Biosynthesis and biotechnological application of the glycine-glucolipid from the marine bacterium *Alcanivorax borkumensis*

Dissertation

zur

Erlangung des Doktorgrades (Dr. rer. nat.)

der

Mathematisch-Naturwissenschaftlichen Fakultät

der

Rheinischen Friedrich-Wilhelms-Universität Bonn

vorgelegt von

Jiaxin Cui

aus

Hebei, China

Bonn, 2024

Angefertigt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät

der Rheinischen Friedrich-Wilhelms-Universität Bonn

Gutachter / Betreuer: Prof. Dr. Peter Dörmann

Gutachter: Prof. Dr. Lukas Schreiber

Tag der Promotion: 20.09.2024

Erscheinungsjahr: 2024

Declaration

I hereby declare that I have written the here presented work without unauthorized assistance and I have no included text passages, graphics, or other materials from third parties without identifying them. Only listed sources and resources were used and all citations from other works are highlighted in accordance with the citation rules for academic writing.

The thesis has not been submitted in any form as part of any other of my examinations.

Bonn, July 2024

Figures

Abbreviation

% (v/v)	Percent volume per volume (ml per 100 ml)
% (w/v)	Percent weight per volume (g per 100 ml)
A domain	Adenylation domain of NRPS
ATP	Adenosine triphosphate
BLAST	Basic Local Alignment Tool
BSA	Bovine serum albumin
C domain	Condensation domain of NRPS
CTAB	Cetyltrimethylammonium bromide
ddH_2O	Double deionized water
dH_2O	Deionized water
DAD	Diode-array detection
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
dNTP	Deoxyribonucleotide triphosphate
EDTA	Ethylenediaminetetraacetic acid
FAME	Fatty acid methyl ester
g	Standard gravity ($g = 9.81\ m\ s^{-2}$)
GC	Gas chromatography
HCB	Hydrocarbonoclastic Bacteria
HPLC	High-performance liquid chromatography
kb	Kilobase pairs
LC	Liquid chromatography
MALDI-TOF	matrix-assisted laser desorption/ionization with time-of-flight detection
MGD	Monogalactosyldiacylglycerol
MS	Mass spectrometry
m/z	Mass-to-charge
NCBI	National Center for Biotechnology Information

NRPS	Non-ribosomal peptide synthetase
OD_{600}	Optical density measured at a wavelength of 600 nm
PC	Phosphatidylcholine
PCR	Polymerase chain reaction
PE	Phosphatidylethanolamine
PG	Phosphatidylglycerol
Q-TOF	Quadrupole-time of flight
rpm	Revolutions per minute
SEM	Scanning electron microscope
T domain	Thiolation domain of NRPS
TE domain	Thioesterase domain of NRPS
TEM	Transmission electron microscopy
TLC	Thin layer chromatography
UV	Ultra violet

Fatty acids were abbreviated as X:Y, where X indicates the number of carbon atoms and Y of the double bonds in the acyl chain. Hydroxy fatty acid were abbreviated as n-OH-X:Y, where n indicates the position of hydroxyl group in the acyl chain.

1. Introduction

1.1. Hydrocarbonoclastic Bacteria

Hydrocarbons and their derivatives, including fossil carbon deposits and compounds of recent biological origin, such as pentadecane and heptadecane produced from fatty acids by marine cyanobacteria, *Prochlorococcus* or *Synechococcus*, are frequently found in marine environments (Lea-Smith, Biller et al. 2015, Love, Arrington et al. 2021). Besides the natural seeps of hydrocarbons, the marine ecosystem is also challenged by the hydrocarbons released by human activity, such as marine oil transport accidents or operational discharges from tankers. According to current estimates, there are over 8 million tons of hydrocarbons released into the ocean each year. The 2010 Gulf of Mexico Deepwater Horizon was the largest accidental oil spill in history. With over 300 million liters of oil combined with dispersants presumed to remain in the Gulf, it is one of the worst pollution events ever (Chakrabarty, Sheehy et al. 2024). Huge discharges from accidents have caused large-scale environmental disasters with extensive damage to the marine ecosystem.

The marine ecosystem represents a complex and highly diverse community of organisms. Oil spills can cause tremendous damage to the biodiversity of the oceans, and the removal of petroleum from contaminated sites is highly challenging. Indeed, petroleum can be degraded by naturally occurring microorganisms (Yakimov, Bargiela et al. 2022). A diverse range of hydrocarbonoclastic bacteria (HCB) has been recognized over the past few years as a new and ecophysiologically unusual group of hydrocarbon-degrading bacteria, and less than 25 % were obtained from marine sources. Unlike most bacteria, these marine HCB occupy a special niche among marine heterotrophic bacteria, as their carbon sources are the chemically stable saturated and aromatic hydrocarbon compounds. Therefore, they have the potential to play a significant role in the biological removal of petroleum-derived hydrocarbons from the oil spill area (Yakimov, Timmis et al. 2007).

Although the isolation and molecular characterization of the marine HCB are rather straightforward, their ecophysiology has not been studied extensively. Nonetheless, a few successful attempts have been tried. For example, one of the best characterized marine HCB is *Alcanivorax borkumensis*, whose genomic analysis was completed in 2006 by S.Schneiker

(Schneiker, Dos Santos et al. 2006).

1.2. *Alcanivorax borkumensis*

Alcanivorax borkumensis is an unusual marine γ-proteobacterium without flagella firstly isolated from sediments collected near Borkum Island, in the North Sea of Germany (Yakimov, Golyshin et al. 1998). This bacterium is classified as HCB as it is able to grow on alkanes or a few soluble substrates such as sodium pyruvate or sodium acetate (Karmainski, Dielentheis-Frenken et al. 2023). The hydrocarbon spectrum of substrates of *A. borkumensis* is exceptionally broad, including linear alkanes, cycloalkanes, isoprenoids (Wright, Bosch et al. 2020). In oil-polluted sea areas, most bacteria become nutrient-starved except *A. borkumensis* that can proliferate up to 80-90 % of the oil-degrading microbial community (Ian M. Head 2006). Since the first description of *A. borkumensis,* it has been detected worldwide. The genomic analysis gave insights into the capacity of *A. borkumensis* for alkane degradation including metabolism, biosurfactant production and biofilm formation (Schneiker, Dos Santos et al. 2006).

Ninety-seven proteins are differentially expressed between alkane-grown and pyruvate-grown cells. These 97 proteins include both cytoplasmic and membrane-associated ones. The genes encoding these proteins are localized to 46 putative operon structures. Membrane proteins up-regulated in alkane-grown cells include 3 enzyme systems that are capable to convert the alkanes into fatty acids via terminal oxidation: the well-known alkB1, two new alkane hydroxylating systems (a P450 cytochrome monooxygenase and a putative flavin-binding monooxygenase) and enzymes mediating β-oxidation of fatty acids (Sabirova, Ferrer et al. 2006). These three different alkane-oxidation systems allow *A. borkumensis* to live on a broad spectrum of hydrocarbons. Cytoplasmic proteins up regulated in alkane-grown cells include the enzymes for fatty acid degradation, synthesis of riboflavin, unsaturated fatty acids and cardiolipin. The up-regulation of these central metabolism genes indicates that the cells may restructure the membrane to adapt to the massive influx of alkane oxidation products. (Schneiker, Dos Santos et al. 2006).

1.3. Biosurfactants in Bacteria

The complex biochemistry inside bacterial systems results in the production of several classes of molecules with complex composition and structure (Tiso, Thies et al. 2017). The molecules may have a specific biological activity, oftentimes surface activity. The HCB are capable of degrading the hydrocarbons that have low solubility in water (Alaa A. Al-Seraih and Ghadban 2022). It is

essential for the bacteria to have access to the hydrophobic hydrocarbons to enable the membrane-bound oxygenase to degrade the hydrocarbons. One strategy is that the cells can emulsify the hydrocarbons by releasing a surfactant and then enhance the contact with the oil phase. For example, *P. putida* cells solubilize the crude oil by producing rhamnolipids when crude oil serves as sole carbon source (Liu, Zeng et al. 2017). The other strategy is that the cells can adjust the hydrophobicity of their membrane and then have better attachment to the hydrocarbons. Therefore, these alkane-grown bacteria usually produce a large variety of surfactants. These bacterial surfactants are named biosurfactants. The group of biosurfactants are oftentimes secondary metabolites with diverse structure including small molecules like lipopeptides, acylated amino acids, glycolipids and polymeric compounds with high molecular weight (Kishimoto, Adachi et al. 1993, Banat, Makkar et al. 2000, Timmis, McGenity et al. 2010, Soberón-Chávez and Maier 2011, Marchant and Banat 2017). The biosurfactants have large potential for applications in industry (Nievas, Commendatore et al. 2008, Nguyen and Sabatini 2011), pharmacology (Vecino, Cruz et al. 2017), bioremediation (Sen 2008) or agriculture (Vatsa, Sanchez et al. 2010, Sachdev and Cameotra 2013) because they are biodegradable, provide low toxicity and have antibacterial and antifungal effects (Banat, Makkar et al. 2000, Marchant and Banat 2017). The class of glycolipids is one of the best-known biosurfactants produced by different fungal and bacteria (Das, Mukherjee et al. 2009).

A. borkumensis produces a specific glycolipid whose structure is composed of four D-3-hydroxydecanoic acids with three of the carboxy groups being esterified to the hydroxyl group of the other 3-hydroxydecanoic acid, while the terminal carboxy group forms an amide linkage to glycine, and the first hydroxyl group is bound to glucose in β-glycosidic linkage. The function of this glycine-glucolipid remained unclear, but it has been suggested that it is involved in the oil degradation process (Abbasi, Bothun et al. 2018, Godfrin, Sihlabela et al. 2018). Commonly, glycolipid biosurfactants are recovered from the extracellular medium of producing bacterial cultures (Kubicki, Bator et al. 2020). Previous studies suggested that the glycine-glucolipids were associated with the bacterial envelope. Only after the cleavage of the glycine moiety, the glycine-free compound would be secreted into the medium, where it might associate with hydrocarbons and aid in the solubilization, uptake, and metabolization of alkanes (Abraham, Meyer et al. 1998, Timmis, McGenity et al. 2010, Godfrin, Sihlabela et al. 2018, Kubicki, Bollinger et al. 2019). However, the respective enzyme for the cleavage and all other enzymes involved in the biosynthesis, remain unknown.

The glycine-glucolipid of *A. borkumensis* was shown to exhibit remarkably low ecotoxicity in comparison to synthetic surfactants and even to other glycolipid biosurfactants (Kubicki, Bollinger et al. 2019). Therefore, it is a promising candidate with significant potential applications in agriculture, the cosmetic and pharmaceutical industry. However, the biosynthetic pathway of the glycine-glucolipid from *A. borkumensis* is unknown. There are two possibilities of glycine-glucolipid biosynthesis. On the one hand, its structure is related to the two rhamnolipids of *Pseudomonas aeruginosa*. The monorhamnolipid and dirhamnolipid of *P. aeruginosa* consist of one or two rhamnose moieties bound to a diester of 3-D-hydroxydecanoic acids, respectively (Burger, Glaser et al. 1963, Ochsner, Koch et al. 1994). The diester is produced from two molecules of 3-D-hydroxydecanoyl-ACP by the *RhlA* gene product (Zhu and Rock 2008). 3-D-hydroxydecanoyl-ACP is derived from fatty acid *de novo* synthesis via the *FabG* gene product. Finally, two rhamnoses are successively transferred onto the acyl diester by two rhamnosyltransferases, RhlB and RhlC, using TDP-rhamnose as sugar donors (Ochsner, Koch et al. 1994, Rahim, Ochsner et al. 2001). It is likely that a glycyltransferase, an acyltransferase and a glucosyltransferase are involved in the glycine-glucolipid biosynthesis pathway as shown for the rhamnolipid. On the other hand, the glycine-glucolipid contains a glycine linked in amide linkage to an acyl group, which is a typical product synthesized by the non-ribosomal peptide synthetase (NRPS) dependent pathway. Therefore, it is possible that a non-ribosomal peptide synthetase (NRPS) is involved in the biosynthetic pathway. NRPS modules form linkages between the amino group of an amino acid with the acyl group of a second amino acid or fatty acid.

1.4. Non-ribosomal Peptide Synthetase

Non-ribosomal peptide synthetases (NRPSs) are large multimodular enzymes that have evolved in bacteria and fungi to assemble highly complex, bioactive secondary metabolites. NRPSs can process hundreds of monomers, including amino acids and fatty acids (Marahiel, Stachelhaus et al. 1997, Konz and Marahiel 1999, Finking and Marahiel 2004), via an array of large, catalytic units so-called modules comprised of several catalytic domains which are covalently linked. The bioactive metabolites have broad applications, for example as antibiotics (daptomycin), antitumor (bleomycin), antifungal agents or immunosuppressants (cyclosporin) (Walsh 2008). This diverse bioactivity can be explained by the way how nature synthesizes these molecules. As shown in Figure 1A, an NRPS elongation module contains at least three essential domains: 1) an adenylation (A) domain selecting and activating the amino acid, and transferring it to the -SH

group of the 4'-phosphopantetheine (P-pant) arm at the 2) thiolation (T) domain, also named peptidyl carrier protein (PCP) domain. 3) A condensation (C) domain forms amide bonds between the -COOH of the intermediate from the upstream T domain or acid donor, and the -NH_2 of the amino acid on the downstream T domain. After the elongation, the mature product will be cleaved and released by the 4) thioesterase (TE) domain either by hydrolysis or by cyclization (Konz and Marahiel 1999).

The A domain is one of the best-characterized NRPS domains. The reaction of the A domain can be divided into two half-reactions: activation of the amino acid with ATP to form an amino acyl adenylate (adenylation), and transfer of this amino acyl residue to the 4'-phosphopantetheine arm of the downstream T domain (thioesterification) (Tan, Dai et al. 2015, Drake, Miller et al. 2016, Miller, Drake et al. 2016, Reimer, Aloise et al. 2016, Tarry, Haque et al. 2017). The A domain operates as a gatekeeper that selects the substrate for transfer onto the T domain and for channeling into the further assembly process in the NRPS (Figure 1B). Besides the 20 proteinogenic amino acids, additional nonproteinogenic amino acids can be the substrates of A domains (Walsh, O'Brien et al. 2013). The phenylalanine-specific A domain is the NRPS domain whose crystal structure was determined by Elena Conti (Conti, Stachelhaus et al. 1997). In the GrsA A domain, the strictly invariant positively charged Lys517 residue compensates for the negative charge of the amino acid carboxylate which is brought into proximity with the ATP-α-phosphate, providing a crucial catalytic driving force for the adenylation reaction (Figure 1C) (Conti, Stachelhaus et al. 1997). The A domain consists of two subunits, the larger N-terminal core (A_{core}) domain and the smaller C-terminal subdomain (A_{sub}) (Figure 1B). Large conformational changes happen during the catalytic process, because the A_{sub} moves towards the A_{core}. To allow for the binding of the amino acid and ATP, the A domain shows an open conformation where the A_{sub} is away from the active site. Afterwards, the A_{sub} moves close to the active site and turns into the catalytic conformation for the adenylation reaction. The A_{sub} rotates by around 140°, after the amino acyl adenylate bond has been produced. The P-pant arm of the downstream T domain enters the A domain from the open space established by the rotation. Finally, the A domain loads the amino acid onto the T domain and returns back to the open conformation (Conti, Stachelhaus et al. 1997). Stachelhaus further determined a group of around 10 amino acids in the amino acid binding pocked of the A domain that dictates the specificity of the amino acid, and deciphered the determining code of these 10 amino acids (Figure 1C) (Stachelhaus, Mootz et al. 1999). This makes it possible to predict the monomer substrate of the A domain reliably, for example through

the PKS/NRPS Analysis Web-site, http://nrps.igs.umaryland.edu/.

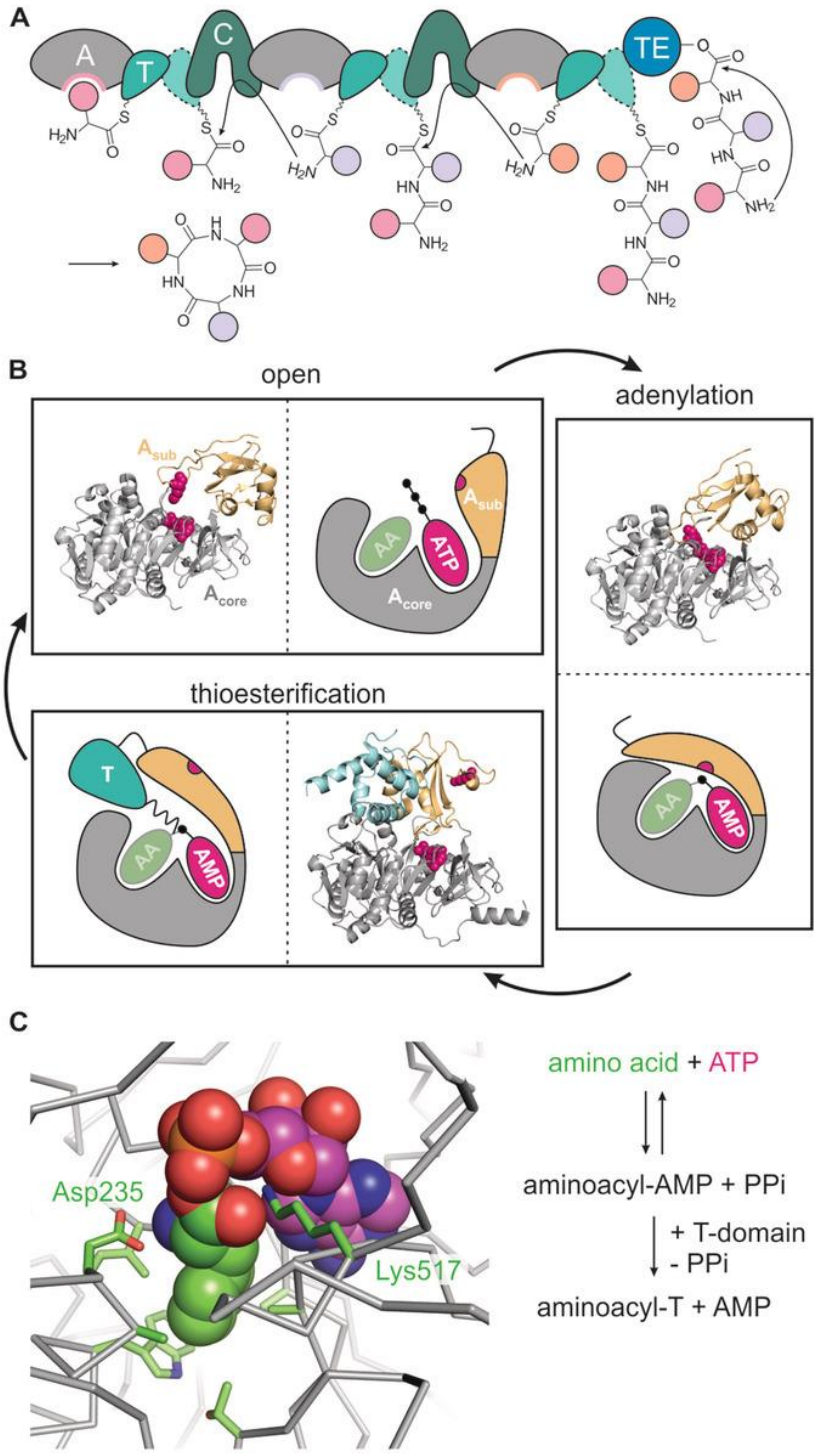

Figure 1 **Structure of the NRPS A domain**

A) A hypothetical NRPS assembles a cyclic tripeptide. B) during adenylation and thioesterification, the A domain goes through a series of conformational states. C) the specificity code residues in the GrsA-A domain (Stanišić and Kries 2019).

The C domain catalyzes the condensation reaction between the amino group and the carboxyl group for the chain elongation. A breakthrough in C domain knowledge came in 2002, when Keating et al. first published the crystal structure of an NRPS C domain, VibH, a standalone C domain involved in the synthesis of the siderophore vibriobactin (Keating, Marshall et al. 2002). Their work showed that the C domain is pseudo-dimeric consisting of two lobes (N-terminal lobe, N-lobe; C-terminal lobe, C-lobe) forming a head-to-tail V-shaped structure (Figure 2) (Keating, Marshall et al. 2002). Each lobe consists of a large αβα sandwich structure (Keating, Marshall et al. 2002). The N-lobe consists of five β-strands surrounded by five α-helices and a 2-strand β-sheet. The C-lobe consists of a core sheet of six β-strands, extended by two additional strands and covered on one side by nine α-helices. One of the five β-strands of N-lobe and the sequence of C-lobe form a feature named latch (VibH 338-360, GglsA 396-419) covering the active site of the C domain. There is a second cross-over region named floor loop at the interface of two lobes. The catalytic site defined by a highly conserved HHxxxDG motif is at the center of the tunnel covered by a loop connecting the central strand of the N-lobe and one of the α-helices (Dekimpe and Masschelein 2021). The tunnel feature where the catalytic site is located, provides access to the second histidine of the HHxxxDG motif from two different faces for both the acid donor and the peptide or amino acid carried by the downstream T domain (Keating, Marshall et al. 2002). The nucleophilic attack n of the α-amino group takes place at this tunnel site.

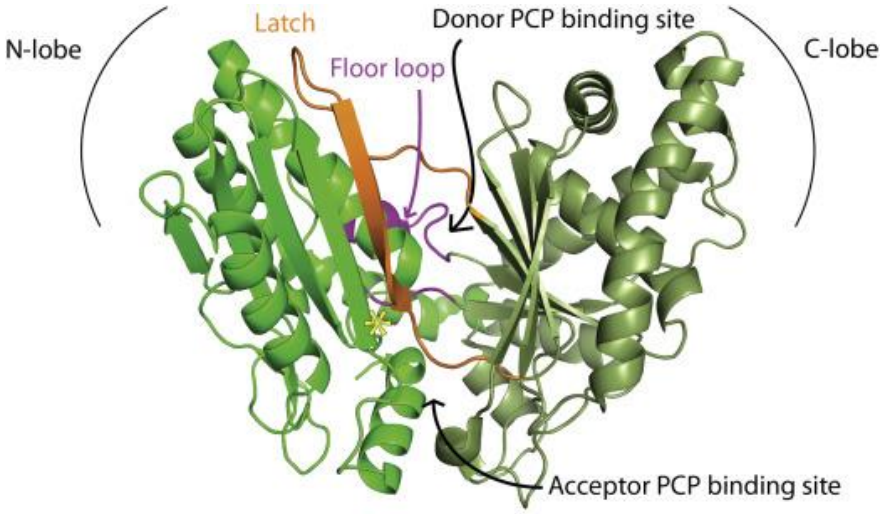

Figure 2 **Crystal structure of VibH.**

The C domain consists of the N-lobe (light green) and the C-lobe (dark green). The active site (HHxxxDG motif) is denoted by a yellow asterisk. The Latch feature and floor loop feature are labeled in orange and magenta, respectively (Bloudoff and Schmeing 2017).

Post-translational modifications serve as a major way to regulate protein structure, activity, localization and function in the cell. The NRPSs activity is regulated by a covalent modification of the serine side chain of the conserved GGXS motif located at the T domain with P-pant by phosphopantetheinyl transferases. The T domain is always associated with the A domain and does not have its own catalytic activity. The thiolation reaction was found to be catalyzed by the A domain(Tan, Dai et al. 2015, Drake, Miller et al. 2016, Miller, Drake et al. 2016, Reimer, Aloise et al. 2016, Tarry, Haque et al. 2017). The long and floppy P-pant prosthetic group attached at the conserved serine residue serves as a carrier and shuttles the intermediate among each catalytic domain (Stanišić and Kries 2019).

The TE domain, located at the final module of NRPSs, catalyzes the cleavage of the product from the NRPS (Schneider and Marahiel 1998). The cleavage can be separated into two half-step reactions, the first half reaction is the transfer of the product from the T domain onto the active site serine to form an O-linked intermediate. There are three different types of the second half-reactions resulting in three different products: 1) cleavage by hydrolysis and release of a linear compound; 2) cleavage by cyclization and release of a cyclic compound; 3) cleavage by oligomerization and release of an oligomeric compound. Individual TE domains perform oftentimes only one kind of cleavage reaction (Hoyer, Mahlert et al. 2007, Alonzo, Magarvey et al. 2015).

The T-TE di-domain study by Dominique P. Frueh, et al. shows that the TE domain controls the chain length via a mobile flap formed by two α-helices (α4TE–α5TE, green) (Figure 3) (Dominique P. Frueh 2008). The relative orientation is well-defined. Within the T-TE domain, there is a primarily hydrophobic interface resulting in a 1,300 $Å^2$ large buried surface area. Besides, the linker between the T (red) and TE (blue and green) domains resulted in a further 200 $Å^2$ area (Frueh, Arthanari et al. 2008). The globular core between the thioesterase and two α-helices (α4TE–α5TE) with two reaction sites (S48 of T domain and S180 of TE domain, yellow balls) is wedged by three helices of the T domain (Frueh, Arthanari et al. 2008). The 17 Å-distance between S48 and S180 can be extended up to ~ 20 Å, which allows the P-pant arm tethered at S48 to enter and reach the S180 for the cleavage reaction (Frueh, Arthanari et al. 2008). The two α-helices (α4TE–α5TE) are flexible and form a mobile flap which provides the conformational possibility to accommodate the various growing intermediates on the P-pant arm (Frueh, Arthanari et al. 2008). Attachment of the C domain to the T-TE di-domain resulted in specific chemical shift changes at the C-domain-binding face on the T domain detected by NMR (Frueh,

Arthanari et al. 2008). Only a minor effect was observed on the TE domain. This effect was observed at the two α-helices (α4TE–α5TE) region (Frueh, Arthanari et al. 2008).

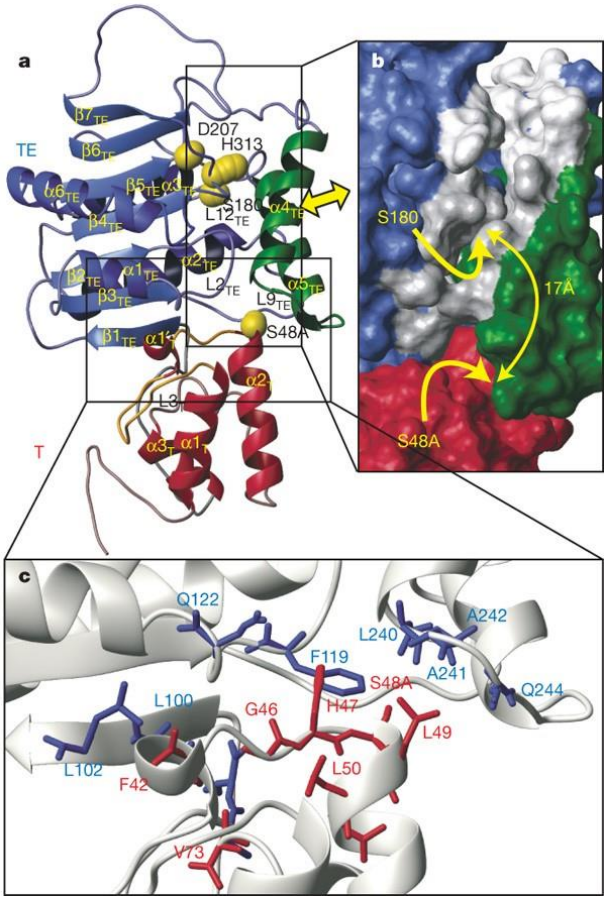

Figure 3 **Crystal structure of EntF T-TE domain.**
The T domain (red) is wedged between the thioesterase (blue) and the two α-helices (α4TE–α5TE) (green) of the TE domain. The activate site S48 and S180 is in yellow (Frueh, Arthanari et al. 2008).

1.5. Phosphopantetheinyl Transferase

Phosphopantetheinyl transferases (PPTases) are essential for cell viability across all three domains of life: bacteria, archaea, and eukaryotes. They play a critical role in the activation of polyketide synthases (PK), fatty acid synthases (FAS), and non-ribosomal peptide synthetases (NRPS) by posttranslational modification (Lambalot, Gehring et al. 1996). PPTases catalyze the transfer of the P-pant moiety from coenzyme A (CoA) to the hydroxyl group of a conserved serine residue within the acyl carrier protein domain, also known as the thiolation domain (Figure 4) (Crosby and Crump 2012). This modification converts the inactive apoproteins into their catalytically active holo-forms. The thiolation domain acts as a scaffold, tethering the growing intermediates through a reactive thioester linkage to a P-pant arm, which extends approximately 20 Å (Frueh,

Arthanari et al. 2008). The P-pant arm is often described as a "prosthetic arm" because it transiently holds all substrates and intermediates during the enzymatic modifications involved in the biosynthesis of polyketides, fatty acids, and non-ribosomal peptides. This allows for the orderly progression of these pathways, facilitating the elongation and modification of the growing chain (Figure 4) (Beld, Sonnenschein et al. 2014).

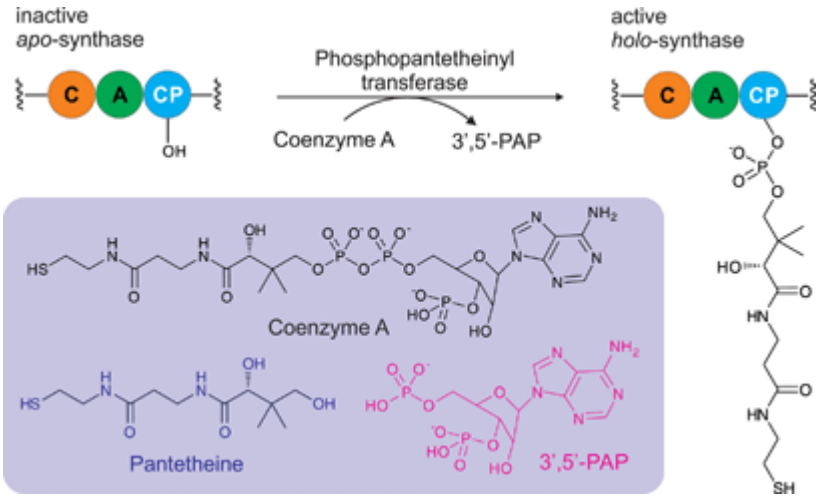

Figure 4 **General reaction scheme of post-translational phosphopantetheinylation by a PPTase.**
The PPTase transfers the P-pant moiety from CoA to a conserved serine residue on the apo-CP to produce holo-CP, here showcased by a typical NRPS module containing the C (condensation), A (adenylation), and CP (carrier protein) domains. 3',5'-PAP is 3',5'-phosphoadenosine phosphate (Beld, Sonnenschein et al. 2014).

Two common types of PPTase can be found in various organisms, classified on the basis of their structural organization. *Holo*-ACP synthase (AcpS) is a typical enzyme of the first family of PPTases recognized (Tpye-I). Type-I PPTases are homotrimers (trimer colored in brown, green and blue) and are generally thought to activate acyl carrier proteins (ACPs) (CoA colored in red) of FASs carrying out primary lipid metabolism (Figure 5). Surfactin phosphopantetheinyl transferase (Sfp) represents the second family of PPTases (Tpye-II). Type-II PPTases exist as a pseudo-dimer (dimer colored in brown and blue) and are generally thought to activate carrier protein/thiolation domains of PKSs and NRPSs (CoA colored in red) involved in secondary metabolism (Parris, Lin et al. 2000, Beld, Sonnenschein et al. 2014). The pseudo-homodimer structure exists like two AcpS monomer with one active site at the pseudo-dimer interface (Figure 5). This structure results in a much broader substrate acceptance. The genes encoding the type-II PPTase often reside in close proximity to, or part of, a synthase gene cluster (Beld, Sonnenschein et al. 2014).

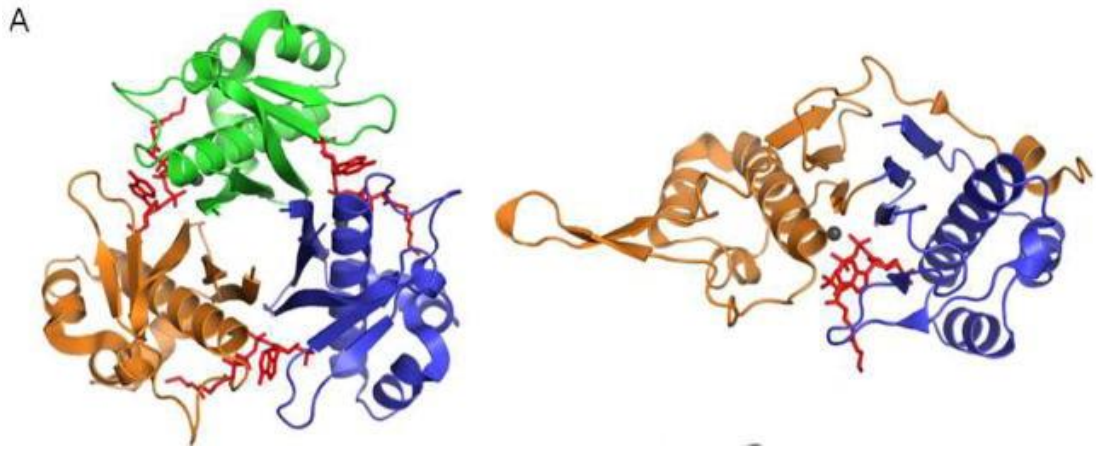

Figure 5 **Comparison of *B. subtilis* AcpS and Sfp.**
Top view of both enzymes. CoA is colored red, and divalent cations are shown as black spheres. Each trimer of AcpS is colored differently to show arrangement of monomers, and each "pseudo-dimer" of Sfp half is colored differently for comparison to AcpS (Beld, Sonnenschein et al. 2014).

1.6. Glycosyltransferase

Glycosyltransferases are enzymes that catalyze the transfer of sugar moieties from activated donor substrates to acceptor molecules. The most common forms of these activated donor sugars are nucleoside diphosphate sugars (e.g., UDP-Glc) (Lairson, Henrissat et al. 2008). The acceptor substrates utilized by glycosyltransferases are most commonly other sugars but can also include proteins, lipids or other small molecules (Lairson, Henrissat et al. 2008). Glycosyltransferases are classified into families based on amino acid sequence similarities, with the number of distinct families having grown to 110 (Lombard, Golaconda Ramulu et al. 2014). The diversity of three-dimensional (3-D) folds observed for glycosyltransferases is very limited, despite the large number of distinct families. Structural characterization of representatives from a large number of the 110 glycosyltransferase families has revealed that the GT-A and GT-B folds are predominant (http://www.cazy.org). These two folds have been observed in all structures of nucleotide-sugar-dependent glycosyltransferases solved to date (Ünligil and Rini 2000, Bourne and Henrissat 2001, Hu and Walker 2002).

Both of these folds are characterized by two associated β/α/β domains (Rossmann-like domains), typical of nucleotide-binding proteins (Lairson, Henrissat et al. 2008). The GT-A fold was first described for the inverting enzyme SpsA from *Bacillus subtilis*, these two domains abut closely, forming a continuous central β-sheet (Figure 6a) (Charnock and Davies 1999). The sizes of these β/α/β domains can vary, but their close association is a common feature, contributing to the

structural integrity and function of the glycosyltransferases. For this reason, some describe the GT-A fold as a single-domain fold. Like the GT-A fold, GT-B enzymes consist of two β/α/β Rossmann-like domains (Figure 6b). However, in GT-B enzymes, these two domains are less tightly associated and face each other, creating a cleft where the active site resides (Campbell, Mosimann et al. 2000).

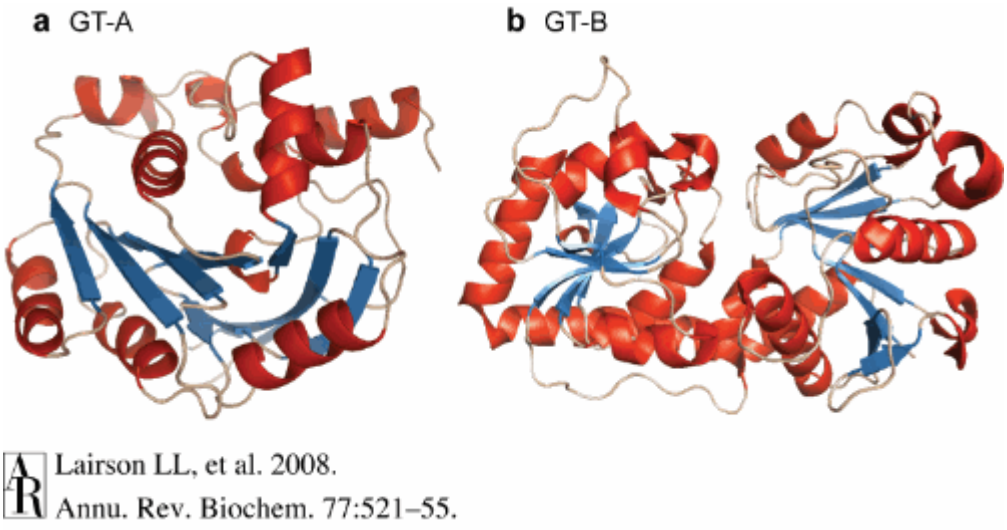

Figure 6 **Overall folds observed for glycosyltransferase enzymes.**
(a) The GT-A fold is represented by the inverting enzyme SpsA from *Bacillus subtilus*, Protein Data Bank (pdb) 1qgq, and (b) the GT-B fold, by bacteriophage T4 β-glucosyltransferase, pdb 1jg7 (Lairson, Henrissat et al. 2008).

Glycosyltransferases from family GT1 adopt the GT-B fold and are responsible for the glycosylation of various important small molecules. This functional diversity makes GT1 one of the most intensely studied families of glycosyltransferases (Lairson, Henrissat et al. 2008). Most glycosyltransferases preferentially select one specific sugar donor due to a set of highly conserved amino acid residues in the C-terminal domain (Zhang, Zhang et al. 2020). GT1 family glycosyltransferases are found across almost all domains of life and are crucial for the biosynthesis of glycosylated natural products. They play a vital role in the final step of molecular assembly (Luzhetskyy, Mendez et al. 2008), producing glycosylated products with activities against various chemical scaffolds, including polyketides (Hoffmeister, Wilkinson et al. 2002) and glycolipopeptides (Mulichak, Losey et al. 2003).

1.7. Acyltransferase

Long-chain hydrophobic acyl residues are crucial for a variety of essential biological structures and processes. They form the hydrophobic core of biological membranes, serve as intracellular storage compounds, and modify protein properties or function as membrane anchors, among other roles (Röttig and Steinbüchel 2013). (Hydroxy-)fatty acids are generally activated for subsequent reactions by esterification of their carboxyl groups with the thiol group of CoA or the acyl carrier protein (ACP), resulting in acyl-thioesters (Röttig and Steinbüchel 2013). There are three basic pathways of long-chain fatty acid production: 1) *de novo* fatty acid synthesis from the central metabolite acetyl-CoA, yielding acyl-ACPs; 2) direct uptake of exogenous fatty acids; 3) other compounds that are converted to fatty acids, e.g. alkanes and their conversion to acyl-CoAs by acyl-CoA synthetases (Black and DiRusso 2003, White, Zheng et al. 2005).

Those activated acyl chains are used by acyltransferases and transferred to various substrates or to polymerize them. AtfA is the key enzyme responsible for neutral lipid accumulation. It catalyzes the synthesis of triacylglycerol (TAG) or wax esters (WE) from acyl-CoAs and diacylglycerol (DAG) or fatty alcohols, respectively, as illustrated in Figure 7 (Kalscheuer and Steinbüchel 2003). Enzymes similar to AtfA are identified as critical for the synthesis of WEs or TAGs in several types of bacteria, including marine species of *Alcanivorax* (*ABO_2742, ABO_1804*). The AtfA type is the prototypical acyltransferase essential for bacterial lipid storage, highlighting its significance in lipid metabolism (Kalscheuer, Stöveken et al. 2007). Inactivation of atfA1 drastically reduced both WS and DGAT activities, but the disruption of atfA2 affected both activities to only a minor degree (Kalscheuer, Stöveken et al. 2007).

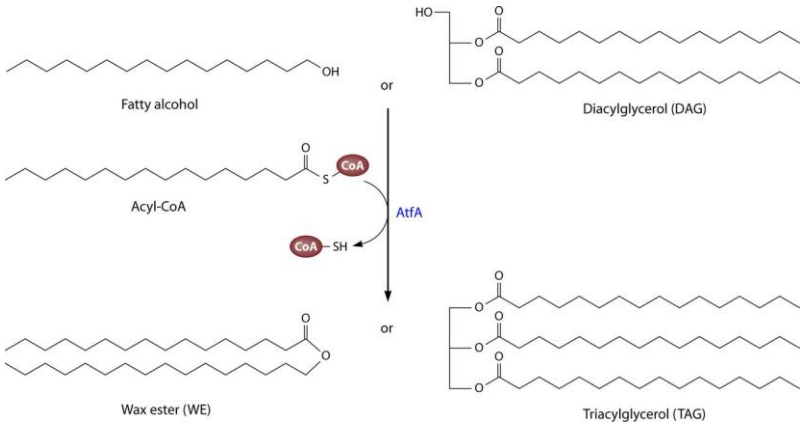

Figure 7 **Synthesis of wax esters or triacylglycerols from acyl-CoA and fatty alcohol or diacylglycerol, respectively, catalyzed by AtfA**. (Röttig and Steinbüchel 2013)

1.8. Carbon Storage Allocation in *A. borkumensis*

For the survival during periods of starvation, bacteria accumulate intracellular carbon storage compounds like storage lipid. These storage lipids occur particularly frequently among the HCB including *A. borkumensis*. Most of the HCB species accumulate specialized polymeric lipids like polyhydroxyalkanoic acids (PHAs) (Alvarez, Pucci et al. 1997). In addition to PHAs, which serve as the most abundant bacterial lipophilic storage carbon, triacylglycerols (TAGs) and wax esters (WEs) were also identified in bacteria as carbon storage compounds with less frequency (Steinbüchel 1991). Only in some species of gram-positive bacteria such as *Mycobacterium*, *Nocardia*, *Rhodococcus* and *Streptomyces*, TAGs was found in large quantities (Alvarez and Steinbüchel 2002). Within the gram-negative species of *Acinetobacter*, WEs of long-chain primary fatty alcohols and long-chain fatty acids were found with small amounts of TAGs. According to the genome sequencing results, *A. borkumensis* is also capable of synthesis and accumulation of TAGs and WEs. In 2007, A. Steinbüchel's group showed that *A. borkumensis* primarily accumulates TAG when cultivated with pyruvate. WE, besides a greater proportion of TAG, are synthesized only when hexadecane is provided (Kalscheuer, Stöveken et al. 2007). This inability is due to the lack of a fatty acyl-CoA reductase, so that during growth with unrelated carbon sources, there are no fatty alcohols present that could serve for WE synthesis (Kalscheuer, Stöveken et al. 2007). The *A. borkumensis* genome comprises two different putative PHA synthetase genes, *phaC1* and *phaC2*, indicating that the cells are capable to synthesize PHAs (Sabirova, Ferrer et al. 2006). But no 3-hydroxy fatty acids were detected after hydrolysis of polymers from *A. borkumensis* by A. Steinbüchel's group (Kalscheuer, Stöveken et al. 2007). Therefore, PHAs are present only in trace amounts in *A. borkumensis*. Large amounts of WEs accumulate as storage lipids when hexadecane serves as sole carbon. TAGs serve as the main storage lipids regardless of the carbon sources used.

1.9. Objectives

It has been speculated that the glycine-glucolipid produced by *A. borkumensis* is required as a biosurfactant during alkane degradation. It had been suggested that the glycine of the glycine-glucolipid was cleaved off, and the glycine-free form was secreted into the extracellular environment for the alkane solubilization. In the initial part of this thesis, different analytical methods were developed for the exact quantification of the glycine-glucolipid from *A. borkumensis* in combination with mass spectrometric analyses of lipid extracts from cells grown

under different conditions. The occurrence of the glycine-glucolipid in its glycine-containing or glycine-free forms within the cells or the culture supernatant was further analyzed.

The glycine-glucolipid exhibits remarkably low ecotoxicity in comparison to synthetic surfactants and even to other glycolipid biosurfactants. As a promising candidate with significant potential applications, the biosynthetic pathway of the glycine-glucolipid remained unknown. In the second part of this thesis, efforts were made to elucidate the molecular origin of the glycine-glucolipid in *A. borkumensis*. The genes involved in glycine-glucolipid synthesis were firstly identified by producing deletion mutant strains of *A. borkumensis*. The function of each gene involved was further explored by single or multiple heterologous expression in *E. coli* followed by lipid analysis via LC-MS/MS. The substrate studies were conducted with heterologous expression in *E. coli* and *in vitro* assays.

The key enzyme involved in the biosynthesis of glycine-glucolipid was identified as an NRPS encoded by *GglsA*. The specificity and role of individual domains of GglsA was investigated by producing a hybrid protein or variant proteins with single amino acid mutations. The hybrid protein was employed to study the possibility of producing an artificial tailor-made glycine-glucolipid. Indeed, a serine-glucolipid was produced in *E. coli* by swapping the A domain of GglsA with EntF of *E. coli*.

