How do marine bacteria respond to nutrient limitation?: A lipidomics approach.

Alastair Smith

A thesis submitted to the School of Life Sciences in fulfilment of the requirements for the degree of Doctor of Philosophy

June 2017

University of Warwick,
Coventry, UK

Table of Contents

List of Figures

List of Tables

Acknowledgements

I would like to thank my supervisors, Dr. Yin Chen and Prof. David Scanlan, for their invaluable guidance and support during the course of this work. Thanks go to all lab members, both past and present. In particular, the advice of Julie Scanlan and Dr. Ian Lidbury was crucial in culturing these organisms and carrying out molecular genetics work. Thanks to Dr. Yijun Zhu for her assistance with the protein purification work. I would also like to acknowledge the assistance of Dr Joost Brandsma and Prof. Tony Postle from the University of Southampton for agreeing to train me in the use of a mass spectrometer. I would also like to thank Dr. Marta Sebastian for helpful discussions on how to present the data in Chapter 3 for publication, and for providing data from her own analysis of the Tara metagenomes. Lastly, thanks go to Dimitri for his love and support through the course of this degree.

Declarations

This thesis is submitted to the University of Warwick in support of my application for the degree of Doctor of Philosophy. It has been composed by myself and has not been submitted in any previous application for any degree. The work presented (including data generated and data analysis) was carried out by me with the exception of the data for PlcP and RecA abundance in the Tara metagenome data set (Chapter 3), which was obtained by Marta Sebastian (from a publicly accessible database) and analysed by me. Some of the data presented in Chapter 3 have been published previously in a paper of which I was joint first author (Sebastian, Smith *et al.* (2016) Lipid remodelling is a widespread strategy in marine heterotrophic bacteria upon phosphorus deficiency. *ISME J.* **10:** 968-978). Only results which I obtained personally are presented.

Abbreviations

1/2 YTSS	Half strength yeast-tryptone-sea salts medium
bp	Base pairs
C	Carbon
cAMP	Cyclic-adenosine monophosphate
CDO	Cysteine dioxygenase
CDP	Cytidine diphosphate
CL	Cardiolipin
CMP	Cytidine monophosphate
CSAD	Cysteine sulfinic acid decarboxylase
DAG	Diacylglycerol
dATP	Deoxyadenine triphosphate
dCTP	Deoxycytidine triphosphate
DGCC	Diacylglyceryl carboxyhydroxymethylcholine
DGDG	Diglycosyl diacylglycerol
DGTA	Diacylglyceryl hydroxymethyl-trimethylalanine
dGTP	Deoxyguanidine triphosphate
DGTS	Diacylglyceryl-*N*,*N*,*N*-trimethyl homoserine
DMSP	Dimethlysulfoniopropionate
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleoside triphosphate
DOM	Dissolved organic matter
DOP	Dissolved organic phosphorus
dTTP	Deoxythymidine triphosphate
ESI	Electrospray ionisation
GABA	γ-aminobutyric acid
GADG	Glucuronic acid diacylglycerol
GBT	Glycine betaine
GLM	Generalised linear model
Gm	Gentamicin
Gm^R	Gentamicin resistance cassette from p34S-Gm
GNAT	Gcn5-related N-acyltransferase
GOS	Global Ocean Sampling
GTR	Generalised time reversible
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HILIC	Hydrophilic interaction liquid chromatography
HPLC	High performance liquid chromatography
HTL	Homotaurine lipid
IMG	Integrated Microbial Genomes
IPTG	Isopropyl β-D-1-thiogalactopyranoside
KL	Lysine lipid

