2 and Supplementary Text). The degree of labeling is indicated by their $\delta^{13}$C values. The leucine-derived carbon atoms (in green) form a part of the carbon skeleton of the $C_{30}$ iso-DA. All experiments were performed in triplicate, and the error bars indicate ± SD.

to better understand how the divergence of lipid membranes or the "lipid divide," proceeded in all life forms. In addition, iso-DAs are thought to be the main precursors of the branched GDGTs (brGDGTs) (fig. S1) (*23*), which occur widespread in the environment and are widely used for paleoclimatological reconstruction (*24*). However, their biological producers remain unclear. Determining the biosynthetic pathway of bacterial (ether) MSL synthesis will allow for the detection of this capacity in other microbial groups.

Here, we identified and confirmed the activity of an MSL synthase in bacteria. In addition, we confirm the enzymatic activity of a plasmalogen synthase homolog that is involved in the formation of ether bonds in bacterial alkyl glycerol lipids. On the basis of phylogenetic analyses of these enzymes, we identified microbial

groups that have the potential to generate these membrane components and hypothesize how this feature was acquired, with evolutionary implications for understanding the acquisition of membrane lipids in all life forms.

## RESULTS AND DISCUSSION

### Iso-DA production via the condensation of iso-FAs

The biosynthesis of the MSL iso-DA is thought to proceed through the coupling of the tails of iso-branched FAs precursors (Fig. 1A) (*1*). Because growth temperature affects membrane stability and therefore changes in the proportion of MSL are expected, we performed culturing experiments with the iso-DA producer *Thermoanaerobacter ethanolicus* under different growth regimes and determined its lipid profile. The relative abundance of the dominant iso-$C_{15:0}$ FA decreased simultaneously with a marked increase in the $C_{30}$ iso-DA during growth at optimal 60°C temperature (Fig. 1B). A similar observation was made at suboptimal (45°C) temperature, albeit the maximum abundance of the $C_{30}$ iso-DA remained lower (14% versus 24% of total core lipids; table S1). The higher relative abundance of $C_{30}$ iso-DA during the stationary phase at optimal versus suboptimal growth temperature suggests that iso-DA production is regulated by temperature. This agrees with an increased production of GDGTs in archaea at higher temperatures (*25*). Our results strongly suggested that the formation of the $C_{30}$ iso-DA proceeds through a coupling of two iso-$C_{15:0}$ FAs. To confirm this substrate-product relationship, we performed incubations with *T. ethanolicus* with a labeled branched amino acid, $^{13}$C-leucine, required for the synthesis of 3-methylbutyryl–coenzyme A, a key building block of iso-$C_{15:0}$ FA (*26*). Label incorporation was detected in iso-$C_{15:0}$ FA but not in $C_{30:0}$ iso-DA 20 min after the addition of $^{13}$C-leucine, whereas after prolonged incubation (90 min) the label was also incorporated into the $C_{30}$ iso-DA (Fig. 1C and table S2). These results confirm that iso-$C_{15:0}$ FA acts as the precursor for $C_{30}$ iso-DA.

### The search for potential proteins for the biosynthesis of membrane-spanning and ether lipids in bacteria

Our results with *T. ethanolicus* indicate that the biosynthetic reaction leading to iso-DA is growth phase dependent (Fig. 1 and table S3), as recently also shown for the formation of the MSL DA in *T. maritima* [Fig. 1B, table S3, and (*19*)]. In both *T. ethanolicus* and *T. maritima*, the percentage of MSLs substantially increased during the stationary phase of growth, suggesting that experimental conditions allow for the detection of the activation of the genes coding for the proteins involved in the synthesis of MSLs. To test this, we analyzed the transcriptomic and proteomic response of these two bacterial species and compared them between different growth phases and at optimal and suboptimal growth temperatures (tables S4 to S14 and Supplementary Text). Supporting the usability of this approach, the gene encoding the modified *plsA* (Tmari_0479), suspected to be an alkyl ether lipid synthase based on protein homology (*19*), was indeed found to be up-regulated by >0.4-fold as revealed by the transcriptomic analysis (Table 1), coinciding with a higher proportion of alkyl ether lipids (table S14 and fig. S2). This is fully in line with its presumed role in the production of ether lipids.

We also searched for potential genes encoding MSL synthases. The anticipated biochemical mechanism for MSL synthesis is based on the dimerization of the FA building blocks through the formation of a carbon–carbon bond between either the ω-1 carbon of $C_{16}$ FAs or the ω carbon of iso-$C_{15:0}$ FAs to form DA or iso-DA, respectively (*1*, *7*). A radical reaction mechanism for the formation of MSLs in bacteria has been previously proposed (*27*) and recently confirmed for the synthesis of the isoprenoidal MSLs in archaea (*8*, *9*). This reaction would involve a radical intermediate formed by a hydrogen extraction at the tail of one of the FAs, followed by a condensation with the other tail, involving the loss of another hydrogen, resulting in the formation of a carbon–carbon bond in the absence of an activated intermediate. These unique reactions are commonly catalyzed by radical proteins, a group that shares an unusual Fe-S cluster associated with the generation of a free radical by the reduction of SAM (*28*). We performed homology searches (protein blast) using the confirmed Tes homolog of the archaeon *Methanococcus aeolicus* (Maeo_0574, ABR56159.1) as query (*8*) but did not detect any Tes homologs in the genomes of *T. ethanolicus* and *T. maritima*, indicating that other enzymes must be involved in MSL synthesis. On the basis of the above considerations, we defined selection criteria for the detection of potential MSL synthases in the pool of genes/proteins found to be activated either in *T. maritima* transcriptome or proteome or in the *T. ethanolicus* transcriptome (Table 1). The gene/protein that met any of the following criteria were selected: (i) code for a radical-SAM protein containing the cysteine-rich motif (CxxxCxxC) usually found in the active site of radical enzymes, (ii) code for an oxidoreductase using an [Fe4 4S] cluster acting on CH or $CH_2$ groups, (iii) code for a membrane-bound protein, as most of the proteins known to be involved in the formation of core membrane lipids are membrane-associated, and/or (iv) are homologous to proteins in other MSL producing bacteria (see figs. S3 to S5 tables S15 to S19, and Supplementary Text for details). This resulted in a list of potential genes encoding MSL synthases (tables S16 and S17).

### Confirmation of the activity of the potential MSL synthases and ether-lipid forming enzyme

To test the activity of the up-regulated radical proteins selected as potential MSL synthases in *T. maritima* and in *T. ethanolicus*, the genes were cloned in an inducible expression vector and expressed in *Escherichia coli* BL21 DE3. This *E. coli* strain produces phosphatidylglycerol intact polar lipids (*29*), thought to be the required building blocks for the MSL synthase of *T. maritima* (*19*). This *E. coli* strain does not produce the iso-FA building blocks for iso-DA but has *n*-$C_{16:0}$, *n*-$C_{18:1}$, and the cyclopropyl FAs, cy-$C_{17:0}$ and cy-$C_{19:0}$, as the most abundant FAs. The heterologous gene expression experiments led to the confirmation of the presence of a gene (named here *mss*), encoding an MSL synthase enabling the coupling of the tails of two FAs, in both *T. ethanolicus* and *T. maritima* genomes (see tables S16 to S18 and S20). Expression of *mss* of *T. ethanolicus* (WP_129545148.1) led to the formation of two dicarboxylic acids that were absent in the control experiment (Fig. 2A) but only when the experiments were performed under anaerobic conditions. The products were identified as $C_{33}$ and $C_{34}$ diacids containing one and two cyclopropyl moieties, respectively. Their formation can be envisaged by ω-ω coupling of two abundant FAs of the *E. coli* strain: an *n*-$C_{16:0}$ FA with a cy-$C_{17:0}$ FA and two cy-$C_{17:0}$ FAs, respectively (Fig. 2C, Supplementary Text, and figs. S6 and S7). Other diacids were also formed albeit in lower relative

**Table 1. Gene expression rates of the membrane-spanning lipid synthase (*mss*) and glycerol ester reductase (*ger*) in *T. ethanolicus* and *T. maritima* during growth phases and at optimal and supoptimal temperature of growth.** Expression changes are shown as $\log_2$ fold values for transcriptomic analyses. exp., exponential.

| | | | *T. ethanolicus* | | *T. maritima* | |
|---|---|---|---|---|---|---|
| **Gene name** | **Gene** | **Growth phase stage/temperature** | **60°C** | **45°C** | **80°C** | **55°C** |
| *Membrane-spanning lipid synthase* | *mss* | Mid*/early-exp to late*-exp/exponential | 2.87 | 2.12 | 0.68 | |
| | | Late-exp*/exponential to stationary | | 2.95 | | |
| | | Mid*/early-exp to stationary | 2.64 | 5.07 | | |
| *Glycerol ester reductase* | *ger* | Exponential to stationary | | | 0.55 | 0.48 |
| | | Early-exp to stationary | | | 0.74 | 0.42 |

*Growth phase conditions for *T. ethanolicus*.

abundance (fig. S8). Similarly, heterologous gene expression of *mss* of *T. maritima* (Tmari_0825) in *E. coli* resulted in the formation of $C_{32}$ and $C_{33}$ DAs only under anaerobic conditions (Fig. 2B and fig. S9). This confirmed that the expression of *T. maritima mss* gene product catalyzes the synthesis of $C_{32}$ and $C_{33}$ DA by joining two *n*-$C_{16:0}$ FAs and one *n*-$C_{16:0}$ FA with a cy-$C_{17:0}$ FA, respectively, at the ω-1 position (Fig. 2C).

Using the same experimental design, we also tested whether the expression of the modified-*plsA* gene (Tmari_0479 in *T. maritima*) led to the formation of ether lipids under aerobic and anaerobic conditions. Only under anaerobic conditions, the induction of the expression of Tmari_0479 in *E. coli* led to the detection of 1-alkyl glycerol monoether (Fig. 3 and fig. S10). A series of 1-alkyl glycerol monoethers, where the alkyl chains reflected the major FAs present in the *E. coli* host were also detected upon expression of the modified-*plsA* homolog (accession number SHJ90043) present in the genome of *Desulfatibacillum alkenivorans* (Fig. 3 and fig. S11), a nonplasmalogen ether lipid–producing bacterium (*14*). These experiments confirm the earlier proposed function of the modified plsA in converting an *sn*-1 ester bond into an *sn*-1 ether bond (*19*), with defined specificity toward the *sn*-1 position, we will further refer to this enzyme as glycerol ester reductase (Ger).

Our findings confirm the enzymatic activity of key enzymes in the production of what are considered to be unusual bacterial membrane-spanning ether lipids.

## The widespread occurrence of genes encoding MSL synthases

We screened selected genomes of bacteria with a confirmed the presence/absence of DA or iso-DA for the presence of homologs of the two confirmed MSL synthases and Tes. We performed protein Position-Specific Iterative Basic Local Alignment Search Tool (PSI)–BLAST searches and selected protein sequences with an *e* value of $\leq 1 \times 10^{-50}$ and identity of ≥30%, which were considered as homologs. These are stringent search criteria but required because MSL synthases belong to the radical SAM family, a large protein family containing many different functions. Less stringent criteria could lead to matches with homologs not involved in the formation of MSLs. The 107 bacterial genomes examined included species of the class Clostridia (phylum Firmicutes), members of the Thermotogales, members of different SDs of the phylum Acidobacteria, as well as species within the Proteobacteria, Chloroflexi, Verrucomicrobia, superphylum Fibrobacteres–Chlorobi–Bacteroidetes (FCB), Dictyoglomi, Planctomycetes-Verrucomicrobia-Chlamydiae (PVC) group, Aquificae, and Fusobacteria. These were selected on the basis of their reported membrane lipid composition (the presence/absence of MSLs and ether lipids; table S21) and new lipid analyses of bacterial strains of interest, which led to the detection of DA biosynthetic capability outside of the Thermotogales and Firmicutes, specifically in species within the Dictyoglomi (table S21).