The function of the glycine-glucolipid in *A. borkumensis* remained unclear, but it had been suggested that it is involved in the oil degradation process (Abbasi, Bothun et al. 2018, Godfrin, Sihlabela et al. 2018). The function of the glycine-glucolipid in *A. borkumensis* was elucidated via deletion mutant strains lacking the glycine-glucolipid or containing only the aglycone instead. The growth of *A. borkumensis* grown in both pyruvate and hexadecane medium was monitored and the impact of the glycine-glucolipid on hydrophobicity was tested. Additionally, observations of the ultrastructure of the cells were conducted via SEM and TEM.

As a substance with significant potential applications, the glycine-glucolipid was further explored as a surfactant for the application of a foliar fertilizer.

2. Material and Methods

2.1. Materials

2.1.1. Equipment

6530 Accurate-mass quadrupole time-of- flight (Q-TOF) LC/MS	Agilent, Böblingen (DE)
6545 Accurate-mass quadrupole time-of- flight (Q-TOF) LC/MS	Agilent, Böblingen (DE)
7890 gas chromatography (GC) with flame ionization detector (FID)	Agilent, Böblingen (DE)
7890 gas chromatography (GC) with mass spectrometry (MS)	Agilent, Böblingen (DE)
Autoclave	Tuttnauer Systec, Kirchseeon-Buch (DE)
Balance 770	Kern, Balingen-Formmern (DE)
Balance PG503-S Delta Range	Mettler Toledo, Gießen (DE)
Block heater SBH130D/3	Stuart, Bibby Scientific, Staffordshire (USA)
Centrifuge 5810 R	Eppendorf, Hamburg (DE)
Centrifuge RC 5B Plus Sorvall	Kendro Laboratory Products, Osterode am Harz (DE)
Chemiluminescence documentation system	Bio-Rad, California (USA)
Confocal microscope, IX71	Olympus Optical Co. Ltd., Tokyo (JP)
Confocal scanner unit	Yokogawa Electric Corporation., Tokuo (JP)
Drop shape analyser, DAS25S	KRÜSS GmbH, Hamburg (DE)
Electrophoresis power supply, EPS 301	Amersham Pharmacia Biotech Inc., Buckinghamshire (UK)
Filter wheel changer, Lambda 10-3	Sutter Instrument, Novato (USA)

Gel caster, Mighty small II	GE Healthcare Europe GmbH, Freiburg (DE)
Gel documentation system, FastGene FAS-DIGI PRO	Nippon Genetic EUROPE GmbH, Düren (DE)
Growing cabinet, Rumed	Rubarth Apparate GmbH (DE)
Heating block	Bioer, Hangzhow (CHN)
High-performance liquid chromatography (HPLC) 1200 series	Agilent, Böblingen (DE)
Incubator, Kelvitron®t	Thermo Scientific Heraeus®, Waltham (USA)
Incubation shaker, Multitron 28570	INFORS, Einsbach (DE)
Linear laser system, 400 Series	Oxford Instruments GmbH, Wisbaden (DE)
Magnetic stirrer MR30001	Heidolph Instruments, Schwabach (DE)
Micro pulser electroporator	BioRad Laboratories, München (DE)
pH meter inoLab pH Level 1	WTW, Weilheim (DE)
Photometer, Specord 205	Analytik Jena, Jena (DE)
Phytochamber SIMATiC OP17	York International, York (USA)
PowerPac Basic electrophoresis power supply	BioRad Laboratories, München (DE)
Quantus Fluorometer	Promega (USA)
Rotary evaporator LABOROTA 4001	Heidolph Instruments, Schwabach (DE)
Running chamber for gel electrophoresis	Cti, Idstein (DE)
Sample concentrator	Techne (Bibby Scientific), Stone (UK)
Sterile bench model Mars 1800 LMS	Scanlaf, Allerød (DK)
Thermocycler Tpersonal 48	Biometra, Göttingen (DE)
Ultracentrifuge Optima L 90K	Beckman Coulter, Krefeld (DE)
UV-transilluminator DP-001 T1A	Vilber Lourmat GmbH, Eberhardzell (DE)

UV VIS Spectrophotometer, Specord 205	Analytik Jena AG, Jena (DE)
Vortex Cetromat® MV	Braun Biotech, Melsungen (DE)
Water bath TW20	Julabo, Seelbach (DE)
Water purification system ELIX® 35	Merck Millipore, Merck KGaA, Darmstadt (DE)

2.1.2. Consumables

Amicon Ultra-2 centrifugal Filter Unit (3 and 30 kDa)	Sigma-Aldrich, Taufkirchen (DE)
Auto sampler vials with inlets and screw caps with PTEF septa	VWR, Farmstadt (DE)
Coverslips, 24 mm x 60 mm	Marienfeld, Lauda-Königshofen (DE)
Cryo vials 2 mL	STARLAB GmbH, Hamburg (DE)
Cuvettes, semi-micro, PS	Ratiolab GmbH, Dreieich (DE)
Disposable glass vials (100 x 12 mm)	Assistant™ Fischer Scientific GmbH, Schwerte (DE)
Electroporation cuvettes	Bio-Budget Technologies GmbH, Krefeld (DE)
Falcon tubes (15 and 50 mL)	Greiner Bio-One, Frickenhausen (DE)
Glass beads	Sigma-Aldrich Chemie GmbH, Taufkirchen (DE)
Glass inlets (conical and flat)	VWR, Farmstadt (DE)
Glass pipettes	Brand, Wertheim (DE)
Glass tubes, 8 mL	Fisher Scientific, Schwerte (DE)
Glass tubes (15.5 x 160 mm)	Schott, Mainz (DE)
MicroAmp Optical Adhesive Film, PCR compatible	Thermo Fisher Scientific, Karlsruhe (DE)
Microcentrifuge tubes (1.5 and 2 mL)	Greiner Bio-One, Frickenhausen (DE)

Microliter pipette tips type 3 series 1700	Labomedic, Bonn (DE)
Petri dishes 94 x 16 mm	Greiner bio-One, Frickenhausen (DE)
Petri dishes 145 x 20 mm	Greiner bio-One, Frickenhausen (DE)
Pots for plant cultivation, 10 cm	Pöppelmann, Lohne (DE)
PTFE screw caps (13.3 and 22.4 mm)	Schott, Mainz (DE)
PTFE septa for screw caps (13.3 and 22.4 mm)	Schmidlin, Neuheim (CH)
Quartz sand	Quarzwerke Witterschlick GmbH, Alfter-Witterschlick (DE)
SPE columns (1 and 6 mL)	Phenomenex Ltd., Aschaffenburg (DE)
Sterile filter (0.22 µm, 25 mm and 500 mL)	Labomedic GmbH, Bonn (DE)
Syringe 50 mL	Terumo Corporation, Tokio (JP)
Trays for plant cultivation	Pöppelmann, Lohne (DE)
Vermiculite, grain size 2-3 mm	Rolfs, Siegburg (DE)

2.1.3. Chemicals

2-Bromoacetophenone	Sigma-Aldrich, Taufkirchen (DE)
2-Morpholino-4-ylethanesulfonic acid monohydrate (MES)	ChemCruzTM/Bio-Connect B. V., Huissen (NL)
3-Hydroxydecanoic acid	Biotrend, Köln (DE)
3-Hydroxydodecanoic acid	Larodan, Malmö (SE)
3-Hydroxymyristic acid	Santa Cruz Biotech. Texas (USA)
3-Hydroxyoctanoic acid	Larodan, Malmö (SE)
Acetic acid	VWR Chemicals, Darmstadt (DE)
Acetone	Prolabo VWR, Darmstadt (DE)

Acetonitrile	VWR Chemicals, Darmstadt (DE)
Agarose	VWR Chemicals, Darmstadt (DE)
Ammonium acetate	Sigma-Aldrich, Taufkirchen (DE)
Ammonium chloride	AppliChem, Darmstadt (DE)
Ammonium nitrate	AppliChem, Darmstadt (DE)
Ammonium persulfate	AppliChem, Darmstadt (DE)
Ammonium sulfate	Sigma-Aldrich, Taufkirchen (DE)
Ampicillin	Roth, Karlsruhe (DE)
Bacto agar	Duchefa Biochemie, Haarlem (NL)
Boric acid	Grüssing, Filsum (DE)
Bovine Serum Albumin (BSA)	Sigma-Aldrich, Taufkirchen (DE)
Bromophenol blue, sodium salt	SERVA Electrophoresis GmbH, Heidelberg (DE)
Calcium chloride dihydrate	AppliChem, Darmstadt (DE)
Chloramphenicol	AppliChem, Darmstadt (DE)
Chloroform	VWR, Darmstadt (DE)
Cobalt (II)-Chlorid Hexahydrate	Merck, Darmstadt (DE)
Coenzyme A trilithium salt	Sigma-Aldrich, Taufkirchen (DE)
Coomassie Brilliant Blue R250	AppliChem, Darmstadt (DE)
Copper (II) sulfate pentahydrate	Merck, Darmstadt (DE)
Cysteamine	TCI, Eschborn (DE)
Diethyl ether	Prolabo VWR, Darmstadt (DE)
Dipotassium hydrogenphosphate	AppliChem, Darmstadt (DE)
Dimethylsulfoxide (DMSO)	Sigma-Aldrich, Taufkirchen (DE)
Ethanol	Merck, Darmstadt (DE)

Ethyl acetate	Merck, Darmstadt (DE)
Ethylenediaminetetraacetic acid (EDTA)	Roth, Karlsruhe (DE)
Formaldehyde	AppliChem, Darmstadt (DE)
Formic acid	VWR, Darmstadt (DE)
Glycerol	AppliChem, Darmstadt (DE)
Glycine	TH. GEYER, Lohmar (DE)
HEPES	AppliChem, Darmstadt (DE)
Hexadecane	Sigma-Aldrich, Taufkirchen (DE)
Hexane	Merck, Darmstadt (DE)
Isopropyl β-d-1-thiogalactopyranoside (IPTG)	Formedium, Norfolk (UK)
Iron (II) sulfate heptahydrate	Merck, Darmstadt (DE)
Isopropanol	AppliChem, Darmstadt (DE)
Kanamycin	Roth, Karlsruhe (DE)
Magnesium chloride hexahydrate	AppliChem, Darmstadt (DE)
Magnesium sulfate heptahydrate	AppliChem, Darmstadt (DE)
Manganese chloride tetrahydrate	AppliChem, Darmstadt (DE)
Methanol	J.T. Baker, Phillipsburg (USA)
β-Mercaptoethanol	AppliChem, Darmstadt (DE)
N-Hydroxysuccinimide	Sigma-Aldrich, Taufkirchen (DE)
N-Methyl-N-(trimethylsilyl)trifluoracetamide (MSTFA)	CS-Chromatographie-Service GmbH, Langerwehe (DE)
Midori Green Advance	NIPPON Genetics GmbH, Düren (DE)
n-Hexane	VWR, Darmstadt (DE)

N,N’-Dicyclohexylcarbodiimide	Sigma-Aldrich, Taufkirchen (DE)
Potassium bromide	Roth, Karlsruhe (DE)
Potassium chloride	Merck, Darmstadt (DE)
Potassium dihydrogenphosphate	Merck, Darmstadt (DE)
Potassium hydroxide	Merck, Darmstadt (DE)
Potassium nitrate	Grüssing, Filsum (DE)
Pyridine	Sigma-Aldrich, Taufkirchen (DE)
Rotiphorese ® Gel (29:1)	Roth, Karlsruhe (DE)
Sodium bicarbonate	AppliChem, Darmstadt (DE)
Sodium chloride	Duchefa Biochemie, Haarlem (NL)
Sodium dihydrogenphosphate	Sigma-Aldrich, Taufkirchen (DE)
Sodium fluoride	Sigma-Aldrich, Taufkirchen (DE)
Sodium molybdate dihydrate	Merck, Darmstadt (DE)
Sodium nitrate	Roth, Karlsruhe (DE)
Sodium pyruvate	Sigma-Aldrich, Taufkirchen (DE)
Streptomycine	Duchefa Biochemie, Haarlem (NL)
Strontium chloride hexahydrate	Roth, Karlsruhe (DE)
Tetracycline	AppliChem, Darmstadt (DE)
Tris(hydroxymethyl)aminomethane (Tris)	AppliChem, Darmstadt (DE)
Triton-X 100	Roth, Karlsruhe (DE)
Water, HPLC grade	VWR, Darmstadt (DE)
Zink chloride	Sigma-Aldrich, Taufkirchen (DE)

2.1.4. Kits and Enzymes

BbsI-HF	New England BioLabs GmbH, Frankfurt (DE)
Bovine serum albumin (20 mg/mL)	New England BioLabs GmbH, Frankfurt (DE)
BsaI-HFv2	New England BioLabs GmbH, Frankfurt (DE)
CloneJet PCRcloning Kit	Macherey-Nagel, Düren (DE)
Color Prestained Protein Standard (11-245 kDa)	New England BioLabs GmbH, Frankfurt (DE)
CutSmart buffer (5 x)	New England BioLabs GmbH, Frankfurt (DE)
DCS DNA Polymerase	DNA Cloning Service, Hamburg (DE)
DCS Reaction Buffer 10 x	DNA Cloning Service, Hamburg (DE)
dNTP	DNA Cloning Service, Hamburg (DE)
GeneRulerTM 1 kb DNA Ladder	New England BioLabs GmbH, Frankfurt (DE)
Lysozyme	Sigma-Aldrich, Taufkirchen (DE)
NucleoSpin Gel and PCR Clean-up Kit	Macherey-Nagel, Düren (DE)
NucleoSpin Plasmid Mini Kit	Macherey-Nagel, Düren (DE)
Q5 High-Fidelity DNA polymerase	New England BioLabs GmbH, Frankfurt (DE)
Q5 High GC Enhancer 5 x	New England BioLabs GmbH, Frankfurt (DE)
Q5 Reaction Buffer 5 x	New England BioLabs GmbH, Frankfurt (DE)
QuantiFluor ONE dsDNA system	Promega (USA)
T4 DNA ligase with 10 x buffer	New England BioLabs GmbH, Frankfurt (DE)

2.2. Methods

2.2.1. Cultivation and Transformation of Organisms

2.2.1.1. Cultivation and Transformation of *Alcanivorax borkumensis*

Alcanivorax borkumensis SK2 was grown in Marine Broth (Difco 2216, Fisher Scientific) with 1 % (w/v) of pyruvate or 1 % (v/v) of hexadecane as the carbon source. To study the glycine-glucolipid production under different conditions, the cells were grown in ONR7a medium modified by Tobias Karmainski from RWTH Aachen (Karmainski, Dielentheis-Frenken et al. 2023). The composition of ONR7a medium was as follows: 22.79 g/L NaCl, 0.72 g/L KCl, 3.98 g/L Na_2SO_4, 0.083 g/L NaBr, 0.031 g/L $NaHCO_3$, 0.027 g/L H_3BO_3, 0.0026 g/L NaF, 11.18 g/L $MgCl_2$ x 6 H_2O, 1.46 g/L $CaCl_2$ x 2 H_2O, 0.024 g/L $SrCl_2$ x 6 H_2O, 0.01 g/L $FeSO_4$ x 7 H_2O, 0.005 g/L $MnSO_4$ x H_2O, 0.0064 g/L $ZnCl_2$, 0.0004 g/L $CoCl_2$ x 6 H_2O, 0.0002 g/L Na_2MoO_4 x 2 H_2O, 0.01274 g/L Na_2EDTA x 2 H_2O, 0.46 g/L NaH_2PO_4, 2 g/L NH_4Cl, 50 mM HEPES (pH 7 for pyruvate and pH 7.5 for hexadecane) with 1 % (w/v) sodium pyruvate or 1 % (v/v) hexadecane as carbon source. For phosphate deprivation, NaH_2PO_4 was omitted from the medium. Cultures were incubated at 28 °C under shaking with 180 rpm. Due to the agitation during growth, the medium components were homogenously dispersed without any accumulation of nutrient components on the surface, the interior or the bottom of the culture broth. Cells were harvested for lipid extraction after 3 days in late exponential growth phase. To obtain sufficient number of cells accumulate at the oil-water interphase, 5 % of hexadecane was used when growing the cells for TEM and SEM.

2.2.1.2. Cultivation and Transformation of *Escherichia coli*

The ElectroSHOX electrocompetent strain has high efficiency in both transformation of large plasmids and the production of plasmids. The ElectroSHOX strain was used for the cloning to produce recombinant plasmids. For the protein production, the BL21 DE3 electrocompetent strain was used because of its high-level protein expression and easy induction. The XL1 blue electrocompetent strain was chosen for the glycine-glucolipid production. All three different *E. coli* strains were incubated at 37 °C under shaking with 180 rpm in Lysogeny Broth (LB) medium. Antibiotics were applied according to the selection markers on the respective plasmids.

2.2.1.3. Feeding of *Escherichia coli* with Fatty Acids

Molecular cloning constructs were grown in LB medium at 37 °C, 180 rpm overnight. The pre-

culture was subsequently diluted 20 folds and grown at 37 °C, 180 rpm until the cultures reached an OD_{600} > 0.8. Once the desired OD_{600} was achieved, the cultures were induced with IPTG at a final concentration of 0.5 mM and incubated at 16 °C, 180 rpm overnight. For the fatty acid assay, the fatty acids were dissolved in methanol in the concentration of 5 mM. The culture was diluted with an equal volume of LB medium. For each culture, either 3-OH-10:0 or 16:0 in methanol was added to reach a final concentration of 2.5 µM. An equivalent volume of methanol was added as control to exclude any potential growth inhibition caused by fatty acid solvent, methanol. The cultures were incubated at 37 °C, 180 rpm for 5 h. The OD_{600} was measured before harvesting for lipid analysis.

2.2.1.4. Generation of Electrocompetent *Escherichia coli* Cells

The pre-culture of the *E. coli* cells was grown from the prepared competent cells in 20 mL of SOB medium with antibiotics (ElectroSHOX: 50 mg/L streptomycin, Bl21 DE3: 40 mg/L chloramphenicol, XL1 blue: 12.5 mg/L tetracycline) at 37 °C, 180 rpm overnight. The pre-culture was diluted 20 folds and grown at 37 °C, 180 rpm until the cultures reached an OD_{600} of 0.5. The culture was chilled on ice for 30 min at 4 °C before harvesting by centrifugation at 4000 rpm, 10 min. The cells were resuspended in 50 mL of ice-cold 1 mM HEPES buffer. The cells were washed 3 times with 50 mL of ice-cold sterile water. Prior to the final resuspension in 2 mL of 10 % ice-cold glycerol, the cells were initially resuspended in 20 % ice-cold glycerol. Finally, 50 µL of the cells were aliquoted into microcentrifuge tubes and flash-frozen in liquid nitrogen. The aliquots were either used directly or stored at -80 °C.

2.2.1.5. Cultivation of *Nicotiana benthamiana*

N. benthamiana seeds were stratified at 4 °C for 7 days on wet tissue paper in a Petri dish to initiate germination. Then the seeds were sown into pots containing a soil and Vermiculite mixture (3:1) that had been pre-soaked with 1 mM of boric acid (pH 8.0). The pots were placed in a growth chamber set to 21 °C, 55 % humidity, 150 µmol/(m^2 s) light intensity under long-day conditions (16 h of light per day). After 22 days of growth, the plantlets were transferred to hydroponic system using modified Hoagland medium. The modified Hoagland medium contained the following components: 5 mM $Ca(NO_3)_2{\cdot}4H_2O$, 5 mM KNO_3, 0.5 mM K_2HPO_4, 0.5 mM KH_2PO_4, 2 mM $MgSO_4{\cdot}7H_2O$, 45 µM H_3BO_3, 9.1 µM $MnCl_2{\cdot}4H_2O$, 0.7 nM $ZnSO_4{\cdot}7H_2O$, 0.4 nM $CuSO_4{\cdot}5H_2O$, $Na_3MoO_4{\cdot}2H_2O$, 37 mg/L Fe-EDTA and 15 mM MES pH 6.0. The medium was replaced every three days. The phosphate starvation treatment commenced 12 days after plantlets had adapted to the

hydroponic system. For phosphate deprivation, K_2HPO_4 and KH_2PO_4 were omitted from the medium. Phosphate deficiency symptoms in the plants began to appear after 12 days of this treatment. Following the onset of phosphate deprivation symptoms, a phosphate solution (5 mM of K_2HPO_4 and 5 mM of KH_2PO_4) with different surfactant (120 mg/L of decylmaltopyranoside, Triton X 100 and glycine-glucolipid) was applied to leaves every three days for 24 days. The leaves were harvested for lipid extraction. After the extraction, the leaves were air dried for the dry weight measurement.

2.2.2. Methods in Molecular Biology

2.2.2.1. Isolation of the Genomic DNA from *Alcanivorax borkumensis*

For the extraction of genomic DNA, *A. borkumensis* cells were collected and transferred into a 1.5 mL microcentrifuge tube containing 4 to 6 small ceramic beads. The collected material was immediately transferred into liquid nitrogen followed by a homogenization step which was performed 3 times for 30 sec at 5900 rpm with the Precellys homogenizer. After adding 1 mL of CTAB extraction buffer, samples were incubated at 65 °C for 10 min and vortexed every 2 min during the incubation to increase the extraction efficiency. Additional 0.4 mL of chloroform were added before centrifugation at 950 x g for 5 min to achieve phase separation. The aqueous (upper) phase was transferred into a fresh 1.5 mL microcentrifuge tube. 0.7 mL of ice-cold isopropanol was added followed by an incubation at -20 °C for at least 10 min to precipitate the DNA. Afterwards the tubes were centrifuged at 10650 x g for 5 min. The DNA pellet was washed with 70 % (v/v) ethanol and centrifuged again for 5 min at 9300 rpm. The washing step was repeated twice. The supernatant was discarded and the DNA was allowed to dry completely under N_2 flow. In the last step the DNA was resuspended in 20 µL ddH_2O. The Quantus Fluorometer was used for the analysis of DNA concentration and quality. Extracted DNA was stored at -20 °C for further use.

CTAB Extraction Buffer

Sorbitol	140 mM	NaCl	800 mM
Tris-HCl, pH 8	220 mM	Sarkosyl (N-lauroylsarcosine)	1 % (w/v)
EDTA	22 mM	CTAB	0.8 % (w/v)

2.2.2.2. Polymerase Chain Reaction PCR

The Polymerase Chain Reaction (PCR) is a widely-used method for amplification of DNA fragments. The process involves denaturation of the double-stranded DNA by heating it to 95 °C, which separates the strands and makes them single-stranded. The denaturation is followed by annealing of the primers (Integrated DNA Technologies (IDT)), and finally elongation of the sequence of interest by a heat-stable polymerase usually at 72 °C. Q5 polymerase from NEB was used for the amplification of the fragments for cloning. The annealing temperature was calculated by NEB Tm calculator taking into account the alignments to the template, the polymerase used and the concentration of primers. For colony PCR, the DCS polymerase was used and the annealing temperature was same as for the Q5 polymerase.

2.2.2.3. Agarose Gel-Electrophoresis

Agarose gel electrophoresis was performed for the analysis of PCR products. 1.2 % (w/v) agarose in 1 x TAE buffer was used to separate DNA fragments. Prior to gel casting, Midori Green Advance was added to the melted agarose gel to a final concentration of 5 µL/100 mL. Midori Green Advance-DNA complexes were visualized under a Blue/Green LED transilluminator at a wavelength of 470-520 nm. The size of the PCR products was estimated with the GeneRuler 1 kb DNA ladder (Thermo Fisher Scientific). Electrophoresis was carried out by applying an electric voltage of 120 V until the desired separation was achieved.

50 x TAE Buffer

Tris base	2 M
Glacial acetic acid	2 M
EDTA, pH 8	50 mM

2.2.2.4. Purification of PCR Products

DNA fragments were purified with the NucleoSpin Gel and PCR Clean-up Kit (Macherey-Nagel) following the manufacturer's instructions. The purified DNA were eluted in 25 µL ddH_2O instead of the elution buffer indicated in the instructions. The purified DNA was stored at -20 °C or directly used for analysis. The Quantus Fluorometer was used for the analysis of DNA concentration and quality.

2.2.2.5. Isolation of Plasmids from *Escherichia coli*

Plasmid DNA was purified from 4 mL cultures grown at 37 °C using the NucleoSpin Plasmid Mini Kit (Macherey-Nagel) following the manufacturer's instructions. The purified plasmids were eluted in 30 µL ddH_2O instead of the elution buffer indicated in the instructions. The purified plasmids were stored at -20 °C or directly used for analysis. The Quantus Fluorometer was used for the analysis of DNA concentration and quality.

2.2.2.6. Golden Gate Cloning

Golden Gate Cloning was performed in PCR tubes. All the DNA fragments and plasmids were adjusted to equal molarities. Restriction enzymes and buffers were added into the mixture according to the Table 1. T4 ligase was added always after the pre-digestion. The Golden Gate Cloning reactions were performed in a Thermocyler with setup according to the Table 1.

Table 1: Golden Gate mixture and Thermocycler program applied for Golden Gate cloning.

Reaction Mixture		**Thermocycler Program**			
DNA fragments	~100 pmol	**Step**	**Temperature**	**Time**	**Cycle**
BSA (20 mg/mL)	2 µL	Pre-digestion	37 °C	2 h	
ATP	2.5 µL	Ligation	20 °C	20 min	X 25
Restriction enzyme	2 µL	Digestion	37 °C	10 min	
T4 DNA ligase	2 µL	Inactivation of the enzymes	80 °C	15 min	
Cutsmart buffer	2.5 µL	Incubation	15 °C		
ddH_2O	ad 25 µL				

2.2.2.6.1. Triple Expression of *GglsA*, *GglsB*, *GglsC*

The vector pTrcHis2C (Invitrogen) was modified to accommodate three genes via Golden Gate cloning by Dr. Georg Hölzl (IMBIO, University of Bonn). The Tac promoter (Tac-a) was amplified from pGEX-3X (GE Healthcare) using the primers Bn3415, Bn3416. The T5 promoter (T5-a), which includes a His_6 tag, was amplified from pQE-80L (Qiagen, Hilden, Germany) using primers Bn3417,

Bn3418. The kanamycin (Kan^R) cassette was amplified from pGEX-3X using the primers Bn3424, Bn3422. GglsB-a and GglsC-a were obtained by PCR amplification of genomic DNA of *A. borkumensis* using primers Bn4633, Bn4634 and Bn4473, Bn4474 respectively. The *Bsa*I restriction site in the GglsA gene was removed by introducing a silent mutation at position 1368 (GAC to GAT). For this purpose, the 5' and 3' parts of the GglsA gene (GglsA-1, GglsA-2) were separately amplified from genomic DNA using the primer pairs Bn4631, Bn4468 and Bn4469, Bn4632. Golden Gate cloning was performed in a reaction containing the pTrcHis2C vector, the promoters Tac-a, T5-a, the Kan^R cassette and the gene fragments (GglsA-1, GglsA-2, GglsB-a, GglsC-a) to generate the construct pTrc-GglsB-Tac:GglsC-T5:His_6-GglsA-Kan^R.

2.2.2.6.2. Co-Expression of *GglsA*, *GglsC*

To generate the construct pTrc-GglsC-Tac-T5:His_6:GglsA-Kan^R, the T5 promoter (T5-b) was amplified from pQE-80L using the primers Bn3420, Bn3418. Golden Gate cloning was performed with pTrcHis2C, the Tac-a and T5-b promoters, the Kan^R cassette and the DNA fragments (GglsA-1, GglsA-2, GglsC).

2.2.2.6.3. Single Expression of *GglsC*

The GglsC-b gene was amplified by PCR from genomic DNA of *A. borkumensis* using the primers Bn4552, Bn4553. The Tac promoter (Tac-b) was PCR amplified from pTrcHis2C using the primers Bn3419, Bn3416. The Tac-b, T5-b promoters and the GglsC-b gene were ligated into pTrcHis2C via Golden Gate cloning to obtain the construct pTrc-Tac-T5:His_6-GglsC-Kan^R.

2.2.2.6.4. Single Expression of *GglsA*

For the construct pTrc-Tac-T5:His_6-GglsA-Kan^R, Golden Gate cloning was performed with the pTrcHis2C vector, the Tac-b and T5-b promoters, the Kan^R cassette and the DNA fragments (GglsA-1, GglsA-2).

2.2.2.6.5. Co-Expression of *GglsA, GglsB, ENTDLIKE (ABO_2807)*

To generate the construct pTrc-GglsB-Tac:EntDlike-T5:His_6-GglsA-Kan^R, the EntDlike gene was PCR amplified from genomic DNA using the primers Bn4865, Bn4866. Golden Gate cloning was performed with the pTrcHis2C vector, the promoters Tac-a and T5-a, the Kan^R cassette and the genomic fragments (GglsA-1, GglsA-2, GglsB-a, EntDlike).

2.2.2.6.6. Domain Swapping of *GglsA-A* with *EntFA*

To swap the adenylation (A) domains of GglsA with that of EntF (EntFA), the EntFA domain was PCR amplified from genomic DNA of *E. coli* XL1-Blue cells using the primers Bn4690, Bn4691. The condensation (C) domain (GglsA-C) and the sequence containing the phosphopantetheine (T) and thioesterase (TE) domains (GglsA-T-TE) were amplified from the plasmid pTrc-Tac-T5-His_6-GglsA-Kan^R using the primers Bn4631, Bn4689 and Bn4692, Bn4632, respectively. Golden Gate cloning was performed with the pTrcHis2C vector, the promoters Tac-b and T5-b, the KanR cassette, and the amplified DNA fragments (GglsA-C, GglsA-T-TE, EntFA,) to obtain the construct pTrc-Tac-T5:His_6-EntFA-GglsA-Kan^R.

2.2.2.6.7. Domain Swapping of *GglsA-C* with *EntFC* and *SrfAAC1*

To swap the condensation (C) domains of GglsA with that of EntF (EntFC) and SrfAA (SrfAAC1), the EntFC domain was PCR amplified from genomic DNA of *E. coli* XL1-Blue cells using the primers Bn4990, Bn4991. The SrfAAC1 domain was PCR amplified from plasmid pUCIDT-SrfAAC1 purchased from Integrated DNA Technologies (IDT, Leuven, Belgium) using the primers Bn4686, Bn4687. The sequence containing the adenylation (A), phosphopantetheine (T) and thioesterase (TE) domains (GglsA-A-T-TE) were amplified from the plasmid pTrc-Tac-T5-His_6-GglsA-Kan^R using the primers Bn4688, Bn4632. Golden Gate cloning was performed with the pTrcHis2C vector, the promoters Tac-b and T5-b, the Kan^R cassette, and the amplified DNA fragments (EntFC or SrfAAC1 and GglsA-A-T-TE) to obtain the construct pTrc-Tac-T5:His_6-EntFC-GglsA-Kan^R or pTrc-Tac-T5:His_6-SrfAAC1-GglsA-Kan^R, respectively.

2.2.2.6.8. Expression of *GglsA* with Ser1177Ala Mutation in the TE Domain

The 5' and 3' parts of GglsA (GglsA-3 and GglsA-4) were amplified from pTrc-Tac-T5:His_6-GglsA-Kan^R using the primers Bn4631, Bn4789 and Bn4788, Bn4632, respectively, thereby changing two base pairs (TCC to GCG) leading to a Ser1177Ala exchange in the TE domain. Golden Gate cloning was performed with the pTrcHis2C vector, the Tac-b and T5-a promoters, the KanR cassette and the DNA fragments (GglsA-3, GglsA-4, GglsC), to generate the construct pTrc-Tac:GglsC-T5:His_6-GglsA-Ser1177Ala-Kan^R.

2.2.2.6.9. Single Expression of *GglsB*

The gene GglsB-b was amplified from genomic DNA of *A. borkumensis* using primers Bn4471, Bn4472. The gene was ligated into pTrcHis2C via Golden Gate cloning to obtain the construct pTrc-Tac-T5:His_6-GglsB-Kan^R.

2.2.2.6.10. Single Expression of the *GglsA* T domain (*GglsA-T*) and Co-Expression of *GglsA-T* with *GglsC* or *ENTDLIKE (ABO_2807)*

The T domain of GglsA (GglsA-T) was amplified from genomic DNA of *A. borkumensis* using primers Bn4772, Bn4773. Golden Gate cloning was performed with the vector pTrcHis2C, the Tac-b and T5-b promoters, the Kan^R cassette and GglsA-T to obtain the construct pTrc-Tac-T5:His_6:GglsA-T-Kan^R. For co-expression of Ggls-T with GglsC, Golden Gate cloning was performed with the pTrcHis2C vector, the Tac-b and T5-a promoters, the Kan^R cassette, and the genes GglsA-T and GglsC-a to obtain the construct pTrc-Tac:GglsC-T5:His_6-GglsA-T-Kan^R. The construct pTrc-Tac:EntDlike-T5:His_6-GglsA-T-Kan^R was created in the same way, but the ENTDLIKE PCR product was used instead of GglsC-a.

2.2.2.6.11. Mutagenesis of Ser1039 in the T domain of *GglsA* and Co-Expression with *GglsC*

To introduce the mutation of Ser1039Ala in GglsA-T, two base pairs (TCT to GCG) were changed by separate amplification of the 5' and 3' parts of GglsA-T (T1, T2) using the primers Bn4772, Bn4791 and Bn4790, Bn4773. Golden Gate cloning was performed with the pTrcHis2C vector, the Tac-b and T5-b promoters, the Kan^R cassette and the T1 and T2 fragments of GglsA-T, to obtain the construct pTrc-Tac-T5:His_6-GglsA-T-Ser1039Ala-Kan^R.

2.2.2.6.12. Co-Expression of *GglsA-T*-Ser1039Ala with *GglsC*

Golden Gate cloning was performed with the pTrcHis2C vector, the Tac-b and T5-a promoters, the Kan^R cassette, and the genes GglsA-T-Ser1039Ala and GglsC-a to obtain the construct pTrc-Tac:GglsC-T5:His_6-GglsA-T-Ser1039Ala-Kan^R.

Constructs were transferred into *E. coli* BL21(DE3) cells or XL1-Blue cells for enzyme assays or *in vivo* production of lipids, respectively. After growth at 37°C to an OD_{600} of 0.8, protein expression was induced with 0.5 mM isopropyl-β-thiogalactoside (IPTG) at 16°C overnight. Cells were harvested by centrifugation, and lipids isolated for LC-MS analysis or protein extracted for enzyme assays.

2.2.2.7. Generation of *A. borkumensis* Mutants

The mutants *ΔGglsA* (ABO_1784), *ΔGglsB* (ABO_1783), *ΔGglsC-1* (ABO_1782), *ΔEntDike* (ABO_2087), *ΔphaC1* (ABO_2214), *ΔphaC2* (ABO_1418), *ΔatfA1* (ABO_2742) and *ΔatfA2* (ABO_1804) were generated through homologous recombination. *ΔGglsA*, *ΔGglsB*, *ΔGglsC*, *ΔEntDlike* strains were generated by Maximilian Fassl, and *ΔphaC1* (ABO_2214), *ΔphaC2* (ABO_1418) were generated by Maya Marita Dierig under the supervision of Dr. Georg Hölzl (University of Bonn).

The target gene was replaced by an antibiotic resistance gene (kanamycin) as a selection marker for deletion. Homologous regions from the 5' and 3' end of the target gene open reading frame, with a Kan^R cassette in between, were cloned into the pJ-GG-lacZ vector modified by Dr. Georg Hölzl (unpublished). The resulting construct was transferred into ElectroSHOX cells to produce large amount of plasmid for *A. borkumensis* transformation. Wild type competent cells of *A. borkumensis* were prepared by Maximilian Fassl. Transformation of *A. borkumensis* with deletion construct was achieved via electroporation. The *A. borkumensis* strains with successful gene replacement show kanamycin resistance. The colonies were further confirmed via colony PCR.

2.2.3. Methods in Lipid Analysis

2.2.3.1. Lipid Extraction from Bacteria

Prior to lipid extraction, the bacterial cultures were vortexed, an aliquot transferred into cuvettes and the OD_{600} was measured against a blank containing the corresponding medium without cells, to quantify the biomass of the cells.

For lipid extraction, the cells were harvested by centrifugation for 30 min at 7000 rpm or 10 min at 4000 rpm for *A. borkumensis* or *E. coli*, respectively. Then the cell pellets were resuspended in a minimal volume of deionized H_2O and boiled at 100 °C for 15 min. After cooling, 1 nmol of MGD di-18:2 (concentration adjusted by Helga Peisker, University of Bonn) was added as internal standard, along with two volumes of $CHCl_3/CH_3OH$ (1:2). The mixture was shaken at 4 °C for 1 h. After centrifugation for 20 min at 2000 rpm, the organic phase was collected, and the solvent evaporated under N_2 gas. The extraction was repeated two more times with two volumes of (ii) $CHCl_3/CH_3OH$ (2:1) and (iii) $CHCl_3$, respectively. After centrifugation, the organic phases were combined with the first extract. The crude lipid extract was washed with 0.9 % NaCl in a final ratio of $CHCl_3/CH_3OH$/0.9 % NaCl of 2:1:0.75. The organic phase was again concentrated under N_2 gas

and the lipids dissolved in the solvent according to the further usage.

2.2.3.2. Lipid Extraction from *Nicotiana benthamiana*

For the analysis of glycine-glucolipid foliar application, total lipids were extracted from *N. benthamiana* leaves. For all the leaf samples, lipid extraction was performed in glass vials rinsed with chloroform/methanol 2:1 (v/v). The leaf tissue was boiled in deionized H_2O at 100 °C for 20 min to inactivate the lipases. After cooling, the leaf tissue was transferred to a new glass vial with 2 mL of chloroform/methanol 1:2 (v/v). The leaf tissue was extracted for 1 h at 4 °C (E1). All the E1 was harvested and transferred into a new glass vial. The leaf tissue was extracted a second time with 2 mL of chloroform/methanol 2:1 (v/v) for 1 h at 4 °C (E2). The E2 was combined with E1 and the solvent was evaporated under N_2 gas. The crude lipid extract was washed with 0.9 % NaCl in a final ratio of chloroform/methanol/0.9 % NaCl of 2:1:0.75 (v/v/v). The organic phase was again concentrated under N_2 gas and the lipids were dissolved in the solvent according to the further usage of the lipid. The leaf tissue was dried under the hood for the dry weight measurement.

2.2.3.3. Thin Layer Chromatography (TLC)

Lipid characterization and separation was achieved via TLC. A total lipid extract from around 5 mL culture was loaded on a TLC plate using different solvents as mobile phase.

For the 2-dimensional TLC, the plate was firstly developed using chloroform/methanol/water (65:25:4, v/v/v). After evaporating the solvent, the plate was rotated by 90°, and the second dimension was developed using chloroform/methanol/acetic acid/water (90:15:10:3.5, v/v/v/v).

For the 1-dimensional TLC, the plate was developed using chloroform/methanol/acetic acid/water (90:15:10:3.5, v/v/v/v) for the total lipid separation. For the separation of storage (non-polar) lipids with 1-dimensional TLC, hexane/diethylether/acetic acid (80:20:1, v/v/v) was used as mobile phase as described (Kalscheuer, Stöveken et al. 2007). Triolein (Merk/SigmaAldrich) and oleyl oleate (self-made by Dr. Margrit Frentzen, RWTH Aachen) were used as reference substances for TAG and WE, respectively.

After the solvent had reached the end of the plates, the plate was removed from the development chamber, dried, and the lipids were stained with iodine vapor for visualization. The glycolipids were stained in purple color with α-naphthol/sulfuric acid (Hölzl and Dörmann 2021). For further lipid analysis and purification, the lipids were stained with Primuline (White, Bursten et al. 1998)

and observed under UV light. The lipid spots from the TLC plates were scrapped off and lipids were extracted from the silica material similar to the extraction from bacteria cell 2.2.3.1..

2.2.3.4. Lipid Purification

The crude lipid extract, dissolved in chloroform, was purified via solid phase extraction (SPE) on silica columns (Strata Silica SI-1, Phenomenex) that had been equilibrated with chloroform. After loading the crude lipid sample dissolved in chloroform, the neutral lipids were eluted from the column with chloroform. Then the glycine-glucolipid was eluted with acetone/isopropanol (9:1, v/v). The glycolipid fraction was further purified by separation on a silica thin layer chromatography (TLC) plate, which was developed with $CHCl_3$/CH_3OH/CH_3COOH/water (90:15:10:3.5, v/v/v/v) as mobile phase. Analytical plates were stained with 2.4 % (w/v) α-naphthol in 10 % sulfuric acid/80 % ethanol to visualize the glycolipids. For preparative separation, the TLC plates were stained with ammonium-8-anilino-1-naphthalinsulfonate (ANS) and observed under UV light. The silica material containing the glycine-glucolipid was isolated from the TLC plate and lipids were extracted with $CHCl_3$/CH_3OH (2:1, v/v).

To investigate whether the glycine-glucolipid can be excreted, *A. borkumensis* cells were grown for 96 hours in 100 mL ONR7a medium with 1 % (w/v) pyruvate or 1 % (v/v) hexadecane. The culture was then divided into two equal portions. For one portion, half of the culture was extracted twice with an equal volume of ethyl acetate. For the other portion, the remaining half of the culture was centrifuged at 14000 rpm for 15 min. The resulting cell pellet was extracted twice with 50 mL of ethyl acetate. The supernatant, obtained from the second portion after centrifugation, was purified by filtration (Filtropur V50; pore size, 0.45 µm; Sarstedt) and extracted twice with an equal volume of ethyl acetate. The lipids in the ethyl acetate phases were dried under nitrogen gas and dissolved in 200 µL $CHCl_3$. The glycine-glucolipid was purified by SPE, and measured by HPLC-DAD or direct infusion mass spectrometry.

2.2.3.5. Synthesis of Fatty Acid Methyl Esters (FAMEs)

The four 3-hydroxy fatty acids of the glycine-glucolipid can be quantified after methylation and trimethylsilylation by GC-MS in the presence of an internal standard (Shantha and Napolitano 1992). For the analysis of the crude lipid extract or the glycine-glycolipid purified by SPE and TLC, 50 or 5 µg, respectively, of internal standard (3-hydroxy-dodecanoic acid, 3-OH-12:0) were added, and the solvent was removed with N_2 gas. The glycine-glucolipid was dissolved in 1 mL of 3 N methanolic HCl. Methylation was carried out at 80°C for 5 h to convert all 3-hydroxy-fatty acids

into the methyl esters. The samples were cooled down and 1 mL of hexane and 1 mL of 0.9 % NaCl were added. After mixing and short centrifugation, the hexane phase was collected and dried with N_2 gas. Free hydroxyl groups were silylated with 50 µL of N-methyl-N-(trimethylsilyl)-trifluoroacetamide (MSTFA) and 150 µL of pyridine by incubation at 80°C for 30 min. The solvent was evaporated under N_2 gas and the derivatized 3-hydroxy-fatty acids dissolved in hexane. The samples were directly measured by GC-MS or stored at – 20 °C.

2.2.3.6. Synthesis of Acyl-CoA

(R,S)-3-Hydroxydecanoyl-CoA was synthesized following published methods (Cronan Jr and Klages 1981, Rehm, Kruger et al. 1998). (R,S)-3-hydroxydecanoic acid (30 µmol) was added to a solution of 30 µmol N-hydroxysuccinimide in 150 µl of dry ethyl acetate. A solution of 30 µmol dicyclohexylcarbodiimide in dry ethyl acetate was added, and the reaction was incubated at room temperature overnight. Dicyclohexylurea was removed by centrifugation and the supernatant was concentrated under N_2 flow to yield crystals of 3-hydroxydecanoic acid N-hydroxysuccinimide ester. The crystals were recrystallized from ethanol and then dissolved in 300 µL of methanol.

3-Hydroxydecanoyl-CoA was synthesized with 10 mU of acyl-CoA synthetase from *Pseudomonas* (Sigma-Aldrich) in 100 µL of 50 mM Tris-HCl, pH 7.5 containing 2 mM ATP, 5 mM $MgCl_2$, 2 mM coenzyme A (lithium salt) and 2 mM 3-hydroxydecanoic acid N-hydroxysuccinimide ester for 60 min at 28 °C. The reaction was stopped by addition of 400 µL hexane that was saturated with isopropanol/water and vortexing. After centrifugation (2000 g, 3 min), 3-hydroxyldeacnoyl-CoA in the lower aqueous phase was harvested. Saturated ammonium sulfate (2 µL) and three volumes of chloroform/methanol 1:2 (v/v) were added, and the samples were incubated at room temperature for 20 min. After centrifugation at 13500 g for 2 min, the pellet containing ammonium sulfate, protein and salts was removed, while the organic supernatant with 3-hydroxydecanoyl-CoA was transferred to a microfuge tube. The solvent was removed in a vacuum concentrator without heating. 3-Hydroxyldecanoyl-CoA was overlaid with N_2 gas and stored at -80 °C. Samples were dissolved in water and quantified by GC-MS. Instead of 5 h incubation at 80 °C, the methylation of acyl-CoA was carried out at 80°C for 30 min.

2.2.3.7. Quantification of FAMEs via Gas Chromatography- Mass Spectrometry (GC-MS)

FAME samples were injected into an Agilent GC-MS (7890A GC system with 5975C inert XL MSD) equipped with an Agilent HP-5MS column. The following temperature gradient was used: starting at 50 °C, increased to 160°C by 25 °C/min, to 170 °C by 2 °C/min, then to 250 °C by 10 °C/min,

and finally decreased to 50 °C by 40 °C/min. The 3-hydroxy-fatty acids of the glycine-glucolipid (mostly HO-10:0 and HO-8:0) were quantified relative to the internal standard HO-12:0 taking into account the different response factors (RF) which were determined from the ratio of the peak area (A) to the amount (n, in nmol) determined by weighing.

$$RF(HO\text{-}8{:}\,0) = \frac{\mathrm{A(HO\text{-}8{:}\,0)}}{\mathrm{n(HO\text{-}8{:}\,0)}}$$

Response factors (area/nmol) for HO-8:0, HO-10:0, HO-12:0 and HO-14:0 were 1726916, 4499168, 6296387 and 6600466 respectively.