Km	Kanamycin
LB	Lysogeny broth
LC	Liquid chromatography
LDS	Lithium dodecylsulfate
LPS	Lipopolysaccharide
LRT	Likelihood ratio test
MB	Marine Broth
MGDG	Monoglycosyl diacylglycerol
miTAG	16S Illumina metagenomic sequencing tag
MS	Mass spectrometry
N	Nitrogen
N*	Excess of nitrogen relative to phosphorus
NMR	Nuclear magnetic resonance
NPP	Net primary production
OD540	Optical density at 540 nm
OG	Orthologous group
OL	Ornithine lipid
OlsB	Ornithine N-acyltransferase
OlsB2	Glutamine N-acyltransferase
OlsF	Ornithine lipid synthase
OM-RGC	Ocean microbial reference gene catalogue
P	Phosphorus
PAGE	Polyacrylamide gel electrophoresis
PC	Phosphatidylcholine
PCR	Polymerase chain reaction
PE	Phosphatidylethanolamine
PG	Phosphatidylglycerol
PGP	Phosphatidylglycerol phosphate
PhoBR	Phosphate response regulator
Pi	Inorganic phosphorus / phosphate
PlcP	Phospholipase C involved in lipid remodelling
pNP	para-nitrophenol
pNPP	para-nitrophenol phosphate
POM	Particulate organic matter
PS	Phosphatidylserine
PTFE	Polytetrafluoroethylene
Q-TOF	Quadrupole-time of flight
QL	Glutamine lipid
RNA	Ribonucleic acid
rRNA	Ribosomal RNA
SDS	Sodium dodecylsulfate
SOC	Super Optimal broth with catabolite repression
SQDG	Sulfoquinovosyl diacylglycerol

TBE	Tris-Borate-EDTA buffer
TLC	Thin layer chromatography
UDP	Uridine diphosphate

Abstract

Microbes inhabiting surface waters of the Earth's oceans are exquisitely adapted to their nutrient-poor environment. Marine phytoplankton, for example, are able to reduce their requirements for phosphorus by replacing membrane phospholipids with alternative non-phosphorus lipids. Heterotrophic bacteria, which can also thrive when phosphorus is scarce, had not, however, been shown to carry out this process – seemingly placing these organisms at a competitive disadvantage. In this thesis, I show that substitution of membrane phospholipids for a variety of non-phosphorus lipids is a conserved response to phosphorus deficiency amongst phylogenetically diverse marine heterotrophic bacteria. By deletion mutagenesis and complementation in the model marine bacterium *Phaeobacter* sp. MED193 and heterologous expression in recombinant *Escherichia coli*, I confirmed the roles of a phospholipase C (PlcP) and a glycosyltransferase in lipid remodelling. Analyses of two large collections of marine metagenomes, the Global Ocean Sampling (GOS) and *Tara* datasets, demonstrate that PlcP is particularly abundant in areas characterised by low phosphate concentrations. To better understand the lipids that potentially replace phospholipids during this remodeling process, I investigated a number of poorly-characterised aminolipids that are prevalent in the globally important marine Roseobacter group. I was able to identify two genes involved in the synthesis of one of these lipids, a glutamine lipid. Subsequent phylogenetic analysis of one of these genes revealed that the capacity to synthesise glutamine lipid appears to be virtually ubiquitous within the Roseobacter group. A further class of aminolipids was present in many Roseobacter strains, which I identified as a novel class of homotaurine-containing lipids using high-resolution, accurate mass spectrometry. These homotaurine lipids were detected in a battery of Roseobacter strains, enabling me to employ a comparative genomics approach to identify genes potentially involved in their biosynthesis. Surprisingly, neither of these aminolipids appeared to play an important role in the response to phosphorus limitation in the Roseobacter strains tested. Together, these results point to a key role for lipid substitution as an adaptive strategy enabling heterotrophic bacteria to thrive in vast, phosphorus-depleted areas of the ocean. Although phosphorus-free lipids play a crucial role in this process of adaptation, my work emphasises that many of these lipids are not simply substitutes for phospholipids but rather appear to have important roles in the cell in their own right.

Chapter 1 Introduction

1.1 The role of microbes in ocean biogeochemistry

The world's oceans form a vast ecosystem in which almost half of the planet's photosynthesis takes place (Falkowski *et al.*, 1998; Field, 1998). Although benthic algae and plants can be important in some near-shore settings (Behringer and Butler IV, 2006), the overwhelming majority of this marine primary production is carried out by microscopic algae collectively known as phytoplankton. These phytoplankton are a diverse mix of unicellular eukaryotes and cyanobacteria. Unlike higher plants on land, they lack substantial structural polymers (such as lignin), resulting in the photosynthetic biomass of the ocean being substantially smaller. Consequently, the ocean's biomass must turn over every 2-6 days as opposed to an average turnover time of 13-16 years for land plants (Behrenfeld and Falkowski, 1997). This rapid turnover implies a tight coupling between the fixation of inorganic carbon into organic matter by phytoplankton and the consumption of this organic matter by heterotrophs (Behrenfeld *et al.*, 2006; Giovannoni and Vergin, 2012). Foremost amongst these consumers are heterotrophic bacteria by virtue of their small size, large numbers and consequently vast active surface area (Azam, 1998). The cycling of organic matter between phytoplankton and heterotrophic bacteria forms a centrepiece of the marine ecosystem, the 'microbial loop' (Azam *et al.*, 1983; Figure 1.1). The substantial fluxes of carbon and other elements passing through this loop mean that marine microbes can have a major influence on the biogeochemistry of the ocean and atmosphere (Falkowski *et al.*, 2008).