Homologs of Tes were not detected in any of the genomes of these DA and iso-DA producers (table S21), indicating that another protein must be responsible for the biosynthesis of bacterial MSLs. Homologs of the *mss* gene were detected in all genomes of the DA-producing phyla Thermotogales and Dictyoglomi and of the MSL-producing species falling in the class Clostridia (mostly producing iso-DAs, except for *Sarcina ventriculi* and *Butyrivibrio fibrisolvens*, both synthesizing DA) (tables S21 and S22). This confirms that *mss* encodes MSL synthase, the enzyme that is responsible for the final step in the biosynthesis of MSLs. In the class Clostridia, the presence of an MSL synthase homolog did not always coincide with confirmed synthesis of MSLs. This may suggest that the identified MSL synthase homologs are not always functional. Alternatively, because we observe that MSL formation is regulated by growth conditions [e.g., up-regulation of the *mss* gene in *T. ethanolicus* across growth phases (Table 1)], the limited experimental setups (i.e., harvesting of the biomass at stationary phase and growth in rich media) may have prevented detection of MSLs in these strains, indicating that the production of MSLs is affected by factors such as culture conditions. No *mss* homologs were detected in the selected genomes of the phylum Acidobacteria, although almost all SD1, SD3, SD4, and SD6 species produce iso-DA (table S21) (*6*). This may be related to the almost exclusive aerobic nature of these strains; *mss* has so far only been detected in genomes of anaerobic bacteria and can only be expressed in *E. coli* under anaerobic conditions. Tes homologs were detected in only 3 of the 14 genomes of acidobacterial species known produce iso-DA, using our stringent search criteria (table S21). This would suggests that acidobacteria use an alternative pathway for MSL synthesis, which would represent a case of convergent evolution.

Alignment of the MSL synthases of all species with confirmed MSL production with the confirmed MSL synthases of *T. ethanolicus* and *T. maritima* revealed the close similarity of specific regions

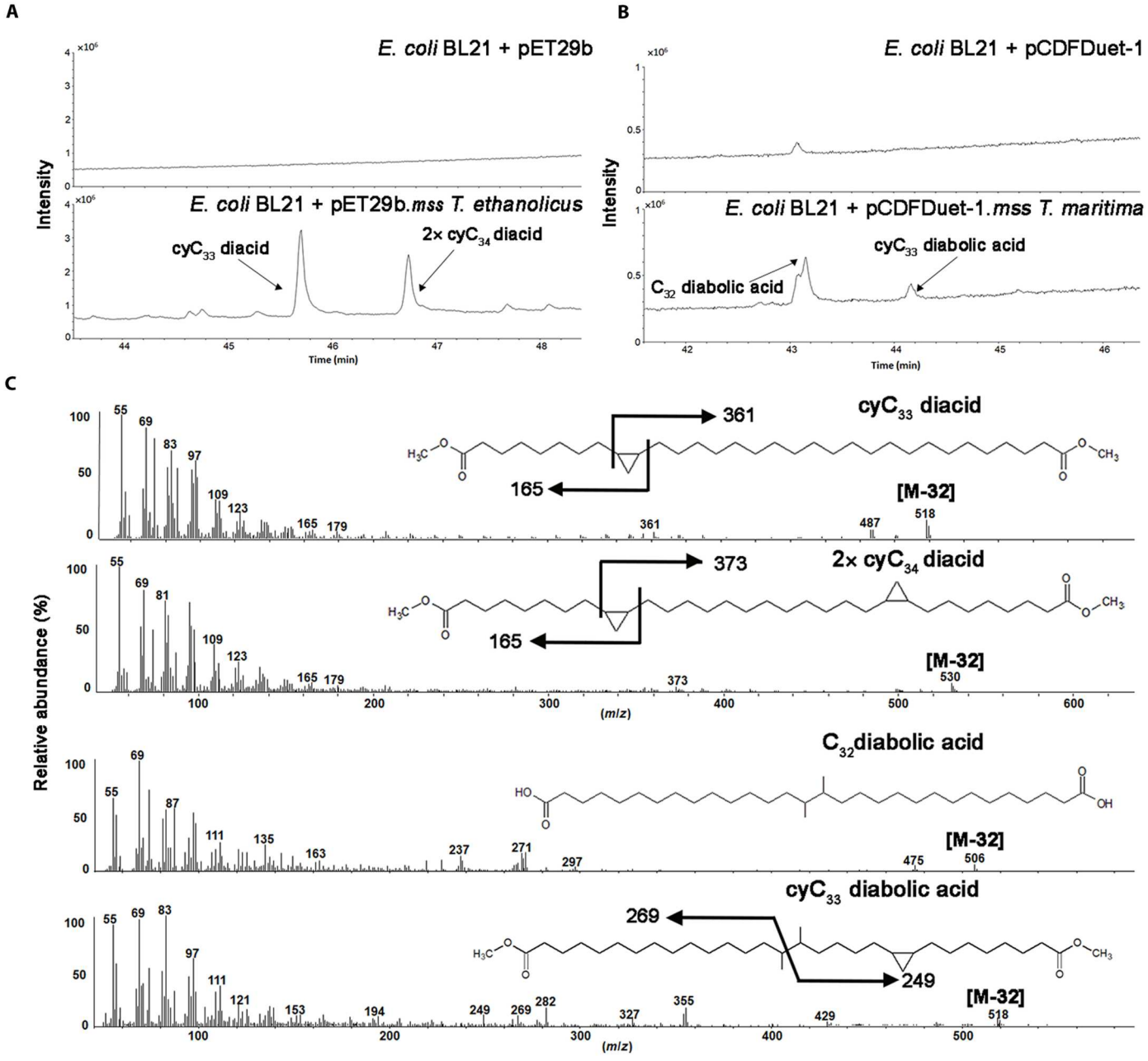


**Fig. 2. Expression of the bacterial MSL synthases of *T. ethanolicus* and *T. maritima* in *E. coli* results in the production of MSLs through condensation of the tails of two FAs at the ω and ω-1 positions, respectively.** (**A**) Partial gas chromatography chromatograms (44 to 48 min) of the base hydrolyzed lipid extract of *E. coli* BL21 (DE3) with "empty" pET29b plasmid (upper trace) or pET29b *mss* (WP_129545148.1) of *T. ethanolicus* plasmid (lower trace), revealing the formation of two major new components, which were identified as $C_{33}$ and $C_{34}$ mono- and bicyclic diacids formed by condensation at the ω positions of the cyclopropane-containing FAs produced by *E. coli*. (**B**) Partial gas chromatography chromatograms (42 to 46 min) of the base hydrolyzed lipid extract of *E. coli* BL21 (DE3) with the empty pCDFDuet-1 plasmid (upper trace) or pCDFDuet *mss* of *T. maritima* (lower trace) show the formation of the $C_{32}$ DAs by C—C bond formation between the ω-1 position of two $C_{16}$ FAs. (**C**) Mass spectra of the diacids and DAs formed. To confirm the presence of the cyclopropyl moieties, the compounds were hydrogenated but remained unaltered. *m*/*z*, mass/charge ratio.

when sequences were grouped according to iso-DA (Fig. 4A) and DA production (Fig. 4B). The overall alignment of all MSL synthases revealed six conserved blocks and hydrophobic regions near their C terminus (fig. S12). The novel enzymatic activity leading to the synthesis of MSL either for iso-DA or DA consists in joining the alkyl tails of two FAs, meaning that the active regions of these enzymes must be hydrophobic in nature. Modeling of the three-dimensional (3D) structures of the confirmed MSL synthases of *T. ethanolicus* (Fig. 4C) and *T. maritima* (Fig. 4D) revealed that the proposed hydrophobic region is localized close to the three conserved cysteines and the SAM binding region (radical SAM core), suggesting that this forms the reaction center for the coupling

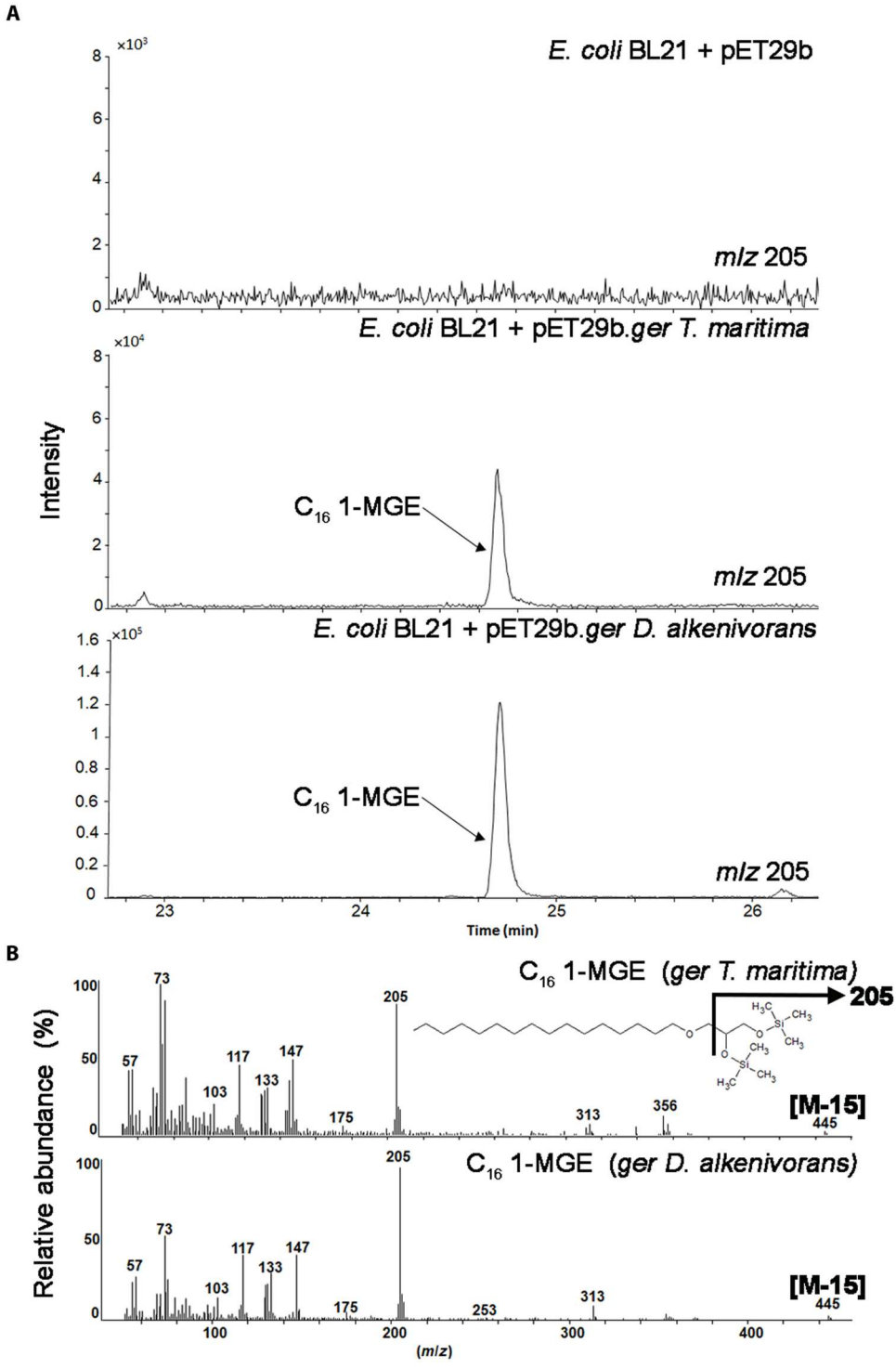


**Fig. 3. Expression of the modified PlsA of *T. maritima and D. alkenivorans* in *E. coli* results in the production of ether lipids.** (**A**) Partial mass chromatograms (*m/z* 205; 23 to 26 min) obtained by gas chromatography–mass spectrometry analysis of the base hydrolyzed lipid extracts of *E. coli* BL21 (DE3) with empty pET29b plasmid (upper trace), pET29b containing the *ger* genes of *T. maritima* (middle trace) (Tmari 0479), and of *D. alkenivorans* (SHJ90043) (bottom trace) revealing the formation of $C_{16}$ glycerol monoether (MGE or 1-*O*-hexadecyl glycerol) by *E. coli*. (**B**) Mass spectra of the $C_{16}$ glycerol monoether (1-MGE) formed.