The relative response factors (RRF) were calculated from the RF of the 3-hydroxy fatty acid relative to the internal standard.

$$RRF(HO\text{-}8{:}\,0/HO\text{-}12{:}\,0) = \frac{\mathrm{RF(HO\text{-}8{:}\,0)}}{\mathrm{RF(HO\text{-}12{:}\,0)}}$$

The RF and RRF values are used to calculate the nmol amounts of 3-hydroxy fatty acids in the glycine-glucolipid. The total amount of glycine-glucolipid equals the sum of HO-8:0 and HO-10:0 divided by four.

$$n(HO\text{-}8{:}\,0) = \frac{\mathrm{A(HO\text{-}8{:}\,0)} * \mathrm{n(HO\text{-}12{:}\,0)}}{\mathrm{A(HO\text{-}12{:}\,0)} * \mathrm{RRF(HO\text{-}8{:}\,0/HO\text{-}12{:}\,0)}}$$

$$n(glycolipid\ GCMS) = \frac{\mathrm{n(HO\text{-}8{:}\,0)} + \mathrm{n(HO\text{-}10{:}\,0)}}{4}$$

2.2.3.8. Quantification of Hexadecane via Gas Chromatography-Mass Spectrum (GC-MS)

The dispersant effectiveness tests were conducted by measuring the hexadecane in the aqueous phase of the culture. *A. borkumensis* WT and mutant cells were grown in hexadecane containing ONR7a medium to early stationary phase as described in 2.2.1.1. The cultures were removed from the shaker and allowed to remain stationary for 10 min. 1 mL of sample was taken from the aqueous phase. The samples were then extracted with equal volume of hexane twice with 0.05 µL tetradecane as internal standard for quantification. The extracts were combined and resuspended in hexane. Hexadecane samples were injected into an Agilent GC-MS (7890A GC system with 5975C inert XL MSD) equipped with an Agilent HP-5MS column. The following temperature gradient was used: starting at 50 °C, increased to 150 °C by 15 °C/min, to 175 °C by 2.5 °C/min with a 5 min hold at 175 °C, then decrease to 50 °C by 40 °C/min.

The hexadecane was quantified relative to the internal standard tetradecane taking into account the different response factors (RF) which were determined from the ratio of the peak area (A) to the volume (V, in µL).

$$RF(hexadecane) = \frac{\mathrm{A(hexadecane)}}{\mathrm{V(hexaadecane)}}$$

Response factors (area/µL) for tetradecane and hexadecane were 90130842.1 and 1237071592 respectively.

The relative response factors (RRF) were calculated from the RF of the hexadecane relative to the tetradecane.

$$RRF(hexadecane/tetradecane) = \frac{\mathrm{RF(hexadecane)}}{\mathrm{RF(tetradecane)}}$$

The RF and RRF values are used to calculate the volume of hexadecane in the aqueous phase of the culture. The dispersant effectiveness (%) of the cells is calculated accordingly:

$$v(hexadecane) = \frac{\mathrm{A(hexadecane)} * \mathrm{v(tetradecane)}}{\mathrm{A(tetradecane)} * \mathrm{RRF(hexadecane/tetradecane)}}$$

$$Effectiveness(\%) = \frac{\mathrm{v(hexadecane)} * \mathrm{v(culture\ volume)}}{\mathrm{v(sample\ volume)} * \mathrm{v(hexadecane\ added\ in\ culture)}}$$

2.2.3.9. Measurement of Lipids via High-Performance Liquid Chromatography with Diode-Array Detection (HPLC-DAD)

The glycine-glucolipid was converted into a phenacyl ester for quantification according to a previous protocol with modifications (Mata-Sandoval, Karns et al. 1999). Undecanoic acid (11:0; 5 µg) was added to the lipid extract as internal standard and the solvent evaporated under N_2 gas. The lipid was dissolved in 100 µL of acetonitrile. Phenaycyl esters were prepared by incubation with 15 mg/mL of 2-bromoacetophenone and 7.5 mg/mL of triethylamine, at 80 °C for 1 h. The samples were centrifuged and the supernatants with the glycine-glucolipid phenacyl esters were transferred to HPLC vials. After injection of 25 µl of the sample, lipids were separated by HPLC (Agilent 1100 series with diode array detector, DAD) on a reversed phase column (Knauer C18, Eurospher II 100-3 C18A, 100 x 3 mm). The HPLC gradient was composed of solvent A (0.01 N H_3P0_4) and solvent B (acetonitrile) with 50 % B for 5 min; from 50 % B to 100 % B in 20 min; from 100 % B to 50 % B in 10 min, at a flow rate of 1.0 mL/min. The absorbance of the glycine-glucolipid phenacyl ester was measured at a wavelength of 244 nm.

The 3-hydroxy-fatty acids of an aliquot of the purified glycine-glucolipid were quantified via GC-MS, and the amount n (glycine-glucolipid GCMS) was calculated according to the equations shown above. After measuring the same amount of the glycine-glucolipid by HPLC-DAD, the response factor in HPLC-DAD was calculated:

$$RF(glycine-glucolipid) = \frac{\mathrm{A(glycine-glucolipid\ HPLC)}}{\mathrm{n(glycine-glucolipid\ GCMS)}}$$

$$RF(11:0) = \frac{\mathrm{A(11:0)}}{\mathrm{n(11:0)}}$$

$$RRF(glycine-glucolipid, 11:0) = \frac{\mathrm{RF(glycine-glucolipid)}}{\mathrm{RF(11:0)}}$$

The response factors (area/nmol) for 11:0 and the glycine-glucolipid were 115 and 170, respectively.

The amount of a molecular species n(Glc(O-10:0)$_4$Gly) of an unknown glycine-glucolipid preparation can be measured by HPLC-DAD and calculated according to the peak area related to the internal standard undecanoic acid (11:0, 5 µg), taking into account the relative response factor of the glycine- glucolipid:

$$n(Glc(O\text{-}10:0)4Gly\ HPLC) = \frac{\mathrm{A(Glc(O\text{-}10:0)4Gly)} * \mathrm{n(11:0)}}{\mathrm{A(11:0)} * \mathrm{RRF\ (glycine-glucolipid, 11:0)}}$$

The total amount of glycine-glucolipid measured by HPLC-DAD equals the sum of the amounts of the individual molecular species:

$$\mathrm{n(glycine\text{-}glucolipid\ HPLC) = n(Glc(O\text{-}10:0)_4Gly) + n((Glc(O\text{-}8:0)(O\text{-}10:0)_3Gly) + n(Glc(O\text{-}8:0)_2(O\text{-}10:0)_2Gly)}$$

2.2.3.10. Measurement of A. borkumensis Lipids via Direct Infusion Q-TOF Mass Spectrometry

Total lipid extracts from *A. borkumensis* or samples collected from HPLC-DAD were dried under nitrogen gas and dissolved in chloroform/methanol/300 mM ammonium acetate (300:665:35, v/v/v). The authentic glycine-glucolipid or its phenacyl esters were measured by direct infusion mass spectrometry (Agilent 6530 Accurate-Mass Q-TOF MS with chipcube nanospray source). Lipids were ionized in the positive mode yielding predominantly NH_4^+ adducts. The instrumental parameters were: drying gas N_2, 8 l/min; fragmentor voltage, 200 V; gas temperature, 300 °C; HPLC-Chip Vcap, 1700 V; scan rate, 1 spectrum/s, fragmentation energy Glc(O-10:0)$_4$Gly, 37.1 V;

Glc(O-10:0)$_4$Gly-phenacyl, 41.5 V. Data were processed using the Agilent MassHunter Qualitative Analysis and Microsoft Office Excel. Glycine-glucolipid abundances were calculated using the peaks of the parental NH_4^+ adducts.

2.2.3.11. Measurement of *N. benthamiana* Lipids via Direct Infusion Q-TOF Mass Spectrometry

The *N. benthamiana* lipids were diluted to 2 mg dw/mL and quantified in relation to internal standard prepared by Helga Peisker (University of Bonn). Each 10 µL of the internal standard contains 10.35 nmol di-14:0 PC, 13.25 nmol di-20:0 PC, 9.33 nmol di-14:0 PE, 12.1 nmol di-20:0 PE, 9.92 nmol di 14:0 PG, 11.25 di 20:0 PG, 6.96 nmol di-14:0 PA, 7.145 nmol di-20:0 PA, 1.41 nmol di-14:0 PS, 15.9 nmol 34:0 PI, 7.278 nmol 34:0 MGDG, 6.75 nmol 36:0 MGDG, 10.134 nmol 34:0 DGDG, 22.2 nmol 36:0 DGDG and 21.2 nmol 34:0 SQDG. 10 µL diluted lipid in chloroform was mixed with 10 µL of internal standard and 80 µL of QTOF solvent (chloroform/methanol/300 mM ammonium acetate (300:665:35, v/v/v)). Lipids were quantified by neutral loss or precursor ion scanning. The target list for lipid quantification is shown in the appendix. Data were processed using the Agilent MassHunter Qualitative Analysis and Microsoft Office Excel.

2.2.3.12. Measurement of Lipids via Q-TOF LC MS/MS

Lipids were isolated from *E. coli* or *A. borkumensis* cells as described (2.2.3.1) and dissolved in chloroform/methanol/300 mM ammonium acetate (300:665:35 v/v/v). The glycine-glucolipid Glc(O-10:0)$_4$Gly and its aglycone H(O-10:0)$_4$Gly were separated on an EC C8 Poroshell 120 Å, 2.7 µm, 50 mm x 2.1 mm, column (Agilent) using the following solvents: 20 mM ammonium acetate (A) and methanol (B). The lipids were measured with an Agilent 6530 Accurate-Mass Q-TOF mass spectrometer in the positive mode without fragmentation. Dioleoyl-monogalactosyl diacylglycerol (di-18:2 MGDG) was used as internal standard after determining the response factors for the glycine-glucolipid and the aglycone. The glycine-glucolipid was separated with a flow rate of 0.25 mL/min using a gradient of 1 min, 80 % B; 12 min, 100 % B; 17 min, 100 % B; 20 min, 80 % B; 25 min, 80 % B. The following $[M+NH_4]^+$ ions were recorded: Glc(O-10:0)$_4$Gly, $C_{48}H_{87}O_{15}N$ $+NH_4^+$, 935.6414 *m/z*; Glc(O-8:0)(O-10:0)$_3$Gly, $C_{46}H_{83}O_{15}N$ $+NH_4^+$, 907.6101 *m/z*; Glc(O-8:0)$_2$(O-10:0)$_2$Gly, $C_{44}H_{79}O_{15}N$ $+NH_4^+$, 879.5788 *m/z*. The aglycone was separated with a flow rate of 0.2 mL/min using a gradient of 1 min, 85 % B; 15 min, 100 % B; 25 min, 100 % B; 28 min, 80 % B; 35 min, 80 % B. The aglycone was recorded using the following $[M+NH_4]^+$ ions: (O-10:0)$_4$Gly, $C_{42}H_{77}O_{10}N + NH_4^+$, 773.5886 *m/z*; (O-10:0)$_4$Ser, $C_{42}H_{79}O_{10}N + NH_4^+$, 803.5991 *m/z*.

2.2.4. Method of Protein Biochemistry

2.2.4.1. Protein Extraction

E. coli cells were harvested by centrifugation at 4000 g for 15 min at 4 °C. The cell pellet from 100 mL culture was suspended in 1 mL of extraction buffer containing 50 mM Tris-HCl (pH 8.0) and 7 % v/v glycerol. The suspension was vortexed for 5 min and sonicated on ice six times with 10 s bursts at 300 W. The supernatant was collected after centrifugation at 10,000 g for 30 min at 4 °C, and the protein was concentrated with an Amicon Ultra-2 Centrifugal Filter Unit (3 kDa). The protein concentration was determined (Bradford 1976).

2.2.4.2. SDS-PAGE and Urea-PAGE

Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) was employed to separate proteins according to their electrophoretic mobility after denaturation (Laemmli 1970). SDS is a detergent which binds to polypeptides and masks their charges, which results in an overall negative charge, allowing a separation according to approximate protein size by migration through the polyacrylamide pores. The separating gel was prepared by mixing all components, but ammonium persulfate (APS) and tetramethylethylenediamine (TEMED) were added at last to induce polymerization. After this, the stacking gel was immediately poured between a glass plate and an alumina plate in a gel caster. To ensure an even top edge of the gel, it was covered with isopropanol (1 mL per gel). After 30 min polymerization, the isopropanol was removed and the stacking gel was poured on top. Immediately prior to pouring the stacking gel, polymerization was initiated by addition of TEMED and APS. After the stacking gel was poured, combs were inserted. After another 30 min, the gels were ready for use and stored at 4°C until use.

For loading, protein samples were mixed with one-third volume of 4x sample loading buffer and incubated for 5 min at 95°C to denature all proteins. Per gel slot, 10-30 µL protein solution was loaded along with the Color Pre-stained Protein Standard, Broad Range (NEB) as a reference for protein size. Electrophoresis was conducted initially at 25 mA and 200 V for a 64 cm^2 gel until the running front approached the separating gel. Then the current was increased to 35 mA until the electrophoresis was stopped.

For Urea PAGE, 10 % of SDS was replaced by 5 M of urea in all buffers.

5 x Loading Buffer		**4 x Loading Buffer**
Tris	500 mM	50 µL ddH_2O
SDS	10 % (w/v)	50 µL β-mercarptoethanol
Glycerol	50 % (v/v)	400 µL 5 x Loading Buffer
Bromophenol Blue	0.025 % (w/v)	
EDTA	5 mM	
Adjust the pH to 6.8		

Stock solution	**4 % Stacking Gel**		**Separating Gel**		
	SDS	**Urea**	**10 % SDS**	**20 % SDS**	**20 % Urea**
40 % Acrylamide	1 mL	1 mL	6.25 mL	12.5 mL	12.5 mL
0.5 M Tris-HCl pH 6.8	2.5 mL	2.5 mL	-	-	-
1.5 M Tris-HCl pH 8.8	-	-	6.25 mL	6.25 mL	6.25 mL
10 % SDS/5 M Urea	0.1 mL	3.003 g	0.25 mL	0.25 mL	7.507 g
H_2O	6.2 mL	6.2 mL	11.85 mL	5.6 mL	5.85 mL
TEMED	5 µL	5 µL	10 µL	10 µL	10 µL
10 % Ammonium persulfate	0.1 mL	0.1 mL	0.15 mL	0.15 mL	0.15 mL
Total volume	10 mL	10 mL	25 mL	25 mL	25 mL

2.2.4.3. Cysteamine Cleavage

Soluble proteins were isolated from *E. coli* cells harboring the construct pTrc-Tac:GglsC-T5:His6GglsA-Ser1177Ala-Kan^R. Cysteamine was added to the isolated proteins at a final concentration of 0.5 mM (Belecki and Townsend 2013, Hobson, Jenner et al. 2022). After 20 h of incubation at 4 °C, the reaction was extracted twice with 1 volume of ethyl acetate each. The combined extracts were evaporated under N_2 gas and the residue dissolved in 100 µL of methanol. The cysteamine adducts were analyzed by LC-MS with an Agilent 6530 Accurate-Mass Q-TOF using a C18 reversed phase column (Knauer Eurospher II, 100 mm x 3 mm). A gradient of 5 % solvent B (0.1% formic acid in acetonitrile)/95% solvent A (0.1% formic acid in water) to 100 % B was used at a flow rate of 0.2 mL/min.

2.2.4.4. Enzyme Assays

2.2.4.4.1. *In Vitro* Assay of Nonribosomal Peptide Synthetase (GglsA)

The GglsA assay was performed with protein extracted from *E. coli* cells harboring the construct pTrc-Tac:GglsC-T5:His6GglsA-Kan^R. This construct was designed for the co-expression of GglsC to increase the degree of phosphopantetheinylation of the T domain of GglsA. The GglsA enzyme assay was conducted with 75 µM 3-hydroxydecanoyl-CoA, 1 mM of glycine, 2 mM of ATP, and 1.6 µg of soluble *E. coli* proteins in 200 µL reaction volume with final concentrations of 50 mM of Tris-HCl, pH 8.0, 10 mM $MgCl_2$, 1 mM EDTA, 1 mM coenzyme A (lithium salt, Larodan) and 3 mM dithiothreitol (DTT). The reaction was carried out at 28 °C for 30 min and was stopped with 2 mL of chloroform/methanol 2:1 (v/v) and 0.5 mL of 10 % ammonium acetate. The lower phase was harvested and evaporated under N_2 flow. The dry residue was dissolved in 100 µL of chloroform/methanol/ 300 mM ammonium acetate (300:665:35; v/v/v) and analyzed by LC-MS.

To unravel if it is possible to terminate the aglycone assembly by adding decanoyl-CoA, the enzyme assay was carried out with 75 µM 3-hydroxydecanoyl-CoA and 7.5 mM decanoyl-CoA as substrate in 1 mL reaction volume. Instead of stopping the reaction with chloroform/methanol 2:1 (v/v), cysteamine was added to the isolated protein at a final concentration of 0.5 mM. The extraction and analysis follow the same steps as shown in 2.2.4.3.

2.2.4.4.2. *In Vitro* Assay of Glucosyltransferase (GglsB)

For GglsB assays, soluble protein was isolated from *E. coli* cells harboring the construct pTrc-Tac-T5:His6-GglsB-Kan^R. The GglsB assays were conducted with 250 µM UDP-glucose (or UDP-galactose) and 250 µM of aglycone of the glycine-glucolipid H(O-10:0)$_4$Gly as substrates (final concentrations). The aglycone was isolated and purified from the *A. borkumensis* ΔgglsB mutant according to 2.2.3.1. Similar to the glycine-glucolipid, the aglycone was eluted with acetone/isopropanol (9:1, v/v) from SPE (2.2.3.4) and quantified by GC-MS (2.2.3.7). The substrates and ~450 ng of *E. coli* soluble proteins were added to a total reaction volume of 200 µL with final concentrations of 7.5 mM of Tricine-KOH pH 8.0, 15 mM $MgCl_2$ and 1.5 mM DTT. The reaction was carried out at 28 °C for 30 min and was stopped by adding 2 mL of chloroform/methanol 2:1 (v/v) and 0.5 mL of 10 % ammonium acetate. The lower phase was harvested and evaporated under N_2 flow. The dry residue was dissolved in 100 µL of chloroform/methanol/ 300 mM ammonium acetate (300:665:35; v/v/v) and analyzed by LC-MS.

2.2.5. Electron Microscopy and Confocal Microscopy

2.2.5.1. Scanning Electron Microscopy (SEM)

To image the *A. borkumensis* cells by SEM, bacteria from a broth culture were injected onto a 0.1 µm sterilization filter, and the filter membrane, now containing bound bacteria, was washed with 3 % NaCl and 3 % NaCl/fixation buffer (1:1, v/v). The fixation buffer contains 4 % paraformaldehyde and 2.5 % glutaraldehyde solution in 0.1 M cacodylate buffer. The filter membrane with bacteria was then fixed in the fixation buffer at 4 °C within a glass vial and the following steps were processed by Dr. Hannes Beckert and Pia Stausberg (Microscopy Core Facility of the Medical Faculty, University of Bonn). The filter membrane is then carefully removed, and the bacterial sample is dehydrated through an ethanol series (30 %, 50 %, 70 %, 90 %, 95 %, absolute ethanol). Subsequently, the sample undergoes two incubations in 100 % hexamethyldisilazane (HMDS) for 15 minutes each. The dehydrated sample is then dried under a fume hood in chambers with desiccant. The filter membrane is mounted on alum stubs using double-sided conductive tape, and silver paint is applied to the edges on the surface of the membrane. Sputter coating with a 2 nm thick layer of platinum is carried out using a Leica ACE600. The final step involves imaging the prepared sample using a Zeiss Crossbeam 550 electron microscope at 0.8 kV acceleration voltage and a variable beam current ranging from 150 to 1000 nA, with detection using a secondary electron (SE2) detector.

2.2.5.2. Transmission Electron Microscopy (TEM)

For the TEM imaging, cells were grown in pyruvate or hexadecane medium and were harvested by centrifugation from the pyruvate medium or from the oil/water interface (biofilm cells) at the top or from the aqueous phase (detached cells) of the hexadecane medium. The sample preparation initialed from two steps of washing with 3 % NaCl and 3 % NaCl/fixation buffer (1:1, v/v). The fixation buffer contains 4% paraformaldehyde and 2.5 % glutaraldehyde solution in 0.1 M cacodylate buffer. The cells were fixed in fixation buffer, stored at 4 °C and the following steps were processed by Dr. Hannes Beckert and Pia Stausberg (Microscopy Core Facility of the Medical Faculty, University of Bonn). After washing with 0.1 M cacodylate buffer, osmification is performed using a solution containing 1 % osmium tetroxide and 0.8 % ferricyanate in cacodylate buffer for 2 hours, followed by additional washes in cacodylate buffer and ddH_2O. Subsequently, cells are embedded in 1 % agarose (melted), hardened on ice, and trimmed into small blocks. Dehydration is carried out through an ethanol series (30 %, 50 %, 70 %), with an intermediate

step of incubation in 0.5 % uranyl acetate in 70 % ethanol for 1 hour. Further dehydration is achieved with ethanol and propylene oxide, and infiltration with EPON resin is conducted. Resin-embedded blocks are cured at 60 °C for 48 hours. Ultrathin sections (~50 nm) are cut from resin blocks using an ultramicrotome and collected on Formvar/carbon-coated TEM Copper slot grids. Finally, imaging is performed using a Zeiss Crossbeam 550 electron microscope at 30 kV acceleration voltage and 150 pA beam current with a STEM detector.

2.2.5.3. Confocal Microscopy

Cultures were washed once and diluted to OD_{600} ~1.0 with carbon free ONR7a medium. 1 mL of sample was stained for 15 min with 1 µL of 25 mM Nile red in dimethylsulfoxide (DMSO) and examined by epifluorescence microscopy with an Olympus BX60 microscope equipped with a high-pressure mercury light source, a 460-490 nm band pass excitation filter and a 560 nm long pass barrier emission filter. 100 µL of hexadecane were added on top of the sample after staining. The samples were vortexed before observation. All images were captured with an exposure time of 1000.0 ms.

2.2.6. Image Analysis via ImageJ

All measurements were conducted using ImageJ Java 1.8.0_345 (64-bit).

2.2.6.1. SEM Cell Length Measurement

Three SEM images from each treatment were selected for analysis. And five cells were chosen from each image. The resulting cell length values were expressed in nanometers.

2.2.6.2. Confocal Fluorescent Intensity Measurement

The quantification of cell attachment to hexadecane droplets was conducted by measuring the fluorescent intensity of the biofilm cell layer. Four confocal images, including the one depicted in Figure 37, from each treatment were selected for analysis. In cases where multiple droplets were present in one image, the largest droplet was selected. Area, integrated intensity and mean grey value were measured for the chosen droplet both with and without the cell layer. The fluorescent intensity of biofilm cell layer was calculated using the formula:

$$\text{fluorescent intensity of biofilm cell layer} = \frac{\text{integrated intensity of Area 2} - \text{integrated intensity of Area 1}}{\text{Area 2} - \text{Area 1}}$$

The resulting fluorescent intensity values were expressed in intensity/μm^2. The Area 2 and Area 1 are shown in Figure 8.

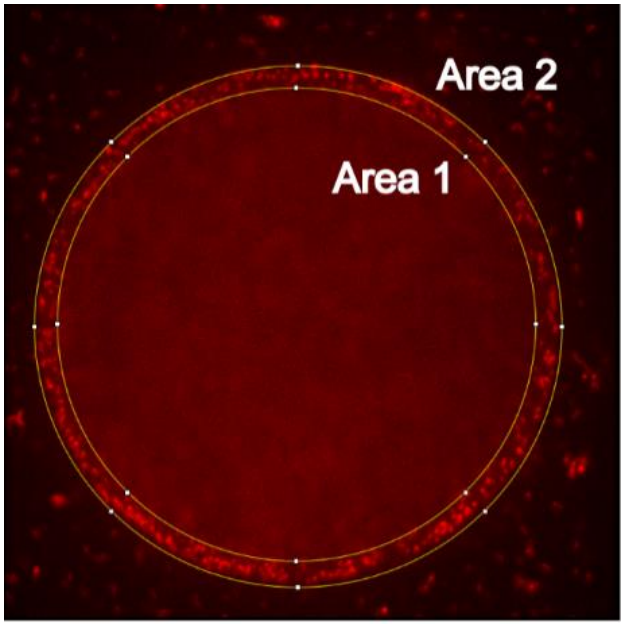

Figure 8 **Confocal image of hexadecane-grown *ΔgglsA* cells attached to hexadecane droplet.**
Area 2 indicates the droplet with biofilm cells, Area 1 indicates the droplet without biofilm.

2.2.7. Statistical Methods

All experiments were conducted with at least three replicas. Means and standard deviation were calculated by GraphPad Prism 9. The figures were produced by GraphPad Prism 9 or Sigmaplot version 11. Statistical significance for all experiments was calculated using Student's T-test. Protein sequence similarity was taken from the NCBI database.

3. Results

3.1. Characterization of the Glycine-Glucolipid in *A. borkumensis*

A total lipid extract from *A. borkumensis* cells grown on pyruvate was loaded onto a two-dimensional TLC plate. After α-naphthol staining, the glycolipids turned purple. As shown in Figure 9, *A. borkumensis* cells contain glycolipids. The lipid spots from the TLC plate were individually purified and analyzed by Q-TOF mass spectrometry. Phosphatidylethanolamine (PE), phosphatidylglycerol (PG) and cardiolipin (CL) are the dominant lipid classes of *A. borkumensis* cells. Besides the glycolipid and phospholipids, *A. borkumensis* also contains triacylglycerol (TAG), two unknown lipids (U1 and U2) and ornithine lipid (OL, identified by Dr. Georg Hölzl).

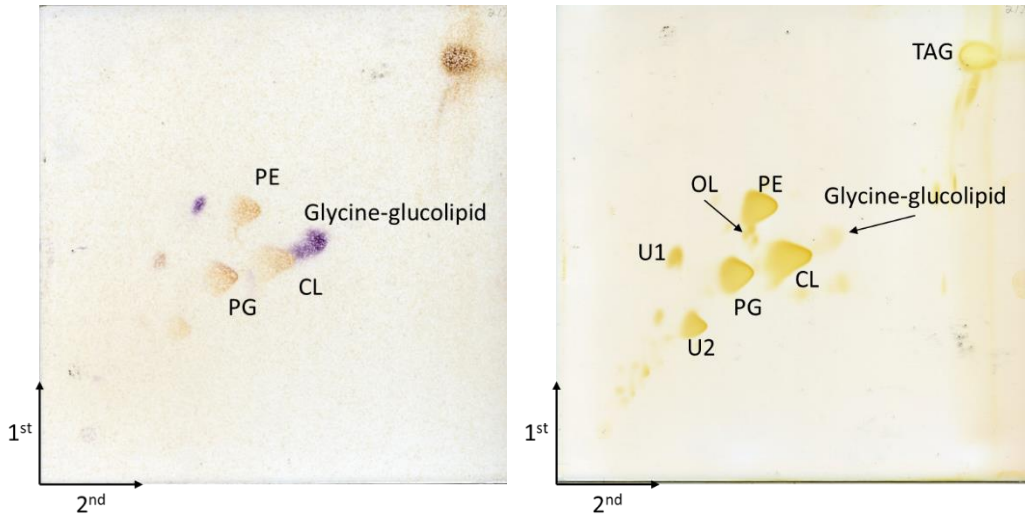

Figure 9 **TLC separation of lipid extracts from *A. borkumensis* cells grown in ONR7a medium with pyruvate.** Lipid extracts were stained with α-naphthol (left) and iodine (right). The glycine-glucolipid was stained in purple with α-naphthol. PE: phosphatidylethanolamine; PG: phosphatidylglycerol; CL: cardiolipin; OL: ornithine lipid; U1, U2: unknown lipid 1, 2; TAG: triacylglycerol.

The molecular species of the glycine-glucolipid were predominantly composed of Glc(O-10:0)$_4$Gly (peak of the ammonia adduct at *m/z* 935.6414) and low amounts of Glc(O-8:0)(O-10:0)$_3$Gly, while Glc(O-8:0)$_2$(O-10:0)$_2$Gly was barely detectable (Figure 10A, Table 2) (Glc, glucose; Gly, glycine; O-8:0, 3-hydroxy-octanoic acid; O-10:0, 3-hydroxy-decanoic acid). The MS/MS spectrum of Glc(O-10:0)$_4$Gly fully agreed with the previously reported structure (Figure 10A and D)(Abraham, Meyer et al. 1998). Fragmentation of the glycine-glucolipid molecular species containing one 3-hydroxy-octanoic acid, Glc(O-8:0)(O-10:0)$_3$Gly, revealed that the 3-HO-8:0 fatty acid is randomly

distributed to the four possible positions of the congener (Figure 10C).

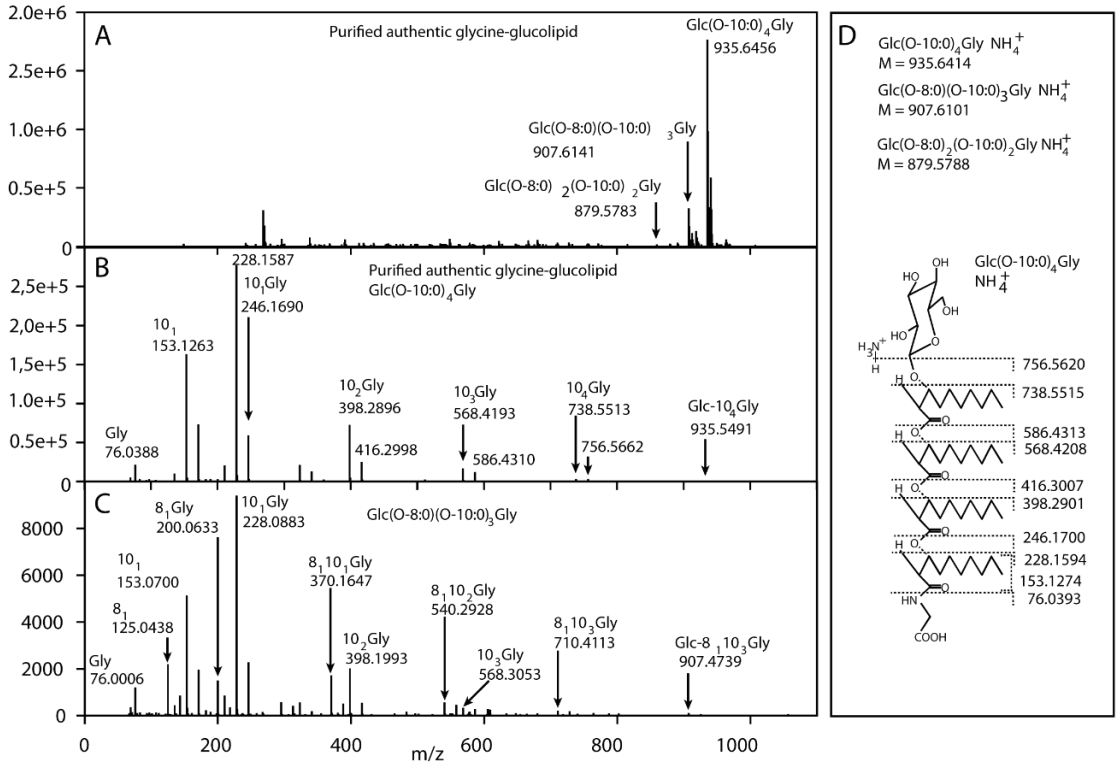

Figure 10 **Analysis of authentic and derivatized *A. borkumensis* glycine-glucolipid by mass spectrometry.**
(A) Total ion count spectrum of the purified authentic glycine-glucolipid. (B) MS/MS spectrum after fragmentation of the main glycine-glucolipid species Glc(O-10:0)$_4$Gly. (C) MS/MS spectrum after fragmentation of the glycine-glucolipid species Glc(O-8:0)(O-10:0)$_3$Gly. (D) Calculated masses of ammonium adducts of different molecular species and fragments generated from Glc(O-10:0)$_4$Gly of the authentic glycine-glucolipid. Glc, glucose; Gly, glycine; O-8:0 or "8," 3-hydroxy-octanoic acid; O-10:0 or "10," 3-hydroxy-decanoic acid. Figure from Cui et al., 2022

3.2. Accurate Quantification of the Glycine-glucolipid from *A. borkumensis*

3.2.1. Quantification of the Glycine-Glucolipid via GC-MS

The determination of the 3-hydroxy fatty acids by gas chromatography-mass spectrometry (GC-MS) was assessed for quantifying the glycine-glucolipid. The glycine-glucolipid contains four 3-hydroxy fatty acids, which are bound to the 3-hydroxy group of the adjacent fatty acid by ester linkages, while the first one is linked to the glucose moiety by a glycosidic linkage, and the terminal 3-hydroxy fatty acid is linked to glycine by an amide linkage. For GC-MS measurements, the 3-hydroxy fatty acids were cleaved and transmethylated with methanolic HCl (Figure 11A).

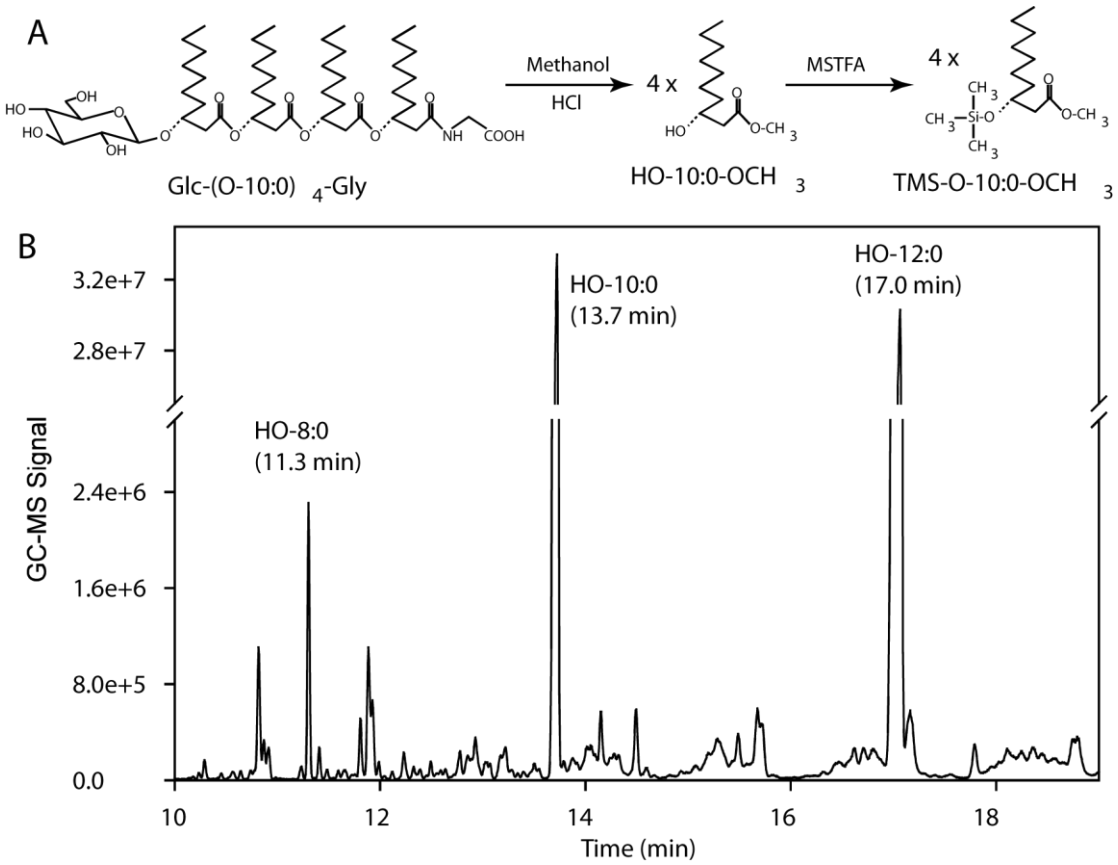

Figure 11 **Analysis of 3-hydroxy-fatty acids derived from the *A. borkumensis* glycine-glucolipid by GC-MS.**
(A) The four 3-hydroxy-fatty acids in the *A. borkumensis* glycine-glucolipid were transmethylated, and the free hydroxy groups were converted into trimethylsilyl (TMS) ethers. (B) GC-MS chromatogram (total ion count) of the *A. borkumensis* glycine-glucolipid, in the presence of the internal standard 3-hydroxy dodecanoic acid (HO-12:0). Figure from Cui et al., 2022.

Because its amide-linked fatty acids are more resistant to transmethylation reactions than ester-linked fatty acids, the reaction was optimized to ensure that all four 3-hydroxy fatty acids were released from the glycine-glucolipid. The transmethylation reaction was monitored by the peak area ratio of hydroxy fatty acid to internal standard, pentadecanoic acid (15:0). As shown in Figure 12, the ratio increased with the increasing derivatization time, until it reached a maximal value at 4 hours. To make sure the transmethylation reactions were complete, the glycine-glucolipids were incubated 5 h with 3 N methanolic HCl.

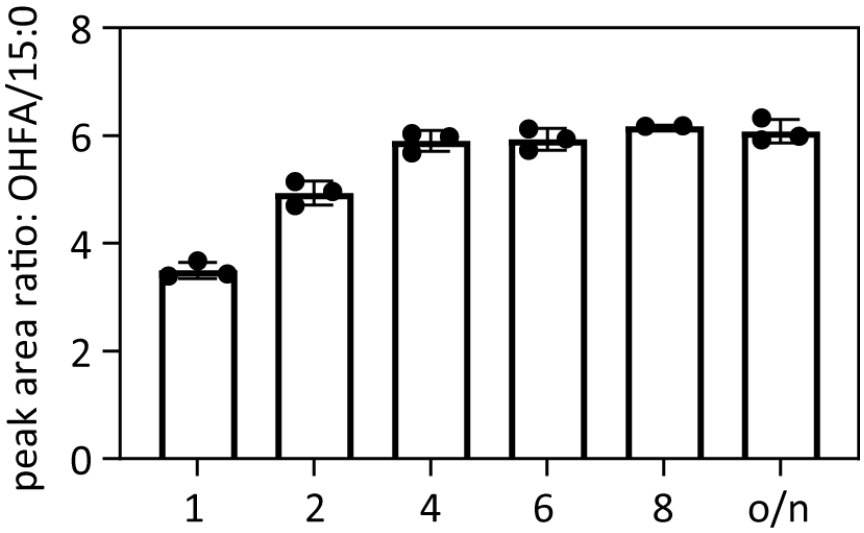

Figure 12 **All 3-hydroxy-fatty acids from glycine-glucolipid were converted into fatty acid methyl esters after 4 h.** Analysis of 3-hydroxy-fatty acids derived from the *A. borkumensis* glycine-glucolipid by GC-MS. After incubated for different times, the degree of transmethylation of 3-hydroxy-fatty acids was increased. Pentadecanoic acid (15:0) was employed as internal standard. After transmethylation, the hydroxyl groups of the 3-hydroxy-fatty acids were silylated, and the products analyzed by GC-MS. Mean ± SD, n=3.

The quantification of glycine-glucolipid by GC-MS was achieved using 3-hydroxy-dodecanoic acid (HO-12:0) as the internal standard. Figure 11B shows that the purified glycine-glucolipid contains 93.0 % HO-10:0 and 7.0 % 3-HO-8:0. For absolute quantification, a mixture of synthesized hydroxy fatty acids including HO-8:0, HO-10:0, HO-12:0, HO-14:0 was measured every time together with the samples to calculate the response factors (Figure 13).

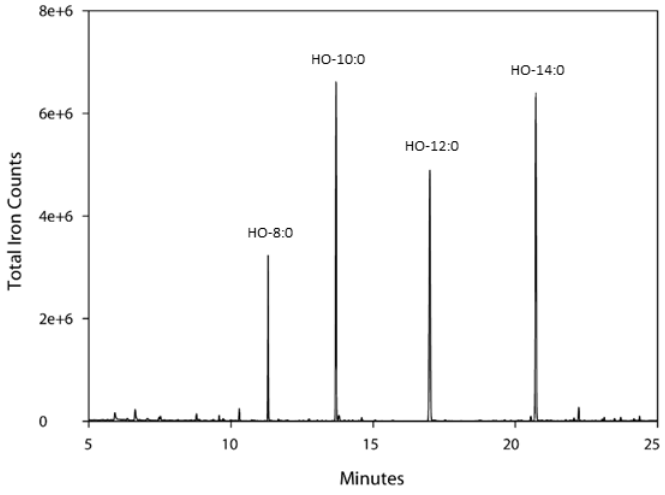

Figure 13 **GC-MS chromatogram (total iron count) of the 3-hydroxy-fatty acids (HO-8:0, HO-10:0, HO-12:0, HO-14:0).** Equal amounts of 3-hydroxy-fatty acids were loaded.

3.2.2. Quantification of the Glycine-Glucolipid via HPLC-DAD

The *A. borkumensis* glycine-glucolipid can barely be detected using a UV/visible light detector, because it has no strong chromophores (Abraham, Meyer et al. 1998). Therefore, an HPLC method was chosen which was previously developed to analyze *P. aeruginosa* rhamnolipids, another class of glycolipids that does not contain strong chromophores for UV/visible light detection (Behrens, Baune et al. 2016). The original protocol included a derivatization step of the terminal fatty acid of the rhamnolipid with p-bromoacetophenone, later modified by replacement of p-bromoacetophenone with 2-bromoacetophenone (Schenk, Schuphan et al. 1995, Mata-Sandoval, Karns et al. 1999). In the *A. borkumensis* glycine-glucolipid, the terminal 3-hydroxy fatty acid is bound to glycine in amide linkage and therefore, the fatty acid cannot be derivatized. However, the terminal glycine of the glycine-glucolipid possesses a free carboxy group, which might be accessible for conversion into its phenacyl ester (Figure 14A). Undecanoic acid (11:0) was selected as internal standard for quantification, a free fatty acid which can be readily converted into its phenacyl ester (Durst, Milano et al. 1975).

Figure 14B and C show the HPLC-DAD chromatograms of a crude lipid extract and the purified glycine-glucolipid from *A. borkumensis* grown on pyruvate after derivatization. Three peaks eluting at 15.5 min, 17.5 min, and 19.5 min were identified, together with the internal standard, which elutes at 12.1 min.

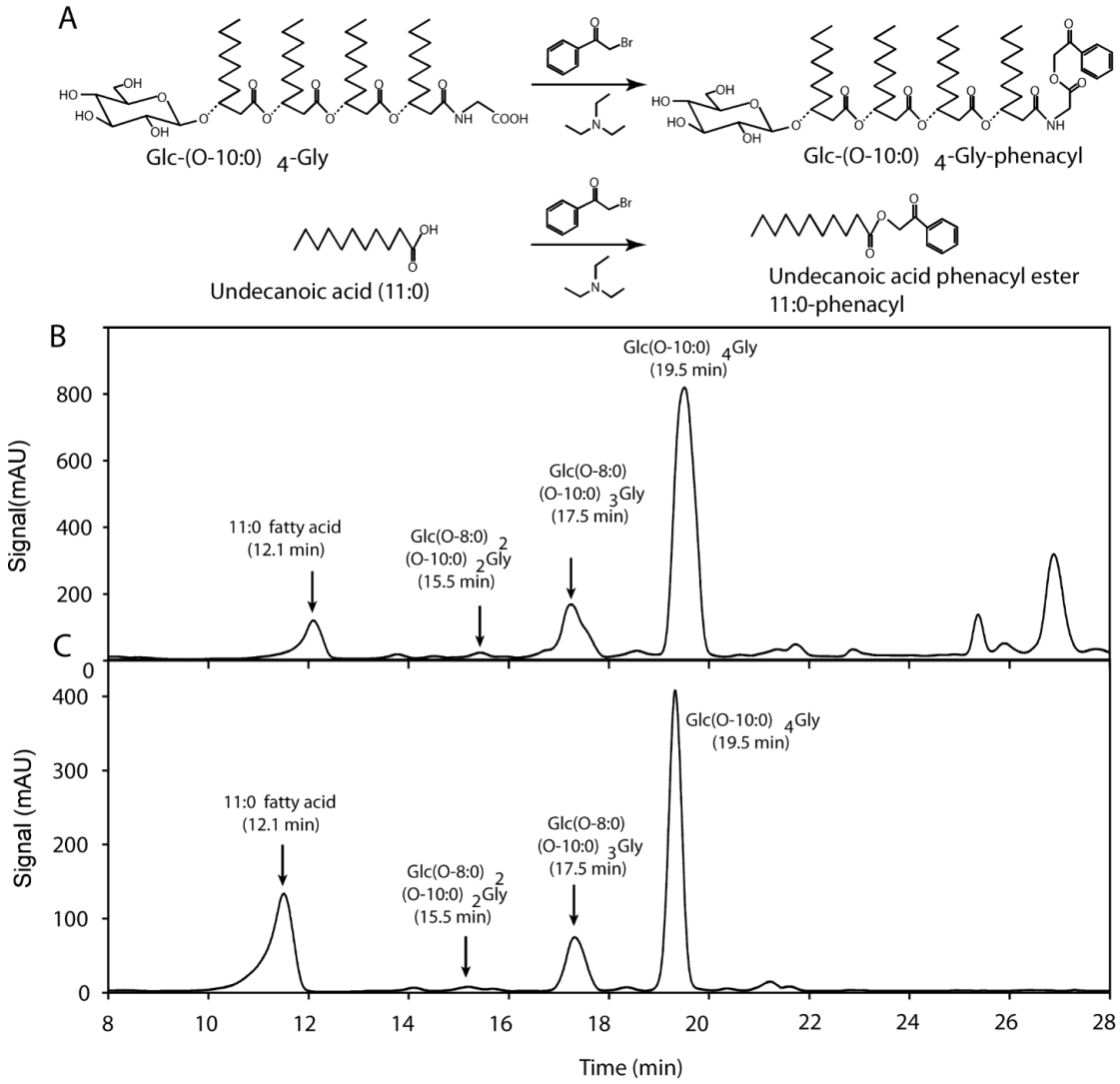

Figure 14 **Detection of the *A. borkumensis* glycine-glucolipid by HPLC-DAD after conversion into phenacyl esters.** (A) The glycine-glucolipid was converted into its phenacyl ester after incubation with 2-bromoacetophenone in the presence of triethylamine. Undecanoic acid (11:0) was employed as an internal standard. (B) HPLC-DAD chromatogram of a crude lipid extract from *A. borkumensis*, after derivatization. (C) HPLC-DAD chromatogram of isolated *A. borkumensis* glycine-glucolipid (purified by SPE and TLC) after derivatization. Glc, glucose; Gly, glycine; O-8:0, 3-hydroxy-octanoic acid; O-10:0, 3-hydroxy-decanoic acid. Figure from Cui et al., 2022.