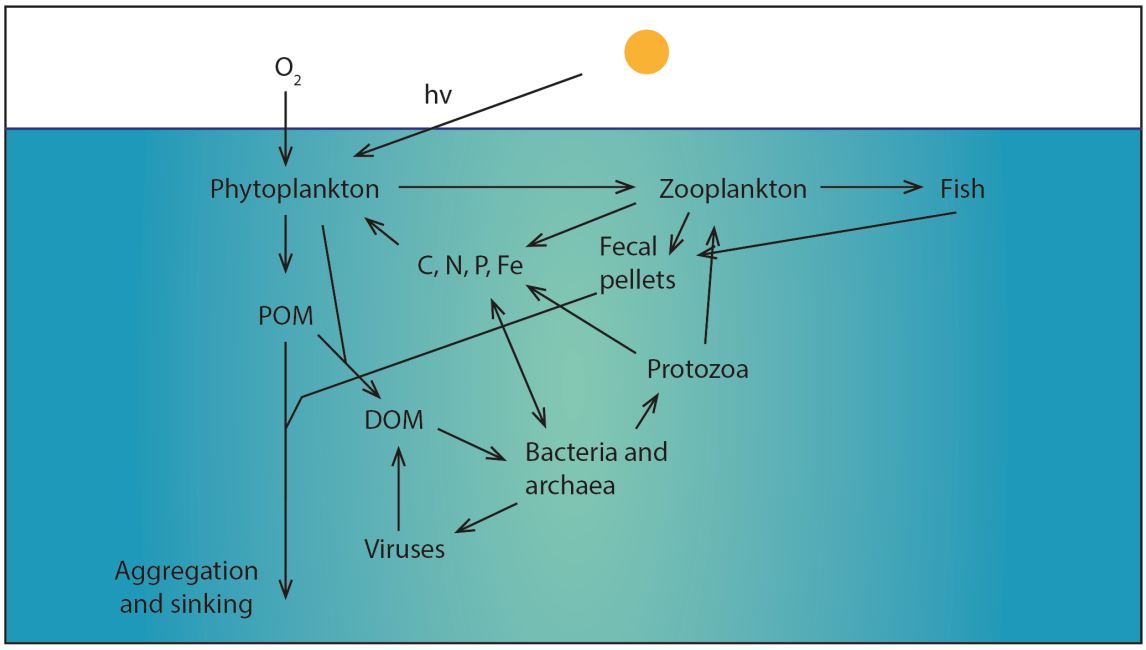


Figure 1.1 Simplified schematic of the ocean surface ecosystem. A large proportion of the organic matter synthesised by phytoplankton is ultimately converted to dissolved organic matter (DOM). The uptake of this is dominated by bacteria. Both bacteria and phytoplankton compete for a pool of inorganic nutrients. Nutrients can be regenerated by the activity of heterotrophic bacteria, as well as by protists and zooplankton. Viral lysis of bacteria returns matter to the pool of DOM. POM: particulate organic matter; hν: light energy.

Although the average depth of the oceans is 3500m, all photosynthesis occurs within the top 100-200m, below which light intensity is insufficient to support photosynthetic activity (Schlesinger and Bernhardt, 2013). This surface layer differs from the depths in other ways as well. Virtually all the energy that enters the oceans does so in this layer that receives irradiation from the sun, resulting in surface temperatures which are, for the most part, substantially warmer than those of deeper waters. Consequently the warmer, more buoyant surface waters sit atop a great volume of cooler, denser deep waters, with only limited mixing occurring between the two. Since most photosynthesis tends to occur in the surface mixed layer, nutrients, such as nitrogen or phosphorus, often become depleted and limiting for phytoplankton growth (Moore *et al.*, 2013). By contrast, the darker waters below the

thermocline (the temperature gradient separating the mixed layer from deeper waters) are relatively nutrient rich and more likely to be limited by energy availability. Nutrients can be introduced to the surface from deep waters by winter mixing at higher latitudes, when lower solar energy inputs and stronger winds result in a deeper mixed layer (Schlesinger and Bernhardt, 2013).