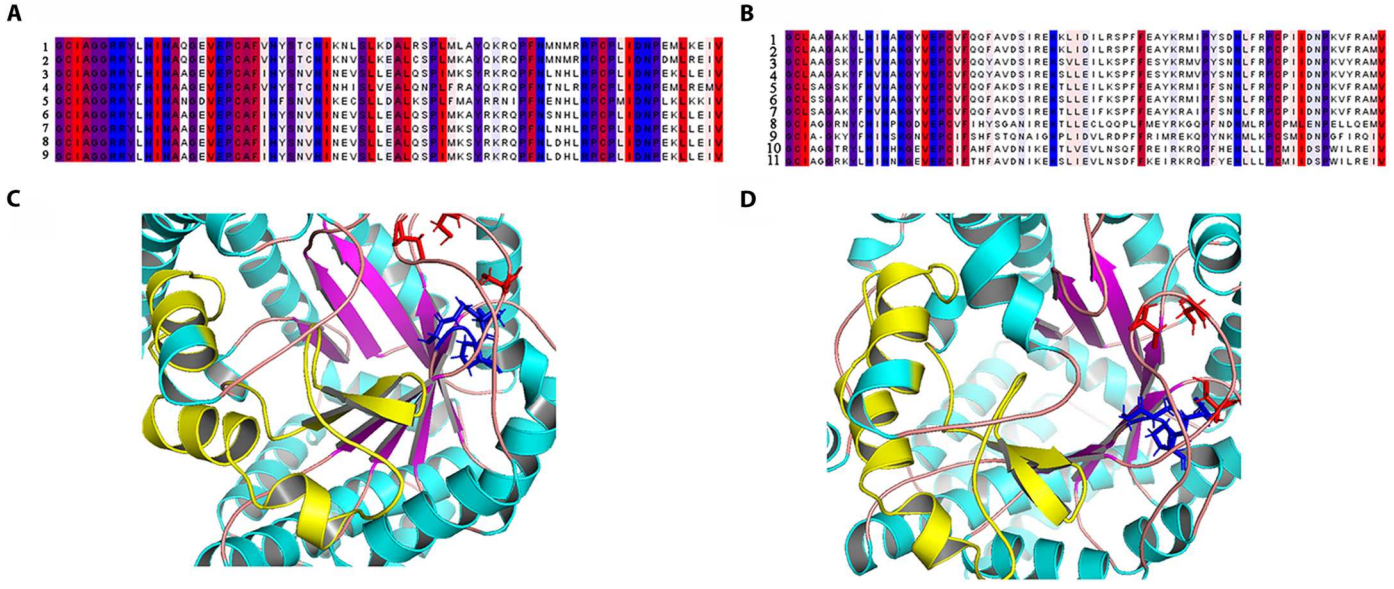


**Fig. 4. Protein alignment of a section of the two types of bacterial MSL synthases and three-dimensional structures of the MSL synthases from *T. ethanolicus* and *T. maritima*.** (**A**) MSLs producing iso-DAs. Key: 1, *Fervidicola ferrireducens*; 2, *Thermosediminibacter oceani*; 3, *Thermoanaerobacter wiegelii*; 4, *Moorella thermoacetica*; 5, *Caldicellulosiruptor owensensis*; 6, *Caldanaerobacter subterraneus*; 7, *Thermoanaerobacter siderophilus*; 8, *Thermoanaerobacter thermohydrosulfuricus*; 9, *T. ethanolicus*. (**B**) MSLs producing DAs. 1, *T. maritima*; 2, *T. neapolitana*; 3, *P. elfii*; 4, *P. hypogea*; 5, *T. africanus*; 6, *T. melanesiensis*; 7, *F. pennivorans*; 8, *B. fibrisolvens*; 9, *S. ventriculi*; 10, *Dictyoglomus thermophilum*; 11, *D. turgidum*. Predicted amino acid regions are colored according to hydrophobicity (*71*). The most hydrophobic residues are red, and the most hydrophilic are colored blue. Bottom: Modeled 3D structures of the MSL synthases of (**C**) *T. ethanolicus* and of (**D**) *T. maritima* zoomed in on the region of the conserved cysteines (in red, top side), the SAM binding motif (in blue, bottom side), and the region corresponding to the alignment proposed to interact with the lipid substrate (in yellow, left side).

of the two hydrophobic tails of the substrates. The key difference between the two types of MSL synthases (i.e., responsible for either DA and iso-DA production) is the lipid substrate (i.e., nonbranched or iso-FA, respectively) and the positions at which the two chains are connected (i.e., ω-1 or ω, respectively). The hydrophobic region of the two groups of MSL synthases likely controls the binding of specific lipid substrates and allows the hydrogen abstraction at the specific position (Fig. 4, C and D, and fig. S13). As would be expected for an enzyme forming iso-DAs (coupling ω carbon from iso-$C_{15}$ FAs), in our expression experiments with the *T. ethanolicus* MSL synthase, several long-chain diacids (Fig. 2, A and C, and fig. S8) were produced by the ω-ω coupling of two FAs, although only the nonbranched FAs produced by the *E. coli* host strain were available for the enzymatic reaction. Consistently, the formation of $C_{32}$ DA in *E. coli* upon expression of the MSL synthase from *T. maritima* confirms the expected enzymatic reaction between the ω-1 carbons from $C_{16}$ FAs. Moreover, only specific FA combinations resulted in the formation of diacids with *T. ethanolicus* MSL synthase, suggesting that only the tails of specific FAs could be accommodated in the hydrophobic region to allow the ω-ω coupling. Thus, the *T. ethanolicus* MSL synthase has a well-defined specificity, which is different from the MSL synthases involved in DA synthesis.

We extended the screening of MSL synthase homologs in the National Center for Biotechnology Information (NCBI) nonredundant protein sequence database using the stringent search criteria (DIAMOND search with an *e*-value of $\leq 1 \times 10^{-50}$ and query coverage of ≥30%) and constructed a maximum-likelihood phylogenetic tree. Homologs were detected widespread throughout the bacterial domain, specifically in species of the phyla Firmicutes, Actinobacteria, Thermotogae, Synergistetes, Spirochaetes, Armatimonadetes, Caldiserica, Calditrichaeota, Coprothermobacterota, Chloroflexi, Nitrospirae, Elusimicrobia, Deferribacteres, Dictyoglomi, Proteobacteria (gamma and delta), Atribacterota, superphylum FCB, Cyanobacteria/Melainabacteria group, PVC group, and in multiple members of the bacteria candidate phyla (see Fig. 5A and data S1), suggesting large-scale gene transfer across the bacterial domain. *Mss* gene homologs were detected in genomes of several classes within the phylum Firmicutes supporting the acquisition of the *mss* gene before the diversification of this phylum. MSL synthase sequences detected in bacteria candidate phyla, superphylum FCB, Spirochaetes, some Chloroflexi, and deltaproteobacteria are closely related, which might suggest that they were acquired by horizontal gene transfer as these groups often coexist in anoxic environments. Unexpectedly, we also detected homologs of MSL synthase in archaeal (meta)genomes of members of the Asgard and Diapherotrites, Parvarchaeota, Aenigmarchaeota, Nanohaloarchaeota, and Nanoarchaeota (DPANN) groups (Fig. 5A, data S1, and Supplementary Text). Biosynthesis of archaeal MSLs (i.e., biphytanes) by the ω-ω coupling of ether-bound phytanyl moieties is performed by Tes (*8*, *9*), a protein remotely related to the MSL synthases described here. In disagreement with the concept of the lipid divide, some archaeal genomes also encode for biosynthesis genes and ester-bond–forming transferases (*30*–*32*), in addition to the archaeal lipid biosynthesis gene machinery. This trait occurs in genomes of uncultured members of the superphylum Asgard (*32*, *33*). Because these archaea also encode a bacterial MSL synthase homolog, this suggests their potential to form MSL membranes. This hypothesis


**Fig. 5. The widespread capability of MSL and ether lipid production in the Bacteria and Archaea domain.** Maximum likelihood phylogenetic tree of the predicted homologs of the confirmed (**A**) MSL synthase and (**B**) ether bond-forming enzyme, Ger, across the tree of life grouped at the phylum or superphylum level as shown with the colors on the branches and indicated in the legend. Similar sequences were clustered on the basis of sequence identity before tree inference. Purple, green, and orange triangles indicate the genomes of microorganisms with homologs that have been seen to synthesize diabolic, iso-diabolic, or ether bonds in screened cultures listed in table S21. Colored circles in the tree indicate archaeal taxa. Full trees can be seen in the data S1 and S2. Scale bars represent the number of substitutions per site.

is exciting because of the position of Asgard archaea as the closest descendants of the archaeal ancestor leading to eukaryotes (*34*) and its significance for better understanding the acquisition of membranes in the membrane transition in eukaryogenesis. Further study is required to investigate these hypotheses and their evolutionary implications.

## The widespread occurrence of genes encoding the Ger enzyme

We screened the genomes of a set of selected bacterial strains, rigorously analyzed for the presence of ether lipids (table S21), for the presence of genes encoding homologs of the Ger proteins identified in *T. maritima* and *D. alkenivorans*. All strains producing alkyl ether-bond membrane lipids also harbored a Ger homolog in their genome, except for the members of the Acidobacteria, strongly supporting the role of Ger in the production of ether lipids in bacteria. The presence of a Ger homolog did not always coincide with confirmed production of alkyl ether lipids in culture (table S21), suggesting that their synthesis may be regulated by specific physiological factors, as discussed for MSL synthases (see above).

The functional domain architecture of the Ger enzyme of *T. maritima* is characterized by the presence of two activation domains identified by the Pfam domain PF01869 (or B; BcrAD_BadFG), one reduction domain PF09989 (or D; DUF2229), followed by four small (ca. 50 amino acids) domains, two reduction (D), and two dehydration domains PF06050 (or H; HGD-D) (BBDHDDH; fig. S14). In *D. alkenivorans*, the functional domain architecture is slightly different (i.e., BBDHDHDH; fig. S14). Analysis of the functional domain architecture of the Ger homologs in all strains with confirmed production of saturated alkyl ethers (table S21) indicated that the two activation and one reduction domains (BBD) are conserved among all alkyl ether lipid producers. However, many variations in the functional domain architecture were detected in the remaining part of the protein (table S21). More studies are needed to unravel which domains are required to enable a functional alkyl ether–synthesis protein.

We detected distant homologs of the gene encoding an alkylglycerone-phosphate synthase (*agps*) of *Myxococcus xanthus* (Mxan_1676), also involved in the biosynthesis of ether lipids (*17*), in some bacterial genomes also harboring Ger homologs (table S21). This suggests that these bacteria have more than one biosynthetic pathway for ether lipid production. Some members of the Acidobacteria SD4 have been proposed to use the *elb* gene cluster for production of ether lipids (*6*). In agreement with this, Ger homologs were not detected in the genomes of these SD4 species, but they were found in genomes of two acidobacteria, *Edaphobacter aggregans* (SD1) and in *Holophaga foetida* (SD8). Both species synthesize alkyl ether bonds (table S21), and of all acidobacterial strains known to synthesize ether lipids, *E. aggregans* and *H. foetida* are either facultatively anaerobic or strict anaerobic, in good agreement with our finding that the occurrence of Ger is limited to anaerobic bacteria. An exception is the detection of the Ger homolog in the genome of the anaerobic acidobacterium *Thermotomaculum hydrothermale*, which has not been reported to synthesize ether lipids (table S21). However, our hypothesis that facultative or strict anaerobic groups within the Acidobacteria could be producers of ether lipids using Ger is also further supported by the detection of Ger homologs in acidobacterial metagenomic assembled genomes from anoxic systems (see Supplementary Text).