Different concentrations and ratios of 2-bromoacetophenone and triethylamine were tested for the derivatization efficiency (10:5, 15:7.5, 20:5, 20:10, 20:15, 20:20; in mg/mL), and the yield of derivatized glycine-glucolipid analyzed by HPLC. Best yields of derivatization were obtained by incubation with 15 mg/mL of 2-bromoacetophenone and 7.5 mg/mL of triethylamine, at 80°C for 1 h (Figure 15), which was applied for further analyses.

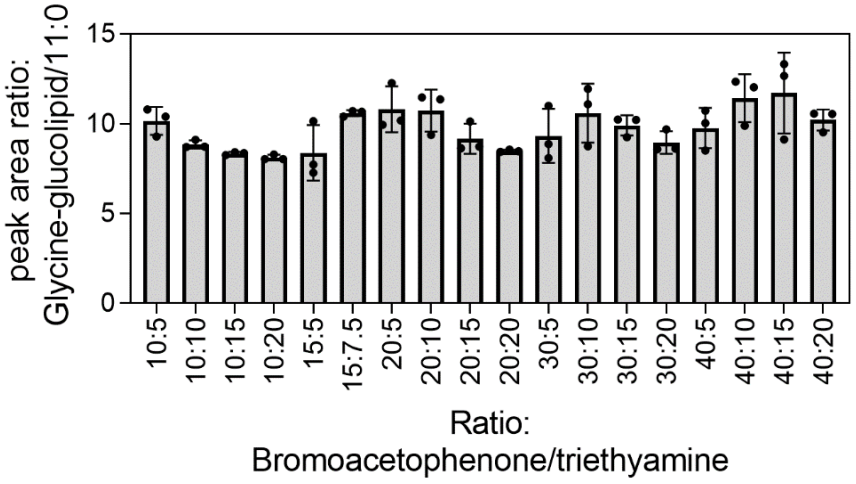

Figure 15 **Best yields of derivatization were obtained by incubation with 15 mg/mL of 2 bromoacetophenone and 7.5 mg/mL of triethylamine.**

Detection of the *A. borkumensis* glycine-glucolipid by HPLC-DAD after conversion into phenacyl esters. After incubation for different times with different concentrations of 2-bromoacetophenone in the presence of triethylamine, the glycine-glucolipid was converted into its phenacyl ester. Data show means and SD (n = 3).

The peak fractions containing the three derivatized lipids were collected from the HPLC column and individually analyzed by direct infusion mass spectrometry. The major peak eluting at 19.5 min showed an *m/z* of 1,053.6836, in agreement with the addition of the phenacyl group on the parental ion [Glc(O-10:0)$_4$Gly-phenacyl+NH_4^+] as compared with the nonderivatized species of [Glc(O-10:0)$_4$Gly+NH_4^+] with an *m/z* of 935.6414 (Figure 16E) (Glc, glucose; Gly, glycine; O-8:0, 3-hydroxy-octanoic acid; O-10:0, 3-hydroxy-decanoic acid). The fragmentation pattern of [Glc(O-10:0)$_4$Gly-phenacy+NH_4^+] indicated that the phenacyl moiety was attached to the carboxylate of the glycine Figure 16F and G). The peaks eluting at 15.5 min and 17.5 min corresponded to the molecular species of Glc(O-8:0)$_2$(O-10:0)$_2$Gly-phenacyl and Glc(O-8:0)(O-10:0)$_3$Gly-phenacyl, respectively.

In addition to the derivatized peaks of the major molecular species Glc(O-10:0)$_4$Gly-phenacyl, a minor nonderivatized peak corresponding to Glc(O-10:0)$_4$Gly was observed in the total ion count spectrum (Figure 16E). Taking the peak sizes as a measure of abundance, the degree of derivatization was calculated to be 97.2% ± 0.8% and 98.9% ± 0.4% for the sum of the three glycine-glucolipid species in the crude lipid extract and the purified glycine-glucolipid, respectively (n = 3).

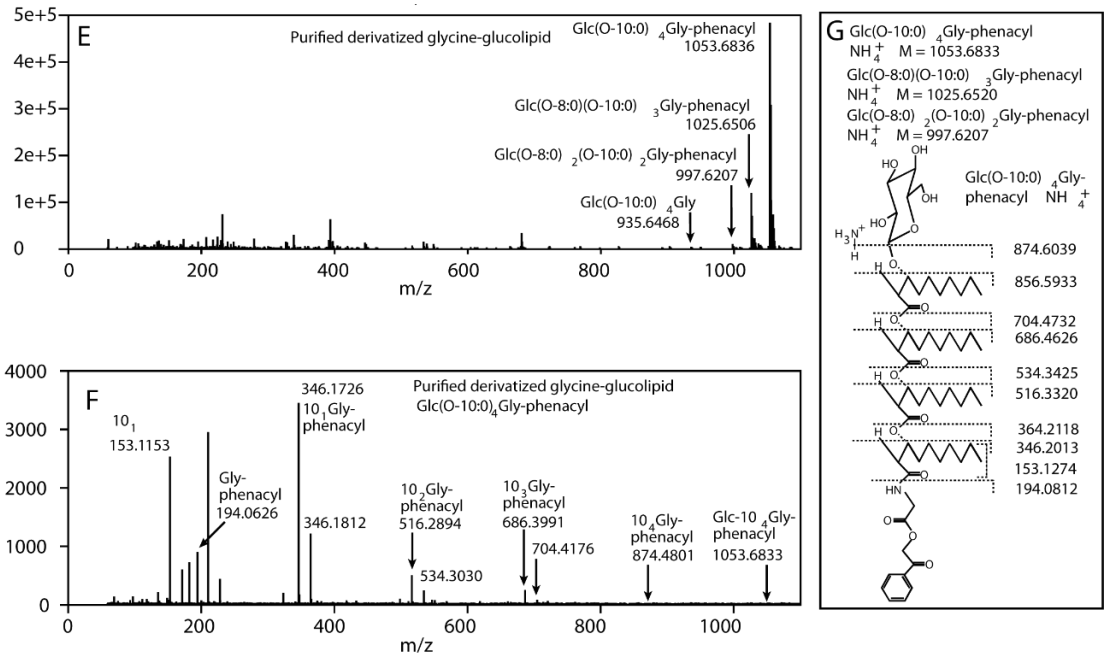

Figure 16 **Analysis of authentic and derivatized *A. borkumensis* glycine-glucolipid by mass spectrometry.**
(E) Total ion count spectrum of the purified glycine-glucolipid after conversion into phenacyl esters. (F) MS/MS spectrum after fragmentation of the main species Glc(O-10:0)$_4$Gly-phenacyl. (G) Calculated masses of ammonium adducts of molecular species and fragments generated from Glc(O-10:0)$_4$Gly-phenacyl. Glc, glucose; Gly, glycine; O-8:0 or "8," 3-hydroxy-octanoic acid; O-10:0 or "10," 3-hydroxy-decanoic acid. Figure from Cui et al., 2022.

3.2.3. Quantification of the Glycine-Glucolipid via LC-MS

After introduction of the genes from *A. borkumensis* into *E. coli* (chapter 3.7), the production of the glycine-glucolipid in *E. coli* was very low. Therefore, the glycine-glucolipid constitutes a relatively small proportion in the crude lipid in transformed *E. coli* cells. The quantitative analysis of the glycine-glucolipid via HPLC-DAD faces challenges due to its limited sensitivity. To overcome this limitation, we devised a rapid and highly sensitive analytical method based on LC-MS (2.2.3.12). This approach was employed for the detection of glycine-glucolipid and its derivatives, including the aglycone, the serine-containing aglycone, and derivatives with the cysteamine dimer.

Quantification of the glycine-glucolipid by LC-MS (without fragmentation) was achieved by using MGD di-18:2 as the internal standard. For absolute quantification, a mixture of purified glycine-glucolipid and internal the standard MGD di-18:2 was measured every time together with the samples to calculate the response factors.

3.3. The Glycine-Glucolipid Accumulates in *A. borkumensis* Cells Grown on Hexadecane

Q-TOF mass spectrometry and HPLC-DAD were employed to quantify the glycine-glucolipid produced by *A. borkumensis* grown under different conditions. *A. borkumensis* was grown with pyruvate or hexadecane as carbon sources (Yakimov, Golyshin et al. 1998, Schneiker, Dos Santos et al. 2006). Cells were harvested and crude lipid extracts were prepared for glycine-glucolipid measurements by HPLC-DAD after conversion into phenacyl esters. The total glycine-glucolipid (sum of Glc(O-8:0)$_2$(O-10:0)$_2$Gly-phenacyl, Glc(O-8:0)(O-10:0)$_3$Gly-phenacyl, and Glc(O-10:0)$_4$Gly-phenacyl) amounted to ~0.8 nmol per 10^{10} cells (equivalent to 4.2 nmol per mg protein) when cells were grown with pyruvate (Figure 17) (Glc, glucose; Gly, glycine; O-8:0, 3-hydroxy-octanoic acid; O-10:0, 3-hydroxy-decanoic acid). The glycine-glucolipid content was higher in cells grown with hexadecane (~1.3 nmol per 10^{10} cells). Therefore, growth on hexadecane resulted in increased glycine-glucolipid accumulation per 10^{10} cells by > 50 %, indicating that its synthesis was stimulated under conditions of alkane metabolization.

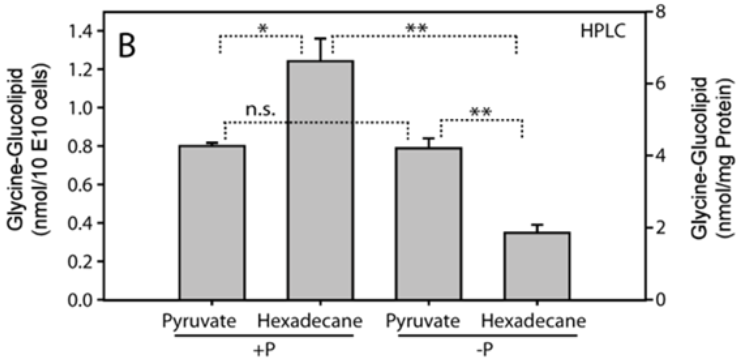

Figure 17 **Production of glycine-glucolipid was stimulated with hexadecane as sole carbon source. The glycine-glucolipid was not involved in the phosphate deprivation response of *A. borkumensis*.**
Quantification of the glycine-glucolipid isolated from cell pellets after growth in different media. The lipid extracts were derivatized with 2-bromoacetophenone and the glycine-glucolipid quantified by HPLC-DAD. All glycine-glucolipid-containing peaks (Glc(O-10:0)$_4$Gly, Glc(O-8:0)(O- 10:0)$_3$Gly, and Glc(O-8:0)$_2$(O-10:0)$_2$Gly) were summarized. The amounts of glycine-glucolipids are expressed in nmol per 10^{10} cells and in nmol per mg protein. Data show means and SD (n = 3). Student T-test, *, P , 0.05; **, P , 0.01; n.s., not significant. Figure from Cui et al., 2022.

Furthermore, the amounts of the three molecular species of glycine-glucolipid from pyruvate-grown cells determined by HPLC was 4.0 ± 0.4%, 25.3 ± 0.3%, and 70.8 ± 0.1% ($Glc(O\text{-}8:0)_2(O\text{-}10:0)_2Gly$, $Glc(O\text{-}8:0)(O\text{-}10:0)_3Gly$ and $Glc(O\text{-}10:0)_4Gly$), respectively (Table 2). The molecular species composition of the glycine-glucolipid from hexadecane-grown cells was shifted in comparison with pyruvate-grown cells, because the relative contents of $Glc(O\text{-}8:0)(O\text{-}10:0)_3Gly$ and $Glc(O\text{-}10:0)_4Gly$ amounted to 11.1% ± 1.0% and 88.9% ± 1.0%, respectively, while $Glc(O\text{-}8:0)_2(O\text{-}10:0)_2Gly$ was not detectable, in hexadecane-grown cells (Table 2).

Table 2: Relative amounts of molecular species of the *A. borkumensis* glycine-glucolipid after growth on pyruvate- or hexadecane-containing medium[a]. Table from Cui et al., 2022.

Glycine-glucolipid molecular species	Relative amount (%)			
	Pyruvate		Hexadecane	
	HPLC	QTOF	HPLC[b]	QTOF
$Glc(O\text{-}8:0)_2(O\text{-}10:0)_2Gly$	4.0 ± 0.4	4.0 ± 0.1	n.d.	0.5 ± 0.2
$Glc(O\text{-}8:0)(O\text{-}10:0)_3Gly$	25.3 ± 0.3	24.5 ± 0.4	11.1 ± 1.0	3.1 ± 1.5
$Glc(O\text{-}10:0)_4Gly$	70.8 ± 0.6	71.4 ± 0.4	88.9 ± 1.0	96.3 ± 1.7

[a] The glycine-glucolipid was extracted from the whole culture broth of bacteria grown on pyruvate or hexadecane. Relative amounts of molecular species were determined by HPLC-DAD or Q-TOF mass spectrometry.

[b] n.d., not detected. Mean ± SD, n = 3.

3.4. The Glycine-glucolipid is not Involved in the P Deficiency Response in *A. borkumensis*

It is known that phosphorus-free glycolipids replace phospholipids during phosphate deprivation in some bacteria and in plants. This phosphate-deprivation response results in the release of phosphate from the membranes which is made available for nucleotide synthesis to support growth under phosphate limiting conditions. Thus, it was possible that the glycine-glucolipid from *A. borkumensis* could be involved in the phosphate-starvation response. The cells were grown in the presence or absence of phosphate, both with pyruvate or hexadecane as a carbon source, to assess whether the glycine-glucolipid accumulation was stimulated not only by the presence of hexadecane but also in response to phosphate deprivation or a combination of both. The glycine-glucolipid content shows no difference between pyruvate-grown cells in the presence and absence of phosphate. The glycine-glucolipid content was higher in hexadecane-grown cells in

the presence of phosphate (~1.3 nmol per 10^{10} cells), but it was only ~0.4 nmol per 10^{10} cells when phosphate was omitted from the hexadecane medium (Figure 17).

The growth of *A. borkumensis* cells was slightly or strongly compromised under phosphate deprivation conditions in pyruvate or hexadecane medium, respectively (Figure 18). Therefore, phosphate deprivation did not increase the glycine-glucolipid content, indicating that it is not involved in the phosphate deprivation response of *A. borkumensis*.

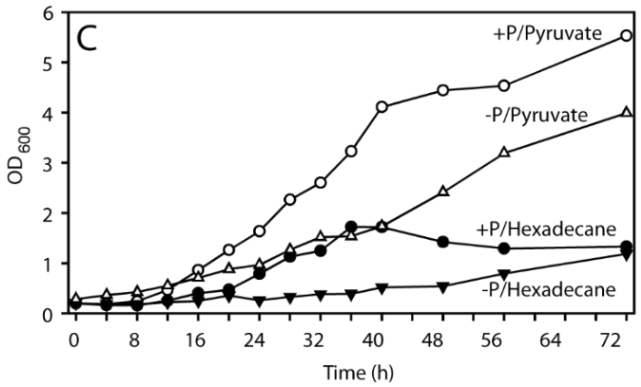

Figure 18 **Phosphate deprivation resulted in growth retardation of *A. borkumensis* cells with different carbon sources.**
Growth curves (OD_{600}) of *A. borkumensis* cells grown with pyruvate and phosphate, with hexadecane and phosphate, with pyruvate without phosphate, or with hexadecane without phosphate. The plot shows individual, representative growth curves repeated at least two times with similar results. Figure from Cui et al., 2022.

It should be noted that the yield per volume of the glycine-glucolipid in pyruvate-grown cells (8 nmol per mL culture) was higher compared with hexadecane-grown cells (3 nmol per mL culture), since the cells grow to a higher density in the stationary phase in pyruvate medium (Figure 18).

3.5. The Glycine-Glucolipid is Resident to the *A. borkumensis* Cell Wall

To determine if the glycine-glucolipid is secreted into the medium, the surface tension of the cell free supernatant was measured with the help of Dr. Viktoria V. Zeisler-Diehl in Prof. Lukas Schreiber's lab (University of Bonn). As shown in Figure 19, the surface tension shows no difference between cell free supernatant and the fresh medium (N.C.). There is no surfactant secreted into the extracellular medium by *A. borkumensis* cells grown either in the pyruvate medium or hexadecane medium.

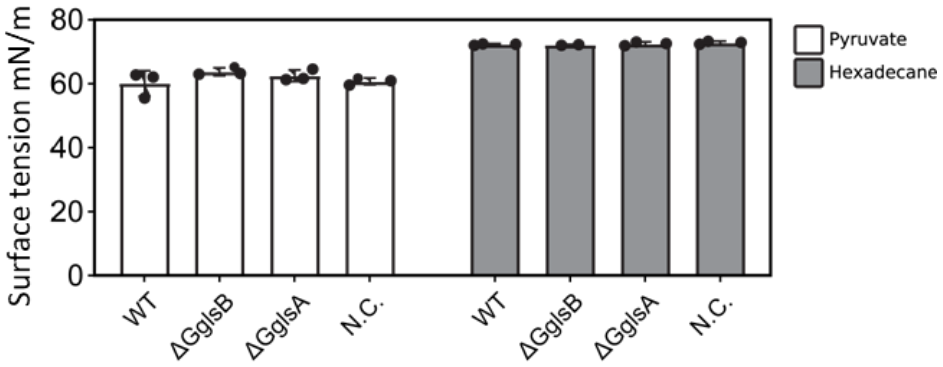

Figure 19 ***A. borkumensis* cells do not secret a surfactant into the environment.**
Analysis of surface tension of cell-free supernatant isolated from different *A. borkumensis* strains grown on various carbon sources. Note that the *ΔgglsA* and *ΔgglsB* mutants are introduced in chapter 3.6. N.C.: negative control, the fresh medium without cells inoculated. Data show means and SD (n = 2 or 3).

To further confirm the surface tension results, the total lipid was extracted and the glycine-glucolipid measured in the whole culture broth, the cell pellet, and the cell-free supernatant. The experiment was conducted with cells grown in the presence of pyruvate or hexadecane as carbon sources, and the glycine-glucolipid was measured by HPLC-DAD and Q-TOF MS. As shown before (Figure 17), the amount of the glycine-glucolipid per cell increased when the cells were grown in the presence of hexadecane. The predominant fraction of the glycine-glucolipid was detected in the cell pellet, while only minute amounts were found in the supernatant, both when the cells were grown on pyruvate or hexadecane (Figure 20).

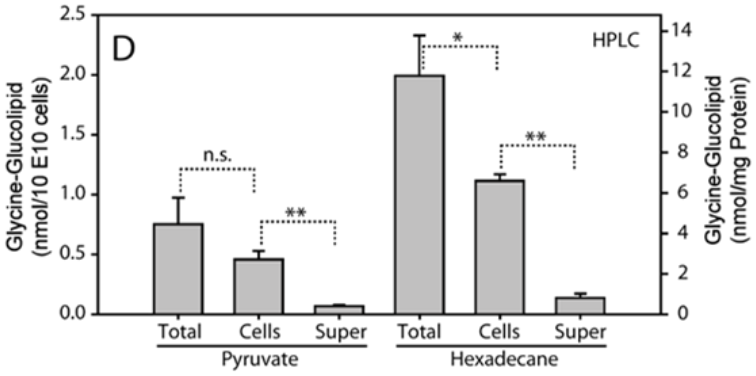

Figure 20 **The glycine-glucolipid was not secreted into the medium.**
Quantification of the glycine-glucolipid after growth in different media and in cell pellets and supernatants of *A. borkumensis* cultures. The lipid extracts were derivatized with 2-bromoacetophenone and the glycine-glucolipid quantified by HPLC-DAD. All glycine-glucolipid-containing peaks ($Glc(O\text{-}10:0)_4Gly$, $Glc(O\text{-}8:0)(O\text{-}10:0)_3Gly$, and $Glc(O\text{-}8:0)_2(O\text{-}10:0)_2Gly$) were summarized. The amounts of glycine-glucolipid are expressed in nmol per 10^{10} cells and in nmol per mg protein (HPLC). Data show means and SD (n = 3). Student T-test, *, P , 0.05; **, P , 0.01; n.s., not significant. Figure from Cui et al., 2022

Analysis of Q-TOF MS shows that the ratio of glycine-glucolipid to the phospholipid PG does not change between cell pellet and supernatant (Figure 21). This indicates that low amounts of PG and glycine-glucolipid, with the same ratio, are released from the cells into the supernatant, probably from broken cells, which further proved that the glycine-glucolipid is resident to the cell, and not specifically secreted. Therefore, most of the glycine-glucolipid accumulates in a form bound to the bacteria, and only very low amounts are found in the supernatant. Those trace amount of glycine-glucolipid can be derived from the cell debris in the supernatant.

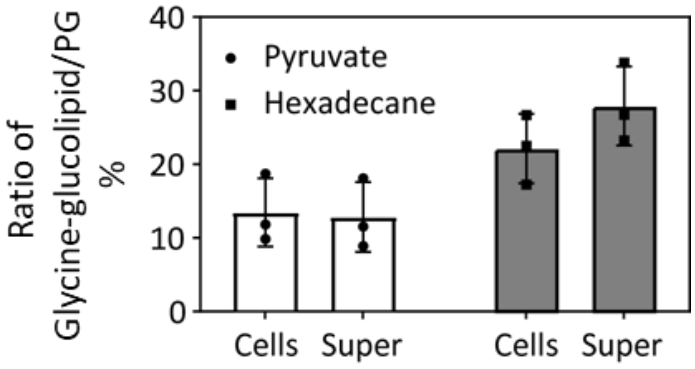

Figure 21 **The ratio of glycine-glucolipid to PG 34:1 was consistent between cell pellet and supernatant.** Analysis of the glycine-glucolipid and PG of *A. borkumensis* cell pellet and supernatant by Q-TOF MS/MS. Data show means and SD (n = 3). Cells: cell pellet, Super: supernatant.

The structure of the glycine-glucolipid in the whole culture broth, the cell pellet, and the supernatant was studied by Q-TOF MS. In all fractions, the main peaks in the mass spectrum showed the glycine-containing form of the glycine-glucolipid carrying four fatty acids-Glc(O-10:0)$_4$Gly, Glc(O-8:0)(O-10:0)$_3$Gly, and Glc(O-8:0)$_2$(O-10:0)$_2$Gly as NH_4^+ adducts with *m/z* 935.6414, 907.6101, and 879.5788, respectively (Glc, glucose; Gly, glycine; O-8:0, 3-hydroxy-octanoic acid; O-10:0, 3-hydroxy-decanoic acid). No evidence was found for the occurrence of the glycine-glucolipid form lacking the glycine residue (Glc[O-10:0]$_4$ as NH_4^+ adducts with *m/z* 878.6199) (Passeri, Schmidt et al. 1992). Therefore, the glycine-glucolipid only exists in the glycine-containing form, and no glycine-free structure could be found, in contrast to previous reports.

3.6. Identification of Genes Involved in Glycine-Glucolipid Biosynthesis in *A. borkumensis*

Based on the structure of glycine-glucolipid, the attachment of glucose at the first hydroxy group requires the involvement of a glycosyltransferase. The amide linkage between glycine and terminal carboxyl group of 3-OH-10:0 is a structure typically formed by NRPS. Therefore, both an NRPS and a PPTase are likely involved in the biosynthesis of glycine-glucolipid. The assembly of the other three acyl moieties may be catalyzed by an acyl transferase. A gene cluster reported for siderophore synthesis includes a PPTase (ABO_1782), a glycosyltransferase (ABO_1783), and an NRPS (ABO_1784). These three genes were developed by Dr. Stephan Thies and Prof. Dr. Karl-Erich Jaeger from Forschungszentrum Jülich. However, no acyltransferase was found in close proximity to this gene cluster. Therefore, acyl transferases involved in lipid metabolism were tested to identify their potential roles in glycine-glucolipid biosynthesis (Schneiker, Dos Santos et al. 2006).

The candidate genes involved in glycine-glucolipid biosynthesis in *A. borkumensis* cells were deleted by inserting a kanamycin cassette into the corresponding locus by homologous recombination with the help from Dr. Georg Hölzl (University of Bonn). Ten deletion mutants were generated, including *ΔABO_1782* (PPTase), *ΔABO_1783* (glucosyltransferase), *ΔABO_1784* (NRPS), *ΔABO_2087* (PPTase), *ΔABO_1112* (acyltransferase), *ΔABO_1397* (acyltransferase) *ΔABO_1418* (acyltransferase), *ΔABO_1804* (acyltransferase), *ΔABO_2214* (acyltransferase), *ΔABO_2742* (acyltransferase). Among these mutants, both *ΔABO_1784* and *ΔABO_1783* mutants exhibited an absence of glycine-glucolipids. *ΔABO_1783* accumulate aglycone H(O-10:0)$_4$Gly instead of glycine-glucolipid. *ABO_1783* encode proteins with sequence similarities to glycosyltransferase. *ABO_1784* encode proteins with sequence similarities to a non-ribosomal peptide synthase (NRPS). The NRPS necessitates a PPTase for post-translational modification, this role is fulfilled by ABO_1782.

3.6.1. ABO_1784 is an NRPS Responsible for the Production of the Aglycone of the Glycine-Glucolipid

The *ABO_1784* gene encodes a protein with 1361 amino acids which displays sequence similarity to bacterial NRPS, e.g. the enterobactin synthetase (EntF) from *Escherichia coli* (31.6% sequence identity) (Figure 50, Appendix 7.1). The ABO_1784 protein contains the typical NRPS domains, including the condensation (C) domain (amino acids 52-467), the adenylation (A) domain (479-999) with the conserved AMP binding site (VIYTSGSTGRP), the thiolation (T) domain (1011-1075), with Ser1039 predicted to bind the phosphopantetheine group (in the conserved domain ALGGHSLLA), and the thioesterase (TE) domain (1097-1359) with the predicted active Ser1177 (in the lipase consensus sequence LLGYSLG)(Konz and Marahiel 1999, Etchegaray, Silva-Stenico et al. 2004).

ΔABO_1784 mutant was grown in ONR7a medium containing pyruvate. A total lipid extract was subjected to LC-MS/MS. The peaks present in wild type with *m/z* of 935.7250 and 907.6905 correspond to the two congeners of the glycine-glycolipid Glc(O-10:0)$_4$Gly and Glc(O-8:0)(O-10:0)$_3$Gly (Glc, glucose; Gly, glycine; O-8:0, 3-hydroxy-octanoic acid; O-10:0, 3-hydroxy-decanoic acid). These two peaks were absent from the *ΔABO_1784* mutant (Figure 22). This suggests that ABO_1784 is the initial enzyme in the glycine-glucolipid biosynthesis pathway. ABO_1784 was designated GglsA for glycine-glucolipid synthetase A.

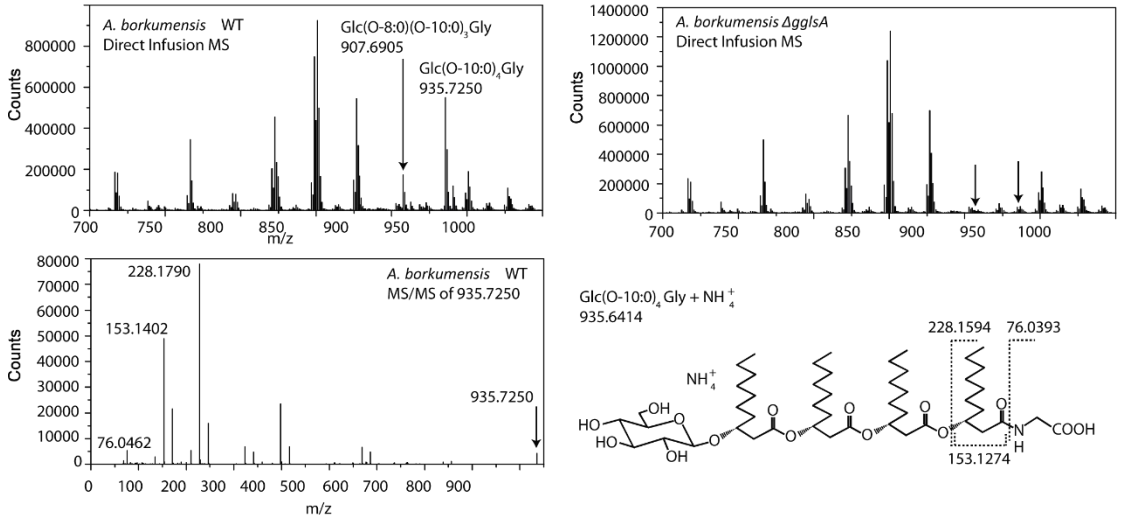

Figure 22 **The *A. borkumensis ΔGglsA* mutant was devoid of the glycine-glucolipid.**
Direct infusion Q-TOF MS spectra of lipids from the WT and *ΔGglsA* (upper panel). The MS/MS spectrum shows the fragmentation of the glycine-glycolipid Glc(O-10:0)$_4$Gly (*m/z* 935.7250) isolated from WT (lower panel).

3.6.2. ABO_1783 is Responsible for the Glycosylation of Aglycone of the Glycine-Glucolipid

The *ABO_1783* encodes a protein of 433 amino acids and shares 25 % amino acid identity with the glucosyltransferase RhlB from *P. aeruginosa* (Figure 53, Appendix 7.4). The ABO_1783 glycosyltransferase shows sequence similarity to glycosyltransferase family GT1 of the GT-B type (www.cazy.org). For this reason, the enzyme is predicted to follow an inverting mechanism resulting in the production of β-glycosidic linkages. This finding is in agreement with the structure of the glycine-glycolipid which carries a β-glucose linked to the hydroxy group of the terminal 3-hydroxydecanoic acid (Figure 22) (Passeri, Schmidt et al. 1992).

The *ΔABO_1783* mutant was grown in ONR7a medium containing pyruvate. A total lipid extract was subjected to LC-MS/MS. The peaks present in wild type with *m/z* of 935.7250 and 907.6905 correspond to the glycine-glucolipid Glc(O-10:0)$_4$Gly and Glc(O-8:0)(O-10:0)$_3$Gly. These two peaks were absent from the *ΔABO_1783* mutant. A new peak emerged with *m/z* of 773.5886. As shown in Figure 23, the presence of the new peak at 773.5886 indicates the accumulation of the aglycone H(O-10:0)$_4$Gly. In agreement with the hypothesis that ABO_1783 is responsible for the final glycosylation step during glycine-glucolipid synthesis. ABO_1783 was designated GglsB.

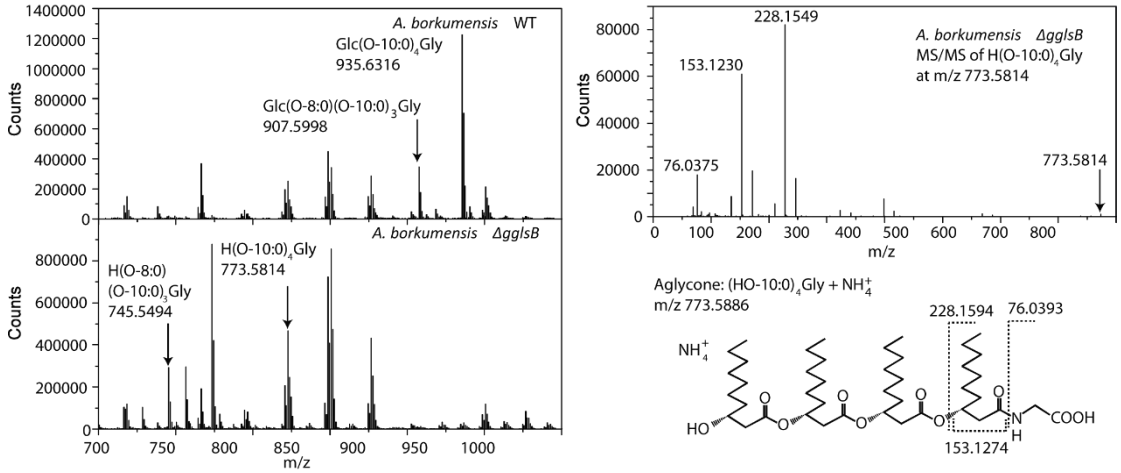

Figure 23 **The *A. borkumensis ΔGglsB* mutant was devoid of the glycine-glucolipid, but accumulates the corresponding aglycone.**

Direct infusion MS/MS spectra of lipids from the WT and *ΔGglsB* (left). The MS/MS spectrum shows the fragmentation of aglycone H(O-10:0)$_4$Gly (*m/z* 773.5814) isolated from *ΔGglsB* (right).

3.6.3. ABO_1782 is Responsible for the Activation of GglsA

A. borkumensis contains two type II PPTases, ABO_1782 is predicted to encode a protein of 253 amino acids, and it shows 20.8 % sequence identity with EntD protein involved in enterobactin synthesis in *E. coli* (Figure 51, Appendix 7.2).

The *ΔABO_1782* mutant was grown in ONR7a medium containing pyruvate. A total lipid extract was subjected to LC-MS/MS. *ΔABO_1782* mutant cells produced equal amounts of glycine-glucolipid as wild-type cells (Figure 24). To study the *in vivo* role of ABO_1782, *ΔABO_1782* mutant cells were grown in the ONR7a medium containing pyruvate or hexadecane. Figure 24 shows that the *ΔABO_1782* mutant cells produced similar amounts of glycine-glucolipid as wild-type cells when grown on pyruvate, but were lower in glycine-glucolipid when grown on hexadecane. ABO_1782 was designated GglsC.

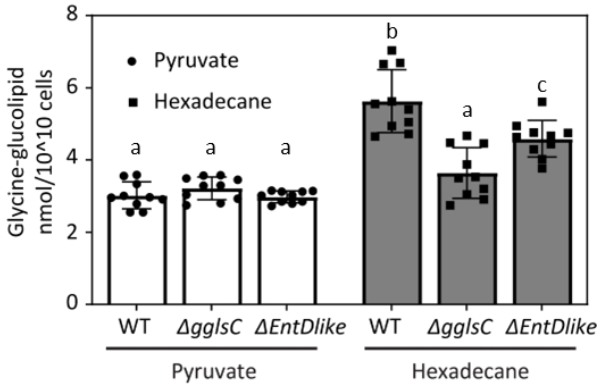

Figure 24 **Glycine-glucolipid production was affected in the *ΔgglsC* and *ΔentDlike* mutants of *A. borkumensis* grown in hexadecane medium.**

LC-MS measurement of glycine-glucolipid isolated from *ΔgglsC* and *ΔentDlike* mutant cells of *A. borkumensis*. Data show means and SD (n = 10). Two-way ANOVA with Tukey's test, different letters indicate significant differences, $p < 0.05$.

3.6.4. GglsC is the Preferred PPTase for the Activation of GglsA

The second type II PPTase is ABO_2087 with 34.06 % sequence identity with ABO_1782 on amino acid level (Figure 52, Appendix 7.3). The ABO_2087 gene is localized close to the gene cluster ABO_2092-ABO_2093, which was suggested to be involved in the synthesis of the siderophore amphibactin (Denaro, Crisafi et al. 2014, Kem, Zane et al. 2014).

To test the hypothesis whether the alternative PPTase ABO_2087 can also contribute to the phosphopantetheinylation of GglsA, the *ΔABO_2087* mutant was grown in ONR7a medium containing pyruvate and hexadecane. A total lipid extract was subjected to LC-MS/MS. The amount of glycine-glucolipid was equal between WT and the *ΔABO_2087* mutant when cells were grown in pyruvate medium. However, the glycine-glucolipid production was slightly reduced in the *ΔABO_2087* mutant grown in hexadecane medium when compared with wild type cells, but it was still higher compared with *ΔgglsC*. Therefore, GglsC is the preferred PPTase for the activation of GglsA (Figure 24). ABO_2087 was designated EntDlike.

3.7. Glycine-Glucolipid production in *E. coli*

3.7.1. GglsA is the First Enzyme Involved in Glycine-Glucolipid Synthesis

To further study the function of the NRPS protein, GglsA was expressed in *E. coli*, lipids were extracted and analyzed by LC-MS/MS. A compound accumulated with a parental mass peak of 773.5554 *m/z* which corresponds with the calculated mass of 773.5886 for the ammonia adduct of the aglycone H(O-10:0)$_4$Gly of the glycine-glucolipid (Figure 25. Fragmentation resulted in peaks at 228.1588 *m/z* characteristic for a monomer of HO-10:0 attached to glycine, while the peaks at 153.1265 *m/z* and 76.0389 *m/z* represent HO-10:0 or glycine, respectively. Therefore, expression of GglsA in *E. coli* resulted in the production of the aglycone H(O-10:0)$_4$Gly. GglsA not only catalyzes the attachment of the first 3-hydroxy fatty acid onto the amino group of the glycine, but also catalyzes the elongation of the intermediate by three esterification reactions of the 3-hydroxy fatty acid onto the free hydroxyl group. Taken together, these results demonstrate that GglsA is involved in glycine-glucolipid synthesis in *A. borkumensis* by producing the aglycone H(O-10:0)$_4$Gly. No other acyltransferase is involved.

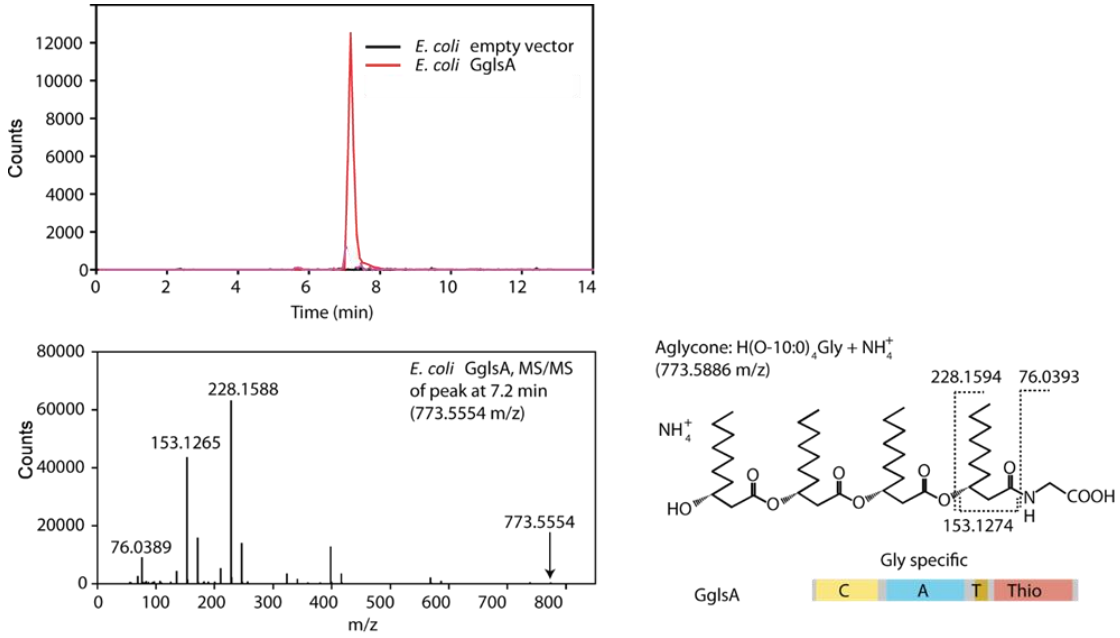

Figure 25 **Expression of GglsA in *E. coli* resulted in production of the aglycone H(O-10:0)$_4$Gly**.
LC-MS chromatograms (upper panel) show a lipid) peak in cells expressing GglsA (red) eluting at 7.2 min, but this peak is absent from empty vector cells (black). The MS/MS spectrum (lower panel) shows the fragmentation pattern of H(O-10:0)$_4$Gly produced by GglsA expression cells (peak at 7.2 min).

3.7.2. The Glycine-Glucolipid is Produced in *E. coli* cells Co-Expressing GglsA and GglsB

For a further insight into the role of GglsB, it was expressed in *E. coli* together with GglsA, and lipids were measured by LC-MS/MS. While the GglsA expressing cells accumulated only the aglycone H(O-10:0)$_4$Gly (Figure 25), the cells co-expressing GglsB synthesized the glycine-glucolipid, Glc(O-10:0)$_4$Gly (Figure 26). Therefore, GglsB is responsible for the glycosylation of H(O-10:0)$_4$Gly produced by GglsA.

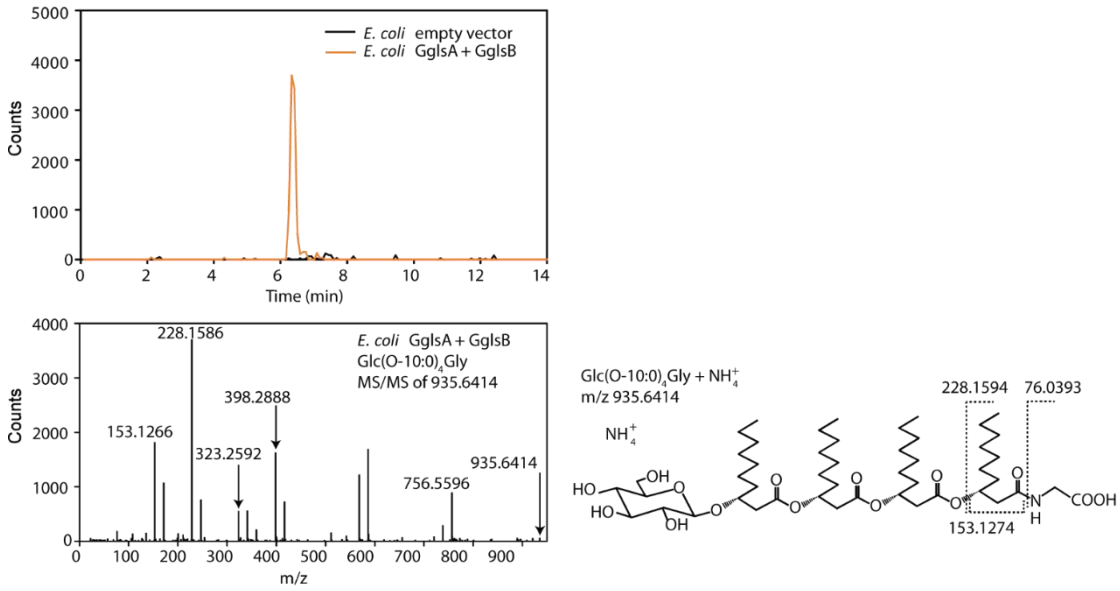

Figure 26 **Co-expression of GglsA and GglsB in *E. coli* resulted in production of the glycine-glucolipid Glc(O-10:0)$_4$Gly.** LC-MS chromatograms (upper panel) show lipids of empty vector cells (black), cells expressing GglsA and GglsB (orange). The peak eluting at 6.3 min only occurs in the GglsA/GglsB expressing cells. The MS/MS spectrum of the peak at 6.3 min (lower panel) shows the fragmentation pattern of Glc(O-10:0)$_4$Gly produced by GglsA and GglsB co-expression cells.

3.7.3. Glycine-Glucolipid Production is Increased by Co-Expression of GglsC Together with GglsA and GglsB

To study the involvement in glycine-glucolipid production, GglsC was co-expressed in *E. coli* with GglsA and GglsB. As shown above, expression of GglsA and GglsB is sufficient for glycine-glucolipid production in *E. coli* (Figure 26). Co-expression of GglsA and GglsB with GglsC resulted in a threefold increase of the amount of glycine-glucolipid produced (Figure 27). Therefore, GglsC is involved in glycine-glucolipid production by activating GglsA.