The majority of primary production is remineralised (i.e. oxidised to inorganic components: CO_2, NO_3^-, PO_4^{2-}) by heterotrophic bacteria (Pomeroy *et al.*, 2007) in the surface ocean. This returns much of the fixed carbon to the atmosphere as CO_2 while replenishing stocks of inorganic nutrients in the mixed layer. In order to be accessible for bacterial metabolism, organic matter must be in the dissolved phase and small enough to pass through the outer membrane - typically smaller than around 600 Daltons (Koebnik *et al.*, 2000). A variable but substantial fraction of primary production is released as dissolved organic matter (DOM) directly by phytoplankton (Carlson and Hansell, 2015). Further DOM can be liberated as a result of viral lysis of phytoplankton or bacterioplankton cells (Suttle, 2005; Breitbart, 2012), and during 'sloppy' feeding by zooplankton (Møller, 2005). Bacteria are also able to extract DOM from particulate organic matter (POM) by secreting extracellular hydrolases. These enzymes are capable of more than satisfying the carbon demand of particle-attached bacteria (Smith *et al.*, 1992), resulting in excess DOM being liberated to the surrounding seawater. This hydrolysis is extremely efficient so that the majority of sinking particles are broken up within the mixed layer, thus enhancing organic matter and nutrient retention within the surface.

By mediating flows of particulate and dissolved organic matter in this way, heterotrophic bacteria exert considerable influence on several components of the marine carbon cycle. The biological carbon pump describes the process by which organic matter can be removed from direct contact with the atmosphere as a result of sinking particles transporting carbon to the deep ocean (Ridgewell and Arndt, 2015). When particles sink out of the mixed layer before being degraded their carbon is

locked away from contact with the atmosphere for a considerable time, possibly up to 400 – 500 years (the average overturning time of the oceans) (Lozier, 2010). The majority of particles that sink into the deep ocean are remineralised before reaching the bottom (Smith *et al.*, 1992; Azam *et al.*, 1994). A small fraction, around 1 - 6% (Dunne *et al.*, 2007), does make it to the sediment and of this fraction, only around 0.3% escapes degradation for long enough to be buried deep enough to avoid direct exchange with the overlying water column. As a result, this carbon is locked away for geological time.

While the majority of DOM is remineralised by heterotrophic bacteria in the surface ocean, a small fraction is seemingly resistant to degradation (Dittmar, 2015). This 'refractory' DOM can survive for thousands of years or multiple oceanic overturning cycles. As a result of its longevity, this pool is the largest reservoir of organic carbon in the oceans – at 630 ± 12 PgC it is similar in size to the entire atmospheric carbon pool (Hansell *et al.*, 2009). It therefore represents a substantial sequestration of carbon away from the atmosphere: the conversion of freshly fixed carbon into refractory DOM has been termed the microbial carbon pump (Jiao *et al.*, 2010; Jiao and Zheng, 2011). The molecular complexity of refractory DOM has proven to be a major obstacle to determining its composition. However, elemental analyses have indicated that it is enriched in carbon relative to nitrogen and phosphorus when compared to more labile DOM pools (Dittmar, 2015). The ways in which microbial metabolism might influence the partitioning of organic matter between refractory and more labile pools has yet to be fully understood but could have significant implications for the global carbon cycle (Jiao and Zheng, 2011).

1.2 Major microbial drivers of marine ecosystems

1.2.1 Marine phytoplankton – An overview

Ocean primary productivity varies by several orders of magnitude from highly productive coastal waters to the extremely oligotrophic subtropical gyres (Behrenfeld