We expanded the investigation of the presence of Ger homologs in prokaryotes using the NCBI nonredundant protein sequence database using stringent search criteria (DIAMOND search with an *e*-value of $\leq 1 \times 10^{-50}$ and a query coverage of $\geq$30%). Ger homologs were detected widely across the bacterial domain, i.e., in species of the phyla Firmicutes, Actinobacteria, Thermotogae, Synergistetes, Spirochaetes, Armatimonadetes, Calditrichaeota, Chloroflexi, Nitrospirae, Elusimicrobia, Deferribacteres, Proteobacteria, Atribacterota, superphylum FCB, Cyanobacteria/Melainabacteria group, PVC group, Nitrospinaea, Thermodesulfobacteria, Aquificae, Acidobacteria, and Chryosiogenetes and in multiple members of the bacteria candidate phyla (Fig. 5B and data S2). This suggests that ether lipid production occurs widespread through the bacterial domain. Ger homologs of Firmicutes are common across the Ger protein tree but often mostly closely related to those of Actinobacteria. In addition, multiple Ger homologs were detected in genomes of deltaproteobacteria, which suggests that the capacity to make ether-bonded lipids in members of this group is much more widespread than originally thought. Although archaea have specific enzymes to produce ether-bound phytanyl lipids [i.e., archaeal prenyl transferases (*35*)], we also encountered homologs of the Ger in genomes of members of the Euryarchaeota (classes Methanosarcinales and Methanomicrobiales), of DPANN archaea (Ca. Woesearchaeota, Pacearchaeota), the Asgard group [Ca. Lokiarchaeota, including *Candidatus* Prometheoarchaeum syntrophicum (*36*)], and the Ca. Heimdallarchaeota (Fig. 5A). In several of these archaeal groups, we have also detected the presence of potential MSL synthase homologs, giving further support to the possibility that specific archaea might have the genetic ability to synthesize (ether-bonded) bacterial-like MSLs.

In conclusion, we have identified the genes responsible for the synthesis of MSL in bacteria and confirmed that the modified-PlsA protein (Ger) is responsible for the conversion of alkyl ester to ether bond alkyl ether lipids in bacteria. The genomes of many bacterial taxa contain homologous genes encoding MSL synthase and Ger. This indicates that the capability of the synthesis of ether bonds and membrane-spanning lipids, traits that once were thought to be characteristic for the archaeal domain, occurs widespread in the bacteria domain. This has important consequences for our understanding of the so-called lipid divide, the idea that membrane lipids of archaea and bacteria/eukaryotes are fundamentally different. Both archaea and bacteria are able to produce membrane-spanning and ether lipids, but the enzymes involved in the synthesis of bacterial (ether-based) MSLs are different from those forming GDGTs in archaea [i.e., Tes enzyme (*8*, *9*) and archaeal prenyl transferases (*35*)], indicating that similar membrane lipid features have been acquired independently in the evolution of lipid membranes but likely emerged due to the same evolutionary pressure (e.g., increase of membrane stability). The different stereochemistry of the glycerol moiety of their membrane lipids and nature of their building blocks (isoprenoidal versus linear/branched alkyl chains) remain discriminating characteristics between bacterial and archaeal lipids, although some archaeal groups have acquired bacterial genes encoding enzymes involved in the biosynthesis of bacterial membrane lipids (including the genes encoding MSL synthase and Ger), likely through lateral gene transfer. The potential synthesis of bacterial-type (ether) MSLs by archaeal members of the Asgard group, together with archaeal lipids, could be an adaptation leading to the hypothetical membrane transition from archaeal to bacterial type during eukaryogenesis. These findings warrant follow-up studies to determine which physiological/environmental conditions trigger the formation of MSL and ether lipids and when and how this evolutionary trait was acquired and retained.

Our findings have also wide implications for our understanding of the biological sources of bacterial brGDGTs (based on iso-DA), which are extensively used for paleoclimate reconstruction (*24*). Acidobacteria have long been thought to be the biological sources of brGDGTs in specific environments [e.g., (*20*)], which have been detected in a few species of Acidobacteria (*6*). However, several studies have pointed to multiple other bacterial sources of brGDGTs, especially in anoxic settings, where most MSL-producing acidobacteria cannot thrive (*37*). In the current study, we predict the synthesis of bacterial ether-bound MSLs by bacterial groups outside of the Acidobacteria, as well as yet-uncultured (facultative) anaerobic acidobacteria, being those commonly present in soils, freshwater, and marine systems. This information will be essential to further constrain the interpretations based on brGDGT distributions.

## MATERIALS AND METHODS

### Strains, media, and growth conditions

*T. maritima* [strain MSB8, DSMZ (Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH), DSM 3109)] was cultivated in basal media under anaerobic conditions in 120-ml batch cultures in 250-ml serum bottles. Cultures were incubated either at optimal (80°C) or lower (55°C) growth temperatures, and after three passages, the batch cultures were inoculated from the acclimated bottle. Growth was monitored by measuring the optical density at 600 nm ($OD_{600}$) and by fluorescent microscopy. Samples for proteomic and transcriptomic analysis were derived from five biological replicates at each condition, for each analysis. Lipid analysis was derived from three biological replicates. The cells were harvested at early exponential, exponential, and stationary phases by centrifugation at 3500 rpm for 10 min at 4°C. The

supernatant was discarded, and the remaining pellet was immediately frozen at −20°C until further processing for lipid, protein, or RNA extraction.

*T. ethanolicus* JW200 (DSM 2246) (*38*) was cultivated under anaerobic conditions in 80-ml batch cultures in 250-ml serum bottles. The medium was composed of $(NH_4)_2SO_4$ (1.3 g/liter), $KH_2PO_4$ (0.375 g/liter), $K_2HPO_4$ (0.75 g/liter), $MgCl_2 \times 6H_2O$ (0.4 g/liter), $CaCl_2 \times 2H_2O$ (0.13 g/liter), $FeSO_4 \times 7H_2O$ [0.001% (w/v) in $H_2SO_4$ 2 mM], yeast extract (4 g/liter), and resazurin standard stock solution 100×, and the pH was adjusted to 6.7 with 5 M NaOH. The medium was anaerobically dispensed in 250-ml serum bottles, and a gas phase of $N_2$ was applied. After sterilization, individual bottles were supplemented with 20× cellobiose solution (100 g/liter) and 100× dilution of L-cysteine hydrochloride monohydrate stock. Cultures were grown at 45° or 60°C for 1 subcultivation growth cycle to ensure acclimatization. Growth was monitored by measuring $OD_{600}$. Cells were harvested at the early exponential ($OD_{600}$ of 0.25), mid-exponential ($OD_{600}$ of 0.4 to 0.5), late-exponential ($OD_{600}$ 0.65 to 0.73), and stationary ($OD_{600}$ 0.62 to 0.67) growth phases. The experiments were performed in triplicate. At given $OD_{600}$, cultures of 80 ml were split in two and immediately centrifuged 15 min at 4600*g* at 4°C. Half of the cultures were used for RNA sequencing, the other half were used for lipid extractions.

For the time-course $^{13}C$ labeling experiments with *T. ethanolicus* with labeled leucine, cells were anaerobically grown in 40 ml of media with either L-leucine or $^{13}C_6$-labeled leucine (Merck Sigma-Aldrich) added to a final concentration of 0.191 mM in the growth media, and samples were collected after 2, 20, and 90 min of incubation at 55°C and harvested according to the above-described procedure.

Strains of the phylum Firmicutes, class Clostridia analyzed for their lipid composition were cultured by DSMZ with their preferred media to stationary phase. Cultures were harvested by centrifugation and cell pellets freeze-dried before lipid analysis.

### Lipid extraction and analysis

FAs and MSLs (DA and iso-DA) were extracted from freeze-dried pellets and analyzed as previously described (*19*). Core lipids were identified on the basis of literature data and library mass spectra. The double-bond positions present in the identified diacids were determined by derivatization with dimethyl disulfide (*39*) with some modifications (incubation at 40°C and then overnight, addition of 400 μl of *n*-hexane and 200 μl of 5% solution $Na_2S_2O_3$, aqueous layer extracted two times with *n*-hexane), before mass spectrometric identification by gas chromatography–mass spectrometry as described in (*19*).

For the time-course $^{13}C$-labeling experiments with leucine, harvested cells of *T. ethanolicus* were frozen and freeze-dried. Lyophilized cells were hydrolyzed with 1.5 N HCl in methanol by refluxing for 3 hours, and the pH was adjusted to neutral by KOH addition. The FAs and iso-DA in the obtained extracts were derivatized to their methyl esters using $BF_3$-methanol solution and analyzed by Gas Chromatography isotope ratio monitoring Mass Spectrometry (GC-irmMS) as previously described (*40*, *41*).

### RNA extraction and transcriptomic analysis

*T. maritima* cell pellets were defrosted on ice and washed with 500 μl of deoxyribonuclease (DNase)/ribonuclease (RNase)–free water (New England Biolabs). Cells were resuspended in 700 μl of RLT buffer, and 50 mg of acid-washed glass beads (0.1 μm in diameter) was added to a safe-lock tube containing the cell suspension. Cell disruption was performed with the Omni bead mill homogenizer (6.3 m/s) 2×. After centrifugation and separation of the cell lysate, the RNA was extracted with the RNeasy Mini Kit (QIAGEN). The RNA concentration was determined using the Qubit HS protocol (Thermo Fisher Scientific). Eighteen microliters (2 μg) of total RNA was hybridized with the Pan-Prokaryote riboPOOL 3′-biotinylated probe (siTOOLs Biotech). The ribosomal RNA (rRNA) was subsequently depleted with hydrophilic streptavidin magnetic beads (New England Biolabs). The depleted rRNA sample was purified before library preparation with RNA Clean & Concentrator (Zymo Research) and eluted in 15 μl of DNase/RNase-free water. Samples were quality control–analyzed with the Agilent RNA 6000 Nano Chips and the Agilent 2100 Bioanalyzer system and stored at −80°C until library preparation.

For *T. ethanolicus*, cells cleared from the medium were snap-frozen in dry ice and stored at −80°C until usage. For RNA extraction, the frozen cell pellets were resuspended in 4 ml of RNAlater (Invitrogen) while maintained on ice. A 200-μl sample was taken from the cell suspension for RNA isolation. The cells were incubated at room temperature for 5 min. The cells were then pelleted for 10 min at 4700*g*. The cell pellet was resuspended in 100 μl of Tris-EDTA (TE) buffer (pH 7.5). Five hundred microliters of TRIzol reagent was added and mixed. The subsequent suspension was transferred to a precooled 2-ml screw-cap tube supplemented with 250-μl sterile 0.1- and 1-mm glass beads. Cells were lysed by beating six rounds of 25-s pulse at 6500 rpm and 5-min pause on watery ice. Ice-cold chloroform (200 μl) was subsequently added and mixed by shaking. Two phases were produced by centrifugation for 15 min at 12,000*g*. The transparent aqueous phase was transferred to a new RNAse-free precooled tube without disturbing the interphase. Ethanol was added at 1 volume and mixed. The isolate was transferred to an RNeasy column and centrifuged 15 s at 8000*g* to bind nucleic acids to the column. The column was rinsed with 350 μl of RLT buffer by centrifugation. An 80-μl RNAse-free DNAse reaction mixture was added to the column and incubated 20 min at 30°C to digest DNA contaminants. The column was rinsed with 350 μl of RW1, incubated 5 min, and centrifuged. The column was then washed with 700 μl of RPE, centrifuged, and washed with 80% ethanol. The column was dried by a 2-min spin at 21,000*g*. To elute the RNA, 35 μl of nuclease-free water was added and eluted by 1-min centrifugation at 8000*g*. For rRNA depletion, 18 μl (1 μg) of total RNA was hybridized with the Pan-Prokaryote riboPOOL 3′ biotinylated probe following the depletion protocol and stored at −80°C until library preparation.