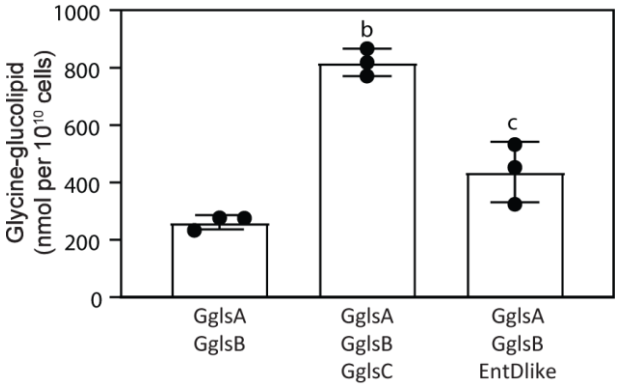

Figure 27 **Co-expression of GglsC together with GglsA and GglsB increased the production of glycine-glucolipid in *E. coli*.**
The glycine-glucolipid production was measured by LC-MS. Data show means and SD (n = 3). One-way ANOVA with Tukey's test, different letters indicate significant differences, $p < 0.05$.

3.7.4. Glycine-Glucolipid Production is Decreased when EntDlike Serves as PPTase Instead of GglsC

In addition, the EntDlike protein from *A. borkumensis* was co-expressed with GglsA and GglsB in *E. coli*. As shown in Figure 27, the *E. coli* cells co-expressing EntDlike produced almost two times the amount of glycine-glucolipid compared with those only expressing GglsA and GglsB. Therefore, the co-expression with GglsC and, to a lower extent with EntDlike, increased the GglsA activity presumably due to an increased level of phosphopantetheinylation of the T domain of GglsA.

3.7.5. Exogenous Fatty Acid Feeding Decreased the Glycine-glucolipid Production in *E. coli*

To improve the glycine-glucolipid production, *E. coli* cells carrying double (GglsA and GglsB) and triple (GglsA, GglsB and GglsC) expression constructs were supplemented with 3-hydroxydecanoic acid or palmitic acid. An equal volume of methanol was added to the control to exclude the impact of solvent. As shown in Figure 28, both 3-hydroxydecanoic acid and palmitic acid supplementation moderately or strongly reduced the glycine-glucolipid production, respectively. The reduced glycine-glucolipid production was observed in both double and triple expression cells. Therefore, exogenous fatty acid feeding cannot improve the glycine-glucolipid production in transformed *E. coli* cells.

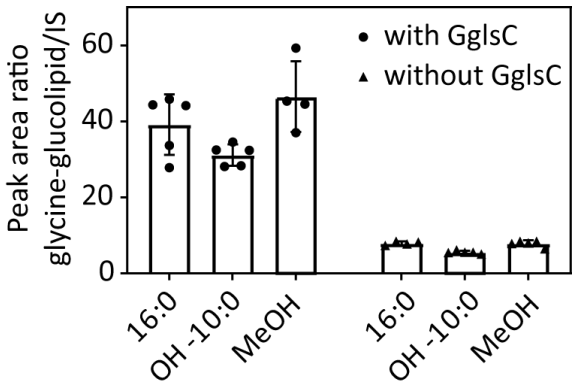

Figure 28 **Exogenous fatty acid feeding did not increase the glycine-glucolipid production in *E. coli*.**
The glycine-glucolipid production was measured by LC-MS. Data show means and SD (n = 5). No significant differences within the two sets of experiments (with GglsC, without GglsC).

3.8. Characterization of GglsA after Heterologous Expression in *E. coli*

3.8.1. C Domain Swapping of GglsA Leads to an Inactive Hybrid Enzyme

In NRPS proteins, the C domain is responsible for the transfer of a single amino acid, fatty acid or peptidyl chain onto the amino acid bound at the T domain. To study the specificity of the GglsA C domain, it was replaced with the C domain of the EntF protein which transfers 2,3-dihydroxybenzoate during enterobactin synthesis in *E. coli* (Frueh, Arthanari et al. 2008) or with the C1 domain of the SrfAA protein which transfers 3-hydroxymyristic acid during surfactin synthesis in *B. subtilis* (Kraas, Helmetag et al. 2010). The chimeric proteins of GglsA containing the C domain of EntF (GglsA-EntFC) or the C1 domain of SrfAA (GglsA-SrfAAC1) were expressed in *E. coli*, lipids were extracted and measured by LC-MS/MS. No new peak occurs in lipid extracts from *E. coli* cells expressing the chimeric GglsA-EntFC or GglsA-SrfAAC1 constructs compared with the *E. coli* cells expressing the empty vector. Therefore, the hybrid enzymes were inactive. The lack of activity of the hybrid enzymes indicates that the C domain is critical for maintaining the enzymatic activity of GglsA. The critical role of the C domain makes it impossible to produce tailor-made glycine-glucolipid with an alternative acid by swapping the C domain with other NRPSs.

3.8.2. A Serine-Glucolipid can be Produced by Domain Swapping with EntF

In NRPS proteins, the A domain is responsible for the activation of a specific amino acid or carboxylic acid. To study the specificity of the GglsA A domain, it was replaced with the A domain of the EntF protein which activates serine during enterobactin synthesis in *E. coli* (Frueh, Arthanari et al. 2008). The chimeric protein of GglsA containing the A domain of EntF (GglsA-EntFA) was expressed in *E. coli*, lipids were extracted and measured by LC-MS/MS (Figure 29). A new peak with 803.6448 *m/z* occurred in lipid extracts from *E. coli* cells expressing the chimeric GglsA-EntFA construct. This peak differs from the glycine-glucolipid peak (773.5554 *m/z*) of wild-type GglsA by 30 Da, equivalent to the mass of one CH_2O unit (calculated 30.0106 Da), representing the difference between the amino acids glycine and serine. This difference is also observed for other fragments, in particular for the peaks at 106.0436 *m/z* (chimeric GglsA-EntFA) and 76.0389 *m/z* (wild type GglsA) representing the amino acids serine and glycine, respectively. Therefore, swapping of the A domains between GglsA and EntF demonstrates that a tailor-made serine-containing aglycone with the structure $H(O\text{-}10:0)_4Ser$ (O-10:0, 3-hydroxy-decanoic acid; Ser, serine) can be produced by the chimeric GglsA-EntFA protein, and that the A domain of GglsA specifically attaches the amino acid glycine to the phosphopantetheine group of the T domain.

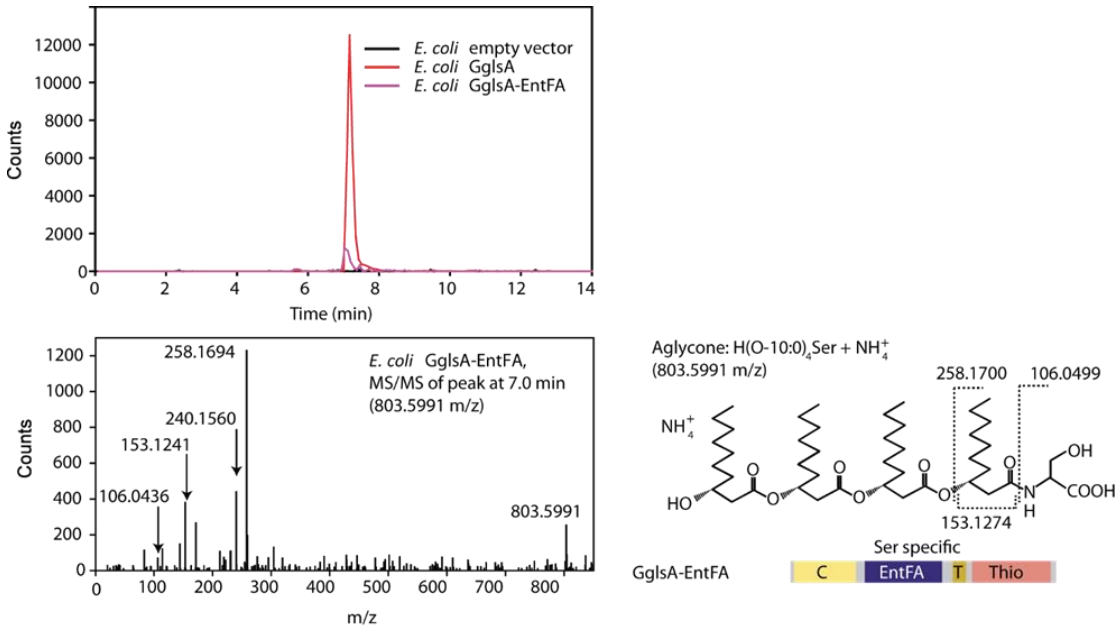

Figure 29 **Expression of GglsA-EntFA in *E. coli* resulted in production of serine containing aglycone $H(O\text{-}10:0)_4Ser$.** LC-MS chromatograms (upper panel) show lipid peaks in cells expressing GglsA (red) and cells expressing GglsA-EntFA (purple) eluting at 7.2 min, but absent from empty vector cells (black). While the peak from GglsA expressing cells represents the aglycone of the glycine-glucolipid, the MS/MS spectrum (lower panel) shows the fragmentation pattern of $H(O\text{-}10:0)_4Ser$ produced by GglsA-EntFA cells.

3.8.3. The GglsA Protein with a Ser1039Ala Mutation is Inactive

To study role of the post-translational modification site of Ser 1039 which carries the phosphopantetheine group, the GglsA protein was mutagenized by replacing the serine 1039 of the thiolation domain with alanine (Ser1039Ala). The mutated protein of GglsA Ser1039Ala was expressed in *E. coli*, lipids were extracted and measured by LC-MS/MS. No aglycone was detected. This means that the Ser1039Ala mutation abolishes the activity of GglsA completely by stopping the activation by phosphopantetheinylation.

3.8.4. No Congeners of the Aglylcone with More or Less than Four Acyl Groups are Produced

To capture potential intermediates of the acylation reaction, the GglsA protein was mutagenized by replacing the serine of the thioesterase domain with alanine (Ser1177Ala). The mutagenized GglsA enzyme is devoid of thioesterase activity and cannot release the intermediates from the T domain. After expression in *E. coli*, a crude protein extract containing the GglsA-Ser1177Ala mutant protein was isolated and the intermediates released from the phosphopantetheine group with cysteamine (Belecki and Townsend 2013). The cysteamine derivatives were analyzed by LC-MS/MS (Figure 30). The extracted ion chromatogram of 890.5956 *m/z* displays a peak for the proton adduct of the di-cysteamine derivative of the aglycone $H(O\text{-}10:0)_4Gly$. No further intermediates were identified in the chromatograms for the putative aglycones with three acyl groups ($(O\text{-}10:0)_3Gly\text{-}cysteamine_2+H^+$, 720.4650 *m/z*), two acyl groups ($(O\text{-}10:0)_2Gly\text{-}cysteamine_2+H^+$, 550.3343 *m/z*) or one acyl group ($(O\text{-}10:0)Gly\text{-}cysteamine_2+H^+$, 380.2036 *m/z*). In agreement with this finding, no congeners of the glycine-glucolipid with more or less than four acyl groups were identified in *A. borkumensis* (Passeri, Schmidt et al. 1992, Abraham, Meyer et al. 1998).

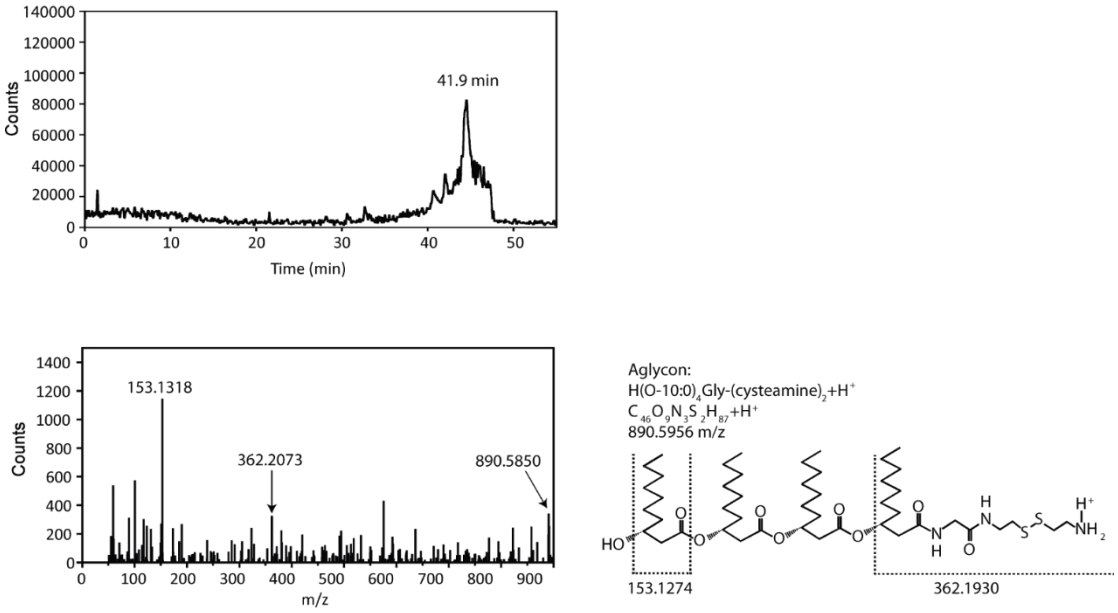

Figure 30 **The GglsA protein with the Ser1177Ala mutation lacks thioesterase activity.**
The glycine-glucolipid intermediates bound to the phosphopantetheine group of the GglsA mutant protein Ser1177Ala were released with cysteamine and analyzed by LC/MS. The LC-MS chromatogram (upper panel) shows the cleavage product eluted at 41.9 min. The MS/MS spectrum (lower panel) shows the fragmentation pattern of the di-cysteamine derivative of the aglycone $H(O\text{-}10:0)_4Gly$.

3.9. *In vitro* Assays of GglsA

3.9.1. GglsA is Directly Involved in Glycine-Glucolipid Production

To demonstrate that GglsA is directly involved in glycine-glucolipid production, the GglsA and GglsC proteins were expressed in *E. coli* and the crude protein extract was used for *in vitro* assays with 3-hydroxy-decanoyl-CoA, ATP and glycine. Lipids were extracted and analyzed by LC-MS/MS. Figure 31 shows that GglsA was capable of *in vitro* synthesis of the aglycone $H(O\text{-}10:0)_4Gly$, i.e. to catalyze the attachment of four 3-hydroxydecanoyl moieties from 3-hydroxy-decanoyl-CoA to glycine. No other aglycones with three or five acyl groups were detected.

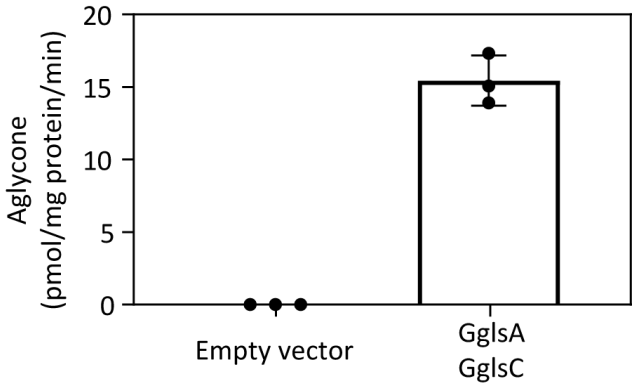

Figure 31 **GglsA can synthesize the aglycone from 3-hydroxydecanoyl-CoA, ATP and glycine as substrates.**
GglsA was co-expressed with GglsC in *E. coli* for a better phosphopantetheinylation of GglsA. The crude protein extract was used for the *in vitro* assay. The formation of aglycone H(O-10:0)$_4$Gly of the glycine-glucolipid was measured by LC-MS. Data show means and SD (n = 3).

3.9.2. GglsA Does Not Use Acyl-CoA Substrates Lacking the 3-Hydroxy Group

As shown in 3.9.1, GglsA catalyzes the attachment of four 3-hydroxydecanoyl moieties from 3-hydroxy-decanoyl-CoA to glycine. To study if GglsA is able to attach decanoyl moieties from decanoyl-CoA to the glycine moiety or the free hydroxyl group of the intermediate, an *in vitro* assay was conducted with decanoyl-CoA as an additional substrate to 3-hydroxy-decanoyl-CoA. To increase the availability of the non-optimal substrate, the concentration of decanoyl-CoA was 5 times higher than 3-hydroxydecanoyl-CoA.Instead of the WT GglsA protein, the Ser1177Ala mutant variant of GglsA was used for the assay. Thus, cysteamine was used to hydrolyze the product of the TE domain. Hydrolysates were extracted and analyzed by LC-MS/MS. The extracted ion chromatogram of 890.5956 *m/z* displays a peak for the proton adduct of the di-cysteamine derivative of the aglycone H(O-10:0)$_4$Gly. No further intermediates with terminal decanoic acid were identified in the chromatograms for the putative aglycones with four acyl groups ((10:0)(O-10:0)$_3$Gly-cysteamine$_2$+H$^+$, 874.6077 *m/z*), three acyl groups ((10:0)(O-10:0)$_2$Gly-cysteamine$_2$+H$^+$, 704.4701 *m/z*), two acyl groups ((10:0) (O-10:0) Gly-cysteamine$_2$+H$^+$, 534.3394 *m/z*) or one acyl group ((10:0)Gly-cysteamine$_2$+H$^+$, 364.20387 *m/z*) (Gly, glycine; O-10:0, 3-hydroxy-decanoic acid; 10:0, decanoic acid). Therefore, decanoyl-CoA lacking the 3-hydroxy group cannot be used as a substrate by GglsA.

3.10. GglsB has High Specificity for UDP-Glucose

To study the substrate specificity, GglsB was expressed in *E. coli* and employed for *in vitro* glycosyltransferase assays. The GglsB protein was incubated with the aglycone H(O-10:0)$_4$Gly isolated from *A. borkumensis ΔgglsB* mutant cells, in combination with different sugar donors, i.e. UDP-glucose, or UDP-galactose. Lipids were isolated and analyzed by LC-MS/MS. The GglsB protein displayed high activity with UDP-glucose, leading to the production of Glc(O-10:0)$_4$Gly, while the activity with UDP-galactose was negligible (Figure 32). Therefore, the glycosyltransferase GglsB attaches a glucose moiety in β-glycosidic linkage to the terminal hydroxyl group using UDP-glucose as donor.

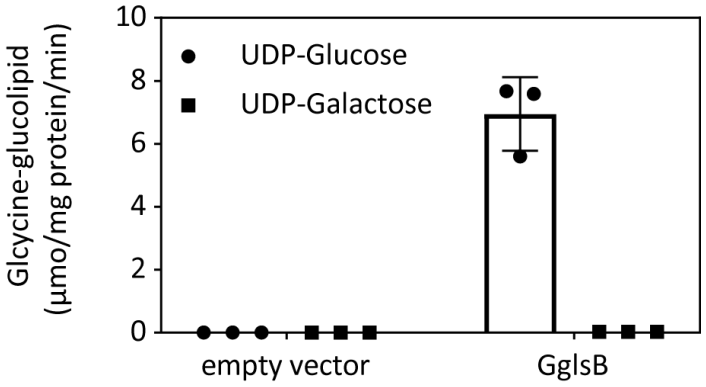

Figure 32 **GglsB glycosylated the aglycone H(O-10:0)$_4$Gly with UDP-glucose as substrate.**
GglsB was expressed in *E. coli*. UDP-Glc and UDP-galactose were tested as sugar donor. The crude protein extract was used for the *in vitro* assay. The formation of glycine-glucolipid Glc(O-10:0)$_4$Gly was measured by LC-MS. Data show means and SD (n = 3).

3.11. *In Vitro* Experiments of the PPTases from *A. borkumensis* to Phosphopantetheinylate the Standalone T Domain of GglsA

To test whether GglsC transfers a phosphopantetheine group to GglsA, the T domain of GglsA fused to a His tag was expressed alone or in combination with GglsC or EntDlike in *E. coli*. In addition, the T domain with the Ser1039Ala mutation was expressed in *E. coli*. Next, the T domain polypeptide was isolated and the purity was checked by 20 % SDS PAGE (Figure 56). The purified T domain was subjected to mass determination. The calculated mass for the apo form of the T domain is 14.46 kDa, and it is 14.80 kDa for the holo form. We observed a mass of ~14.3 kDa in all samples in MALDI-TOF analysis, close to the calculated mass of the apo form, including the Ser1038Ala mutant of the T domain which cannot be phosphopantetheinylated. The mass

difference of 340 Da caused by the attachment of the phosphopantetheine group was not observed via MALDI-TOF analysis (with the help of Dr. Mark Sylvester, Proteomics Core Facility, University of Bonn) (Figure 33). However, all the samples showed same mass, indicating that both GglsC or EntDlike did not catalyze the phosphopantetheinylation reaction on the standalone T domain of GglsA. The reason for the lack of phosphopantetheinylation activity remains unknown.

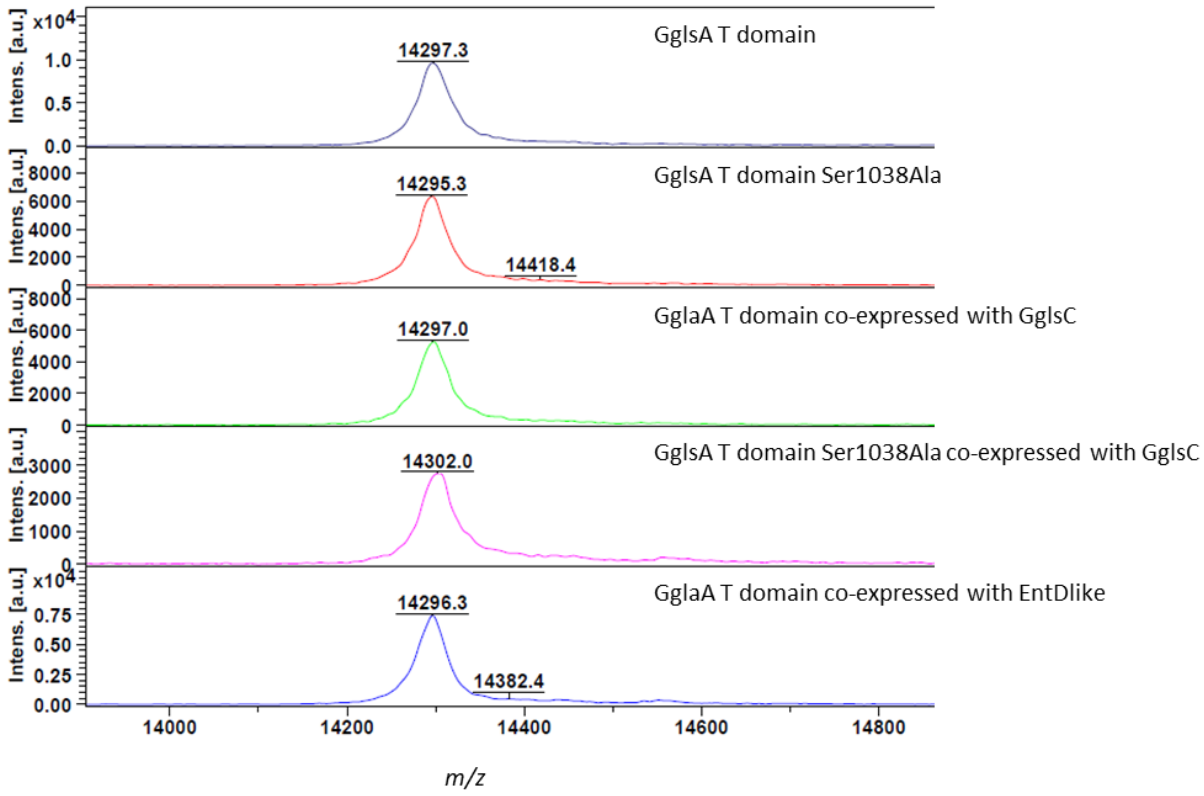

Figure 33 **MALDI-TOF analysis of purified GglsA T domain.**

The GglsA T domain with or without the Ser1039Ala mutation was expressed in *E. coli* alone or co-expressed with GglsC or EntDlike. The calculated size of the T domain without phosphopanteteinylation was 14.46 kDa, phosphopantetheinylated T domain was 14.80 kDa. A mass difference of 340 Da caused by the attachment of the phosphopantetheine group was not observed.

3.12. The Production of Glycine-Glucolipid in *A. borkumensis* Depends on the Capacity of Carbon Storage

To study whether the capacity to store carbon is a prerequisite for the stimulation of production of the glycine-glucolipid, four acyltransferase gene deletion mutant strains were analyzed. These four acyltransferases are involved in triacylglycerol synthesis (*ΔatfA1*, *ΔatfA2*), and *ΔphaC1* and *ΔphaC2* are involved in PHA synthesis (Sabirova, Ferrer et al. 2006, Kalscheuer, Stöveken et al. 2007). All four mutant strains were grown in both ONR7a pyruvate and hexadecane medium. The amounts of glycine-glucolipid were measured by LC-MS/MS. Figure 34 shows that glycine-glucolipid contents were very similar in wild type and all mutants grown on pyruvate. When the cells were grown on hexadecane, the amounts of glycine-glucolipid in WT increased by a factor of two, but the increase was strongly compromised in all mutants.

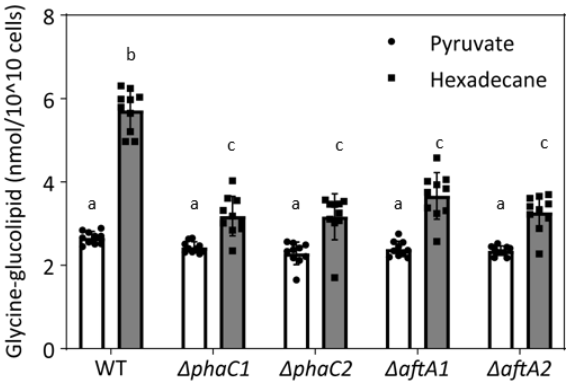

Figure 34 **The glycine-glucolipid production is compromised in *A. borkumensis* mutants affected in storage compound allocation.**

The glycine-glucolipid was quantified by LC-MS. *A. borkumensis* cells were grown in pyruvate and hexadecane medium. The WT and mutant strains affected in synthesis of PHA (*ΔphaC1*, *ΔphaC2*) or TAG (*ΔatfA1, ΔatfA2*) were analyzed. Data show means and SD (n = 10). Two-way ANOVA with Tukey's test, different letters indicate significant differences, $p < 0.05$.

3.13. The Glycine-Glucolipid Increases the Attachment of *A. borkumensis* Cells to the Oil/Water Interphase

It has previously been shown that *A. borkumensis* cells grown in a hexadecane pre-culture display improved growth on hexadecane compared with pyruvate-grown cells (Passeri, Schmidt et al. 1992, Godfrin, Sihlabela et al. 2018), suggesting that the cells specifically adapt to hexadecane as a carbon source, possibly by accumulating the glycine-glucolipid. It had been speculated that the glycine-glucolipid might be required as a surfactant for the attachment of the cells to the alkane-water interphase and for biofilm formation on the oil droplet. The glycine-glucolipid deficient mutants were used to study this hypothesis experimentally.

3.13.1. Absence of the Glycine-Glucolipid Affects the Growth of *A. borkumensis*

The growth of *A. borkumensis* WT, the *ΔgglsA* and *ΔgglsB* mutants on pyruvate or hexadecane medium was compared. As shown in Figure 35B, growth of the mutant cells was very similar to WT on pyruvate medium. Therefore, deletion of the entire glycine-glucolipid biosynthetic pathway (*ΔgglsA*), or replacement with the corresponding aglycone (*ΔgglsB*), had a minor impact on growth on a soluble carbon source, pyruvate. However, the *ΔgglsA* and *ΔgglsB* mutant cells displayed strongly or moderately reduced growth, respectively, on hexadecane medium. For these experiments, cultures were grown in flasks on a rotary shaker with 180 rpm rotation (Figure 35A). To further investigate if the glycine-glucolipid deficiency can be compensated by increasing the oxygen supply and hexadecane bioavailability, *A. borkumensis* WT, the *ΔgglsA* and *ΔgglsB* mutants were grown on hexadecane medium with 0 rpm and 300 rpm rotation (Figure 35C and D). The higher speed of rotation resulted in a better oxygen supply and improved the bioavailability of hexadecane. Both WT and mutant cells showed reduced growth without rotation, especially *ΔgglsB* mutant cells displayed strongly reduced growth. Faster rotation increased the growth of WT cells in the early stationary phase. The growth of mutant cells was not improved by rotation faster than 180 rpm. This result demonstrates that the glycine-glucolipid is crucial for the growth on water-insoluble carbon sources under oxygen limitation, and that the replacement of the glycine-glucolipid with the aglycone partially complements the absence of glycine-glucolipid. Mutant cells deficient in glycine-glucolipid are less tolerant to low oxygen supply and hexadecane bioavailability.

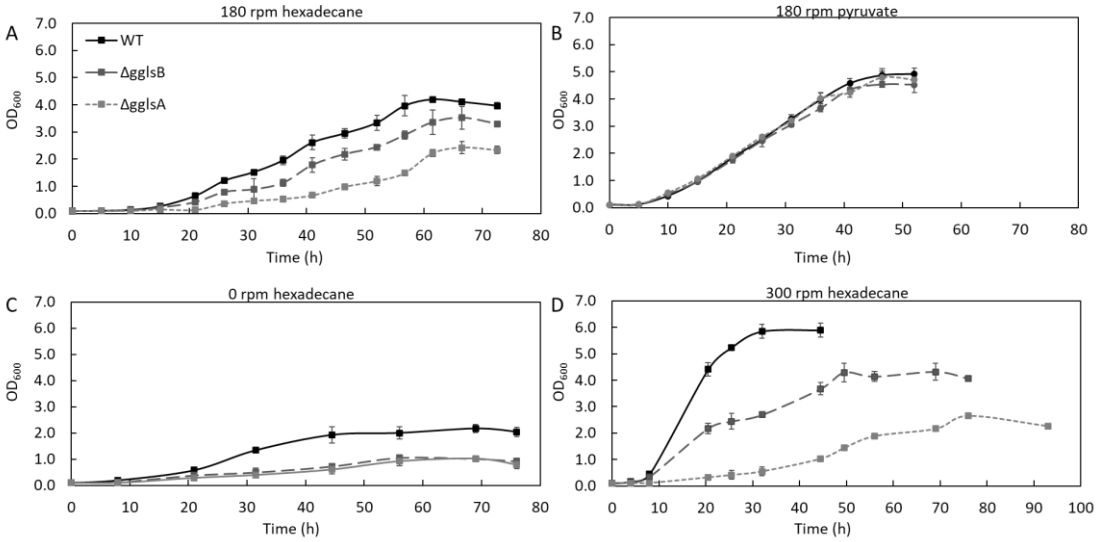

Figure 35 **Growth of *A. borkumensis* mutants deficient in glycine-glucolipid production.**
The growth of *A. borkumensis* cells was monitored by measuring OD_{600} with corresponding medium as blank. Data show means and SD (n = 3).

3.13.2. Hexadecane-Grown Mutant Cells Lacking the Glycine-Glucolipid Display a Compromised Attachment to the Oil/Water Interphase

To study the attachment of the cells to hexadecane directly, WT and mutant cells were diluted to OD_{600} ~1.0 with carbon free ONR7a medium. 900 µL of cell broth and 100 µL of hexadecane were added into the photometer cuvette in the 'bacterial adhesion to hydrocarbons (BATH) assay' (Godfrin, Sihlabela et al. 2018). The cells attach to the alkane/water interface, accumulating as a top layer. The movement of the cells to the hexadecane layer was monitored photometrically by recording the decline of OD_{600} in the middle of the cuvette. When the bacteria were pre-cultured in pyruvate medium, only a low number of cells (~30 % after 10 min) were attached to the hexadecane layer, and no differences between WT, the *ΔGglsA* and *ΔGglsB* mutants were observed (Figure 36 left). Pre-culturing of the WT cells on hexadecane resulted in a stronger attachment to the hexadecane layer (~80 % of cells). However, the attachment of *ΔgglsB* mutant cells to hexadecane was compromised (~60 %), and it was even further decreased in the *ΔgglsA* mutant (~30 %).

The movement of the cells to the hexadecane layer was monitored by recording the decline of OD_{600} in the middle of cuvette every 30 s for 10 min. The OD_{600} was converted into cell number according to the counting of colony forming units (CFU) (Figure 55). As shown in Figure 36 (right),

pre-culturing of the WT cells on hexadecane resulted in a faster attachment to the hexadecane layer (~2.0 * 10^9 cells/min). However, the attachment speed of *ΔgglsA* and *ΔgglsB* mutant cells to hexadecane was slowed down (~0.43 * 10^9 cells/min). The bacteria pre-cultured in pyruvate medium show the slowest attachment to the hexadecane layer (<0.3 * 10^9 cells/min). Therefore, lack of the glycine-glucolipid strongly affects the capacity of the cells to attach to the oil/water interphase, while the aglycone can in part functionally compensate for the loss of the glycine-glucolipid.

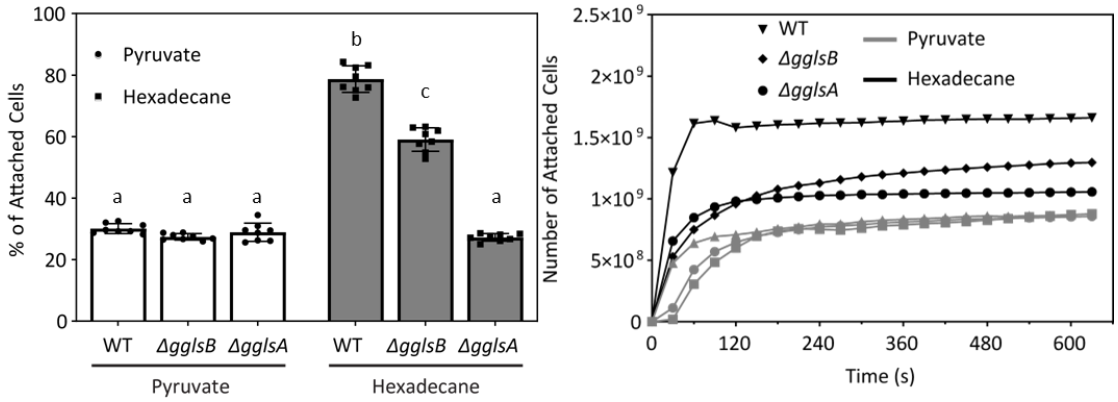

Figure 36 **Attachment to the oil/water interface of the *A. borkumensis* mutants deficient in glycine-glucolipid.** Bacterial adhesion to hydrocarbons (BATH) assay of *A. borkumensis* WT and mutant cells (*ΔgglsA, ΔgglsB*) grown on pyruvate or hexadecane. The attachment of the cells to the oil/water interface at the top was recorded by measuring the decline in OD_{600} (OD_{600} = 1.0 at 0 s) at the center of the cuvette, and converted into number of cells (right, time course) and % of cells (left, after 10 min). Data show means and SD (n = 8). Two-way ANOVA with Tukey's test; different letters indicate significant differences; $p < 0.05$.

Attachment of the cells to hexadecane droplets was visualized by confocal fluorescence microscopy. The cells were diluted with fresh medium to OD_{600} ~1.0 after growth in pyruvate or hexadecane medium, stained with Nile Red and after staining, 900 µL of cell broth were mixed with 100 µL of hexadecane. Cells attached to the hexadecane droplets were then observed by confocal fluorescence microscopy. Figure 37 shows that only very few cells of WT or the mutants were attached to the hexadecane droplets when the cells were grown in pyruvate. In contrast, WT cells grown in hexadecane showed a very strong attachment to the hexadecane droplets, while attachment of hexadecane-grown *ΔgglsA* and *ΔgglsB* mutant cells was strongly reduced. A slightly higher number of hexadecane-grown cells of the *ΔgglsB* mutant which produces the aglycone was attached to the hexadecane droplets compared with the *ΔgglsA* mutant which is completely deficient in the glycine-glucolipid.

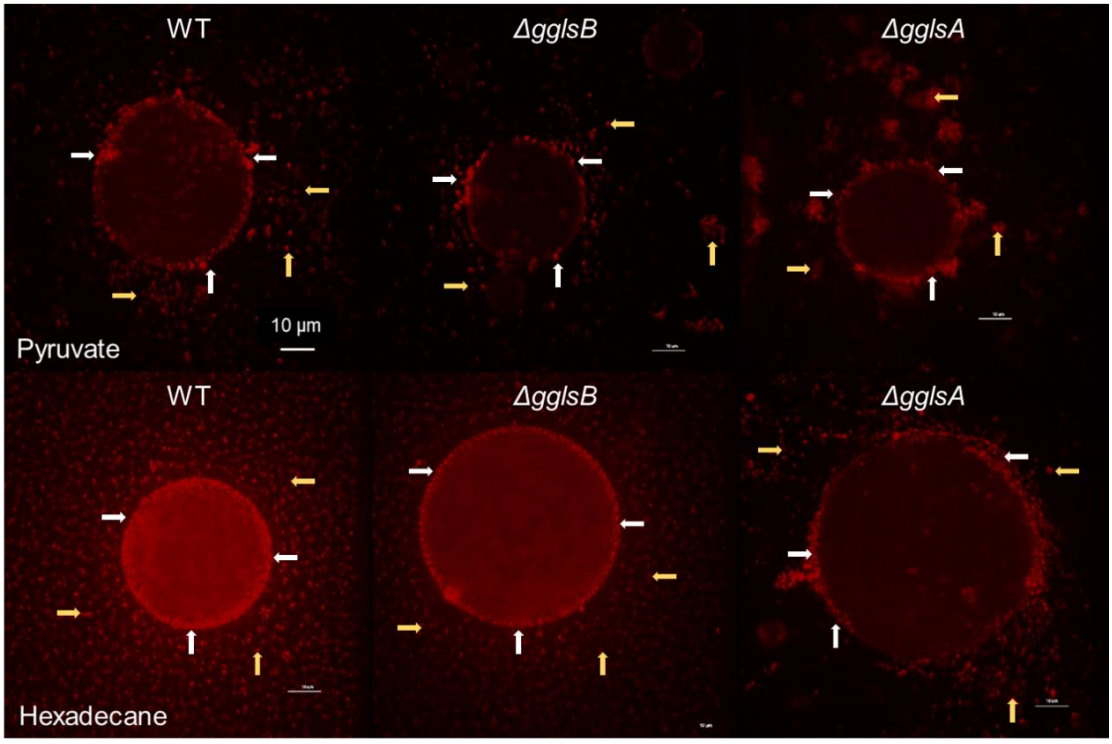

Figure 37 **Attachment to the oil/water interface of the *A. borkumensis* mutants deficient in glycine-glucolipid visualized via confocal fluorescence microscopy.**

A. borkumensis WT and mutant cells (*ΔgglsA, ΔgglsB*) grown on pyruvate or hexadecane were stained with Nile Red and suspended in medium with hexadecane. Red fluorescent cells were attached to oil droplet (biofilm cells, white arrows), and not associated (detached cells, yellow arrows). Bar, 10 µm

The biomass of cells attached to the hexadecane droplet observed by confocal fluorescence microscopy was quantified by integrating the red fluorescence intensity of the bacterial cross-sectional layer around the oil droplets. To ensure that the fluorescence was solely attributed to cells, all photos for fluorescence integration were taken with the same exposure time and normalized to the area of the bacterial layer. Figure 38 shows that pyruvate-grown cells of WT or the mutants attached to the hexadecane droplets resulted in the fluorescence of ~4,500 arbitrary unit/µm^2 (AU/µm^2). In contrast, WT cells grown in hexadecane resulted in the fluorescence over 10,000 AU/µm^2, indicating a very strong attachment to the hexadecane droplets. The fluorescence of hexadecane-grown *ΔgglsA* and *ΔgglsB* mutant cells which were attached to the droplets was reduced to 4,400 AU/µm^2 and 8,300 AU/µm^2, respectively.

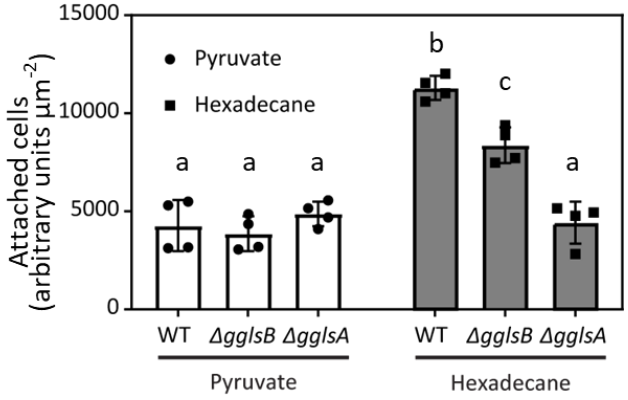

Figure 38 **Attachment to the oil/water interface of the *A. borkumensis* mutants deficient in glycine-glucolipid.** The biomass of cells attached to the oil droplet was determined by integrating the intensity of the red fluorescence of the cross-sectional bacterial layer around the oil droplet, normalized to the area of the bacterial layer. Data show means and SD (n = 4). Two-way ANOVA with Tukey's test; different letters indicate significant differences; $p < 0.01$.

3.13.3. Hexadecane-grown Mutant Cells without Glycine-Glucolipid show Lower Dispersant Effectiveness than WT

Because the glycine-glycolipid remains attached to the bacteria, the effectiveness of emulsification of WT and mutant cells was tested. The cells were diluted into OD_{600} ~ 0.5 with ONR7a medium without N and P but containing 1 % of hexadecane. The cells were incubated at 28 °C 180 rpm for 30 min. The proportion of hexadecane in the aqueous phase was measured by GC-MS. At 10 min after mixing, ~6% of the hexadecane was found in the aqueous medium in the presence of pyruvate-grown cells, independent of the mutations, and similar to the cell-free control (Figure 39). In contrast, ~15% and ~11% of hexadecane was emulsified by hexadecane-grown WT or *ΔgglsB* mutant cells, respectively, but only 6% by the *ΔgglsA* mutant which was similar to the control. Therefore, the glycine-glycolipid is required for hexadecane emulsification by the cells. Replacement with the aglycone (*ΔgglsB*) or loss of the glycine-glycolipid (*ΔgglsA*) decreases or abolishes the effectiveness of emulsification, respectively.

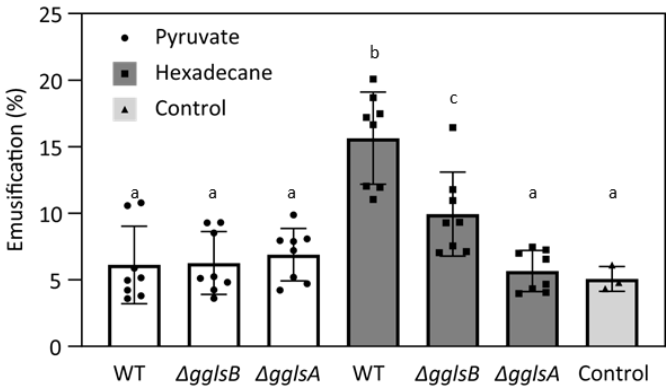

Figure 39 **Effectiveness of emulsification of hexadecane by *A. borkumensis* cells.**
WT and mutant cells grown on pyruvate or hexadecane were resuspended in ONR7a medium without N and P with 1 % hexadecane. Data show means and SD (n = 8). Two-way ANOVA with Tukey's test; different letters indicate significant differences; $p < 0.05$.

3.13.4. Glycine-Glucolipid Deficient Mutant Cells Grown in Hexadecane Show Decreased Cell Sizes

For ultrastructural analyses, the WT and mutant cells were grown in pyruvate or hexadecane medium. After growth in hexadecane medium, two populations of cells were separated, i.e. cells attached to the oil/water interface at the top of the medium ('biofilm cells'), or cells existing in the aqueous phase ('detached cells'), as previously described (Klein, Grossi et al. 2008). The cells were collected on 0.1 µm filters, they were fixed and analyzed by SEM (Figure 40). The hexadecane-grown cells formed circular aggregates which presumably derived from collapsed oil droplets.