et al., 2006; Figure 1.2). These productivity gradients are driven primarily by differences in the availability of nutrients, although other factors, such as temperature and light intensity may also play a role. The composition of the phytoplankton community also varies considerably across both time and space. Cyanobacteria, numerically the most abundant phytoplankton, are represented in the oceans almost entirely by two genera, *Prochlorococcus* and *Synechococcus* (Scanlan *et al.*, 2009). Both are comprised of a number of ecotypes favouring particular ecological niches, thus allowing these groups to thrive over large areas of the oceans (Johnson *et al.*, 2006; Zwirglmaier *et al.*, 2007). For reasons that are not fully understood, *Prochlorococcus* appears to be entirely absent from polar seas, while *Synechococcus* is also generally absent from waters with temperatures below 0 °C (Flombaum *et al.*, 2013), though see Paulsen *et al.* (2016). Cyanobacteria, and particularly *Prochlorococcus*, tend to dominate photosynthesis in the open oceans of the tropics and subtropics; in more productive coastal waters and in high latitude seas, larger eukaryotic algae such as diatoms tend to be more abundant (Follows *et al.*, 2007; Armbrust, 2009). During the summer season, when light levels are sufficient to support photosynthesis, diatoms dominate the polar oceans, and particularly the Southern Ocean (Arrigo, 1999). At temperate latitudes, fast growing eukaryotic phytoplankton, including diatoms and the coccolithoporid *Emiliana huxleyi*, often dominate spring blooms, which can be large enough to be visible from space (Behrenfeld and Boss, 2014). These are thought to be stimulated by the presence of high levels of nutrients in surface waters following winter mixing coupled with an initial absence of predation pressure (Behrenfeld and Boss, 2014).