For *T. maritima*, we sequenced 30 RNA samples, five biological replicates across three growth phases (early exponential, exponential, and stationary) and two temperature conditions (55° and 80°C). For *T. ethanolicus*, we sequenced 18 RNA samples, three biological replicates across three growth phases (mid-exponential, late exponential, and stationary), and two temperature conditions (45° and 60°C). rRNA-depleted RNA samples were used to prepare sequencing libraries with the TruSeq RNA stranded kit and sequenced at the Utrecht Sequencing Facility (USeq, The Netherlands) on an Illumina NextSeq 500 sequencing platform in single-end mode with a read length of 75 nucleotides.

Both transcriptomics libraries were treated equally unless specified otherwise. FASTQ files containing the Illumina reads were

quality-filtered, and standard Illumina adapters were removed using Trimmomatic v0.36 (*42*) (parameter settings: IlluminaClip: TruSeq3-SE.fa:2:30:10, leading: 20, trailing: 20, sliding window: 5: 20 min length: 40); poly-G tails were removed using cutadapt v1.16 with the --nextseq-trim paremeter set to 20 (*43*). Gene counts were calculated using Salmon v1.1.0 (*44*) for the *T. maritima* MSB8 library and Salmon v1.3 for the *T. ethanolicus* JW 200, both with the mapping-based mode against the reference genome, being *T. maritima* MSB8 (*45*) (NCBI Reference Sequence: NC_021214.1) and *T. ethanolicus* JW 200 (NCBI Reference Sequence: CP033580.1), respectively. Quantification estimates were imported into R/Bioconductor with tximport (*46*), and differential gene expression between growth phases and temperatures was assessed with DESeq2 v1.26.0 (*47*) using the default values. Genes with an adjusted *P* value of ≤0.05 were considered to demonstrate significant differential gene expression between the indicated sample groups. We report $\log_2$ fold change values for these significantly up-regulated (>0.2) or down-regulated (≥0.2) transcripts. The RNA sequencing (RNA-seq) data have been deposited in NCBI's Gene Expression Omnibus (GEO) (*48*) and are accessible through GEO Series accession number GSE211548.

### Protein extraction and proteomic analysis

After harvesting the cultures, the *T. maritima* frozen cell pellets were defrosted on ice, washed twice with 10 ml of 50 mM (tris, pH 8), and resuspended in 100 μl of the same buffer. Cell suspensions were sonicated for 20 s (4×) to lyse the cells. Cell debris and unbroken cells were removed by centrifugation at 10,000 rpm for 10 min at 4°C. The cell-free protein extracts were transferred to a 1.5-ml LoBind tube (Eppendorf) for further processing. Protein concentrations were determined with the Qubit protocol for protein quantification (Thermo Fisher Scientific). Proteins (60 μg) were loaded to 12% Mini-PROTEAN TGX Precast Protein Gels, 10 wells, 50 μl (Bio-Rad) and run for 20 min at 120 V. Proteins were visualized by staining the gels for 3 hours at room temperature with the Colloidal Blue Staining Kit (Thermo Fisher Scientific), washed with ultrapure (UP) water, and destained for 18 hours in UP water. Disulfide bridges were reduced with 20 mM dithiothreitol in 50 mM ammonium bicarbonate (ABC) for 1 hour at room temperature. Gels were washed with UP water (3×), followed by alkylation with 20 mM acrylamide in ABC. Each gel lane containing the samples was cut individually and sliced into smaller pieces of ca. 1 $mm^2$ and transferred to 0.5-ml Protein LoBind tubes (Eppendorf). Samples were incubated at room temperature for 15 hours in 200 μl of 0.05 ng/μl of trypsin solution. The enzymatic digestion was stopped and acidified by adding 10% trifluoroacetic acid until the pH decreased between 2 and 4. Peptides were extracted by loading the samples onto an activated clean-up μ-column containing two C18 Attract SPE (Affinisep) disks and ca. 2 mg of Lichoprep RP-18 (Merck) as column material (25 to 40 μm). The column was washed with 100 μl of HCOOH (1 ml/1iter) in water and eluted with 100 μl of acetonitrile: formic acid (1 ml/liter) in water (1:1). Samples were concentrated with an Eppendorf concentrator at 45°C for about 2 hours. The volume of each sample was adjusted to 50 μl and stored at −20°C until they were injected in the nLC 1000 (Thermo EASY nLC) tandem mass spectroscopy (MSMS) as described in (*49*). The MS/MS spectra were analyzed with MaxQuant 1.6.3.4 (*50*, *51*) with default settings for the Andromeda search engine completed by on-default variable modification settings for the deamidation on N and O. The protein sequence database for *T. maritima* MSB8 (downloaded from UniProt, March 2019) together with a contaminants custom database that contains sequences of common contaminants like trypsins (P00760, bovine and P00761, porcine) and human keratins [Keratin K22E (P35908), Keratin K1C9 (P35527), Keratin K2C1 (P04264), and Keratin K1CI (P35527)] was used to identify the protein's identities on the basis of the detected peptides. The "label-free quantification" and the "match between runs" options were enabled. Deamidated peptides were allowed to be used for protein quantification, and all other quantification settings were kept default. An intensity-based label-free quantification (LFQ) method (*52*) was used for statistical comparisons (*t* test) of normalized intensities of the protein groups between growth phases and temperature analysis. Filtering and statistical analyses were performed with Perseus v1.6.1. LFQ intensities were used for the analyses. For protein identification, groups were filtered to contain only proteins with at least two peptides of which at least one should be unique and at least one should be unmodified. Reverse hits and contaminants were filtered from the dataset. LFQ values were transformed to $\log_{10}$. For calculations, LFQ missing values were replaced by 6.8, a value slightly below the lowest measured value. A two-sample *t* test with a false discovery rate threshold set to 0.05 and significance: S0 = 0.05 were applied for comparisons. Proteins with a *P* value of ≤0.05 were considered to be significantly regulated between the sample groups. We report the protein abundance ratio ($\log_{10}$) values for those significantly up-regulated (>0.05) or down-regulated (>−0.05) proteins. The MS proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE (*53*) partner repository with the dataset identifier PXD035794.

### Selection of potential MSL synthases

The selection criteria for the detection of MSL synthases included those genes found to be activated either in the *T. maritima* transcriptome and/or proteome and in the *T. ethanolicus* transcriptome, respectively, following the criteria specified above. The selected genes and proteins for the MSL synthase were retrieved after screening their annotation and assigned biosynthetic pathways with Kyoto Encyclopedia of Genes and Genomes and UniProt. Searches for potential homologous proteins in *B. fibrisolvens* and *Clostridium ventriculi* were performed with the PSI-BLAST algorithm (*54*) at the protein level search using the UniProt ID from the selected radical or membrane proteins as a query (table S15). For *T. ethanolicus*, we performed a blast search of the candidates against NCBI non-redundant NR database (downloaded in November 2020) removing proteins belonging to any species from the *Thermoanaerobacter* genus (taxid: 1754). We achieved this by providing a negative sequence ID list (data S3), the search was done with a minimum *e*-value of 0.0001 and 50 maximum target sequences, and we kept the best hit for each query sequence.

We downloaded all genomes from species that are known to produce DAs from the PATRIC genome database (*55*) on 29 January 2021. Apart from *T. maritima* (*4*), other DA producers include *B. fibrisolvens*, *C. ventriculi*, *Fervidobacterium pennivorans*, *Pseudothermotoga elfii*, *Pseudothermotoga hypogea*, *Pseudothermotoga lettingae*, *Thermosipho africanus*, *Thermosipho melanesiensis*, and *Thermotoga neapolitana* (table S21). No genomes of *Pseudothermotoga subterranea* were present in the PATRIC database. For *T. maritima*, these include the assemblies (*45*,56)*46* on which the

GenBank and UniProt annotations are based and that were used for our transcriptomic and proteomic analyses [PATRIC ID 243274.17 (GCA_000390265.1_ASM39026v1) and 243274.5, respectively].

Completeness and contamination were estimated with CheckM v1.1.3 (*57*) in lineage_wf mode, and genomes whose completeness – 5× contamination was lower than 70% -- were excluded, as were plasmids. The resulting 50 genomes were annotated with Prokka v1.14.6 (*58*) with the --kingdom flag set to Bacteria. Orthologs were identified with Roary v13.3.0 (*59*), with the minimum percentage identity for blastp (−i) set to 10% and MCL inflation value (−iv) set to 1.5 to allow for sequence divergence of the highly diverse set of species, and the -s flag (do not split paralogs) set. Transmembrane helices were identified in the predicted proteins with TMHMM v2.0 (*60*). Proteins predicted by Prokka were linked to those in the GenBank and UniProt annotation files by bidirectional best blastp hit (the GenBank proteins to the proteins predicted from PATRIC identifier 243274.17 and the UniProt proteins to the proteins predicted from ID 243274).

### Recombinant production of candidate MSL synthases and Ger coding genes

To examine the potential enzymatic activity of the candidate MSL synthases and Ger coding genes (see tables S16 and S17 and Supplementary Text), they were commercially synthesized (Eurofins, Germany), subcloned in pET29b or pCDFDuet-1 and expressed in the *E. coli* BL21 DE3 strain. For the case of the potential MSL synthase genes (see tables S16 and S17), liquid cultures (50 ml) of exponentially growing *E. coli* harboring the empty expression vector or the vector including the coding gene were grown in 2× Yeast extract tryptone (YT) medium at 37°C and induced with 0.2 mM isopropyl-β-D-thiogalactopyranoside (IPTG) for 3 and 16 hours both anaerobically and aerobically. To test the activity of the potential modified-plsA genes (Ger-coding gene), liquid cultures of exponentially growing *E. coli* harboring an empty pET29b or pET29b-coding gene were grown in 2× YT media at 37°C, induced with 0.2 mM IPTG, and incubated at 25°C for 3 and 16 hours both anaerobically and aerobically. The expression of the proteins was verified using 7.5 or 12% Mini-PROTEAN TGX Precast Gels (Bio-Rad), stained with Colloidal Blue staining (Invitrogen).

### Homology searches of MSL synthases and Ger enzymes and phylogenetics

Confirmed MSL synthase of *T. ethanolicus* and Ger of *T. maritima* MSB8 were queried with DIAMOND v2.0.6 (*61*) against the NCBI non-redundant protein sequence database (NR) (*62*) downloaded on 7 January 2021. Hits with an *e*-value of $\leq 1 \times 10^{-50}$ and query coverage of ≥30% were selected. Proteins were clustered with cd-hit v4.8.1 (*63*) using a sequence identity threshold of 0.7, and representative sequences were aligned with Clustal Omega v1.2.4 (*64*). Gaps in the alignments were removed with trimAl v1.4.rev15 in -gappyout mode. Phylogenetic trees were constructed with IQ-TREE v2.1.2 (*65*) with 1000 ultrafast bootstraps. Model selection (*66*) was based on nuclear models, and the best-fit model was chosen according to Bayesian Information Criterion (BIC) (LG + R10 for both genes). Maximum-likelihood trees were visualized in Interactive Tree of Life (iTOL) (*67*).