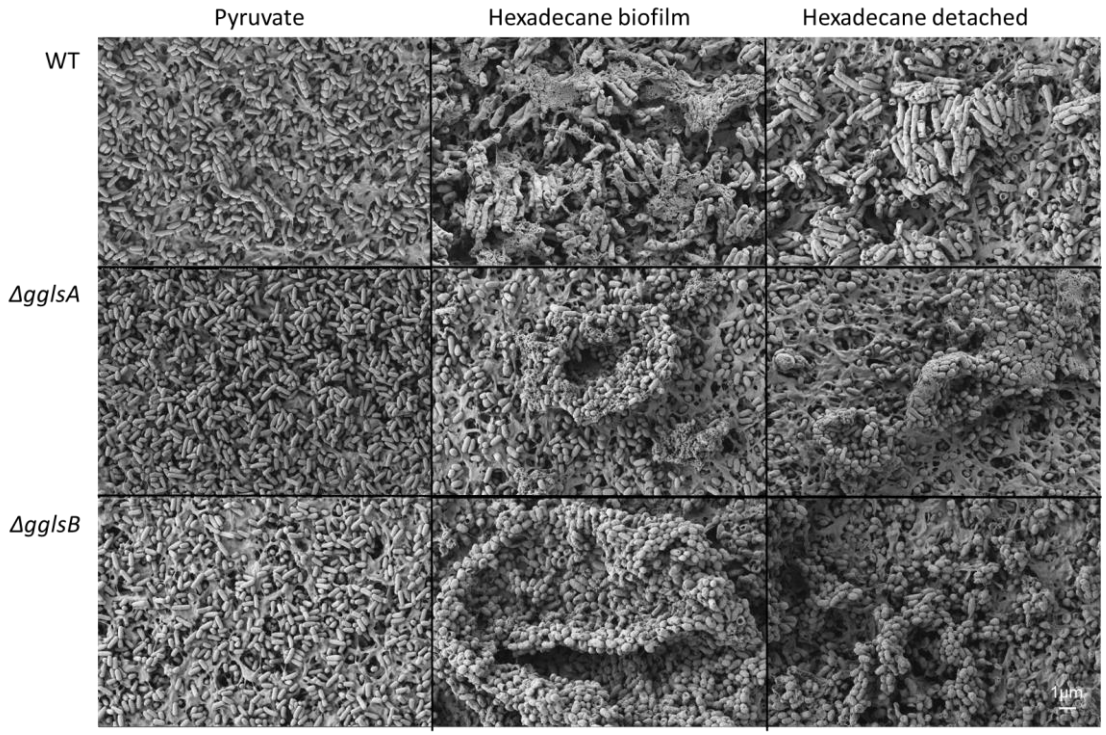

Figure 40 **Ultrastructure of *A. borkumensis* mutants deficient in glycine-glucolipid.**
Scanning electron microscopic observation of *A. borkumensis* WT, *ΔgglsA and ΔgglsB* mutants grown on pyruvate or hexadecane medium. The hexadecane grown cells were separated into cells attached to oil/water layer at the top (biofilm cells) and growing in the aqueous phase (detached cells). Bar, 1 µm

The lengths of the cells which were mostly fixed in longitudinal orientation were determined. While WT and mutant cells grown on pyruvate revealed similar cell lengths (Figure 41), WT cells (biofilm and detached) grown on hexadecane displayed a large size and small inward invaginations presumably corresponding to inclusions of storage lipid that were removed during fixation and dehydration.

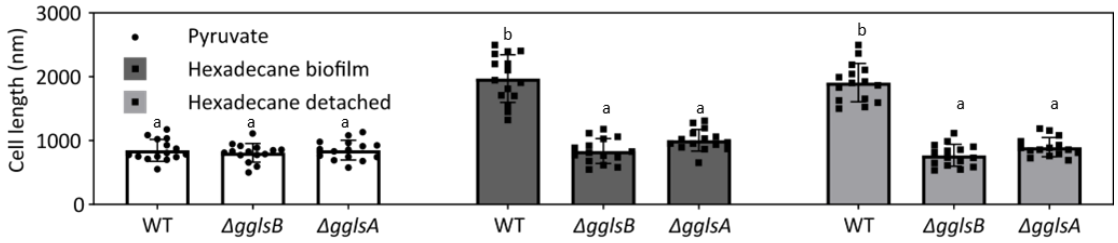

Figure 41 **Morphological characteristics of *A. borkumensis* mutants deficient in glycine-glucolipid.**
Length of cells identified in SEM images. Data show means and SD (n = 3 images, ca. 5 cells per image). Two-way ANOVA with Tukey's test; different letters indicate significant differences; $p < 0.01$.

3.13.5. Glycine-Glucolipid Deficient Mutants Show Decreased Numbers of Inclusions

For the TEM analysis, the WT and mutant cells were grown as described for SEM (3.13.4). After harvesting by centrifugation, cells were fixed and embedded for ultrathin sectioning. Observation by TEM revealed that the size and morphology of the pyruvate-grown WT and *ΔgglsA* and *ΔgglsB* cells were similar (Yakimov, Golyshin et al. 1998). The ultrastructure of the cells revealed translucent inclusions, representing storage compounds (Kalscheuer, Stöveken et al. 2007, Klein, Grossi et al. 2008, Manilla-Pérez, Lange et al. 2011). The abundance of the inclusions was dependent on the growth conditions (Figure 42). During growth in pyruvate medium, the numbers of cells with inclusions were not significantly different between WT and mutant cells.

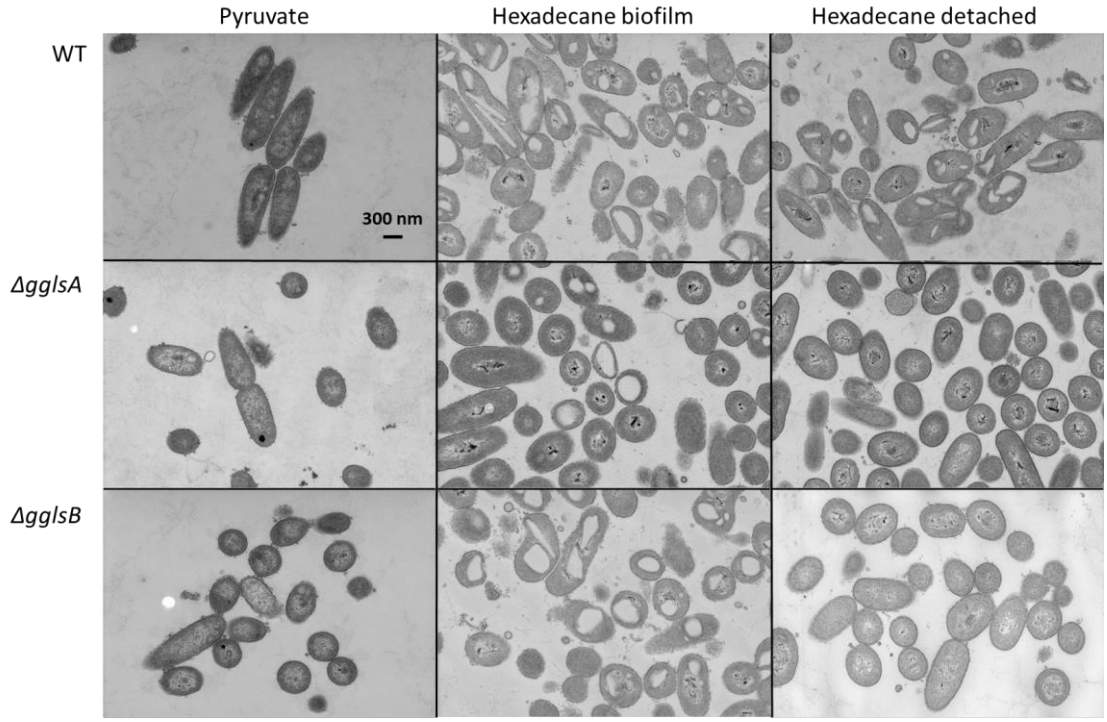

Figure 42 **Ultrastructure of *A. borkumensis* mutants deficient in glycine-glucolipid.**
Transmission electron microscopic observation of *A. borkumensis* WT, *ΔgglsA and ΔgglsB* mutants grown on pyruvate or hexadecane medium. The hexadecane grown cells were separated into cells attached to oil/water layer at the top (biofilm cells) and growing in the aqueous phase (detached cells). Bar, 300 nm

As shown in Figure 43, about 90 % of the biofilm cells of WT contained inclusions, indicating that these cells accumulated large amounts of storage compounds. Much lower numbers of the *ΔgglsB* cells and to an even stronger extent, the *ΔgglsA* mutant cells had inclusions, indicating that the glycine-glucolipid is important for hexadecane uptake, metabolization and allocation to storage compounds. Similarly, a larger number of the detached WT cells displayed inclusions compared with the *ΔgglsA* and *ΔgglsB* cells, which showed no inclusions. Detached cells are derived from the biofilm cells which detach from the oil/water interface, but not from growth and division of detach cells in the aqueous medium (Klein, Grossi et al. 2008).

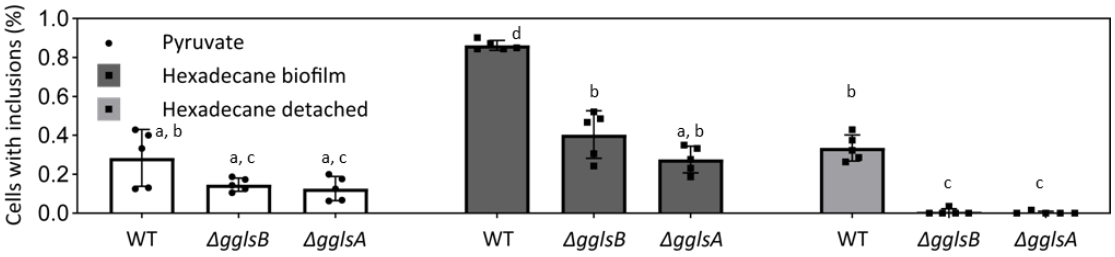

Figure 43 **Morphological characteristics of *A. borkumensis* mutants deficient in glycine-glucolipid.**
Number of inclusion-containing cells identified in TEM images (in % of the total number of cells per image). Data show means and SD (n = 5 images, ca. 5-50 cells per image). Two-way ANOVA with Tukey's test; different letters indicate significant differences; $p < 0.05$.

3.13.6. Glycine-Glucolipid Deficient Mutants Grown in Hexadecane Show Decreased Amounts of Storage Lipids

The inclusions found by TEM could be storage lipids. According to Kalscheuer et al. (2007), the storage lipids could be fatty acids, triacylglycerol (TAG) or wax esters (WEs). The storage lipids in *A. borkumensis* grown in both pyruvate and hexadecane were analyzed via TLC. To make sure that the lipids loaded onto the TLC plate originated from the same number of cells, the total lipid extracts were diluted according to the OD_{600}, and equal volumes of the cell extracts were loaded.

The cells for lipid analysis were not separated into biofilm cells and detached cells, because it is hard to determine the biomass of biofilm cells. Therefore, the total lipid extracts for the storage lipid analysis were originated from the whole culture.

As shown in the Figure 44, the storage lipids in pyruvate-grown cells are mainly TAG, WEs were not observed. The deficiency of glycine-glucolipid has no impact on the accumulation of these storage lipids. Both TAG and WEs content of hexadecane-grown mutants decreased compared with WT. Therefore, glycine-glucolipid deficiency mostly affects carbon storage as TAG and WEs.

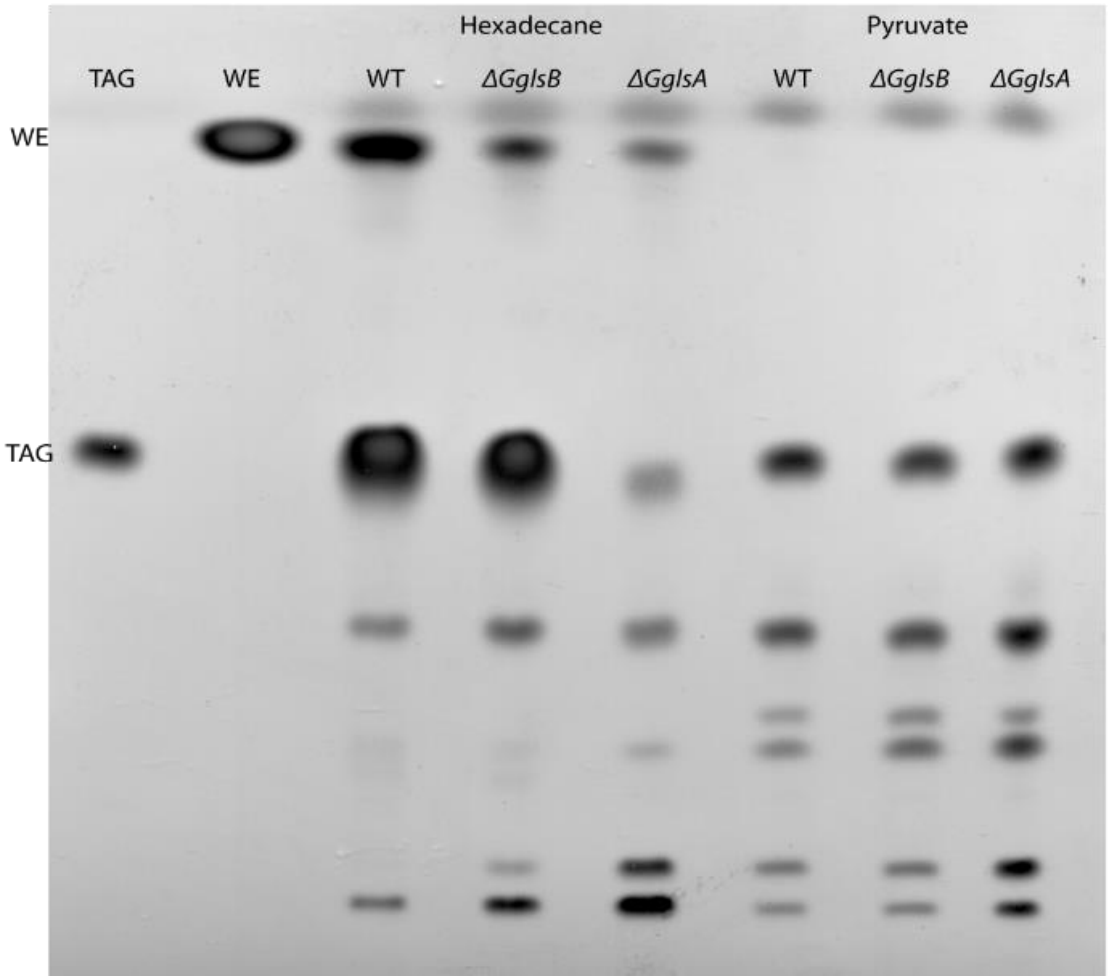

Figure 44 **TLC separation of storage lipids isolated from *A. borkumensis* WT and glycine-glucolipid deficient mutant cells grown in ONR7a medium with pyruvate or hexadecane.**

The storage lipids were separated together with triolein (TAG) and wax ester (WE) for band identification. Storage lipids were isolated from equal amounts of *A. borkumensis* cells according to the OD_{600} of the cultures.

Taken together, the result shows that the glycine-glucolipid is crucial for *A. borkumensis* cells during alkane degradation, because it is required for the attachment to the biofilm at the oil/water interphase and the emulsification of the oil. The absence of glycine-glucolipid affected further the accumulation of storage compounds and the capacity to reach full size during growth on hexadecane.

3.14. The Glycine-glucolipid can be Used as Surfactant for Foliar Fertilization

The glycine-glucolipid was studied as a candidate for a biosurfactant which can be used as additive of foliar fertilization. For this study, the lipid material was isolated by Tobias Karmainski in Prof. Lars M. Blank's lab (RWTH Aachen). As a surfactant, the first step is to study the critical micelle concentration of the surfactant. The critical micelle concentration (CMC) of the glycine-glucolipid was analyzed via measuring the surface tension with a drop shape analyzer together with Dr. Viktoria V. Zeisler-Diehl in Prof. Lukas Schreiber's lab (University of Bonn). As shown in Figure 45, after reaching the concentration of 120 mg/L, the surface tension remains relatively constant at around 53 mN/m and decreases with a lower slope. Therefore, the CMC of glycine-glucolipid is measured as 120 mg/L. This concentration of glycine-glucolipid was applied further for foliar phosphate fertilization.

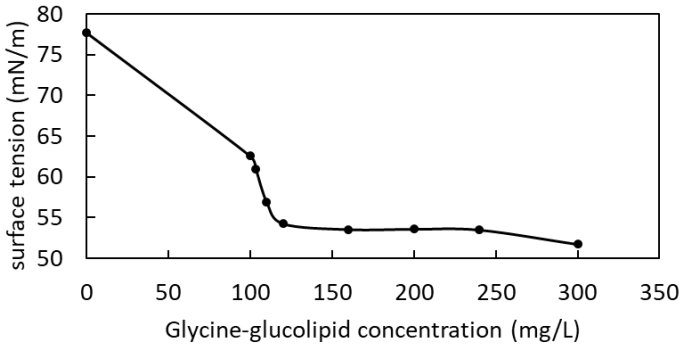

Figure 45 **Surface tension measurement of the glycine-glucolipid via drop shape analyzer (DSA).**
The critical micelle concentration of the glycine-glucolipid is measured as 120 mg/L

To study the improvement of foliar phosphorus uptake by adding the surfactant, equal concentrations (120 mg/L) of Triton X-100, decylmaltopyranoside and glycine-glucolipid were tested. Phosphate fertilizer, with or without surfactant, was applied to the leaves of *N. benthamiana*. In *N. benthamiana* as in all other plants, phospholipid levels decrease during phosphate deprivation as glycolipids replace phospholipids in the membranes (Minnikin, Abdolrahimzadeh et al. 1974, Hölzl and Dörmann 2007). Therefore, the total phospholipid content per dry weight of leaves was measured to record phosphate deficiency and the extent of recovery after foliar phosphate application.

As shown in Figure 46, in the absence of surfactant (water control) resulted in a total phospholipid content similar to that observed when only water was applied (negative control, NC), indicating minimal phosphate uptake from the leaves without surfactant aid. Both glycine-glucolipid (GG) and Triton X-100 (Triton) resulted in a comparable improvement in foliar phosphorus uptake, as evidenced by increased phospholipid content in the leaves.

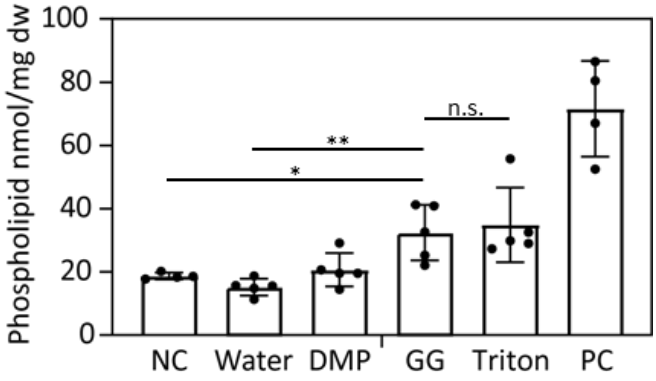

Figure 46 **The improvement of phosphorus fertilization b different surfactants on leaves of *N. benthamiana*.** Phospholipid was measured as indicator of the degree of phosphate deprivation via Q-TOF direct infusion MS and normalized to mg dry weight (dw). NC: negative control, no phosphate applied; Water: phosphate dissolved in water without surfactant; DMP: phosphate dissolved in water with decylmaltopyranoside as surfactant; GG: phosphate dissolved in water with glycine-glucolipid as surfactant; Triton: phosphate dissolved in water with Triton X 100 as surfactant; PC: positive control, plants grown in medium with phosphate. Data show means and SD (n = 5). Student test; *: $p < 0.05$, **: $p < 0.01$; n.s.: not significantly different.

4. Discussion

4.1. The Glycine-Glucolipid is Resident in the Bacterial Cell Wall and Facilitates Attachment to the Oil/Water Interface

In some bacteria and plants, the amounts of glycolipids, primarily glucose or galactose moieties linked to diacylglycerol, increase during growth in phosphate-deprived medium, because the glycolipids replace phospholipids in the membranes (Minnikin, Abdolrahimzadeh et al. 1974, Hölzl and Dörmann 2007). Accumulation of the glycine-glucolipid, which is composed of glucose attached to four 3-hydroxy-fatty acids with a glycine bound to the terminal fatty acid in amide linkage (Yakimov, Golyshin et al. 1998), was not stimulated by phosphate deprivation. Instead, its production was correlated with the carbon source, specifically alkanes, indicating its role in the alkane degradation process rather than in the replacement of phospholipids (Godfrin, Sihlabela et al. 2018).

Biosurfactant production can be regarded as a response to alkanes present in the environment and may allow the bacteria to access these water-insoluble compounds as carbon sources. Two different forms of the glycolipid were previously described in *A. borkumensis*, the glycine-glucolipid (carrying a glycine residue in amide linkage at the terminal 3-hydroxy decanoic acids) and glycine-free form (lacking a glycine moiety the terminal 3-hydroxy decanoic acids). The glycine-free form was suggested to be present in the culture broth (Schulz, Passeri et al. 1991, Passeri, Schmidt et al. 1992, Abraham, Meyer et al. 1998) and to be involved in the emulsification of the water-insoluble alkanes. The differences in the structures of glycine-glucolipid in the cell pellet and culture broth led to the speculation that the glycine-glucolipid in the cell wall of *A. borkumensis* might be a precursor for the secreted, glycine-free form, needed for alkane solubilization (Abraham, Meyer et al. 1998, Timmis, McGenity et al. 2010, Godfrin, Sihlabela et al. 2018). However, only the glycine-containing form was identified predominantly in the cell pellet; the abundance was very low in the cell-free supernatant (Figure 20). A further lipid analysis via Q-TOF mass spectrometry was performed on the lipid extracts of pyruvate-grown culture and hexadecane-grown cultures. In particular, phosphatidylglycerol (PG), an abundant phospholipid in *A. borkumensis*, shows a very similar distribution to that of the glycine-glucolipid, i.e., most PG was found in the cell pellet and only minute amounts in the supernatant (Figure 21). PG in the supernatant is presumably derived from damaged cell membranes or outer membrane vesicles

rather than from active secretion. The supernatant surface tension measurement showed that the supernatants of both pyruvate-grown cells and hexadecane-grown cells have the same surface tension as cell free medium (Figure 19). These findings support the hypothesis that the *A. borkumensis* cells are directly attached to the lipid-water interphase (Abbasi, Bothun et al. 2018, Godfrin, Sihlabela et al. 2018), and that the glycine-glucolipid is not secreted for the solubilization of alkanes, but that it adjusts the hydrophobicity of the cells for better attachment to the alkanes (Schulz, Passeri et al. 1991).

The most critical step during alkane degradation is the movement to and attachment at the oil/water interface. *A. borkumensis* cells form a biofilm with tubular extensions around the oil droplets, thereby increasing the interphase where the cells take up the alkanes for intracellular degradation (Prasad, Obana et al. 2023). The physiological experiments including the BATH experiment, hexadecane emulsification, visualization via confocal microscopy and integration of fluorescence, were performed with *ΔgglsB* and *ΔgglsA* mutants grown in pyruvate or hexadecane containing medium. The hexadecane-grown mutant cells without glycine-glucolipid and pyruvate-grown cells showed reduced adhesion to hexadecane and poorer emulsification compared to hexadecane-grown WT cells. The absence of glycine-glucolipid impacted the attachment to hexadecane, further reducing growth on hexadecane medium by limiting hexadecane utilization. Previous reports indicate that fatty alcohols used for WEs synthesis are produced through alkane oxidation in *A. borkumensis* (Kalscheuer, Stöveken et al. 2007). Therefore, storage lipid accumulation, especially WE, is likely directly affected by hexadecane uptake. Analysis of storage lipids supports this scenario, showing reduced WE in the mutant (Figure 44, in chapter 3.13.6). The reduced storage lipid accumulation further results in smaller cell size. Besides, it was reported before that the water contact angle of pyruvate-grown cells (~75°) was considerably lower than that of hexadecane-grown cells (~110°) (Naether, Slawtschew et al. 2013). These results demonstrate that the glycine-glucolipid plays a crucial role in the first step of oil degradation which is the attachment to the oil and the effectiveness of emulsification of the oil (Abbasi, Bothun et al. 2018, Kubicki, Bator et al. 2020).

The replacement of the glycine-glucolipid with the aglycone complements the absence of glycine-glucolipid with less efficiency. Therefore, the interphase where the cells take up the water-insoluble alkanes is limited and finally resulted in a strong reduction of growth (Prasad, Obana et al. 2023).

4.2. GglsA, GglsB and GglsC are Involved in the Biosynthesis of the Glycine-Glucolipid in *A. borkumensis*

Two different theories were proposed about the biosynthesis pathway of the glycine-glucolipid in *A. borkumensis*, the rhamnolipid-like pathway via glycyltransferase, acyltransferase and glucosyltransferase or the NRPS-dependent pathway via NRPS, acyltransferase and glucosyltransferase. The deletion mutations in *A. borkumensis* and expression in *E. coli* show that the glycine-glucolipid is produced in *A. borkumensis* via the NRPS-dependent pathway (Figure 47). Contrary to initial speculation, there is no extra acyltransferase involved in assembling the four fatty acids. Instead, only three enzymes are involved: NRPS/GglsA, PPTase/GglsC, and glucosyltransferase/GglsB.

The GglsC activates GglsA by phosphopantetheinylation on the T domain Ser1039. The A domain of GglsA activates and transfers a glycine onto the phosphopantetheine arm tethered on the T domain with the consumption of an ATP molecule. The C domain of GglsA then catalyzes the condensation reaction, transferring a 3-D-hydroxydecanoyl moiety from 3-D-hydroxydecanoyl CoA and forming an amide bond between the amino group of glycine and the acyl group of the 3-D-hydroxydecanoyl moieties. GglsA subsequently attaches three additional 3-D-hydroxydecanoyl moieties onto the hydroxy group of the intermediate tethered on the T domain. Finally, the T domain shuttles the aglycone from the C/T surface to the T/TE globular core, and the aglycone is hydrolyzed from GglsA by the TE domain at catalytic site Ser1177. The aglycone released from GglsA is then glycosylated by GglsB using UDP-glucose as a substrate.

The synthesis of glycine-glucolipid via GglsA is different from enterobactin synthesis via EntF in *E. coli*, where the intermediate is assembled and released after trimerization via cyclization (Miller, Drake et al. 2016). It is more similar to the myxochelin synthesis, where the C domain of MxcG transfers twice dihydroxybenzoic acid to the lysine on the T domain of MxcG (Gaitatzis, Kunze et al. 2001). However, it is also different from the glycine-glucolipid synthesis, because lysine contains two amino group, and the dihydroxybenzoic acid moieties are transferred onto the α and ε amino group and form an amide bond instead of an ester bond (Gaitatzis, Kunze et al. 2001). Therefore, GglsA represents a unique NRPS system involved in glycine-glucolipid synthesis due to its ability to sequentially attach four acyl groups to the amino acid glycine, all catalyzed by a single C domain.

The exact mechanism enabling GglsA to transfer four 3-hydroxydecanoic acids sequentially is unknown. It is hypothesized that the C domain might have a long tunnel with four active sites for condensation, with the thioesterase domain only conducting hydrolysis once the final product H(O-10:0)$_4$Gly has been assembled. Alternatively, the C domain has a tunnel with similar length as other NRPSs, the distances between T and C domain is larger and the TE domain is specific to the aglycone with four fatty acids. Once the glycine is attached onto the T domain, only after all four fatty acids are attached, the T domain is able to shuttle the aglycone tethered on the T domain to the TE domain, the size of aglycone with four fatty acids can fulfill the space between thioesterase and two α-helices (α4TE–α5TE), finally accommodates the intermediate to the T/TE domain active sites (Ser1039 and Ser1177) for hydrolysis (Frueh, Arthanari et al. 2008). These hypotheses might be confirmed by protein structure analysis. The sequence analysis of GglsA C domain shows only one active site HHxxxDG, i.e. it is more likely that the number of fatty acids is controlled via C and TE domain.

A. borkumensis contains two type II PPTases: ABO_1782 and ABO_2087. *ABO_1782* is predicted to encode a protein of 253 amino acids, and it shows 20.8 % sequence identity with the EntD protein involved in enterobactin synthesis in *E. coli*. EntD belongs to the type II family of PPTases which activate carrier proteins in NRPSs, in contrast to type I *holo*-ACP synthase like ACPS from *E. coli* (gene P24224) which activates ACP involved during *de novo* synthesis of fatty acids (Jung, Bashiri et al. 2016). EntD from *E. coli* can transfer a phosphopantetheine moiety to the T domain of two substrates, EntF (NRPS involved in enterobactin synthesis) and EntB (which carries the activated 2,3-dihydroxybenzoate during enterobactin synthesis). The second type II PPTase in *A. borkumensis* is ABO_2087 with 34.06 % sequence identity with ABO_1782 on amino acid level. The *ABO_2087* gene is localized close to the gene cluster *ABO_2092-ABO_2093*, which was suggested to be involved in the synthesis of the siderophore amphibactin (Denaro, Crisafi et al. 2014, Kem, Zane et al. 2014).

The lipid analysis of the two PPTase deletion mutation strains, *ΔgglsC* and *ΔentDlike* showed that the glycine-glucolipid production in hexadecane-grown *ΔgglsC* and *ΔentDlike* was decreased strongly or moderately, respectively. GglsC is the preferred PPTase of GglsA, but EntDlike can also activate GglsA. *E. coli* co-expression of the two PPTase with GglsA and GglsB shows the same result. Besides, the type II PPTase EntD from *E. coli* is also able to phosphopantetheinylate GglsA with lower efficiency, as shown (Figure 27), the expression of GglsA and GglsB is sufficient for glycine-glucolipid production. This finding is in agreement with the scenario that the type II

PPTases has boarder substrate acceptance (Beld, Sonnenschein et al. 2014). Notably, *A. borkumensis* only harbors two type-II PPTases, i.e. GglsC and EntDlike, according to the UniProt blast result (Schneiker, Dos Santos et al. 2006). Therefore, it is impossible to delete both PPTases in *A. borkumensis*. Either GglsC or EntDlike are not only required for the activation of NRPSs in secondary metabolism, but also for the production of *holo*-ACP involved in fatty acid *de novo* synthesis in the primary metabolism (Schneiker, Dos Santos et al. 2006). Similarly, the genomes of different bacteria (*Pseudomonas, Synechocystis*) lack a type-I holo-ACP synthase gene, but contain a type-II PPTase with broad substrate specificity which is essential in these bacteria (Mootz, Finking et al. 2001).

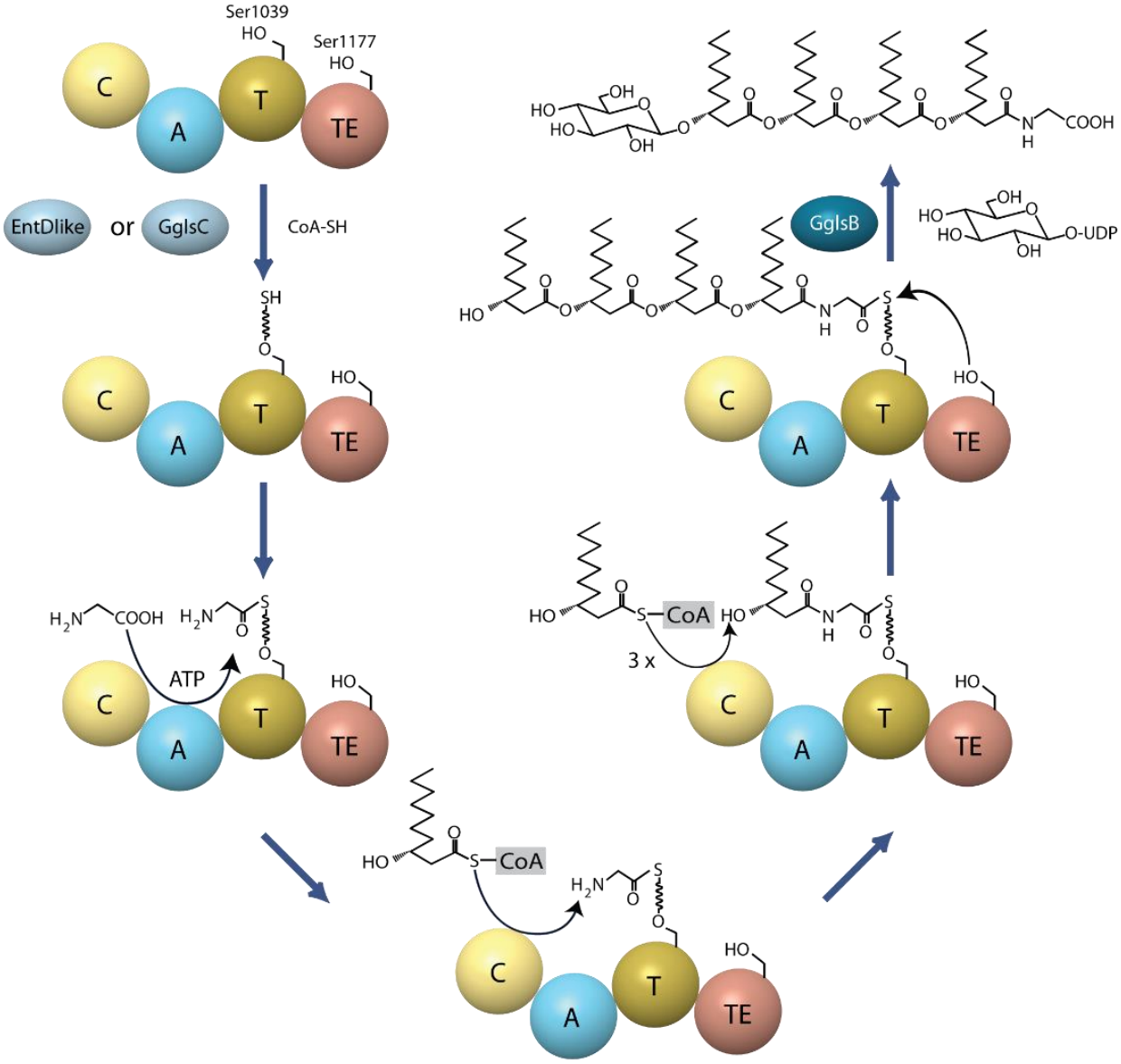

Figure 47 **Proposed pathway of glycine-glucolipid synthesis in *A. borkumensis*.**

The NRPS protein GglsA contains four-domains (C: condensation; A: adenylation; T: thiolation; TE: Thioesterase). A phosphopantetheine group (curled line) is ligated to Ser1039 of the T domain by the phosphopantetheinyl transferases GglsC or EntDlike (ABO_2807). The A domain catalyzes the transfer of a glycine to the T domain. The C domain transfers a D-3-hydroxydecanoic acid from coenzyme A (CoA) to the amino group of glycine. Three additional D-3-hydroxydecanoic moieties are ligated to the hydroxy groups of the D-3-hydroxydecanoic acids, catalyzed by the C domain. The TE domain releases the aglycone $H(O\text{-}10:0)_4Gly$. Finally, a glucose moiety is added onto the terminal hydroxy group of the aglycone by the glycosyltransferase GglsB.

4.3. An Artificial Tailor-Made Glucolipid can be Produced after Editing the GglsA NRPS

In general, the C domains of NRPS proteins catalyze the transfer of a single amino acid, fatty acid or peptidyl chain onto the amino acid bound at the T domain. During lipopeptide synthesis, however, an initiation or starter C domain attaches a lipid moiety (often a fatty acid) to the N-terminus of the N-terminal amino acid. The C domains of NRPS proteins can be organized into different phylogenetic clades reflecting the substrate specificity based on sequence similarity (Etchegaray, Silva-Stenico et al. 2004, Rausch, Hoof et al. 2007). The C domains of *A. borkumensis* GglsA and 63 other NRPS proteins were aligned and a phylogenetic tree was generated (Figure 48). In accordance with previous results, different clades were observed with C domain sequences specific for L-peptidyl donors, D-peptidyl donors or lipid initiation reactions.

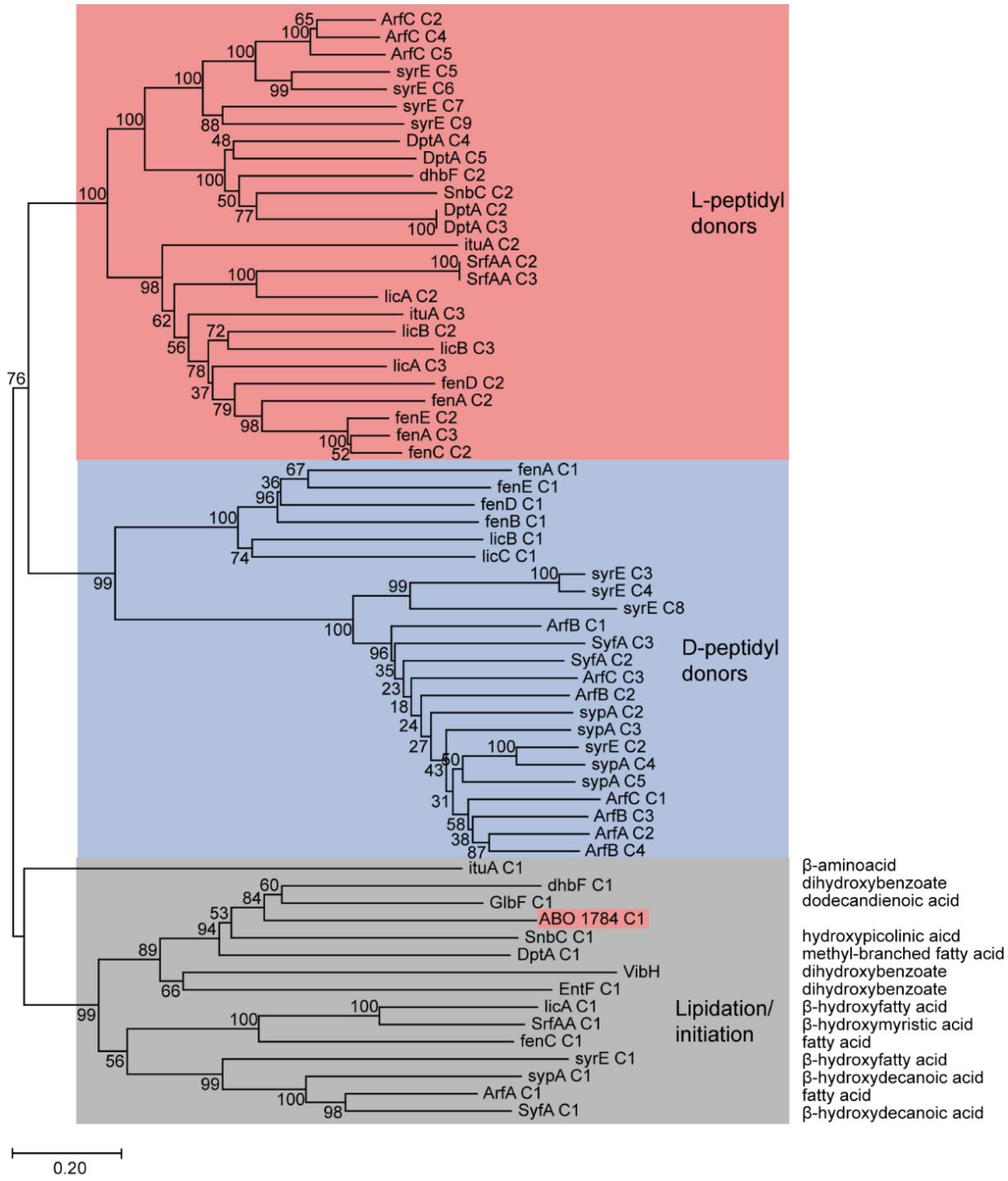

Figure 48 **Phylogenetic tree of C domains of different NRPS proteins**

Sequences were retrieved from Genbank and the domain structures predicted with the PKS/NRPS Analysis Web-site. The amino acid sequences were aligned using ClustalW and a phylogenetic tree generated with the Neighbor Joining method employing the Mega 11 software.

The C domain of GglsA is found in the clade of lipid initiation enzymes, revealing high similarity with C domains of (i) GlbF (involved in the transfer of 2(E),4(E)-dodecadienoic acid during glidobactin A synthesis); (ii) SrfAA (transfer of 3-hydroxymyristic acid during surfactin synthesis); (iii) EntF (2,3-dihydroxybenzoate transfer during enterobactin synthesis) ; (iv) fenC (fatty acid transfer during fengycin biosynthesis); and (v) licA (transfers a β-hydroxyfatty acid during lichenysin synthesis) (Gehring, Mori et al. 1998, Konz, Doekel et al. 1999, Lin, Chen et al. 1999, Imker, Krahn et al. 2010, Kraas, Helmetag et al. 2010). The sequence similarity of the C domain of GglsA with other initiation C domains is in agreement with the scenario that it transfers fatty acids to the glycine at the acceptor site of the T domain.

According to the sequence similarity, we tried to produce tailor-made glycine-glucolipid with different fatty acids instead of 3-hydroxydecanoic acid by swapping the C domain of GglsA with C1 domain of SrfAA and C domain of EntF. But both hybrid enzymes were inactive. The C domain is V-shaped, with the active site located in a tunnel that has one side for the acid donor binding site and the other side for the amino acid receptor binding site (Bloudoff and Schmeing 2017). It is possible that the tunnel of the C domain is not only specific for the acid donor but also selective for the amino acid receptor. Consequently, the SrfAA-C1 and EntF C domains might not accomodate glycine into the tunnel properly, preventing the enzyme from catalyzing the condensation reaction.

In NRPS proteins, the A domain is responsible for the activation of a specific amino acid or carboxylic acid. The specificity for the amino acid can be predicted from residues in the binding pocket of the A domain (Stachelhaus, Mootz et al. 1999). A set of eight amino acids in the A domain of GglsA is predicted to establish the binding pocket, DILQVGLI (PKS/NRPS Analysis Web-site, http://nrps.igs.umaryland.edu/). These eight amino acid residues show an exact match with A domains in four characterized NRPS proteins which are all specific for glycine: (i) the nostopeptolide A gene cluster from *Nostoc sp.* GSV224; (ii) the A1 domain of the dhbF gene involved in the synthesis of the siderophore 2,3-dihydroxybenzaote-glycine from *Bacillus subtilis*; (iii) the cdaPS2 gene (module 8) involved in the synthesis of the calcium-dependent antibiotic CDA from *Streptomyces coelicolor*; and (iv) the safA gene encoding saframycin Mx1 synthetase A from *Myxococcus xanthus* (Pospiech, Bietenhader et al. 1996, May, Wendrich et al. 2001, Hojati, Milne et al. 2002, Hoffmann, Hevel et al. 2003).

According to the sequence similarity, we tried to produce a tailor-made glycine-glucolipid with a different amino acid instead of glycine by swapping the A domain of GglsA with A domain of EntF. The expression of this hybrid enzyme, GglsA-EntFA, resulted in the production of serine-containing aglycone, i.e. it is possible to produce serine-glucolipid by co-expression of the GglsA-EntFA together with GglsB. But the yield in the *E. coli* cells is around 1000 times lower than the WT aglycone. The reason for the low catalytic efficiency could be that the serine does not fit into the tunnel of the C domain properly.

Overall, these results demonstrate that, while domain swapping can offer insights into enzyme functionality and potential for producing new compounds, the specificity of the C and A domains is crucial for the activity of the NRPS enzymes. The failure of activity of the hybrid enzymes with altered C domains suggests that both donor and acceptor specificities are essential for proper enzyme function.

4.4. The Production of the Glycine-Glucolipid in *A. borkumensis* is Dependent on Carbon Storage Capacity

The amount of glycine-glucolipid increases when *A. borkumensis* cells are grown with hexadecane as a sole carbon source (Figure 17). *A. borkumensis* cells primarily store excess carbon in the form of triacylglycerol or wax ester (WE) and, to a lesser extent, polyhydroxyalkanoates (PHA) (Figure 49) (Sabirova, Ferrer et al. 2006, Kalscheuer, Stöveken et al. 2007). When grown on hexadecane, all acyltransferase deletion mutant cells showed compromised glycine-glucolipid production compared to WT cells. The substrate of glycine-glucolipid synthesis, D-3-hydroxydecanoyl-CoA, is produced by fatty acid *de novo* synthesis (Ochsner, Koch et al. 1994, Rahim, Ochsner et al. 2001) rather than through breakdown of hexadecane or storage lipids by β-oxidation, which produces L-3-hydroxyldecanoyl-CoA. Unlike rhamnolipid synthesis in *P. aeruginosa*, glycine-glucolipid synthesis in *A. borkumensis* does not compete with PHA synthesis (Ochsner, Koch et al. 1994). The synthesis of the glycine-glucolipid partially depends on the cells' metabolic capacity to store carbon from hexadecane in the form of lipids or PHA. The cells refrain from upregulating glycine-glucolipid production when they cannot metabolize or store the carbon effectively.