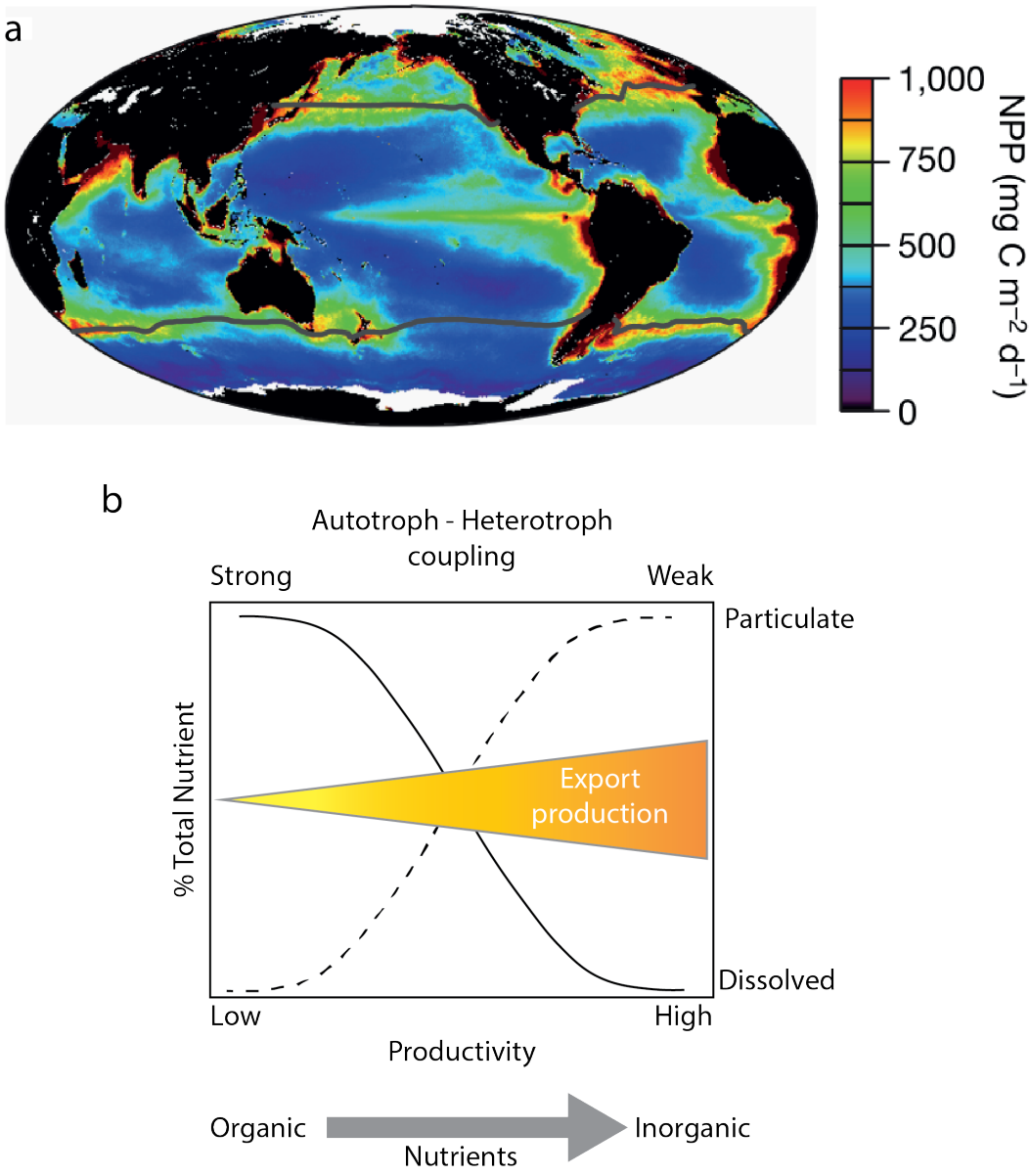


Figure 1.1. Productivity in the oceans. **a)** Annual mean net primary production (NPP), inferred by satellite measurements. From Behrenfeld *et al.* (2006). **b)** Conceptual model showing changes in ecosystem and nutrient characteristics across a productivity gradient in pelagic aquatic ecosystems. In low productivity ecosystems, most nutrients are in the dissolved phase, and most primary production is rapidly consumed by heterotrophs. Nutrients are primarily present in organic molecules. In highly productive systems, a greater proportion of nutrients are present in larger particles, which can be more rapidly transported from the surface. As a result, a smaller fraction of primary production is consumed by heterotrophs. Adapted from Cotner and Biddanda (2002).

1.2.2 Heterotrophic bacteria in the oceans

As is the case for phytoplankton, heterotrophic bacterial abundance varies along broad environmental gradients ultimately determined by factors such as nutrient availability. Bacterioplankton numbers are broadly constrained by primary productivity (Church, 2008), although in oligotrophic regions of the oceans, bacterial production appears to exceed that of phytoplankton (Del Giorgio *et al.*, 1997). In more eutrophic ecosystems, bacteria consume a smaller fraction of primary production (Figure 1.1) due to a combination of competition from phytoplankton for nutrients and carbon being more efficiently removed from the surface as a result of the fast sinking of larger phytoplankton (Cotner and Biddanda, 2002). The combination of varying physical and chemical conditions in the oceans, along with variations in the phytoplankton community, result in heterotrophic bacteria occupying a wide variety of ecological niches (Fuhrman, 2009). Nonetheless, many of the most abundant bacterial groups have strikingly cosmopolitan distributions in the surface ocean, with *Alphaproteobacteria, Gammaproteobacteria,* and *Bacteroidetes* together making up the majority of the bacterioplankton community globally (Wietz *et al.*, 2010). Within these higher phylogenetic groupings, a small number of ubiquitous clades of marine bacteria often dominate the heterotrophic assemblage (Giovannoni and Stingl, 2005).

The most abundant of all microbial groups in the oceans are the SAR11 clade bacteria (Morris *et al.*, 2002). This group of *Alphaproteobacteria* is characterised by an extremely small cell size and highly streamlined genome, rich in high affinity ABC-type transporters but lacking regulatory genes (Giovannoni *et al.*, 2005). This streamlining appears to be an attempt to minimise the energy and nutrient cost of replicating the genome (Giovannoni *et al.*, 2014), and is a strategy which appears to be shared with other abundant groups of bacteria in the oceans as well as the cyanobacterium *Prochlorococcus* (Partensky and Garczarek, 2010). This apparently strong selective pressure to dispense with functions that provide insufficient competitive advantage has resulted in several distinctive features of SAR11

physiology. Despite the high concentrations of sulfate in the oceans, SAR11 are unable to use it as a source of sulfur, instead relying on reduced forms of sulfur produced by other microorganisms (Tripp *et al.*, 2008). A similar theme emerges when examining the vitamin requirements of SAR11 strains. Several SAR11 strains analysed in one study lacked genes required for several steps in the biosynthesis of thiamin and thus required supply of a biosynthetic intermediate (Carini *et al.*, 2014). Similarly there is widespread auxotrophy for vitamin B12 amongst marine plankton, including SAR11 (Sañudo-Wilhelmy *et al.*, 2014). Findings such as these suggest an 'outsourcing' of the production of certain compounds, which may be energetically expensive to produce, to other members of the microbial community. As a result, SAR11 are presumably able to grow more efficiently but are also dependent on the presence of sufficient numbers of these producers (Morris *et al.*, 2012). Similar patterns of adaptive loss of essential functions seem to be widespread amongst the marine microbial community and may explain the difficulty in cultivating many abundant marine bacteria (Rappé and Giovannoni, 2003).