### 3D model and protein domain analysis

Sequences of MSL synthases of iso-DA and DA producers were retrieved with BLASTP (*68*). The alignment was performed with Multiple Alignment using Fast Fourier Transform (MAFFT) (*69*) in the www.ebi.ac.uk/Tools/msa/mafft/ server using the BLOSUM62 substitution matrix, a gap open penalty of 1.53 and a gap extension of 0.12. The multiple-sequence alignment was edited with Jalview (*70*), and the amino acid regions forming the conserved blocks in all MSL synthases were retrieved by coloring 100% percentage identity of conservation between the proteins. Structure-based alignment showing conserved hydrophobic regions (*71*) between iso-DA and DA producers were colored on the basis of the percentage of conservation (90%) of the hydrophobic or hydrophilic residues. The secondary structure prediction showed in the alignment was performed with Jpred. The 3D models of the proteins were calculated with AlphaFold V2.1.0 with the Google Colab platform (https://colab.research.google.com/github/deepmind/alphafold/blob/main/notebooks/AlphaFold.ipynb), accessed in February 2022, with no templates and refined using the relax option. The resulting prediction was visualized using the PyMOL software (*72*).

## Supplementary Materials

**This PDF file includes:**
Supplementary Text
Figs. S1 to S14
Data S1 and S2
References (*73*–106)

**Other Supplementary Material for this manuscript includes the following:**
Data S3
Tables S1 to S22

View/request a protocol for this paper from *Bio-protocol*.

## REFERENCES AND NOTES


1. S. Jung, S. E. Lowe, R. I. Hollingsworth, J. G. Zeikus, *Sarcina ventriculi* synthesizes very long chain dicarboxylic acids in response to different forms of environmental stress. *J. Biol. Chem.* **268**, 2828–2835 (1993).
2. R. A. Klein, G. P. Hazlewood, P. Kemp, R. M. C. Dawson, A new series of long-chain dicarboxylic acids with vicinal dimethyl branching found as major components of the lipids of *Butyrivibrio* spp. *Biochem. J.* **183**, 691–700 (1979).
3. N. M. Carballeira, M. Reyes, A. Sostre, H. Huang, H. M. F. J. M. Verhagen, M. W. W Adams, Unusual fatty acid compositions of the hyperthermophilic archaeon *Pyrococcus furiosus* and the bacterium *Thermotoga maritima*. *J. Bacteriol.* **179**, 2766–2768 (1997).
4. J. S. Sinninghe Damsté, W. I. C. Rijpstra, E. C. Hopmans, S. Schouten, M. Balk, A. J. M. Stams, Structural characterization of diabolic acid-based tetraester, tetraether and mixed ether/ester, membrane-spanning lipids of bacteria from the order Thermotogales. *Arch. Microbiol.* **188**, 629–641 (2007).
5. S. Jung, J. G. Zeikus, A. Hollingsworth, A new family of very long chain α,ω-dicarboxylic acids is a major structural fatty acyl component of the membrane lipids of *Thermoanaerobacter ethanolicus* 39E. *J. Lipid Res.* **35**, 1057–1065 (1994).
6. J. S. S. Damsté, W. I. C. Rijpstra, B. U. Foesel, K. J Huber, J. Overmann, S. Nakagawa, J. J. Kim, P. F. Dunfield, S. N. Dedysh, L. Villanueva, An overview of the occurrence of ether- and ester-linked iso-diabolic acid membrane lipids in microbial cultures of the Acidobacteria: Implications for brGDGT paleoproxies for temperature and pH. *Org. Geochem.* **124**, 63–76 (2018).
7. W. Fitz, D. Arigoni, Biosynthesis of 15,16-dimethyltriacontanedioic acid (diabolic acid) from [16-$^2H_3$]- and [14-$^2H_2$]-palmitic acids. *Chem. Commun.* **20**, 1533–1534 (1992).
8. Z. Zeng, H. Chen, H. Yang, Y. Chen, W. Yang, X. Feng, H. Pei, P. Welander, Identification of a protein responsible for the synthesis of archaeal membrane-spanning GDGT lipids. *Nat. Commun.* **13**, 1–9 (2022).

9. C. T. Lloyd, D. F. Iwig, B. Wang, M. Cossu, W. W. Metcalf, A. K. Boal, S. J. Booker, Discovery, structure, and mechanism of a tetraether lipid synthase. *Nature* **609,** 197–203 (2022).
10. R. Huber, T. Wilharm, D. Huber, A. Trincone, S. Burggraf, H. Konig, R. Reinhard, I. Rockinger, H. Fricke, K. O. Stetter, *Aquifex pyrophilus* gen. nov. sp. nov., represents a novel group of marine hyperthermophilic hydrogen-oxidizing bacteria. *Syst. Appl. Microbiol.* **15**, 340–351 (1992).
11. R. Huber, P. Rossnagel, C. R. Woese, R. Rachel, T. A. Langworthy, K. O. Stetter, Formation of ammonium from nitrate during chemolithoautotrophic growth of the extremely thermophilic bacterium *Ammonifex degensii* gen. nov. sp. nov. *Syst. Appl. Microbiol.* **19**, 40–49 (1996).
12. J. S. Sinninghe Damsté, W. I. C. Rijpstra, J. A. Geenevasen, M. Strous, M. S. Jetten, Structural identification of ladderane and other membrane lipids of planctomycetes capable of anaerobic ammonium oxidation (anammox). *FEBS J.* **272**, 4270–4283 (2005).
13. T. A Langworthy, G. Holzer, J. G. Zeikus, T. G. Tornabene, Iso- and anteiso-branched glycerol diethers of the thermophilic anaerobe *Thermodesulfotobacterium commune*. *Syst. Appl. Microbiol.* **4**, 1–17 (1983).
14. V. Grossi, D. Mollex, A. Vinçon-Laugier, F. Hakil, M. Pacton, Mono- and dialkyl glycerol ether lipids in anaerobic bacteria: Biosynthetic insights from the mesophilic sulfate reducer *Desulfatibacillum alkenivorans* PF2803T. *Appl. Environ. Microbiol.* **81**, 3157–3168 (2015).
15. M. W. Ring, G. Schwär, V. Thiel, J. S. Dickschat, R. M. Kroppenstedt, S. Schulz, H. B. Bode, Novel iso-branched ether lipids as specific markers of developmental sporulation in the myxobacterium *Myxococcus xanthus*. *J. Biol. Chem.* **281**, 36691–36700 (2006).
16. H. Goldfine, The appearance, disappearance and reappearance of plasmalogens in evolution. *Prog. Lipid Res.* **49**, 493–498 (2010).
17. W. Lorenzen, T. Ahrendt, K. A. Bozhüyük, H. B. A. Bode, A multifunctional enzyme is involved in bacterial ether lipid biosynthesis. *Nat. Chem. Biol.* **10**, 425–427 (2014).
18. D. R. Jackson, C. D. Cassilly, D. R. Plichta, H. Vlamakis, H. Liu, S. B. Melville, R. J. Xavier, J. Clardy, Plasmalogen biosynthesis by anaerobic bacteria: Identification of a two-gene operon responsible for plasmalogen production in *Clostridium perfringens*. *ACS Chem. Biol.* **16**, 6–13 (2021).
19. D. X. Sahonero-Canavesi, L. Villanueva, N. J. Bale, J. Bosviel, M. Koenen, E. C. Hopmans, J. S. Sinninghe Damsté, Changes in the distribution of membrane lipids during growth of *Thermotoga maritima* at different temperatures: Indications for the potential mechanism of biosynthesis of ether-bound diabolic acid (membrane-spanning) lipids. *Appl. Environ. Microbiol.* **88**, e0176321 (2022).
20. J. W Weijers, S. Schouten, J. C. van den Donker, E. C. Hopmans, J. S. Sinninghe Damsté, Environmental controls on bacterial tetraether membrane lipid distribution in soils. *Geochem. Cosmochim. Acta* **71**, 703–713 (2007).
21. N. Jimenez-Rojo, H. Riezman, On the road to unraveling the molecular functions of ether lipids. *FEBS Lett.* **593**, 2378–2389 (2019).
22. Y. Koga, Thermal adaptation of the archaeal and bacterial lipid membranes. *Archaea* **2012**, 1–6 ( **2012**).
23. J. S. Sinninghe Damsté, E. C. Hopmans, R. D. Pancost, S. Schouten, J. A. J. Geenevasen, Geenevasen newly discovered non-isoprenoid dialkyl diglycerol tetraether lipids in sediments. *Chem. Commun.* **23**, 1683–1684 (2000).
24. S. Schouten, E. C. Hopmans, J. S. Sinninghe Damsté, The organic geochemistry of glycerol dialkyl glycerol tetraether lipids: A review. *Org. Geochem.* **54**, 19–61 (2013).
25. N. Nemoto, Y. Shida, H. Shimada, T. Oshima, A. Yamagishi, Characterization of the precursor of tetraether lipid biosynthesis in the thermoacidophilic archaeon *Thermoplasma acidophilum*. *Extremophiles* **7**, 235–243 (2003).
26. T. Kaneda, Iso- and anteiso-fatty acids in bacteria: Biosynthesis, function, and taxonomic significance. *Microbiol. Rev.* **55**, 288–302 (1991).
27. P. Galliker, O. Gräther, M. Rümmler, W. Fitz, D. Arigoni, New structural and biosynthetic aspects of the unusual core lipids from archaebacteria, in *Vitamin $B_{12}$ and $B_{12}$-Proteins: Lectures presented at the 4th European Symposium on Vitamin $B_{12}$ and $B_{12}$-Proteins*, B. Kräutler, D. Arigoni, B. Golding, Eds. (Wiley-VCH, 1998), pp. 447–458.
28. J. B. Broderick, B. R. Duffus, K. S. Duchene, E. M. Shepard, Radical *S*-adenosylmethionine enzymes. *Chem. Rev.* **114**, 4229–4317 (2014).
29. A. C. K. Teo, S. C. Lee, N. L. Pollock, Z. Stroud, S. Hall, A. Thakker, A. R. Pitt. T. R. Dafforn, C. M. Spickett, D. I. Roper, Analysis of SMALP co-extracted phospholipids shows distinct membrane environments for three classes of bacterial membrane protein. *Sci. Rep.* **12**, 1813 (2019).
30. A. Gattinger, M. Schloter, J. C. Munch, Phospholipid ether lipid and phospholipid fatty acid fingerprints in selected euryarchaeotal monocultures for taxonomic profiling. *FEMS Microbiol. Lett.* **213**, 133–139 (2002).
31. J. Lombard, P. López-García, D. Moreira, The early evolution of lipid membranes and the three domains of life. *Nat. Rev. Microbiol.* **10**, 507–515 (2012).
32. L. Villanueva, S. Schouten, J. S. Sinninghe Damsté, Phylogenomic analysis of lipid biosynthetic genes of Archaea shed light on the 'lipid divide'. *Environ. Microbiol.* **19**, 54–69 (2017).
33. L. Villanueva, F. A. B. von Meijenfeldt, A. B. Westbye, S. Yadav, E. C. Hopmans, B. E. Dutilh, J. S. Sinninghe Damsté, Bridging the membrane lipid divide: Bacteria of the FCB group superphylum have the potential to synthesize archaeal ether lipids. *ISME J.* **15**, 168–182 (2021).
34. A. Spang, J. H. Saw, S. L. Jørgensen, K. Zaremba-Niedzwiedzka, J. Martijn, A. E. Lind, R. Van Eijk, C. Schleper, L. Guy, T. J. Ettema, Complex archaea that bridge the gap between prokaryotes and eukaryotes. *Nature* **521**, 173–179 (2015).
35. Y. Koga, H. Morii, Biosynthesis of ether-type polar lipids in archaea and evolutionary considerations. *Microbiol. Mol. Biol. Rev.* **71**, 97–120 (2007).
36. H. Imachi, M. K. Nobu, N. Nakahara, Y. Morono, M. Ogawara, Y. Takaki, Y. Takano, K. Uematsu, T. Ikuta, M. Ito, Y. Matsui, M.Mizayaki, K. Murate, Y. Saito, S. Sakai, C. Song, E. Tasumi, Y. Yamanaka, T. Yamaguchi, Y. Kamagata, H. Tamaki, K. Takai, Isolation of an archaeon at the prokaryote-eukaryote interface. *Nature* **577**, 519–525 (2020).
37. Y. Weber, J. S. Sinninghe Damsté, J. Zopfi, C. De Jonge, A. Gilli, C. Schubert, F. Lepori, M. F. Lehmann, H. Niemann. H. Niemann, Redox-dependent niche differentiation provides evidence for multiple bacterial sources of glycerol tetraether lipids in lakes. *Proc. Natl. Acad. Sci. U.S.A.* **115**, 10926–10931 (2018).
38. Y. E. Lee, M. K. Jain, C. Lee, J. G. Zeikus, Taxonomic distinction of saccharolytic thermophilic anaerobes: Description of *Thermoanaerobacterium xylanolyticum* gen. nov., sp. nov., and *Thermoanaerobacterium saccharolyticum* gen. nov., sp. nov.; reclassification of *Thermoanaerobium brockii*, *Clostridium thermosulfurogenes*, and *Clostridium thermohydrosulfuricum* E100-69 as *Thermoanaerobacter brockii* comb. nov., *Thermoanaerobacterium thermosulfurigenes* comb. nov., and *Thermoanaerobacter thermohydrosulfuricus* comb. nov., respectively; and transfer of *Clostridium thermohydrosulfuricum* 39E to *Thermoanaerobacter ethanolicus*. *Int. J. Syst. Bacteriol.* **43**, 41–51 (1993).
39. P. Blokker, S. Schouten, H. van den Ende, J. W. de Leeuw, J. S. Sinninghe Damsté, Cell wall-specific ω-hydroxy fatty acids in some freshwater green microalgae. *Phytochemistry* **49**, 691–695 (1998).
40. N. J. Bale, W. I. C. Rijpstra, I. Y. Oshkin, S. E. Belova, S. N. Dedysh, J. S. Sinninghe Damsté, Fatty acid and hopanoid adaption to cold in the methanotroph *Methylovulum psychrotolerans*. *Front. Microbiol.* **10**, 589 (2019).
41. S. Schouten, W. C. M. K. Breteler, P. Blokker, N. Schogt, W. I. C. Rijpstra, K. Grice, M. Baas, J. S. S. Damsté, Biosynthetic effects on the stable carbon isotopic compositions of algal lipids: Implications for deciphering the carbon isotopic biomarker record. *Geochim. Cosmochim. Acta* **62**, 1397–1406 (1998).
42. A. M. Bolger, M. Lohse, B. Usadel, Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
43. M. Martin, Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.* **17**, 10–12 (2011).
44. R. Patro, G. Duggal, M. I. Love, R. A. Irizarry, C. Kingsford, Salmon provides fast and bias-aware quantification of transcript expression. *Nat. Methods* **14**, 417–419 (2017).
45. H. Latif, J. A. Lerman, V. A. Portnoy, Y. Tarasova, H. Nagarajan, A. C. Schrimpe-Rutledge, R. D. Smith, J. N. Adkins, D. H. Lee, Y. Qiu, K. Zengler, The genome organization of *Thermotoga maritima* reflects its lifestyle. *PLOS Genet.* **9**, e1003485 (2013).
46. C. Soneson, M. I. Love, M. D. Robinson, Differential analyses for RNA-seq: Transcript-level estimates improve gene-level inferences. *F1000Research* **4**, 1521 (2016).
47. M. I. Love, W. Huber, S. Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 1–21 (2014).
48. R. Edgar, M. Domrachev, A. E. Lash, Gene expression omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res.* **30**, 207–210 (2002).
49. Y. Liu, A. de Groot, S. Boeren, T. Abee, E. J. Smid, *Lactococcus lactis* mutants obtained from laboratory evolution showed elevated vitamin K2 content and enhanced resistance to oxidative stress. *Front. Microbiol.* **12**, 746770 (2021).
50. J. Cox, M. Y. Hein, C. A. Luber, I. Paron, N. Nagaraj, M. Mann, Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. *Mo. Cell. Proteomics* **13**, 2513–2526 (2014).
51. N. C. Hubner, A. W. Bird, J. Cox, B. Splettstoesser, P. Bandilla, I. Poser, A. Hyman, M. Mann, Quantitative proteomics combined with BAC TransgeneOmics reveals in vivo protein interactions. *J. Cell Biol.* **189**, 739–754 (2010).
52. C. Smaczniak, N. Li, S. Boeren, T. America, W. van Dongen, S. S. Goerdayal, S. de Vries, G. C. Angenent, K. Kaufmann, Proteomics-based identification of low-abundance signaling and regulatory protein complexes in native plant tissues. *Nat. Protoc.* **7**, 2144–2158 (2012).
53. J. A. Vizcaíno, A. Csordas, N. Del-Toro, J. A. Dianes, J. Griss, I. Lavidas, G. Mayer, Y. Perez-Riverol, F. Reisinger, T. Ternent, Q. Xu, R. Wang, H. Hermjakob, 2016 Update of the PRIDE database and its related tools. *Nucleic Acids Res.* **44**, D447–D456 (2016).