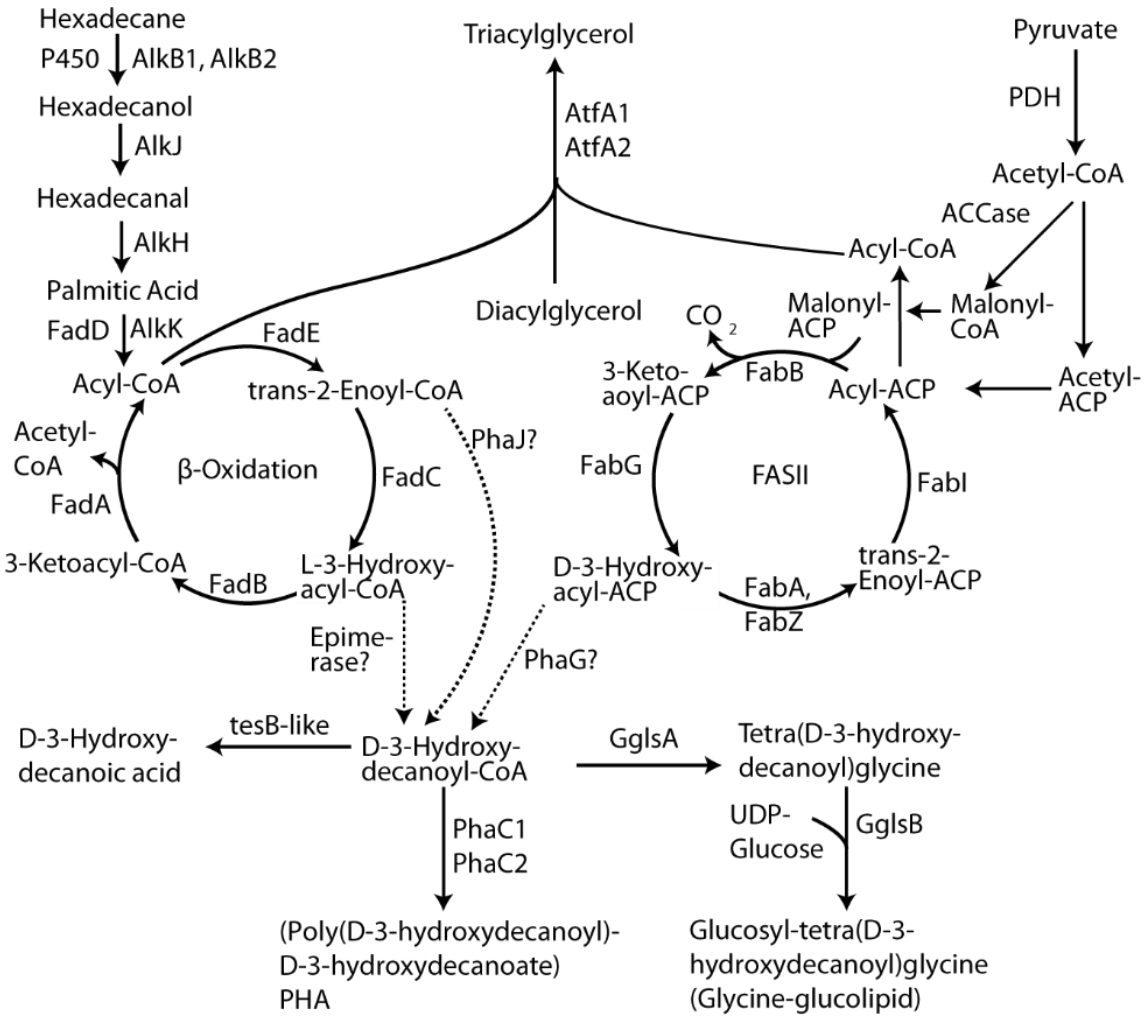

Figure 49 **Biochemical pathways of acyl lipid metabolism in *A. borkumensis* based on (Schneiker, Dos Santos et al. 2006)**

Hexadecane is incorporated and converted into palmitoyl-CoA. Fatty acids can be broken down via β-oxidation with 3-L-hydroxyacyl-CoAs as intermediates, while *de novo* fatty acid synthesis involves D-3-hydroxyacyl-ACP intermediates. Triacylglycerol is produced from diacylglycerol and acyl-CoA, catalyzed by AtfA1 and AtfA2. PHA is synthesized from 3-D-hydroxyacyl-CoA by PhaC1 and PhaC2. The glycine-glucolipid is also derived from 3-D-hydroxydecanoyl-CoA. Reactions and genes that have not been identified in *A. borkumensis* are indicated with question marks and dashed arrows. ACCase, acetyl-CoA carboxylase; AlkB1, ABO_2707, AlkB2, ABO_0122, alkane 1-monooxidase; AlkH, ABO_2709, aldehyde dehydrogenase; AlkJ, ABO_2710, alcohol dehydrogenase; AlkK, ABO_2748, acyl-CoA synthetase; AtfA1, ABO_2742, AtfA2, ABO_1804, diacylglycerol acyltransferase; FabA, ABO_0835, D-3-hydroxyacyl-ACP dehydratase; FabB, ABO_0647, 3-ketoacyl-ACP synthase; FabG, ABO_1069, 3-ketoacyl-ACP reductase; FabI, ABO_1215, trans-3-enoyl-ACP reductase; FabZ, ABO_1154, D-3-hydroxyacyl-ACP dehydratase; FadA, ABO_1653, 3-ketoacyl-CoA thiolase; FadB, ABO_1566, 3-L-hydroxyacyl-CoA dehydrogenase; FadC, ABO_1645, trans-2-enoyl-CoA hydratase; FadD, ABO_0367, acyl-CoA synthetase; FadE, ABO_1121, acyl-CoA dehydrogenase; GglsA, ABO_1784, nonribosomal peptide synthase involved in glycine-glucolipid synthesis; GglsB, ABO_1783, UDP-glucose:tetra(D-3-hydroxydecanoyl)glycine glucosyltransferase; PDH, pyruvate dehydrogenase; PhaC1, ABO_2214, PhaC2, ABO_1418, PHA synthase; PhaG, acyl-ACP:CoA transacylase; PhaJ, trans-2-enoyl-hydratase (D-hydroxyacyl-CoA specific); tesB-like, ABO_1111, acyl-CoA thioesterase.

4.5. The Application of Glycine-Glucolipid as Biosurfactant during Foliar Fertilization is Limited by the Production Capacity

The expression of the key enzyme GglsA leads to a very poor growth in glycine-glucolipid-producing *E. coli* strain. It is possible that GglsA is competitive to the enzymes involved in primary metabolism in *E. coli*, because it might remove D-3-hydroxydecanoyl-ACP units from the fatty acid de novo synthesis (Figure 49). Although, the glycine-glucolipid production per 10^{10} cells of *E. coli* is higher than *A. borkumensis*, the low cell density results in a yield of only ~65 nmol/L culture in in *E. coli* compared to ~8 µmol/L culture in *A. borkumensis* when using pyruvate. Unfortunately, there is no measure right now to circumvent the growth inhibition of *E. coli* caused by GglsA expression, making *A. borkumensis* the best strain for glycine-glucolipid production. However, *A. borkumensis* relies on expensive and unsustainable carbon sources. According to Tobias Karmainski et al, pyruvate in the *A. borkumensis* medium can be replaced by acetate, which is a sustainable carbon source that can be produced from lignocellulose, syngas or CO_2 and methyl formate (Karmainski, Dielentheis-Frenken et al. 2023). Alternatively, instead of using artificial carbon sources, co-cultivation with hydrocarbon-producing cyanobacteria such as *Prochlorococcus* or *Synechococcus* could offer a more sustainable approach for cultivating *A. borkumensis* (Lea-Smith, Biller et al. 2015). The hydrocarbons produced by cyanobacteria originate from photosynthesis, utilizing CO_2 and sunlight (McGenity, McKew et al. 2021). Furthermore, the insoluble carbon produced could stimulate the production of glycine-glucolipid in *A. borkumensis* as shown in chapter 3.3. The experiment in chapter 3.14 shows the importance of surfactants in enhancing nutrient uptake through foliar fertilization. The effectiveness of glycine-glucolipid is particularly noteworthy, as it shows potential as a bio-based surfactant alternative to chemical surfactants like Triton X-100, whose degradation product is ecotoxic as it possesses hormone-like (estrogeno-mimetic) activity that may act on wildlife (White, Jobling et al. 1994). This aligns with the ongoing search for sustainable and environmentally friendly agricultural practices. However, the high cost of glycine-glucolipid currently limits its potential application for agricultural use.

5. Summary

This study investigated the production, biosynthesis pathway, and physiological function of the glycine-glucolipid produced by the marine bacterium *Alcanivorax borkumensis*. An absolute quantification method using HPLC was developed, which involves converting the glycine-glucolipid to phenacyl esters, followed by measurement with HPLC with UV detection. Different molecular species were separated by HPLC and identified via mass spectrometry (MS) as glucosyl-tetra(3-hydroxy-acyl)-glycine, containing varying numbers of 3-hydroxy-decanoic acid or 3-hydroxy-octanoic acid groups. HPLC and MS results show that the glycine-glucolipid accumulates in cells, it is not secreted into the medium, and its production is correlated with the alkane degradation process, rather than a response to phosphate deprivation.

The deletion mutations in *A. borkumensis* and expression studies in *E. coli* demonstrate that the glycine-glucolipid is produced in *A. borkumensis* via a non-ribosomal peptide synthetase (NRPS)-dependent pathway involving the enzymes ABO_1784 (GglsA, NRPS), ABO_1783 (GglsB, glycosyltransferase), and ABO_1782 (GglsC, phosphopantetheinyl transferase, PPTase). GglsC activates GglsA through phosphopantetheinylation on the Ser1039 of the thiolation (T) domain. The adenylation (A) domain of GglsA activates and transfers a glycine onto the phosphopantetheine arm tethered to the T domain, consuming ATP in this process. The condensation (C) domain of GglsA catalyzes the condensation reaction, transferring a 3-D-hydroxydecanoyl moiety from 3-D-hydroxydecanoyl-CoA, forming an amide bond between glycine's amino group and the acyl group of the 3-D-hydroxydecanoyl moiety. GglsA subsequently attaches three additional 3-D-hydroxydecanoyl moieties onto the hydroxy groups of the intermediate on the T domain. Finally, the T domain shuttles the aglycone from the C/T surface to the T/TE globular core, where the aglycone is hydrolyzed from GglsA by the thioesterase (TE) domain at catalytic site Ser1177. The aglycone released from GglsA is then glycosylated by GglsB using UDP-glucose as a substrate.

Physiological experiments, including the bacterial adhesion to hydrocarbons (BATH) assay, hexadecane emulsification, and observation via confocal microscopy, were performed with *ΔgglsB* and *ΔgglsA* mutants grown in pyruvate or hexadecane-containing medium. The hexadecane-grown mutant cells lacking the glycine-glucolipid and pyruvate-grown cells showed reduced adhesion to hexadecane and poorer emulsification efficiency compared to hexadecane-grown wild-type cells. The absence of the glycine-glucolipid not only impacted the attachment to

hexadecane, but also reduced growth on hexadecane medium, decreased storage lipid accumulation, and resulted in smaller cell size.

Expression of GglsA in *E. coli* severely limits growth and glycine-glucolipid production. Thus, *A. borkumensis* is required for glycine-glucolipid production. However, *A. borkumensis* can only be grown with expensive carbon sources. Alternatively, co-cultivation with hydrocarbon-producing cyanobacteria could offer a more sustainable approach for cultivating *A. borkumensis*. These cyanobacteria produce hydrocarbons through photosynthesis using CO_2 and sunlight, and the insoluble carbon produced could stimulate glycine-glucolipid production in *A. borkumensis*. The glycine-glucolipid can potentially be employed as a biosurfactant in agriculture, replacing ecotoxic chemical surfactants like Triton X-100. However, its high cost currently limits widespread agricultural application despite its environmental benefits.

6. Reference

Abbasi, A., G. D. Bothun and A. Bose (2018). "Attachment of *Alcanivorax borkumensis* to hexadecane-in-artificial sea water emulsion droplets." Langmuir **34**(18): 5352-5357.

Abraham, W.-R., H. Meyer and M. Yakimov (1998). "Novel glycine containing glucolipids from the alkane using bacterium Alcanivorax borkumensis." Biochimica et Biophysica Acta **1393**(1): 57-62.

Alaa A. Al-Seraih, W. A. S., Murtakab Y. Al-hejjaj, Fatima H. Al-Laibai, and A. K. Ghadban (2022). "Isolation and Partial Characterization of Glycolipopeptide Biosurfactant Derived from A Novel *Lactiplantibacillus plantarum* Lbp_WAM." Basrah Journal of Agricultural Sciences **35**(2): 78-98.

Alonzo, D. A., N. A. Magarvey and T. M. Schmeing (2015). "Characterization of cereulide synthetase, a toxin-producing macromolecular machine." PLOS ONE **10**(6): e0128569.

Alvarez, H., O. Pucci and A. Steinbüchel (1997). "Lipid storage compounds in marine bacteria." Applied Microbiology and Biotechnology **47**: 132-139.

Alvarez, H. and A. Steinbüchel (2002). "Triacylglycerols in prokaryotic microorganisms." Applied Microbiology and Biotechnology **60**: 367-376.

Banat, I. M., R. S. Makkar and S. S. Cameotra (2000). "Potential commercial applications of microbial surfactants." Applied Microbiology and Biotechnology **53**: 495-508.

Behrens, B., M. Baune, J. Jungkeit, T. Tiso, L. M. Blank and H. Hayen (2016). "High performance liquid chromatography-charged aerosol detection applying an inverse gradient for quantification of rhamnolipid biosurfactants." Journal of Chromatography A **1455**: 125-132.

Beld, J., E. C. Sonnenschein, C. R. Vickery, J. P. Noel and M. D. Burkart (2014). "The phosphopantetheinyl transferases: catalysis of a post-translational modification crucial for life." Natural Product Reports **31**(1): 61-108.

Belecki, K. and C. A. Townsend (2013). "Biochemical determination of enzyme-bound metabolites: preferential accumulation of a programmed octaketide on the enediyne polyketide synthase CalE8." Journal of the American Chemical Society **135**(38): 14339-14348.

Black, P. N. and C. C. DiRusso (2003). "Transmembrane movement of exogenous long-chain fatty acids: proteins, enzymes, and vectorial esterification." Microbiology and Molecular Biology Reviews **67**(3): 454-472.

Bloudoff, K. and T. M. Schmeing (2017). "Structural and functional aspects of the nonribosomal peptide synthetase condensation domain superfamily: discovery, dissection and diversity." Biochimica et Biophysica Acta **1865**(11, Part B): 1587-1604.

Bourne, Y. and B. Henrissat (2001). "Glycoside hydrolases and glycosyltransferases: families and functional modules." Current Opinion in Structural Biology **11**(5): 593-600.

Bradford, M. M. (1976). "A rapid and sensitive method for the quantitation of microgram quantities

of protein utilizing the principle of protein-dye binding." Analytical Biochemistry **72**(1-2): 248-254.

Burger, M. M., L. Glaser and R. M. Burton (1963). "The enzymatic synthesis of a rhamnose-containing glycolipid by extracts of *Pseudomonas aeruginosa*." Journal of Biological Chemistry **238**(8): 2595-2602.

Campbell, R. E., S. C. Mosimann, M. E. Tanner and N. C. Strynadka (2000). "The structure of UDP-N-acetylglucosamine 2-epimerase reveals homology to phosphoglycosyl transferases." Biochemistry **39**(49): 14993-15001.

Chakrabarty, P., A. J. Sheehy, X. Clute, S. B. Cruz and B. Ballengée (2024). "Ten years later: An update on the status of collections of endemic Gulf of Mexico fishes put at risk by the 2010 Oil Spill." Biodiversity Data Journal **12**.

Charnock, S. J. and G. J. Davies (1999). "Structure of the nucleotide-diphospho-sugar transferase, SpsA from *Bacillus subtilis*, in native and nucleotide-complexed forms." Biochemistry **38**(20): 6380-6385.

Conti, E., T. Stachelhaus, M. A. Marahiel and P. Brick (1997). "Structural basis for the activation of phenylalanine in the non-ribosomal biosynthesis of gramicidin S." The EMBO Journal **16**(14): 4174-4183.

Cronan Jr, J. E. and A. L. Klages (1981). "Chemical synthesis of acyl thioesters of acyl carrier protein with native structure." Proceedings of the National Academy of Sciences of the United States of America **78**(9): 5440-5444.

Crosby, J. and M. P. Crump (2012). "The structural role of the carrier protein – active controller or passive carrier." Natural Product Reports **29**(10): 1111-1137.

Das, P., S. Mukherjee and R. Sen (2009). "Biosurfactant of marine origin exhibiting heavy metal remediation properties." Bioresource Technology **100**(20): 4887-4890.

Dekimpe, S. and J. Masschelein (2021). "Beyond peptide bond formation: the versatile role of condensation domains in natural product biosynthesis." Natural Product Reports **38**(10): 1910-1937.

Denaro, R., F. Crisafi, D. Russo, M. Genovese, E. Messina, L. Genovese, M. Carbone, M. Ciavatta, M. Ferrer and P. Golyshin (2014). "*Alcanivorax borkumensis* produces an extracellular siderophore in iron-limitation condition maintaining the hydrocarbon-degradation efficiency." Marine Genomics **17**: 43-52.

Dominique P. Frueh, H. A., Alexander Koglin, David A. Vosburg, Andrew E. Bennett, Christopher T. Walsh , Gerhard Wagner (2008). "Dynamic thiolation–thioesterase structure of a non-ribosomal peptide synthetase." Nature **454**(7206): 903-906.

Drake, E. J., B. R. Miller, C. Shi, J. T. Tarrasch, J. A. Sundlov, C. Leigh Allen, G. Skiniotis, C. C. Aldrich and A. M. Gulick (2016). "Structures of two distinct conformations of holo-non-ribosomal peptide synthetases." Nature **529**(7585): 235-238.

Durst, H., M. Milano, E. Kikta, S. Connelly and E. Grushka (1975). "Phenacyl esters of fatty acids via crown ether catalysts for enhanced ultraviolet detection in liquid chromatography." Analytical chemistry **47**(11): 1797-1801.

Etchegaray, A., M. E. Silva-Stenico, D. H. Moon and S. M. Tsai (2004). "In silico analysis of nonribosomal peptide synthetases of *Xanthomonas axonopodis* pv. citri: identification of putative siderophore and lipopeptide biosynthetic genes." Microbiological Research **159**(4): 425-437.

Finking, R. and M. A. Marahiel (2004). "Biosynthesis of Nonribosomal Peptides1." Annual Review of Microbiology **58**(Volume 58, 2004): 453-488.

Frueh, D. P., H. Arthanari, A. Koglin, D. A. Vosburg, A. E. Bennett, C. T. Walsh and G. Wagner (2008). "Dynamic thiolation–thioesterase structure of a non-ribosomal peptide synthetase." Nature **454**(7206): 903-906.

Gaitatzis, N., B. Kunze and R. Müller (2001). "*In vitro* reconstitution of the myxochelin biosynthetic machinery of *Stigmatella aurantiaca* Sg a15: Biochemical characterization of a reductive release mechanism from nonribosomal peptide synthetases." Proceedings of the National Academy of Sciences of the United States of America **98**(20): 11136-11141.

Gehring, A. M., I. Mori and C. T. Walsh (1998). "Reconstitution and characterization of the *Escherichia coli* enterobactin synthetase from EntB, EntE, and EntF." Biochemistry **37**(8): 2648-2659.

Godfrin, M. P., M. Sihlabela, A. Bose and A. Tripathi (2018). "Behavior of marine bacteria in clean environment and oil spill conditions." Langmuir **34**(30): 9047-9053.

Hobson, C., M. Jenner, X. Jian, D. Griffiths, D. M. Roberts, M. Rey-Carrizo and G. L. Challis (2022). "Diene incorporation by a dehydratase domain variant in modular polyketide synthases." Nature Chemical Biology **18**(12): 1410-1416.

Hoffmann, D., J. M. Hevel, R. E. Moore and B. S. Moore (2003). "Sequence analysis and biochemical characterization of the nostopeptolide A biosynthetic gene cluster from *Nostoc sp.* GSV224." Gene **311**: 171-180.

Hoffmeister, D., B. Wilkinson, G. Foster, P. J. Sidebottom, K. Ichinose and A. Bechthold (2002). "Engineered urdamycin glycosyltransferases are broadened and altered in substrate specificity." Chemistry & Biology **9**(3): 287-295.

Hojati, Z., C. Milne, B. Harvey, L. Gordon, M. Borg, F. Flett, B. Wilkinson, P. J. Sidebottom, B. A. Rudd and M. A. Hayes (2002). "Structure, biosynthetic origin, and engineered biosynthesis of calcium-dependent antibiotics from *Streptomyces coelicolor*." Chemistry & Biology **9**(11): 1175-1187.

Hölzl, G. and P. Dörmann (2007). "Structure and function of glycoglycerolipids in plants and bacteria." Progress in Lipid Research **46**(5): 225-243.

Hölzl, G. and P. Dörmann (2021). Thin-layer chromatography. Plant Lipids: Methods and Protocols, Springer**:** 29-41.

Hoyer, K. M., C. Mahlert and M. A. Marahiel (2007). "The iterative gramicidin S thioesterase catalyzes peptide ligation and cyclization." Chemistry & Biology **14**(1): 13-22.

Hu, Y. and S. Walker (2002). "Remarkable structural similarities between diverse glycosyltransferases." Chemistry & Biology **9**(12): 1287-1296.

Ian M. Head, D. M. J., Wilfred F. M. Röling (2006). "Marine microorganisms make a meal of oil." Nature Reviews Microbiology **4**(3): 173-182.

Imker, H. J., D. Krahn, J. Clerc, M. Kaiser and C. T. Walsh (2010). "N-acylation during glidobactin biosynthesis by the tridomain nonribosomal peptide synthetase module GlbF." Chemistry &Biology **17**(10): 1077-1083.

Jung, J., G. Bashiri, J. M. Johnston and E. N. Baker (2016). "Mass spectral determination of phosphopantetheinylation specificity for carrier proteins in *Mycobacterium tuberculosis*." FEBS Open Bio **6**(12): 1220-1226.

Kalscheuer, R. and A. Steinbüchel (2003). "A novel bifunctional wax ester synthase/acyl-CoA: diacylglycerol acyltransferase mediates wax ester and triacylglycerol biosynthesis in *Acinetobacter calcoaceticus* ADP1." Journal of Biological Chemistry **278**(10): 8075-8082.

Kalscheuer, R., T. Stöveken, U. Malkus, R. Reichelt, P. N. Golyshin, J. S. Sabirova, M. Ferrer, K. N. Timmis and A. Steinbüchel (2007). "Analysis of storage lipid accumulation in *Alcanivorax borkumensis*: evidence for alternative triacylglycerol biosynthesis routes in bacteria." Journal of Bacteriology **189**(3): 918-928.

Karmainski, T., M. R. Dielentheis-Frenken, M. K. Lipa, A. N. Phan, L. M. Blank and T. Tiso (2023). "High-quality physiology of *Alcanivorax borkumensis* SK2 producing glycolipids enables efficient stirred-tank bioreactor cultivation." Frontiers in Bioengineering and Biotechnology **11**.

Keating, T. A., C. G. Marshall, C. T. Walsh and A. E. Keating (2002). "The structure of VibH represents nonribosomal peptide synthetase condensation, cyclization and epimerization domains." Nature Structural Biology **9**(7): 522-526.

Kem, M. P., H. K. Zane, S. D. Springer, J. M. Gauglitz and A. Butler (2014). "Amphiphilic siderophore production by oil-associating microbes." Metallomics **6**(6): 1150-1155.

Kishimoto, N., K. Adachi, S. Tamura, M. Nishihara, K. Inagaki, T. Sugio and T. Tano (1993). "Lipoamino acids isolated from *Acidiphilium organovorum*." Systematic and Applied Microbiology **16**(1): 17-21.

Klein, B., V. Grossi, P. Bouriat, P. Goulas and R. Grimaud (2008). "Cytoplasmic wax ester accumulation during biofilm-driven substrate assimilation at the alkane–water interface by *Marinobacter hydrocarbonoclasticus* SP17." Research in Microbiology **159**(2): 137-144.

Konz, D., S. Doekel and M. A. Marahiel (1999). "Molecular and biochemical characterization of the protein template controlling biosynthesis of the lipopeptide lichenysin." Journal of Bacteriology **181**(1): 133-140.

Konz, D. and M. A. Marahiel (1999). "How do peptide synthetases generate structural diversity?" Chemistry & Biology **6**(2): R39-R48.

Kraas, F. I., V. Helmetag, M. Wittmann, M. Strieker and M. A. Marahiel (2010). "Functional dissection of surfactin synthetase initiation module reveals insights into the mechanism of lipoinitiation." Chemistry & Biology **17**(8): 872-880.

Kubicki, S., I. Bator, S. Jankowski, K. Schipper, T. Tiso, M. Feldbrügge, L. M. Blank, S. Thies and K.-E. Jaeger (2020). "A straightforward assay for screening and quantification of biosurfactants in microbial culture supernatants." Frontiers in Bioengineering and Biotechnology **8**: 958.

Kubicki, S., A. Bollinger, N. Katzke, K.-E. Jaeger, A. Loeschcke and S. Thies (2019). "Marine biosurfactants: biosynthesis, structural diversity and biotechnological applications." Marine Drugs **17**(7): 408.

Laemmli, U. K. (1970). "Cleavage of structural proteins during the assembly of the head of bacteriophage T4." Nature **227**(5259): 680-685.

Lairson, L. L., B. Henrissat, G. J. Davies and S. G. Withers (2008). "Glycosyltransferases: Structures, Functions, and Mechanisms." Annual Review of Biochemistry **77**(Volume 77, 2008): 521-555.

Lambalot, R. H., A. M. Gehring, R. S. Flugel, P. Zuber, M. LaCelle, M. A. Marahiel, R. Reid, C. Khosla and C. T. Walsh (1996). "A new enzyme superfamily — the phosphopantetheinyl transferases." Chemistry & Biology **3**(11): 923-936.

Lea-Smith, D. J., S. J. Biller, M. P. Davey, C. A. Cotton, B. M. Perez Sepulveda, A. V. Turchyn, D. J. Scanlan, A. G. Smith, S. W. Chisholm and C. J. Howe (2015). "Contribution of cyanobacterial alkane production to the ocean hydrocarbon cycle." Proceedings of the National Academy of Sciences of the United States of America **112**(44): 13591-13596.

Lin, T.-P., C.-L. Chen, L.-K. Chang, J. S.-M. Tschen and S.-T. Liu (1999). "Functional and transcriptional analyses of a fengycin synthetase gene, fenC, from *Bacillus subtilis*." Journal of Bacteriology **181**(16): 5060-5067.

Liu, Y., G. Zeng, H. Zhong, Z. Wang, Z. Liu, M. Cheng, G. Liu, X. Yang and S. Liu (2017). "Effect of rhamnolipid solubilization on hexadecane bioavailability: enhancement or reduction?" Journal of Hazardous Materials **322**: 394-401.

Lombard, V., H. Golaconda Ramulu, E. Drula, P. M. Coutinho and B. Henrissat (2014). "The carbohydrate-active enzymes database (CAZy) in 2013." Nucleic Acids Research **42**(D1): D490-D495.

Love, C. R., E. C. Arrington, K. M. Gosselin, C. M. Reddy, B. A. Van Mooy, R. K. Nelson and D. L. Valentine (2021). "Microbial production and consumption of hydrocarbons in the global ocean." Nature Microbiology **6**(4): 489-498.

Luzhetskyy, A., C. Mendez, J. A. Salas and A. Bechthold (2008). "Glycosyltransferases, important tools for drug design." Current Topics in Medicinal Chemistry **8**(8): 680-709.

Manilla-Pérez, E., A. B. Lange, H. Luftmann, H. Robenek and A. Steinbüchel (2011). "Neutral lipid production in *Alcanivorax borkumensis* SK2 and other marine hydrocarbonoclastic bacteria." European Journal of Lipid Science and Technology **113**(1): 8-17.

Marahiel, M. A., T. Stachelhaus and H. D. Mootz (1997). "Modular Peptide Synthetases Involved in Nonribosomal Peptide Synthesis." Chemical Reviews **97**(7): 2651-2674.

Marchant, R. and I. M. Banat (2017). "Protocols for measuring biosurfactant production in microbial

cultures." Hydrocarbon and Lipid Microbiology Protocols: Activities and Phenotypes: 119-128.

Mata-Sandoval, J. C., J. Karns and A. Torrents (1999). "High-performance liquid chromatography method for the characterization of rhamnolipid mixtures produced by *Pseudomonas aeruginosa* UG2 on corn oil." Journal of Chromatography A **864**(2): 211-220.

May, J. J., T. M. Wendrich and M. A. Marahiel (2001). "The *dhb* operon of *Bacillus subtilis* encodes the biosynthetic template for the catecholic siderophore 2, 3-dihydroxybenzoate-glycine-threonine trimeric ester bacillibactin." Journal of Biological Chemistry **276**(10): 7209-7217.

McGenity, T. J., B. A. McKew and D. J. Lea-Smith (2021). "Cryptic microbial hydrocarbon cycling." Nature Microbiology **6**(4): 419-420.

Miller, B. R., E. J. Drake, C. Shi, C. C. Aldrich and A. M. Gulick (2016). "Structures of a Nonribosomal Peptide Synthetase Module Bound to MbtH-like Proteins Support a Highly Dynamic Domain Architecture." Journal of Biological Chemistry **291**(43): 22559-22571.

Minnikin, D., H. Abdolrahimzadeh and J. Baddiley (1974). "Replacement of acidic phospholipids by acidic glycolipids in *Pseudomonas diminuta*." Nature **249**(5454): 268-269.

Mootz, H. D., R. Finking and M. A. Marahiel (2001). "4'-Phosphopantetheine transfer in primary and secondary metabolism of *Bacillus subtilis*." Journal of Biological Chemistry **276**(40): 37289-37298.

Mulichak, A. M., H. C. Losey, W. Lu, Z. Wawrzak, C. T. Walsh and R. M. Garavito (2003). "Structure of the TDP-*epi*-vancosaminyltransferase GtfA from the chloroeremomycin biosynthetic pathway." Proceedings of the National Academy of Sciences of the United States of America **100**(16): 9238-9243.

Naether, D. J., S. Slawtschew, S. Stasik, M. Engel, M. Olzog, L. Y. Wick, K. N. Timmis and H. J. Heipieper (2013). "Adaptation of the hydrocarbonoclastic bacterium *Alcanivorax borkumensis* SK2 to alkanes and toxic organic compounds: a physiological and transcriptomic approach." Applied and Environmental Microbiology **79**(14): 4282-4293.

Nguyen, T. T. and D. A. Sabatini (2011). "Characterization and emulsification properties of rhamnolipid and sophorolipid biosurfactants and their applications." International Journal of Molecular Sciences **12**(2): 1232-1244.

Nievas, M., M. Commendatore, J. Esteves and V. Bucalá (2008). "Biodegradation pattern of hydrocarbons from a fuel oil-type complex residue by an emulsifier-producing microbial consortium." Journal of Hazardous Materials **154**(1-3): 96-104.

Ochsner, U. A., A. K. Koch, A. Fiechter and J. Reiser (1994). "Isolation and characterization of a regulatory gene affecting rhamnolipid biosurfactant synthesis in *Pseudomonas aeruginosa*." Journal of Bacteriology **176**(7): 2044-2054.

Parris, K. D., L. Lin, A. Tam, R. Mathew, J. Hixon, M. Stahl, C. C. Fritz, J. Seehra and W. S. Somers (2000). "Crystal structures of substrate binding to *Bacillus subtilis holo*-(acyl carrier protein) synthase reveal a novel trimeric arrangement of molecules resulting in three active sites." Structure **8**(8): 883-895.

Passeri, A., M. Schmidt, T. Haffner, V. Wray, S. Lang and F. Wagner (1992). "Marine biosurfactants. IV.

Production, characterization and biosynthesis of an anionic glucose lipid from the marine bacterial strain MM1." Applied Microbiology and Biotechnology **37**: 281-286.

Pospiech, A., J. Bietenhader and T. Schupp (1996). "Two multifunctional peptide synthetases and an O-methyltransferase are involved in the biosynthesis of the DNA-binding antibiotic and antitumour agent saframycin Mx1 from *Myxococcus xanthus*." Microbiology **142**(4): 741-746.

Prasad, M., N. Obana, S.-Z. Lin, S. Zhao, K. Sakai, C. Blanch-Mercader, J. Prost, N. Nomura, J.-F. Rupprecht and J. Fattaccioli (2023). "*Alcanivorax borkumensis* biofilms enhance oil degradation by interfacial tubulation." Science **381**(6659): 748-753.

Rahim, R., U. A. Ochsner, C. Olvera, M. Graninger, P. Messner, J. S. Lam and G. Soberón-Chávez (2001). "Cloning and functional characterization of the *Pseudomonas aeruginosa* rhlC gene that encodes rhamnosyltransferase 2, an enzyme responsible for di-rhamnolipid biosynthesis." Molecular Microbiology **40**(3): 708-718.

Rausch, C., I. Hoof, T. Weber, W. Wohlleben and D. H. Huson (2007). "Phylogenetic analysis of condensation domains in NRPS sheds light on their functional evolution." BMC Evolutionary Biology **7**: 1-15.

Rehm, B. H., N. Kruger and A. Steinbuchel (1998). "A new metabolic link between fatty acid *de novo* synthesis and polyhydroxyalkanoic acid synthesis." Journal of Biological Chemistry **273**(37): 24044-24051.

Reimer, J. M., M. N. Aloise, P. M. Harrison and T. Martin Schmeing (2016). "Synthetic cycle of the initiation module of a formylating nonribosomal peptide synthetase." Nature **529**(7585): 239-242.

Röttig, A. and A. Steinbüchel (2013). "Acyltransferases in bacteria." Microbiology and Molecular Biology Reviews **77**(2): 277-321.

Sabirova, J. S., M. Ferrer, D. Regenhardt, K. N. Timmis and P. N. Golyshin (2006). "Proteomic insights into metabolic adaptations in *Alcanivorax borkumensis* induced by alkane utilization." Journal of Bacteriology **188**(11): 3763-3773.

Sachdev, D. P. and S. S. Cameotra (2013). "Biosurfactants in agriculture." Applied Microbiology and Biotechnology **97**: 1005-1016.

Schenk, T., I. Schuphan and B. Schmidt (1995). "High-performance liquid chromatographic determination of the rhamnolipids produced by *Pseudomonas aeruginosa*." Journal of Chromatography A **693**(1): 7-13.

Schneider, A. and M. A. Marahiel (1998). "Genetic evidence for a role of thioesterase domains, integrated in or associated with peptide synthetases, in non-ribosomal peptide biosynthesis in *Bacillus subtilis*." Archives of Microbiology **169**: 404-410.

Schneiker, S., V. A. Dos Santos, D. Bartels, T. Bekel, M. Brecht, J. Buhrmester, T. N. Chernikova, R. Denaro, M. Ferrer and C. Gertler (2006). "Genome sequence of the ubiquitous hydrocarbon-degrading marine bacterium *Alcanivorax borkumensis*." Nature Biotechnology **24**(8): 997-1004.

Schulz, D., A. Passeri, M. Schmidt, S. Lang, F. Wagner, V. Wray and W. Gunkel (1991). "Marine biosurfactants, I. Screening for biosurfactants among crude oil degrading marine microorganisms from the North Sea." Zeitschrift für Naturforschung C **46**(3-4): 197-203.

Sen, R. (2008). "Biotechnology in petroleum recovery: the microbial EOR." Progress in Energy and Combustion Science **34**(6): 714-724.

Shantha, N. and G. E. Napolitano (1992). "Gas chromatography of fatty acids." Journal of Chromatography A **624**(1-2): 37-51.

Soberón-Chávez, G. and R. M. Maier (2011). Biosurfactants: a general overview. Heidelberg, Springer Berlin.

Stachelhaus, T., H. D. Mootz and M. A. Marahiel (1999). "The specificity-conferring code of adenylation domains in nonribosomal peptide synthetases." Chemistry & Biology **6**(8): 493-505.

Stanišić, A. and H. Kries (2019). "Adenylation domains in nonribosomal peptide engineering." ChemBioChem **20**(11): 1347-1356.

Steinbüchel, A. (1991). Polyhydroxyalkanoic acids. Biomaterials: novel materials from biological sources, Springer**:** 123-213.

Tan, X.-F., Y.-N. Dai, K. Zhou, Y.-L. Jiang, Y.-M. Ren, Y. Chen and C.-Z. Zhou (2015). "Structure of the adenylation-peptidyl carrier protein didomain of the *Microcystis aeruginosa* microcystin synthetase McyG." Acta Crystallographica Section D **71**(4): 873-881.

Tarry, M. J., A. S. Haque, K. H. Bui and T. M. Schmeing (2017). "X-Ray Crystallography and Electron Microscopy of Cross- and Multi-Module Nonribosomal Peptide Synthetase Proteins Reveal a Flexible Architecture." Structure **25**(5): 783-793.e784.

Timmis, K. N., T. McGenity, J. R. Van Der Meer and V. de Lorenzo (2010). Handbook of hydrocarbon and lipid microbiology, Springer Berlin.

Tiso, T., S. Thies, M. Müller, L. Tsvetanova, L. Carraresi, S. Bröring, K.-E. Jaeger and L. M. Blank (2017). Rhamnolipids: production, performance, and application.

Ünligil, U. M. and J. M. Rini (2000). "Glycosyltransferase structure and mechanism." Current Opinion in Structural Biology **10**(5): 510-517.

Vatsa, P., L. Sanchez, C. Clement, F. Baillieul and S. Dorey (2010). "Rhamnolipid biosurfactants as new players in animal and plant defense against microbes." International Journal of Molecular Sciences **11**(12): 5095-5108.

Vecino, X., J. Cruz, A. Moldes and L. Rodrigues (2017). "Biosurfactants in cosmetic formulations: trends and challenges." Critical Reviews in Biotechnology **37**(7): 911-923.

Walsh, C. T. (2008). "The chemical versatility of natural-product assembly lines." Accounts of Chemical Research **41**(1): 4-10.

Walsh, C. T., R. V. O'Brien and C. Khosla (2013). "Nonproteinogenic amino acid building blocks for

nonribosomal peptide and hybrid polyketide ccaffolds." Angewandte Chemie International Edition **52**(28): 7098-7124.

White, R., S. Jobling, S. A. Hoare, J. P. Sumpter and M. G. Parker (1994). "Environmentally persistent alkylphenolic compounds are estrogenic." Endocrinology **135**(1): 175-182.

White, S. W., J. Zheng, Y.-M. Zhang and C. O. Rock (2005). "The structural biology of type II fatty acid biosynthesis." Annual Review of Biochemistry **74**: 791-831.

White, T., S. Bursten, D. Federighi, R. A. Lewis and E. Nudelman (1998). "High-resolution separation and quantification of neutral lipid and phospholipid species in mammalian cells and sera by multi-one-dimensional thin-layer chromatography." Analytical Biochemistry **258**(1): 109-117.

Wright, R. J., R. Bosch, M. I. Gibson and J. A. Christie-Oleza (2020). "Plasticizer degradation by marine bacterial isolates: a proteogenomic and metabolomic characterization." Environmental Science Technology **54**(4): 2244-2256.

Yakimov, M. M., R. Bargiela and P. N. Golyshin (2022). "Calm and Frenzy: marine obligate hydrocarbonoclastic bacteria sustain ocean wellness." Current Opinion in Biotechnology **73**: 337-345.

Yakimov, M. M., P. N. Golyshin, S. Lang, E. R. Moore, W.-R. Abraham, H. Lünsdorf and K. N. Timmis (1998). "*Alcanivorax borkumensis* gen. nov., sp. nov., a new, hydrocarbon-degrading and surfactant-producing marine bacterium." International Journal of Systematic Evolutionary Microbiology **48**(2): 339-348.

Yakimov, M. M., K. N. Timmis and P. N. Golyshin (2007). "Obligate oil-degrading marine bacteria." Current Opinion in Biotechnology **18**(3): 257-266.

Zhang, P., Z. Zhang, L. Zhang, J. Wang and C. Wu (2020). "Glycosyltransferase GT1 family: Phylogenetic distribution, substrates coverage, and representative structural features." Computational and Structural Biotechnology Journal **18**: 1383-1390.

Zhu, K. and C. O. Rock (2008). "RhlA converts β-hydroxyacyl-acyl carrier protein intermediates in fatty acid synthesis to the β-hydroxydecanoyl-β-hydroxydecanoate component of rhamnolipids in *Pseudomonas aeruginosa*." Journal of Bacteriology **190**(9): 3147-3154.

7. Appendix

7.1. Alignment of GglsA with EntF

Sequence ID: HAL0431670.1 Length: 1293 Number of Matches: 1

Range 1: 3 to 1291 GenPept Graphics Next Match Previous Match

Score	Expect	Method	Identities	Positives	Gaps
602 bits(1553)	0.0	Compositional matrix adjust.	423/1340(32%)	676/1340(50%)	81/1340(6%)

```
GglsA  50    KSLPLASAQRGLWFGEQIGPKDAVYNIAEYCEILGEIKVDIFIAALKQITQEAETTRVEI  109
             + LPL +AQ G+W  E++    + +++A Y E+ GE+   + + A+     +A+T R+
EntF   3     QHLPLVAAQPGIWMAEKLSDLPSAWSVAHYVELTGEVDAPLLVRAVVAGLAQADTLRMRF  62

GglsA  110   HDTADGPKQIILDRYKSEFP-VIDLSGATDPRSEAEKWMMDELHRPVD-HACDPLWFCAL  167
              +      Q + D    E P +IDL    DP   A   M  +L + +   +  PL F  L
EntF   63    TEDNGEVWQWVDDAQTFELPEIIDLRTNIDPHGTARALMQADLQQDLRVDSGKPLVFHQL  122

GglsA  168   FKISEQCYFWYQRSHHLVLDGFSAGIISRRCAEIYNALLEHRDIGPSPFLPLRIQHEQEQ  227
              ++++  ++WYQR HHL++DGFS   I+R+ A IY ALL       SPF P     E+ Q
EntF   123   IQVADNRWYWYQRYHHLLVDGFSFPAITRQIANIYCALLRGEPTPASPFTPFADVVEEYQ  182

GglsA  228   RYRESARYERDRNYWINQLSGLPDPVSLTRHNAVRAGGMRRSTAQLSIDTSN-TLRNLGK  286
             +YRES  ++RD  +W  Q   LP P SL+    +           +L ++ +N   R L
EntF   183   QYRESEAWQRDAAFWAEQRRQLPPPASLS-PTPLPGRSASADILRLKLEFTNGEFRQLAT  241

GglsA  287   D-TKASLPQILIGLFSAYIYRMTGADDLVFGMPVSARPNRDMRSIPSMMANAVSIRLAMN  345
               +      + + L + ++ R+    D   G     R      +    + N + + + +
EntF   242   QLSGVQRTDLALALAALWLGRLCNRMDYAAGFIFMRRLGSAALTATGPVLNVLPLGIHIA  301

GglsA  346   AELNLSTLTAQVSKAVRQALRHQMFRYEDLRRELGLLGQGQQISWVGVNIEPFDYDLRFG  405
             A+  L  L  +++  +++  RHQ +  E + R+ G     + +    +NI+ FDY L
EntF   302   AQETLPELATRLAAQLKKMRRHQRYDAEQIVRDSGRAAGEEPLFGPVLNIKVFDYQLDIP  361

GglsA  406   GHQCLAHNISNGTVEDLTVFIYDRGAHAGLRIDLDANPALYSQHDLDTHRDRFVRLISSV  465
             G Q   H ++ G V DL + ++    H  L I++ AN   Y +  L  H +R   LI+
EntF   362   GIQAQTHTLATGPVNDLELALFP-DEHGDLSIEILANKQRYDEPTLIQHAERLKMLIAQF  420

GglsA  466   VNDPSAPIGNISLLSEFERKCILDDWNKTEKPLPDTNLLALFDSQAIKTPDSIALTSGDK  525
               DPS   G++ ++   E    L   N T+  +P+T L AL   QA KTPD+ AL
EntF   421   AADPSLLCGDVDIMLPGE-YAQLAQINATQVEIPETTLSALVAEQAAKTPDAPALADARY  479

GglsA  526   KLTYQQLQSEVVRLSHALIQRGIGSGDIVAVALPRDESMPIALLAIMRTGAAYLPLDPTA  585
             + +Y++++ +VV L++ L +RG+  GD VAVALPR   + +AL AI+  GAA+LPLD
EntF   480   QFSYREMREQVVALANLLRERGVKPGDSVAVALPRSVFLTLALHAIVEAGAAWLPLDTGY  539

GglsA  586   PAERLALVLGEGKPTLILSA-SSISRALFKDDVAVLNLDTVDTGLAETSINPAPI---HG  641
             P +RL ++L + +P+L+++    + R  F D   + NL ++      T    AP+
EntF   540   PDDRLKMMLEDARPSLLITTDDQLPR--FSD---IPNLTSLCYNAPLTPQGSAPLQLSQP  594

GglsA  642   NTSAYVIYTSGSTGRPKGVEISHRALLNFLLAMQDELQVHANDKLLALTTVAFDIAALEL  701
             + +AY+I+TSGSTGRPKGV +  +A++N LL MQ+   +   D +   T  +FD++  E
EntF   595   HHTAYIIFTSGSTGRPKGVMVGQQAIVNRLLWMQNHYPLTGEDVVAQKTPCSFDVSVWEF  654

GglsA  702   YLPLISGAEVVIASREIAKDPEKLAALIADEKINLMQATPSHWQALLADYADQISN----  757
             + P I+GA++V+A  E  +DP  +    AD  +      PS   A +A    Q +
EntF   655   FWPFIAGAKLVMAEPEAHRDPLAMQHFFADYGVTTTHFVPSMLAAFVASLTPQTARQSCA  714

GglsA  758   --VRPLVGGEALPAQLAHKMRKL-GHPIINLYGPTETTIWSTIMKLDGDDL-----DAPP  809
                +    GEALPA L  + ++L G P+ NLYGPTE  +  +     G++L       + P
EntF   715   TLKQVFCSGEALPADLCREWQQLTGAPLHNLYGPTEAAVDVSWYPAFGEELAEVRGSSVP  774
```

Active site HHxxxDG motif

C domain

A domain

AMP binding

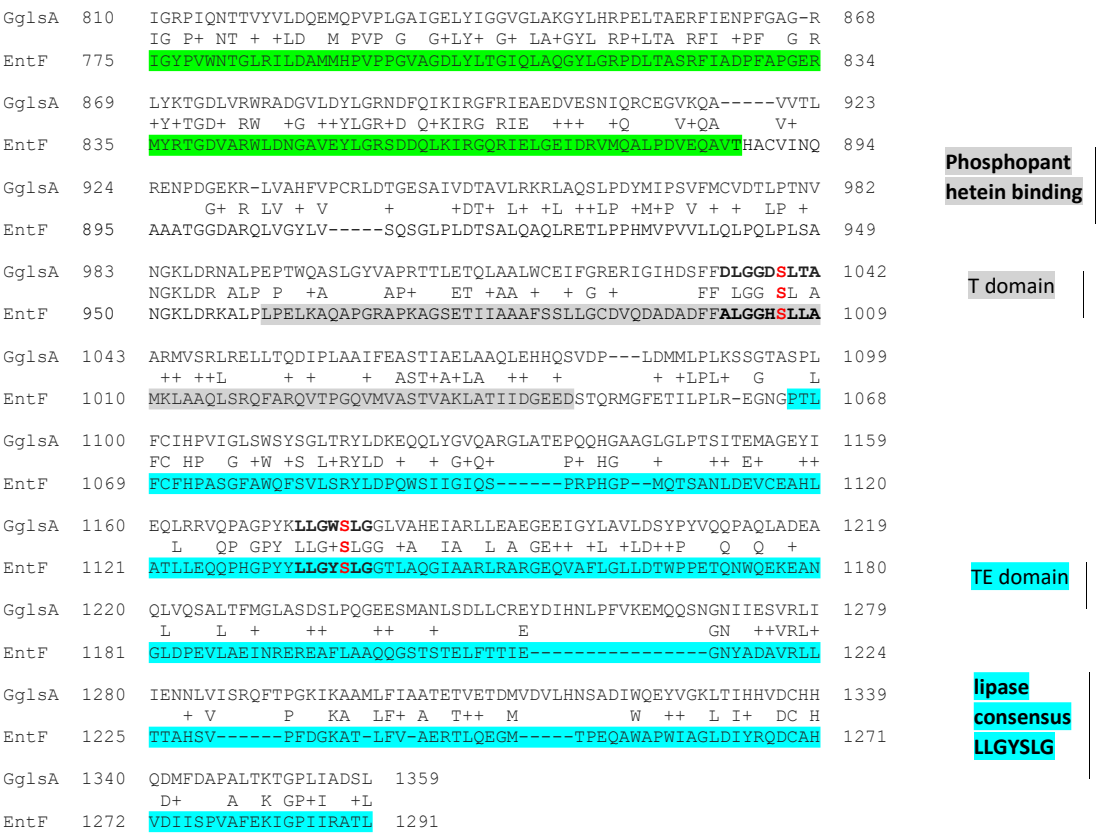

Figure 50 **The protein sequence alignment of the nonribosomal peptide synthetases (NRPS) ABO_1784 (GglsA) and EntF**.