Despite individual SAR11 strains having little phenotypic plasticity, the clade as a whole is a major component of the bacterioplankton community in seemingly all ocean surface waters, down to several hundred meters depth (Pommier *et al.*, 2007; Schattenhofer *et al.*, 2009). This ability to thrive in a range of niches arises from the enormous genetic diversity contained within the clade: several sub-clades have been described which appear to have distinct niches (Carlson *et al.*, 2009; Vergin *et al.*, 2013). In a study conducted in the Sargasso Sea, within the oligotrophic North Atlantic Subtropical Gyre, SAR11 subclade Ia dominated the surface bacterioplankton community during the summer, when the site was highly stratified (Vergin *et al.*, 2013). By contrast, representatives from the subclades Ib and IIa dominated the community during winter when the mixed layer was deeper (Vergin *et al.*, 2013).

Another important group of marine *Alphaproteobacteria* is the Roseobacter group within the family *Rhodobacteraceae*, which can make up to 20% of the bacterial community, particularly in coastal waters (Buchan *et al.*, 2005; Wagner-Döbler and Biebl, 2006). While often referred to as a clade, the marine members of the Roseobacter lineage within the *Rhodobacteraceae* are not monophyletic (Simon *et al.*, 2017). Unlike SAR11 clade bacteria, which seemingly have entirely planktonic lifestyles, many Roseobacters can attach to surfaces (Slightom and Buchan, 2009) and they are commonly found associated with eukaryotic phytoplankton (Grossart *et al.*, 2005; Amin *et al.*, 2015). As a result, the Roseobacter group is often found to be abundant during phytoplankton blooms (Teeling *et al.*, 2012). Even in the open ocean, however, strains affiliated to the Roseobacter group are often prominent in the microbial community (Selje *et al.*, 2004; Wietz *et al.*, 2010). These pelagic strains appear to have a distinct lifestyle from their relatives in more eutrophic waters, sharing many of the characteristics of genome streamlining with SAR11 (Luo, 2014; Billerbeck *et al.*, 2016). A characteristic of Roseobacters is their ability to utilise a diverse range of substrates, including aromatic and methylated compounds (Newton *et al.*, 2010; Voget *et al.*, 2014). Particular attention has focused on the ability of Roseobacters to degrade the algal metabolite anti-stress molecule dimethylsulfoniopropionate (DMSP), which is produced in vast quantities by marine algae (Howard *et al.*, 2006; Todd *et al.*, 2007, 2009, 2012). One possible product of DMSP catabolism is dimethylsulfide, a volatile organo-sulfur compound whose release represents a major flux of sulfur from the oceans to the atmosphere (Kettle and Andreae, 2000). Roseobacters have also been used to study interactions and metabolic exchange between bacteria and phytoplankton. These studies have revealed mutualistic exchanges, such as production of vitamin B12 by the Roseobacter in exchange for algal metabolites (Durham *et al.*, 2014; Wang *et al.*, 2014; Amin *et al.*, 2015). Some Roseobacter strains can also switch from a mutualistic to an antagonistic relationship as the algae begin to age, producing algicidal compounds (Seyedsayamdost, Carr, *et al.*, 2011; Wang *et al.*, 2014). Thus Roseobacters may play a role in both the growth and collapse of phytoplankton blooms (Buchan *et al.*, 2014).

Although less well studied than their *Alphaproteobacterial* counterparts, *Gammaproteobacteria* are also abundant and globally distributed in the oceans. The SAR86 clade, in particular, is widely distributed across the ocean, typically accounting for 2–7% of bacterial cells (Schattenhofer *et al.*, 2009; Wietz *et al.*, 2010). Typically for pelagic bacteria with a planktonic lifestyle, SAR86 have small, streamlined genomes (Dupont *et al.*, 2012; Swan *et al.*, 2013) and lack the ability to synthesise several essential amino acids and vitamins. In line with the extremely dilute nature of seawater, SAR86 genomes encode numerous transporters to aid in the acquisition of sufficient nutrients (Lauro *et al.*, 2009). However, unlike the SAR11 clade, which appears to specialise in the uptake of highly labile, low molecular weight compounds, the SAR86 clade is characterised by a proliferation of TonB-dependent transporters (Dupont *et al.*, 2012). These transporters allow bacteria to transport larger molecules, such as vitamins, siderophores and carbohydrates across the outer membrane. In the case of the SAR86 clade, they are predicted to allow a specialisation in the consumption of lipids and carbohydrates (Dupont *et al.*, 2012), thus allowing SAR86 to avoid competition with SAR11.