54. S. F. Altschul, W. Gish, W. Miller, E. W. Myers, D. J. Lipman, Basic local alignment search tool, *J. Mol. Biol.* **215**, 403–410 (1990).
55. J. J. Davis, A. R. Wattam, R. K. Aziz, T. Brettin, R. Butler, R. M. Butler, P. Chlenski, N. Conrad, A. Dickerman, E. M. Dietrich, J. L. Gabbard, S. Gerdes, A. Guard, R. W. Kenyon, D. Machi, C. Mao, D. Murphy-Olson, M. Nguyen, E. K. Nordberg, G. J. Olsen, R. D. Olson, J. C. Overbeek, R. Overbeek, B. Parrello, G. D. Pusch, M. Shukla, C. Thomas, M. Van Oeffelen, V. Vonstein, A. S. Warren, F. Xia, D. Xie, H. Yoo, R. Stevens, The PATRIC bioinformatics resource center: Expanding data and analysis capabilities: Expanding data and analysis capabilities. *Nucleic Acids Res.* **48**, D606–D612 (2019).
56. R. Singh, J. Gaddnigo, D. White, A. Lipzen, J. Martin, W. Schackwitz, E. Moriyama, P. Blum, Complete genome sequence of an evolved *Thermotoga maritima* isolate. *Genome Announc.* **3**, 9–10 (2015).
57. D. H. Parks, M. Imelfort, C. T. Skennerton, P. Hugenholtz, G. W. Tyson, CheckM: Assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res.* **25**, 1043–1055 (2015).
58. T. Seemann, Prokka: Rapid prokaryotic genome annotation. *Bioinformatics* **30**, 2068–2069 (2014).
59. A. J. Page, C. A. Cummins, M. Hunt, V. K. Wong, S. Reuter, M. T. Holden, M. Fookes, D. Falush, J. A. Keane, J. Parkhill, Roary: Rapid large-scale prokaryote pan genome analysis. *Bioinformatics* **31**, 3691–3693 (2015).
60. A. Krogh, B. Larsson, G. Von Heijne, E. L. L. Sonnhammer, Predicting transmembrane protein topology with a hidden Markov model: Application to complete genomes. *J. Mol. Biol.* **19**, 567–580 (2001).
61. B. Buchfink, K. Reuter, H. G. Drost, Sensitive protein alignments at tree-of-life scale using DIAMOND. *Nat. Methods* **18**, 266–268 (2021).
62. NCBI Resource Coordinators, Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res.* **44**, D7–D19 (2016).
63. L. Fu, B. Niu, Z. Zhu, S. Wu, W. Li, CD-HIT: Accelerated for clustering the next-generation sequencing data. *Bioinformatics* **28**, 3150–3152 (2012).
64. F. Sievers, A. Wilm, D. Dineen, T. J. Gibson, K. Karplus, W. Li, R. Lopez, H. McWilliam, M. Remmert, J. Söding, J. D. Thompson, D. G. Higgins, Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **7**, 539–539 (2011).
65. L. T. Nguyen, H. A. Schmidt, A. Von Haeseler, B. Q. Minh, IQ-TREE: A fast and effective stochastic algorithm for estimating maximum likelihood phylogenies. *Mol. Biol. Evol.* **32**, 268–274 (2015).
66. S. Kalyaanamoorthy, B. Q. Minh, T. K. F. Wong, A. Von Haeseler, L. S. Jermiin. ModelFinder: Fast model selection for accurate phylogenetic estimates. *Nat. Methods* **14**, 587–589 (2017).
67. I. Letunic, P. Bork, Interactive Tree of Life (iTOL) v3: An online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Res.* **44**, W242–W245 (2016).
68. S. F. Altschul, T. L. Madden, A. A. Schäffer, J. Zhang, Z. Zhang, W. Miller, D. J. Lipman, Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res.* **25**, 3389–3402 (1997).
69. K. Katoh, D. M. Standley, MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Mol. Biol. Evol.* **30**, 772–780 (2013).
70. A. M. Waterhouse, J. B. Procter, D. M. A. Martin, M. Clamp, G. J. Barton, Jalview Version 2—A multiple sequence alignment editor and analysis workbench. *Bioinformatics* **25**, 1189–1191(2009).
71. J. Kyte, R. F. Doolittle, A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* **157**, 105–132 (1982).
72. L. Schrödinger, W. DeLano, PyMOL (2020); www.pymol.org/pymol.
73. P. D. Karp, R. Billington, R. Caspi, C. A. Fulcher, M. Latendresse, A. Kothari, I. M. Keseler, M. Krummenacker, P. E. Midford, Q. Ong, W. K. Ong, S. M. Paley, P. Subhraveti, The BioCyc collection of microbial genomes and metabolic pathways. *Brief. Bioinf.* **20**, 1085–1093 (2019)
74. M. Kanehisa, S. Goto, M. Furumichi, M. Tanabe, M. Hirakawa, KEGG for representation and analysis of molecular networks involving diseases and drugs. *Nucleic Acids Res.* **38**, D355–D360 (2010).
75. I. Kolodkin-Gal, D. Romero, S. Cao, J. Clardy, R. Kolter, R. Losick, D-amino acids trigger biofilm disassembly. *Science* **328**, 627–629 (2010).
76. F. Cava, H. Lam, M. A. De Pedro, M. K. Waldor, Emerging knowledge of regulatory roles of D-amino acids in bacteria. *Cell. Mol. Life Sci.* **68**, 817–831(2011).
77. H. Lam, D. C. Oh, F. Cava, C. N. Takacs, J. Clardy, M. A. de Pedro, M. K. Waldor, D-Amino acids govern stationary phase cell wall remodeling in bacteria. *Science* **325**, 1552–1555 (2009).
78. J. Wang, C. Zhao, B. Meng, J. Xie, C. Zhou, X. Chen, K. Zhao, J. Shao, Y. Xue, N. Xu, Y. Ma, S. Liu, The proteomic alterations of *Thermoanaerobacter tengcongensis* cultured at different temperatures. *Proteomics* **7**, 1409–1419 (2007).
79. Z. Chen, B. Wen, Q. Wang, W. Tong, J. Guo, X. Bai, J. Zhao, Y. Sun, Q. Tang, Z. Lin, L. Lin, S. Liu, Quantitative proteomics reveals the temperature-dependent proteins encoded by a series of cluster genes in *Thermoanaerobacter tengcongensis*. *Mol. Cell. Proteomics* **12**, 2266–2277 (2013).
80. R. Huber, T. A. Langworthy, H. König, M. Thomm, C. R. Woese, U. B. Sleytr, K. O. Stetter, *Thermotoga maritima* sp. nov. represents a new genus of unique extremely thermophilic eubacteria growing up to 90°C. *Arch. Microbiol.* **144**, 324–333 (1986).
81. P. Londei, S. Altamura, R. Huber, K. O. Stetter, P. Cammarano, Ribosomes of the extremely thermophilic eubacterium *Thermotoga maritima* are uniquely insensitive to the miscoding-inducing action of aminoglycoside antibiotics. *J. Bacteriol.* **170**, 4353–4360 (1988).
82. C. L. Nesbø, K. S. Swithers, H Dahle, T. H. Haverkamp, N. K. Birkeland, T. Sokolova, I. Kublanov, O. Zhaxybayeva, Evidence for extensive gene flow and *Thermotoga* subpopulations in subsurface and marine environments. *ISME J.* **9**, 1532–1542 (2015).
83. J. DiRuggiero, N. Santangelo, Z. Nackerdien, J. Ravel, F. T. Robb, Repair of extensive ionizing-radiation DNA damage at 95°C in the hyperthermophilic archaeon *Pyrococcus furiosus*. *J. Bacteriol.* **179**, 4643–4645 (1997).
84. Z. Wang, W. Tong, Q. Wang, X. Bai, Z. Chen, J. Zhao, N. Xu, S. Liu, The temperature dependent proteomic analysis of *Thermotoga maritima*. *PLOS ONE* **7**, e46463 (2012).
85. J. Gautam, Z. Xu, Construction and validation of a genome-scale metabolic network of *Thermotoga* sp. strain RQ7. *Appl. Biochem. Biotechnol.* **193**, 896–911 (2021).
86. E. T. Michelini, G. C. Flynn, The unique chaperone operon of *Thermotoga maritima*: Cloning and initial characterization of a functional Hsp70 and small heat shock protein. *J. Bacteriol.* **181**, 4237–4244 (1999).
87. T. Holder, C. Basquin, J. Ebert, N. Randel, D. Jollivet, E. Conti, G. Jekely, F. Bono, Deep transcriptome-sequencing and proteome analysis of the hydrothermal vent annelid *Alvinella pompejana* identifies the CvP-bias as a robust measure of eukaryotic thermostability. *Biol. Direct* **8**, 1–16 (2013).
88. K. B. Zeldovich, I. N. Berezovsky, E. I. Shakhnovich, Protein and DNA sequence determinants of thermophilic adaptation. *PLOS Comput. Biol.* **3**, 62–72 (2007).
89. O. Zhaxybayeva, K. S. Swithers, P. Lapierre, G. P. Fournier, D. M. Bickhart, R. T. DeBoy, K. E. Nelson, C. L. Nesbø, W. F. Doolittle, J. P. Gogarten, K. M. Noll, On the chimeric nature, thermophilic origin, and phylogenetic placement of the Thermotogales. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 5865–5870 (2009).
90. S. M. Pollo, O. Zhaxybayeva, C. L. Nesbø, Insights into thermoadaptation and the evolution of mesophily from the bacterial phylum *Thermotogae*. *Can. J. Microbiol.* **61**, 655–670 (2015).
91. R. Keto-Timonen, N. Hietala, E. Palonen, A. Hakakorpi, M. Lindström, H. Korkeala, Cold shock proteins: A minireview with special emphasis on Csp-family of enteropathogenic *Yersinia*. *Front. Microbiol.* **7**, 1–7 (2016).
92. H. J. Sofia, G. Chen, B. G. Hetzler, J. F. Reyes-Spindola, N. E. Miller, Radical SAM, a novel protein superfamily linking unresolved steps in familiar biosynthetic pathways with radical mechanisms: Functional characterization using new analysis and information visualization methods *Nucleic Acids Res.* **29,** 1097–1106 (2001).
93. C. L. Hemme, H. Mouttaki, Y. J. Lee, G. Zhang, L. Goodwin, S. Lucas, A. Copeland, A. Lapidus, T. Glavina del Rio, H. Tice, E. Saunders, T. Brettin, J. C. Detter, C. S. Han, S. Pitluck, M. L. Land, L. J. Hauser, N. Kyrpides, N. Mikhailova, Z. He, L. Wu, J. D. Van Nostrand, B. Henrissat, Q. He, P. A. Lawson, R. S. Tanner, L. R. Lynd, J. Wiegel, M. W. Fields, A. P. Arkin, S. W. Schadt, B. S. Stevenson, M. J. Mclnerney, Y. Yang, H. Dong, D. Xing, N. Ren, A. Wang, R. L. Huhnke, J. R. Mielenz, S. Y. Ding, M. E. Himmel, S. Taghavi, D. van der Lelie, E. M. Rubin, J. Zhou, Sequencing of multiple clostridial genomes related to biomass conversion and biofuel production. *J. Bacteriol.* **192**, 6494–6496 (2010).
94. K. E. Nelson, R. A. Clayton, S. R. Gill, M. L. Gwinn, R. J. Dodson, D. H. Haft, K. Hickey, J. D. Peterson, W. C. Nelson, K. A. Ketchum, L. McDonald, T. R. Utterback, J. A. Malek, K. D. Linher, M. M. Garrett, A. M. Stewart, M. D. Cotton, M. S. Pratt, C. A. Phillips, D. Richardson, J. Heidelberg, G. G. Sutton, R. D. Fleischmann, J. A. Eisen, O. White, S. L. Salzberg, H. O. Smith, J. C. Venter, C. M. Fraser, Evidence for lateral gene transfer between Archaea and Bacteria from genome sequence of *Thermotoga maritima*. *Nature* **399**, 323–329 (1999).
95. Y. M. Zhang, C. O. Rock, Membrane lipid homeostasis in bacteria. *Nat. Rev.Microbiol.* **6**, 222–233 (2008).
96. M. S. Blevins, D. R. Klein, J. S. Brodbelt, Localization of cyclopropane modifications in bacterial lipids via 213 nm ultraviolet photodissociation mass spectrometry. *Anal. Chem.* **91**, 6820–6828 (2019).
97. J. S. Sinninghe Damsté, W. I. C. Rijpstra, E. C. Hopmans, J. W. Weijers, B. U. Foesel, J. Overmann, S. N. Dedysh, 13, 16-Dimethyl octacosanedioic acid (*iso*-diabolic acid), a common membrane-spanning lipid of *Acidobacteria* subdivisions 1 and 3. *Appl. Environ. Microbiol.* **77**, 4147–4154 (2011).
98. J. L. Vey, C. L. Drennan, Structural insights into radical generation by the radical SAM superfamily. Chem. Rev. **111**, 2487–2506 (2011).