C domain, yellow; A domain, green; T domain, grey; TE domain, blue; Ser1039 (attachment site of phosphopantetheine group) and Ser1138 (active site of TE domain), red.

7.2. Alignment of GglsC with EntD

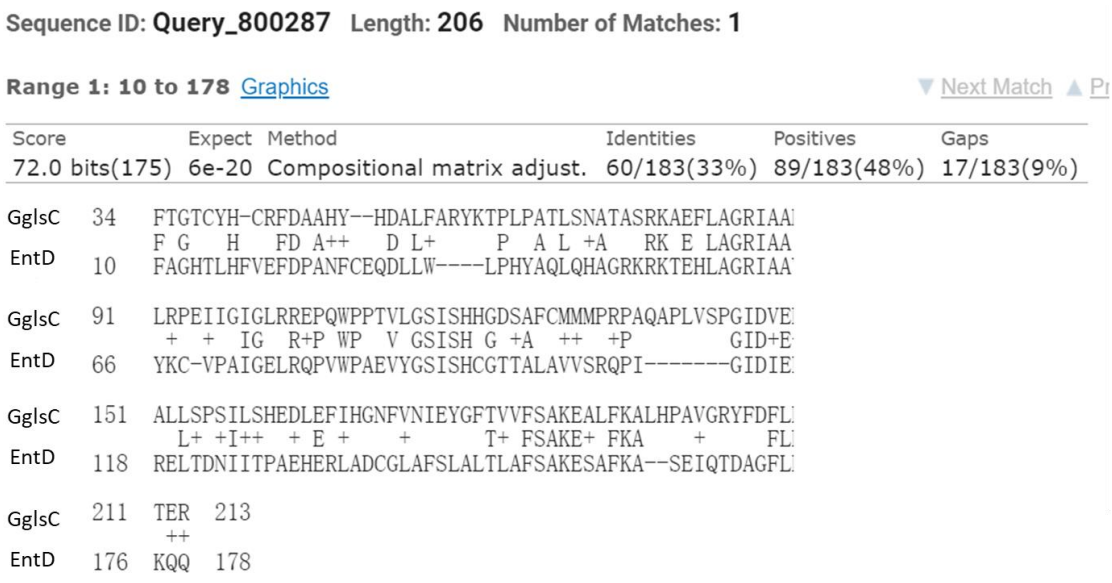

Figure 51 **The protein sequence alignment of the phosphopantetheinyl transferase ABO_1782 (GglsC) and EntD.**

7.3. Alignment of EntDlike with EntD

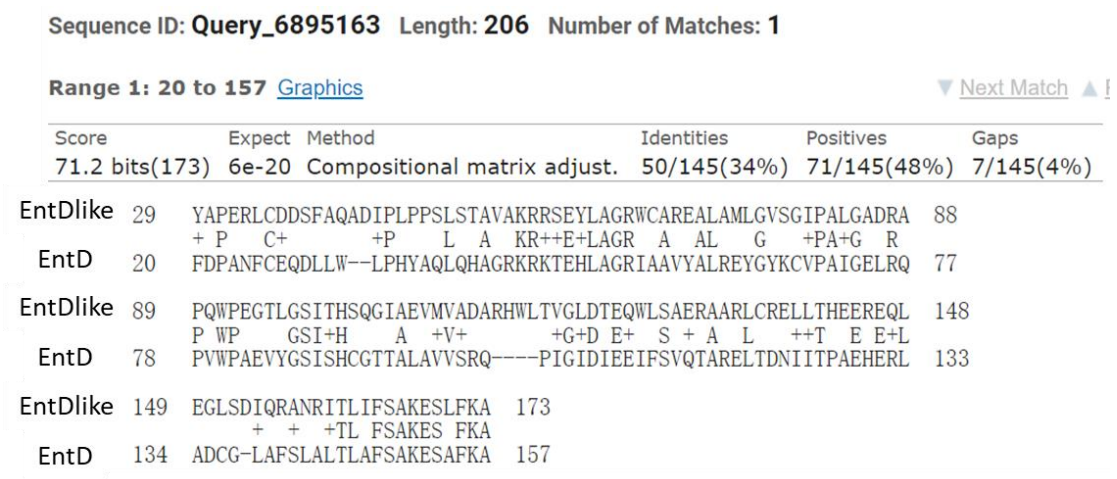

Figure 52 **The protein sequence alignment of the phosphopantetheinyl transferases ABO_2087 (EntDlike) and EntD.**

7.4. Alignment of GglsB with RhlB

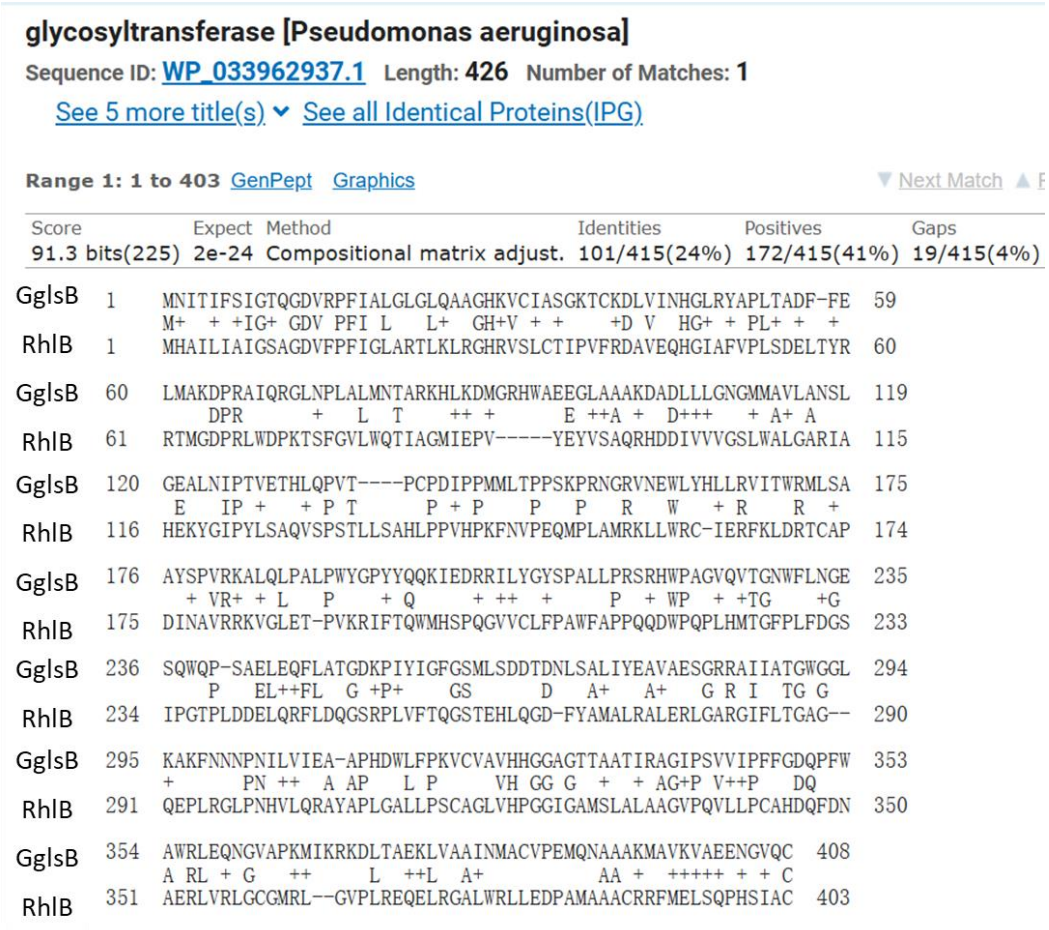

Figure 53 **The protein sequence alignment of the glycosyltransferases ABO_1783 (GglsB) and RhlB.**

7.5. Plasmid Maps

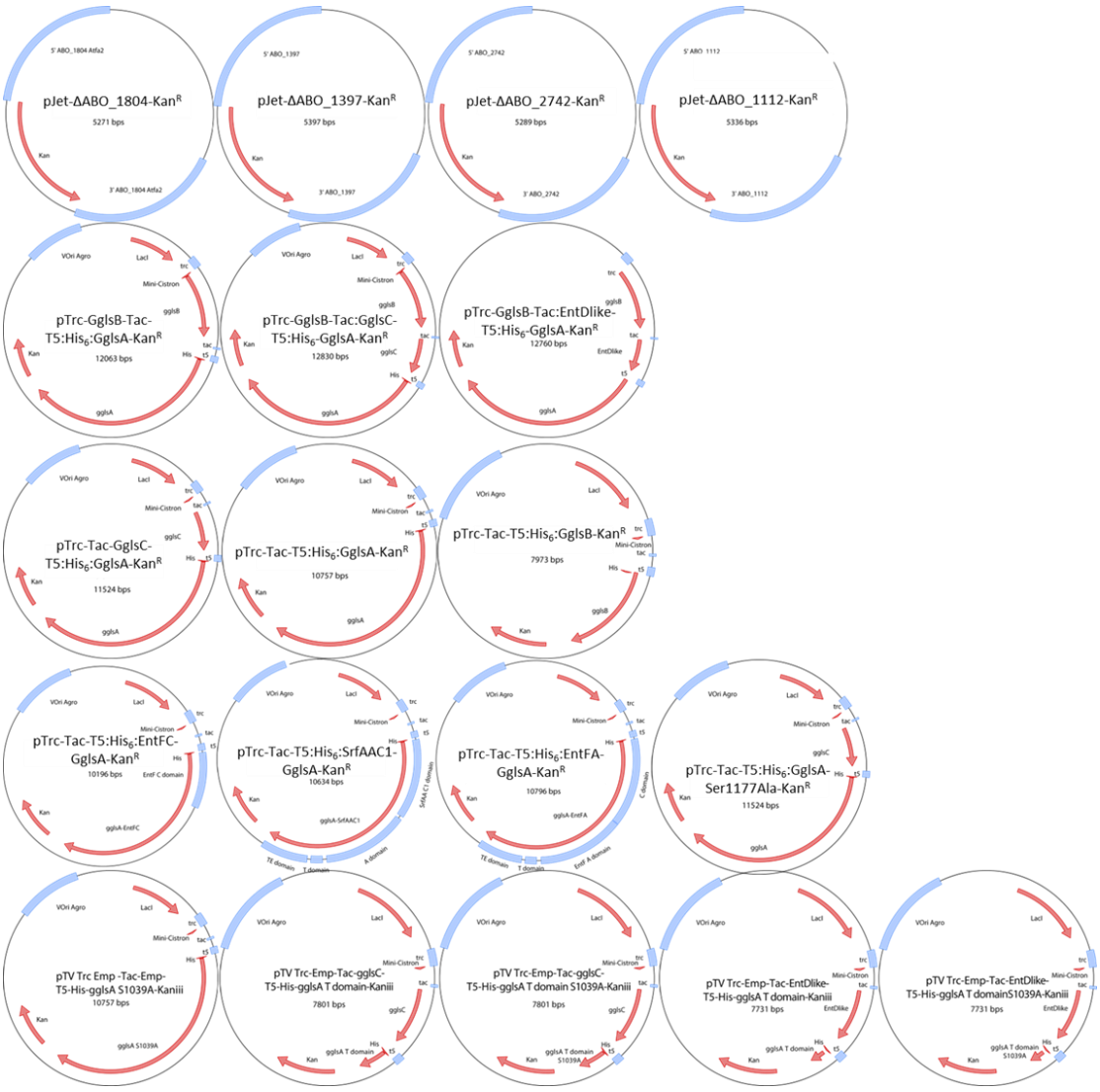

Figure 54 **Plasmid maps used in this study.**

7.6. Synthetic Oligonucleotides Used

Oligonucleotides used were ordered from Integrated DNA Technologies (IDT, Leuven, Belgium) (Table 3).

Restriction sites are indicated in bold face, and the sequence which aligns to the template is underlined. the cleavage site of BsaI (type II restriction enzyme) is depicted in italics

Table 3: Oligonucleotides used in this study. Fow: forward; Rev: reverse

Name	Sequence (5'-3')	Restriction site	Purpose/ function
Bn4631	GTATATTT**GGTCTC**T*AGGA*ATGAAAAAACGTATAGCCGCTC	BsaI	GglsA
Bn4468	GTATATTT**GGTCTC**G*ATCT*CTGTGAGTATCTAAATCA	BsaI	GglsA
Bn4469	GTATATTT**GGTCTC**G*AGAT*CGTTTTGTGCGGCTAA	BsaI	GglsA
Bn4632	GTATTTAT**GGTCTC**G*GTGG*TTATGGCATAAGTGAGTCAGC	BsaI	GglsA
Bn4688	GTATATTT**GGTCTC**G*GATG*ACTGGAATAAAACTGAG	BsaI	GglsA A domain Fow
Bn4689	GTATATTT**GGTCTC**G*CTAG*CGCCAGCAAATTAGTG	BsaI	GglsA C domain Rev
Bn4692	GTATATTT**GGTCTC**G*CAAG*TTAGATCGAAACGCCC	BsaI	GglsA T domain Fow
Bn4772	GTATATTT**GGTCTC**T*AGGA*GCCTCTCTTGGTTATGTCGCAC	BsaI	GglsA T domain
Bn4773	GTATTTAT**GGTCTC**G*GTGG*TTAGCGGGTTAAACCGCTGTAGC	BsaI	GglsA T domain
Bn4791	GTATTTAT**GGTCTC**G*GCGC*GTCGCCGCCTAGGTCAA	BsaI	GglsA Ser1039Ala Rev
Bn4790	GTATATTT**GGTCTC**G*GCGC*TAACAGCTGCAAGAATGGTATC	BsaI	GglsA Ser1039Ala Fow
Bn4789	GTATTTAT**GGTCTC**G*ACGC*CCAGCCCAATAGTTTATACG	BsaI	GglsA Ser1177Ala Rev
Bn4788	GTATATTT**GGTCTC**G*GCGT*TAGGTGGATTAGTGGCACATG	BsaI	GglsA Ser1177Ala Fow
Bn4633	GTATATTT**GGTCTC**T*GGGT*ATGAATATTACCATCTTTAGTATC	BsaI	GglsB
Bn4634	GTATTTAT**GGTCTC**G*CGTG*TCAAGTATCAGCCGATAACGT	BsaI	GglsB
Bn4471	GTATATTT**GGTCTC**T*AGGA*ATGAATATTACCATCTTTAGTATC	BsaI	GglsB
Bn4472	GTATTTAT**GGTCTC**G*GTGG*TCAAGTATCAGCCGATAACGT	BsaI	GglsB
Bn4473	GTATATTT**GGTCTC**G*TATT*ATGAGCGCTATCCACGCAAG	BsaI	GglsC

Bn4474	GTATTTAT**GGTCTC**G*CAAG*TCAATGGCTAGGGAGCACCCAG	BsaI	GglsC
Bn4552	GTATATTT**GGTCTC**T*AGGA*ATGAGCGCTATCCACGCAAG	BsaI	GglsC
Bn4553	GTATTTAT**GGTCTC**G*GTGG*TCAATGGCTAGGGAGCACCCAG	BsaI	GglsC
Bn4865	GTATATTT**GGTCTC**G*TATT*ATGTCATTATTCCATCGCCC	BsaI	EntDlike
Bn4866	GTATTTAT**GGTCTC**G*CAAG*TTAGCGGGGCAGGGCGATC	BsaI	EntDlike
Bn3424	GTATATTT**GGTCTC**G*CCAC*CTGCCGCAAGCACTCAGG	BsaI	Kanamycin cassette expression
Bn3422	GTATTTCC**GGTCTC**T*ACCG*CTCAGAAGAACTCGTCAAGAA	BsaI	Kanamycin cassette expression
Bn3415	TATATTTA**GGTCTC**T*CACG*CAGCTTATCGACTGCACGG	BsaI	Tac Promoter
Bn3416	TATATTTC**GGTCTC**G*AATA*CTGTTTCCTGTGTGAAATTGT	BsaI	Tac Promoter
Bn3419	TATATTTA**GGTCTC**T*GGGT*CAGCTTATCGACTGCACGG	BsaI	Tac Promoter
Bn3417	TATATTTA**GGTCTC**T*CTTG*CTTCTTCTGAGCGGGACTC	BsaI	T5 Promoter
Bn3418	TATATTTC**GGTCTC**A*TCCT*CCGTGATGGTGATGGTGATG	BsaI	T5 Promoter
Bn3420	TATATTTA**GGTCTC**T*TATT*CTTCTTCTGAGCGGGACTC	BsaI	T5 Promoter
Bn4690	GTATATTT**GGTCTC**G*CTAG*TGGCAGAACAAGCGGC	BsaI	EntF A domain
Bn4691	GTATATTT**GGTCTC**G*CTTG*CCGTTGGCGCTAAGTG	BsaI	EntF A domain
Bn4990	GTATATTT**GGTCTC**T*AGGA*ATGAGCCAGCATTTACCTTTGGTC	BsaI	EntF C domain
Bn4991	GTATATTT**GGTCTC**G*CATC*CGCAATGTGAATACCCAACGGTA	BsaI	EntF C domain
Bn4686	GTATATTT**GGTCTC**T*AGGA*ATGGAAATAACTTTTTACCCTTTA	BsaI	SrfAA C1 domain
Bn4687	GTATATTT**GGTCTC**G*CATC*AAGCAGCTTCTCTTTC	BsaI	SrfAA C1 domain
Bn4167	GTATATTT**GGTCTC**T*ACCG*TCGTCGATACCCAGCT	BsaI	Abo_2742 5’ homologous Fow
Bn4168	GTATTTAT**GGTCTC**T*ACCC*ACAGGAACAGTTGATCCAC	BsaI	Abo_2742 5’ homologous Rev
Bn4169	GTATATTT**GGTCTC**A*CGTG*CTGGCAGTTACACCATCG	BsaI	Abo_2742 3’ homologous Fow
Bn4170	GTATTTAT**GGTCTC**A*GCGT*TAGCAGCCTGTGGACCT	BsaI	Abo_2742 3’ homologous Rev
Bn4175	GTATATTT**GGTCTC**T*ACCG*TAGCATCCACTGCCTG	BsaI	Abo_1804 5’ homologous Fow

Bn4176	GTATTTAT**GGTCTC**T*ACCC*AGCCGGAATCCATAATAGA	BsaI	Abo_1804 5’ homologous Rev
Bn4177	GTATATTT**GGTCTC**A*CGTG*ACACGTTGCCTCATATT	BsaI	Abo_1804 3’ homologous Fow
Bn4178	GTATTTAT**GGTCTC**A*GCGT*TGTTTTTGCAGGGCACC	BsaI	Abo_1804 3’ homologous Rev
Bn4183	GTATATTT**GGTCTC**T*CACG*CTGCCGCAAGCACTCAGG	BsaI	Kanamycin cassette deletion
Bn4184	GTATTTCC**GGTCTC**T*GGGT*TCAGAAGAACTCGTCAAGAAG	BsaI	Kanamycin cassette deletion
Bn4419	GTTGGGTAACTGGGCCAC		ABO_1804 knock out screening
Bn4420	GACAGGCTGGTGAGCGTC		ABO_1804 knock out screening
Bn4421	CCGGCTCGTTGGTTACTGC		ABO_2742 knock out screening
Bn4422	ACTGGCTTTCTACGTGTTC		ABO_2742 knock out screening
Bn4341	ATCGCCTTCTATCGCCTTCTT		Knock out screening Kan^R Rev
Bn4342	ACTGGCTTTCTACGTGTTC		Knock out screening Kan^R Fow

7.7. Standard Curves of *A. borkumensis* OD600 vs. Colony Forming Units and Protein Content

To convert the OD_{600} measurement to cell number or total protein content. Serial dilutions of culture aliquots were plated on Marine Broth medium with 1 % (w/v) pyruvate and 0.0005 % (w/v) triphenltetrazolium (TTC) and colony formation units were counted (Figure 55 upper pannel). Similarly, total protein was extracted from serial dilutions of culture aliquot and was measured by Bradford method (2.2.4.1) (Figure 55 lower).

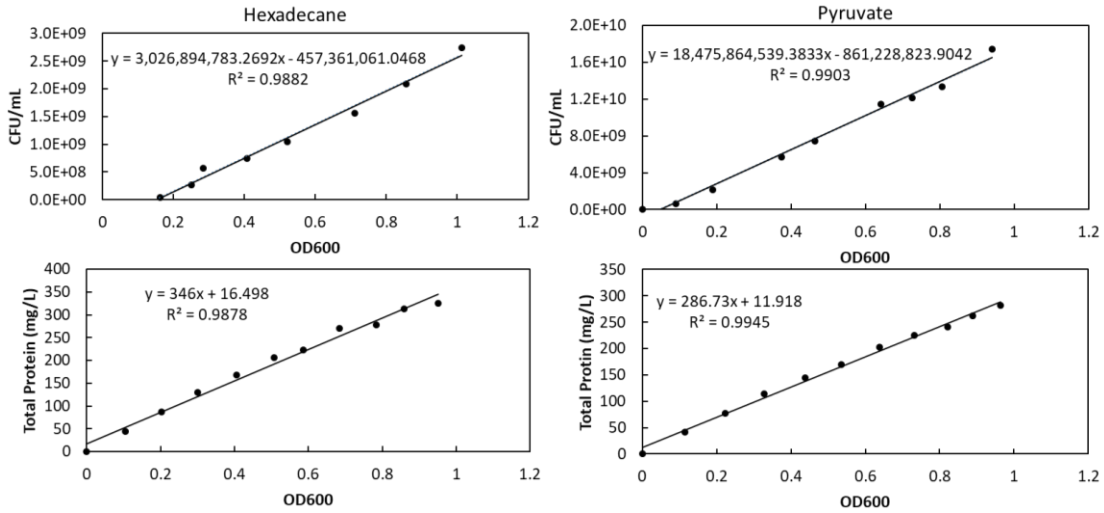

Figure 55 **Standard curves of *A. borkumensis* WT cells grown in hexadecane medium (left) or pyruvate medium (right).**

OD_{600} measurement was converted into colony formation units (CFU) per mL (upper panels) or total protein (mg/L) (lower panels).

7.8. Protein Gel for MALDI-TOF Measurement

The GglsA T domain with or without mutation was expressed in *E. coli* alone or co-expressed with GglsC or EntDlike. The calculated size of the T domain was 14 kDa. The T domain was collected from the flow through of the 30 kDa centrifugal filter and further dialyzed with 3 kDa centrifugal filter. The purified T domain was analyzed via MALDI-TOF or LC-MS.

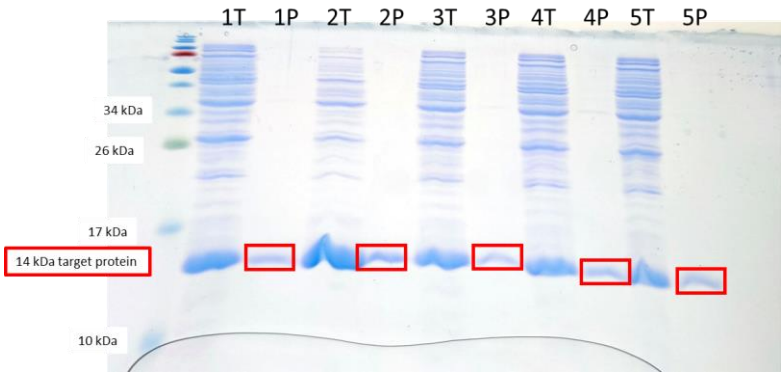

Figure 56 **SDS PAGE analysis of the GglsA T domain purification.**
T: Total protein; P: purified protein; 1: GglsA T domain; 2. GglsA T domain Ser1038Ala, 3: GglaA T domain co-expressed with GglsC; 4: GglsA T domain Ser1038Ala co-expressed with GglsC; 5: GglaA T domain co-expressed with EntDlike

7.9. Protein gel of Expression of GglsA, GglsB, GglsC, EntDlike in *E. coli*

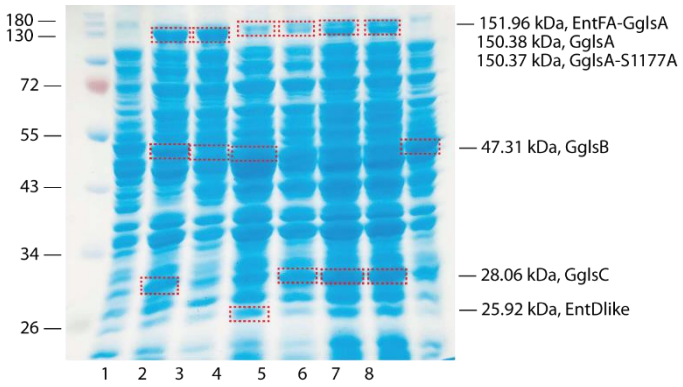

Figure 57 **Expression of proteins involved in glycine-glucolipid synthesis in *E. coli*.**
Detection of proteins by SDS PAGE after *E. coli* expression. 1, empty vector control; 2, pTrc-GglsB-Tac:GglsC-T5:His6-GglsA-Kan^R; 3, pTrc-GglsB-Tac-T5:His6-GglsA-Kan^R; 4, pTrc-GglsB-Tac:EntDlike-T5:His6-GglsA-Kan^R; 5, pTrc-GglsC-Tac-T5:His6:GglsA-Kan^R; 6, pTrc-Tac-T5:His6-EntFA-GglsA-Kan^R; 7, pTrc-Tac:GglsC-T5:His6-GglsA-Ser1177Ala-Kan^R; 8, pTrc-Tac-T5:His6-GglsB-Kan^R. Heterologously expressed protein bands are depicted with dashed boxes. The numbers on the right show the calculated masses.

7.10. 3-Hydroxy-decanoyl-ACP Synthesis via *E. coli* AAS and *V. harveyi* AAS

In addition to 3-hydroxy-decanoyl-CoA, there is another candidate of the GglsA fatty acid donor, 3-hydroxy-decanoyl-ACP. Similar to 3-hydroxy-decanoyl-CoA, there is no commercial 3-hydroxy-decanoyl-ACP available. Therefore, the 3-hydroxy-decanoyl-ACP synthesis was carried out with *E. coli* acyl acyl carrier protein synthetase (AAS) purified by Dr. Mathias Brands (IMBIO, University of Bonn). The purity and identity of the acyl carrier protein were checked by using 20 % urea PAGE Figure 58A. the reaction products were checked by 20 % urea PAGE, with hexadecanoyl-ACP as positive control. The acyl-ACP synthesis was successful for hexadecanoyl-ACP as shown in Figure 58B. The hexadecanoyl-ACP (lower band) band runs further than the ACP band (upper band). However, the 3-hydroxy-decanoyl-ACP was not produced, because there was only the upper ACP band.

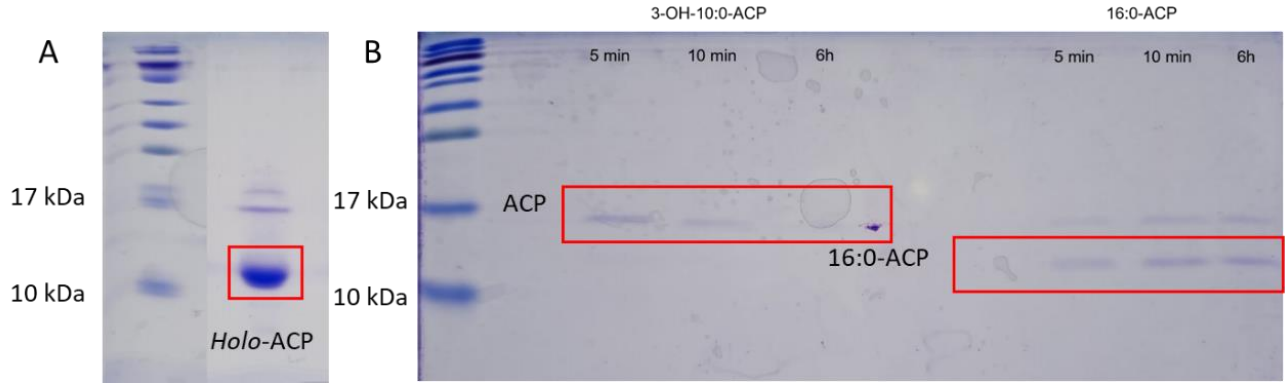

Figure 58 **Experiments for the Preparation of 3-hydroxydecanoyl-ACP.**
(A) Urea PAGE of ACP isolated from *E. coli*. (B) SDS PAGE of acylation reaction of 3-hydroxy decanoic acid (left) with palmitic acid (right) as positive control using AAS isolated from *E. coli*.

Because the *E. coli* AAS might not be able to catalyze the acylation with a hydroxy fatty acid, the *Vibrio harveyi* AAS was expressed in *E. coli* and purified for the hydroxy-acyl-ACP synthesis (Figure 59A). The hydroxy-acyl-ACP synthesis was performed with *V. harveyi* AAS, and ACP served as negative control. As shown in Figure 59B, the reaction products were loaded onto the 20 % UREA PAGE, the band pattern was same as the negative control, which suggests that the hydroxy-acyl-ACP synthesis was also not successful.

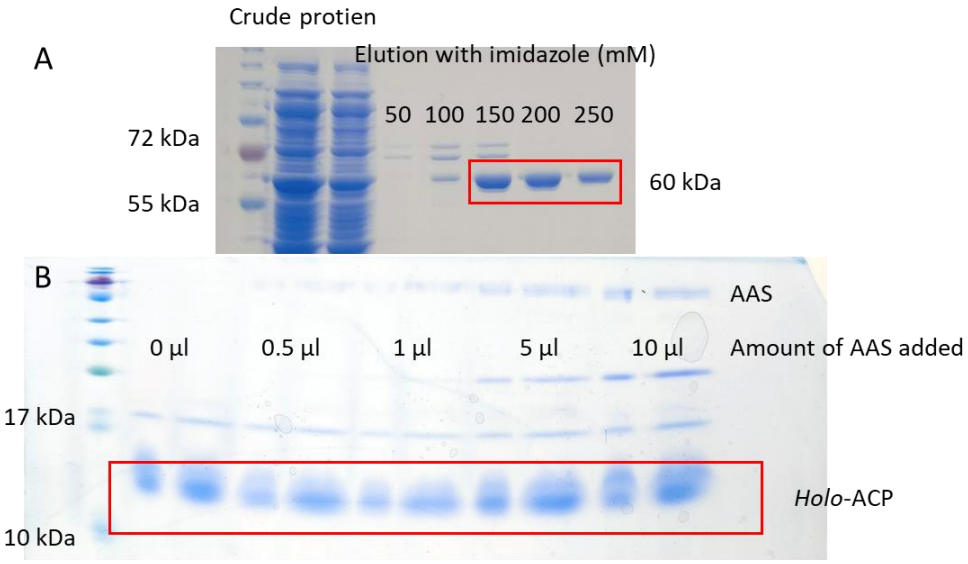

Figure 59 **Preparation of 3-hydroxydecanoyl-ACP with *V. harveyi* AAS.**
(A) SDS PAGE of *E. coli* culture expressing *V. harveyi* AAS and purified via Ni-NTA. The *V. harveyi* AAS was eluted with gradient of imidazole elution buffer. The fraction with 150 mM, 200 mM and 250 mM was combined for further desalting and acyl-ACP synthesis. The calculated size of the *V. harveyi* AAS is 60 kDa. (B) Urea PAGE of acylation reaction of 3-hydroxy decanoic acid with different amount of AAS isolated from *V. harveyi*. The reaction time was 3 h. The reaction without AAS added served as negative control (0 µL).

7.11. 3-Hydroxy-decanoyl-CoA Synthesis with Acyl-CoA Synthetase from *Pseudomonas*

The C domains of GglsA catalyze the transfer of a single 3-hydroxy-fatty acid onto the glycine bound at the T domain. There are two candidates of the fatty acid donor of GglsA, acyl-coenzyme A (acyl-CoA) or acyl-acyl carrier protein (acyl-ACP). As 3-hydroxy-decanoyl-CoA is not available, it was synthesized with the 3-hydroxydecanoate and lithium coenzyme A with acyl-CoA synthetase from *Pseudomonas* according to 2.2.3.6. The production of 3-hydroxydecanoyl-CoA was confirmed via LC-MS/MS. The calculated molecular weight of protonated 3-hydroxy-decanoyl-CoA is 938.2532. The protonated molecule was detected with *m/z* of [M+H$^+$] 938.2621 (Figure 60).

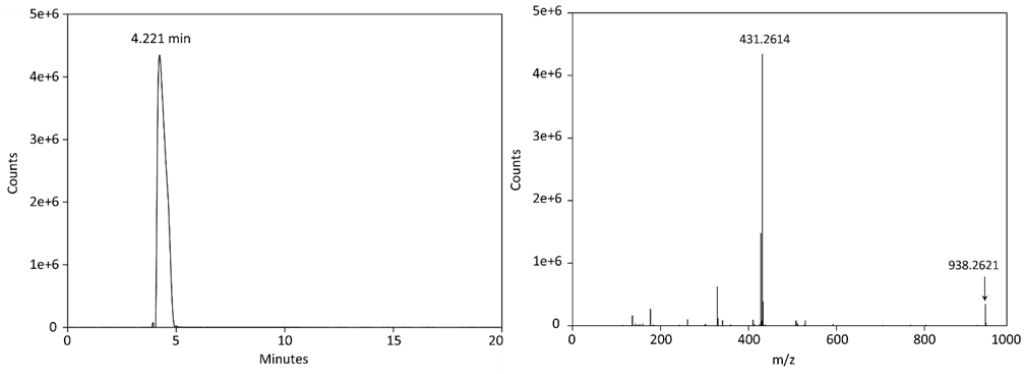

Figure 60 **Analysis of 3-hydroxy-decanoyl-CoA by mass spectrometry.**
LC-MS chromatogram (left) shows 3-hydroxy-decanoyl-CoA eluted at 4.221 min. The MS/MS spectrum (right) shows the fragmentation pattern of 3-hydroxy-decanoyl-CoA.

7.12. List of Strains Used in This Study

Box Nr.	Stock Nr.	Donor species	Strain		Cloning Vector	Resistence
Box 22	1441		Alcanivorax borkume	WT		
Box 24	1506	Alcanivorax borkumensis	Alcanivorax borkume	ΔABO_1112		Kan
Box 24	1507	Alcanivorax borkumensis	Alcanivorax borkume	ΔABO_1397		Kan
Box 24	1508	Alcanivorax borkumensis	Alcanivorax borkume	ΔABO_1804		Kan
Box 24	1509	Alcanivorax borkumensis	Alcanivorax borkume	ΔABO_2742		Kan
Box 24	1535	Bacillus subtilis	E. coli Electoshox	PUCIDT-SrfAA C domain	pUCIDT	Kan
Box 24	1550	Alcanivorax borkumensis	E.coli XL1 Blue	pTv-Trc-Abo_1784-Tac-Abo_1782-T5 His-Abo_1783-Kanii	PTV-GG-AMP	Kan
Box 24	1551	Alcanivorax borkumensis	E.coli XL1 Blue	pTv-Trc-Abo_1784-Tac-EVT5-Abo_1783—Kaniii	PTV-GG-AMP	Kan
Box 24	1559	Alcanivorax borkumensis	E.coli XL1 Blue	pTv-Trc-EV-Tac-Abo_1782-T5 His-Abo_1784-Kaniii	PTV-GG-AMP	Kan
Box 24	1560	Alcanivorax borkumensis	E.coli XL1 Blue	pTv-Trc-Abo_1783-Tac-EV-T5-Abo_1784—Kaniii	PTV-GG-AMP	Kan
Box 24	1561	Alcanivorax borkumensis	E.coli XL1 Blue	pTv-EV-Tac-EV-T5 His-Abo_1784-Kaniii	PTV-GG-AMP	Kan
Box 24	1563	Alcanivorax borkumensis	E.coli XL1 Blue	pTv-Trc-Abo_1783-Tac-Abo_1782-T5 His-Abo_1784-Kanii	PTV-GG-AMP	Kan
Box 25	1608	A. borkumensis	Alcanivorax borkume	ΔABO_1782	PJ-GG-lacZ	Kan
Box 25	1619	A. borkumensis	E.coli XL1 Blue	pTv-ev-Tac-EV-T5 His-Abo_1784-Kaniii Ser1177Ala	PTV-GG-AMP	Kan
Box 25	1620	A. borkumensis	E.coli XL1 Blue	pTv-ev-Tac-EV-T5 His-SrfAA Abo_1784-Kaniii Ser1177Ala	PTV-GG-AMP	Kan
Box 25	1621	A. borkumensis	E.coli XL1 Blue	pTv-ev-Tac-EV-T5 His-Abo_1784-Kaniii Ser1039Ala	PTV-GG-AMP	Kan
Box 25	1622	A. borkumensis	E.coli XL1 Blue	pTv-ev-Tac-EV-T5 His-SrfAA Abo_1784-Kaniii Ser1039Ala	PTV-GG-AMP	Kan
Box 25	1623	A. borkumensis	E.coli XL1 Blue	pTv-ev-Tac-EV-T5 His-EntF Abo_1784-Kaniii Ser1039Ala	PTV-GG-AMP	Kan
Box 25	1624	A. borkumensis	E. coli BL 21 (DE3)	pTv-ev-Tac-EV-T5 His-Abo_1784 T domain-Kaniii	PTV-GG-AMP	Kan
Box 25	1625	A. borkumensis	E. coli BL 21 (DE3)	Ptv-ev-Tac-Abo_1782-T5 His-Abo_1784 T domain-Kaniii	PTV-GG-AMP	Kan
Box 25	1626	A. borkumensis	E. coli BL 21 (DE3)	pTv-ev-Tac-EV-T5 His-Abo_1784 T domain-Kaniii Ser1039Ala	PTV-GG-AMP	Kan
Box 25	1627	A. borkumensis	E. coli BL 21 (DE3)	Ptv-ev-Tac-Abo_1782-T5 His-Abo_1784 T domain-Kaniii Ser103	PTV-GG-AMP	Kan
Box 25	1628	A. borkumensis	E.coli XL1 Blue	pTv-ev-Tac-EV-T5 His-SrfAA Abo_1784-Kaniii	PTV-GG-AMP	Kan
Box 25	1629	A. borkumensis	E.coli XL1 Blue	pTv-ev-Tac-EV-T5 His-EntF A Abo_1784-Kaniii	PTV-GG-AMP	Kan
Box 25	1635	A. borkumensis	E. coli BL 21 (DE3)	pTv-ev-Tac-EV-T5 His-ABO_1783-Kaniii	PTV-GG-AMP	Kan
Box 25	1636	A. borkumensis	E. coli BL 21 (DE3)	pTv-ev-Tac-ABO_2087-T5 His- Abo_1784 T domain-Kaniii	PTV-GG-AMP	Kan
Box 25	1642	Alcanivorax borkumensis	E.coli XL1 Blue	pTv-Trc-Abo_1783-Tac-Abo_2087-T5 His-Abo_1784-Kanii	PTV-GG-AMP	Kan
Box 26	1659	A. borkumensis	E.coli XL1 Blue	Ptv-ev-Tac-ABO_1782-T5 His-Abo_1784-Kaniii Ser1177Ala	PTV-GG-AMP	Kan
Box 26	1660	A. borkumensis	E.coli XL1 Blue	Ptv-ev-Tac-ABO_1782-T5 His-SrfAA Abo_1784-Kaniii Ser1177Al	PTV-GG-AMP	Kan
Box 26	1661	A. borkumensis	E.coli XL1 Blue	Ptv-ev-Tac-ABO_1782-T5 His-SrfAA Abo_1784-Kaniii	PTV-GG-AMP	Kan
Box 26	1671	Alcanivorax borkumensis	E.coli XL1 Blue	pTv-Trc-Abo_1783-Tac-EV-T5-Abo_1784—Kaniii	PTV-GG-AMP	Kan
Box 26	1672	Alcanivorax borkumensis	E.coli XL1 Blue	pTv-ev-Tac-EV-T5 His-EntF C Abo_1784-Kaniii	PTV-GG-AMP	Kan
Box 26	1687	A. borkumensis	E.coli XL1 Blue	pTV-EV-Tac-EV-T5-His-EntF TE ABO_1784-Kan iii	PTV-GG-AMP	Kan

Figure 61 **Bacteria strains used in this study**

8. Publication

Jiaxin Cui, Georg Hölzl, Tobias Karmainski, Till Tiso, Sonja Kubicki, Stephan Thies, Lars M. Blank, Karl-Erich Jaeger, Peter Dörmann. The glycine-glucolipid of *Alcanivorax borkumensis* is resident to the bacterial cell wall. Applied and Environmental Microbiology, 2022, 88(16): e01126-22. doi: 10.1128/aem.01126-22

Under review:

Jiaxin Cui, Maximilian Fassl, Vaisnavi Vasanthakumaran, Maya Marita Dierig, Georg Hölzl, Tobias Karmainski, Till Tiso, Sonja Kubicki, Stephan Thies, Lars M. Blank, Karl-Erich Jaeger, Peter Dörmann. Biosynthesis and role during oil-degradation of the biosurfactant from *Alcanivorax borkumensis*

9. Acknowledgments

I am deeply thankful to my colleagues, family, and friends whose unwavering support, especially during the challenging times of the pandemic, made this work possible. Completing my studies from the master's to doctoral level at the University of Bonn is a great honor.

I extend my sincere thanks to my supervisor, Prof. Dr. Peter Dörmann, for providing me with the opportunity to learn, understand, and explore the field of lipid biosynthesis. Your knowledge and support over the past few years have been invaluable. Being far away from my family abroad, I am particularly grateful for your guidance through my academic journey.

I am grateful to Prof. Dr. Lukas Schreiber for graciously agreeing to co-review my thesis and for his invaluable help with Dr. Viktoria V. Zeisler-Diehl in measuring the surface tension. The lectures during my master's program sparked my interest in ecophysiology, and I thoroughly enjoyed them.

I am deeply thankful to Prof. Dr. Uwe Deppenmeier and Prof. Dr. Gabriel Schaaf for being my examination committee members and for their guidance and expertise.

I would like to express my appreciation to all the members of IMBIO for creating an excellent working atmosphere. Special thanks to Dr. Katharina Gutbrod and Helga Peisker for their expertise in mass spectrometry, Dr. Georg Hölzl for assistance with cloning, and Andreas Ahrens for helping me realize my ideas with his magic toolbox.

I extend my gratitude to Prof. Lars M. Blank, Dr. Till Tiso, and Tobias Karmainski (RWTH Aachen) for their invaluable assistance in isolating glycine-glucolipid for biotechnological applications. I also appreciate the support of Prof. Karl-Erich Jaeger, Dr. Stephan Thies, Dr. Sonja Kubicki, Lea Sundermeyer, Fabio Diaz, and Lisa-Marie Kirschen (Forschungszentrum Jülich) in the initial isolation of the GglsA-GglsB-GglsC gene cluster and their experimental assistance.

I would like to acknowledge Dr. Hannes Beckert and Pia Stausberg (Microscopy Core Facility of the Medical Faculty, University of Bonn) for their support with electron microscopy. My thanks also go to Dr. Mark Sylvester (Proteomics Core Facility, University of Bonn) for his support with MALDI-TOF analysis.