99. Y. Nicolet, C. L. Drennan, AdoMet radical proteins—from structure to evolution—alignment of divergent protein sequences reveals strong secondary structure element conservation. *Nucleic Acids Res.* **32**, 4015–4025 (2004).
100. H. Nishida, T. Beppu, K. Ueda, Whole-genome comparison clarifies close phylogenetic relationships between the phyla Dictyoglomi and Thermotogae. *Genomics* **98**, 370–375 (2011).
101. P. J. Brumm, K. Gowda, F. T. Robb, D. A. Mead, The complete genome sequence of hyperthermophile *Dictyoglomus turgidum* DSM 6724™ reveals a specialized carbohydrate fermentor. *Front. Microbiol.* **7**, 1–20 (2016).
102. M. Y. Galperin, Genome diversity of spore-forming Firmicutes. The bacterial spore: From molecules to systems. *Microbiol. Spectr.* **1**, 1–18 (2016).
103. A. M. Kielak, C. C. Barreto, G. A. Kowalchuk, J. A. Van Veen, E. E. Kuramae, The ecology of *Acidobacteria*: Moving beyond genes and genomes. *Front. Microbiol.* **7**, 1–16 (2016).
104. B. Hausmann, C. Pelikan, C. W. Herbold, S. Köstlbacher, M. Albertsen, S. A. Eichorst, T. Glavina del Rio, M. Huemer, P. H. Nielsen, T. Rattei, U. Stingl, S. G. Tringe, D. Trojan, C. Wentrup, D. Woebken, M. Pester, A. Loy, Peatland *Acidobacteria* with a dissimilatory sulfur metabolism. *ISME J.* **12**, 1729–1742 (2018).
105. S. Hunter, R. Apweiler, T. K. Attwood, A. Bairoch, A. Bateman, D. Binns, P. Bork, U. Das, L. Daugherty, L. Duquenne, R. D. Finn, J. Gough, D. Haft, N. Hulo, D. Kahn, E. Kelly, A. Laugraud, I. Letunic, D. Lonsdale, R. O. Lopez, M. Madera, J. Maslen, C. McAnulla, J. McDowall, J. Mistry, A. Mitchell, N. Mulder, D. Natale, C. Orengo, A. F. Quinn, J. D. Selengut, C. J. A. Sigrist, M. Thimma, P. D. Thomas, F. Valentin, D. Wilson, C. H. Wu, C. Yeats, InterPro: The integrative protein signature database.*Nucleic Acids Res.* **37**, D211–D215 (2009).
106. R. D. Finn, A. Bateman, J. Clements, P. Coggill, R. Y. Eberhardt, S. R. Eddy, A. Heger, K. Hetherington, L. Holm, J. Mistry, E. L. L. Sonnhammer, J. Tate, M. Punta, Pfam: The protein families database. *Nucleic Acids Res.* **42**, D222–D230 (2014).


**Acknowledgments:** We thank M. van der Meer for help in the labeled incubation data interpretation. We thank D. Jackson for discussions on the PlsA protein and two anonymous referees for comments on an earlier version of this paper. **Funding:** This work was supported by The European Research Council (ERC) under the European Union's Horizon 2020 Research and Innovation Program (grant agreement no. 694569-MICROLIPIDS) (to J.S.S.D.). L.V. and J.S.S.D. received funding from the Soehngen Institute for Anaerobic Microbiology (SIAM) through a Gravitation Grant (024.002.002) from the Dutch Ministry of Education, Culture, and Science (OCW). K.F. (PI-LV) was supported by the Moore-Simons Project on the Origin of the Eukaryotic Cell, Simons Foundation 735929LPI, https://doi.org/10.46714/735929LPI. **Author contributions:** D.X.S.-C. designed the experiments, executed the experiments, analyzed the data, interpreted the results, and wrote the paper. M.F.S. designed and executed part of the experiments. J.C.E. and A.A.A. contributed to the bioinformatic analysis. S.B. executed and interpreted the proteomic data. M.K. and N.J.B. performed the lipid analysis and interpretation. F.A.B.v.M. contributed to the bioinformatic and phylogenetic analyses. L.S.v.S. and K.F. contributed to the gene expression assays. J.S.S.D. and L.V. acquired funding and contributed to the design of the experiments, interpretation of the data, and writing of the manuscript. All coauthors have read and approved the manuscript. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** Raw reads of transcriptomic data (RNA-seq) can be found in NCBI's GEO repository (https://ncbi.nlm.nih.gov/geo) accessible through GEO series accession number GSE211548. The MS proteomics data files are available from the ProteomeXchange Consortium via the PRIDE partner repository with the dataset identifier PXD035794. Lipid analysis, phylogenetic trees raw files, and the data S1 and S2 files can be found in Zenodo repository https://doi.org/10.5281/zenodo.6966462. All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. Code availability is not applicable.

Submitted 6 May 2022
Accepted 11 November 2022
Published 16 December 2022
10.1126/sciadv.abq8652

ScienceAdvances

# Disentangling the lipid divide: Identification of key enzymes for the biosynthesis of membrane-spanning and ether lipids in Bacteria

Diana X. Sahonero-Canavesi, Melvin F. Siliakus, Alejandro Abdala Asbun, Michel Koenen, F.A.Bastiaan von Meijenfeldt, Sjef Boeren, Nicole J. Bale, Julia C. Engelman, Kerstin Fiege, Lora Strack van Schijndel, Jaap S. Sinninghe Damst, and Laura Villanueva

*Sci. Adv.*, **8** (50), eabq8652.
DOI: 10.1126/sciadv.abq8652



*Science Advances* (ISSN ) is published by the American Association for the Advancement of Science. 1200 New York Avenue NW, Washington, DC 20005. The title *Science Advances* is a registered trademark of AAAS.