Other abundant marine Gammaproteobacteria, such as the OM60 group are also predicted to specialise in the consumption of carbohydrates and lipids (Fuchs *et al.*, 2007). Unlike SAR86, this group tends to be more abundant in coastal waters (Yan *et al.*, 2009), and indeed *Gammaproteobacteria* are commonly associated with phytoplankton blooms (Schattenhofer *et al.*, 2009; Teeling *et al.*, 2012). Several gammaproteobacterial groups are also common in polar waters, particularly in the Southern Ocean, where SAR86 are seemingly less active (Ghiglione *et al.*, 2012; Nikrad *et al.*, 2014).

Flavobacteria, from the *Flavobacteriaceae* family within the *Bacteroidetes* are another important component of the marine bacterioplankton community (Pommier *et al.*, 2007; Schattenhofer *et al.*, 2009). Although found globally, they tend to be especially important in productive coastal environments (Cottrell and Kirchman, 2000; Alonso

et al., 2007) and in polar waters (Williams *et al.*, 2013). A large number of TonB-dependant transporters is predicted to enable marine flavobacteria to degrade biopolymers such as peptides and polysaccharides (Cottrell and Kirchman, 2000; Kabisch *et al.*, 2014; Xing *et al.*, 2014). In so doing they may render this organic material more available for other members of the bacterial community (Williams *et al.*, 2013).

A variety of other groups of bacteria show cosmopolitan distributions in the oceans, and are likely to play important roles in surface ocean ecology. These include *Actinobacteria* of the uncultured OM1 clade, which are commonly found in the open ocean at abundances of up to 5% of the bacterial community (Jensen and Lauro, 2008; Morris *et al.*, 2012). Little is known about the physiology of these organisms, although the genome of one marine actinobacterium was partially reconstructed using metagenomic data (Ghai *et al.*, 2013). This revealed a streamlined, low GC genome characteristic of adaptation to a pelagic lifestyle. Among other bacteria, *Verrucomicrobia* also display a global distribution (Freitas *et al.*, 2012), while planctomycetes are commonly detected, though usually minor, members of the microbial community (Orsi *et al.*, 2016). Annamox *Planctomycetales* do play an important role in the marine nitrogen cycle within oxygen minimum zones (Ganesh *et al.*, 2014).

1.3 Nutrient limitation in the oceans

1.3.1 The Redfield ratio

All life requires a number of essential elements, or nutrients, to survive (Merchant and Helmann, 2012). Of the macronutrient elements, needed to make bulk components of the cell, nitrogen (N) and phosphorus (P) are present in the oceans at the lowest concentrations relative to the requirements for life (Moore *et al.*, 2013). Early studies revealed a remarkably consistent relationship between the ratio of N:P in plankton biomass and that of inorganic N:P in the deep ocean (Redfield, 1958).

This ratio of C:N:P (106:16:1) has become known as the 'Redfield ratio' and is a key tenet of ocean biogeochemistry. Subsequent work has shown that the N:P of planktonic biomass is not constant across the oceans but varies systematically with latitude (Martiny *et al.*, 2013), with the Redfield ratio being a global average. The origins of varying N:P stoichiometry in living organisms lies in differing abundances of various macromolecular pools in the cell, which differ in their elemental composition (Sterner and Elser, 2002). For example, proteins and RNA have similar N contents but RNA also contains P whereas proteins contain a negligible amount (Geider and La Roche, 2002). Therefore, organisms with higher contents of RNA, of which ribosomal RNA is usually the major component, will tend to be richer in P relative to N so will have a lower N:P (Sterner and Elser, 2002). Since the ribosomes are the biosynthetic centres of the cell, these organisms would be expected to grow faster relative to organisms with less allocation towards ribosomes. Hence the 'growth rate hypothesis' proposes that growth rate increases as N:P decreases (Elser *et al.*, 1995; Sterner and Elser, 2002). Subsequent tests of this theory have shown that the predicted positive relationship between P content and growth rate holds for a diverse range of organisms, from bacteria to insects (Elser *et al.*, 2003). Inspired by this approach, theoretical approaches to understanding phytoplankton stoichiometry have conceptualised N:P as resulting from a trade-off between N-rich transport proteins ('resource acquisition') versus P-rich ribosomes (Arrigo, 2004; Klausmeier *et al.*, 2004). Intriguingly, a recent mathematical model of this trade-off derived an optimal ratio of protein to rRNA of 3 ± 0.7 under nutrient-replete conditions, translating to approximately the 16:1 of the Redfield ratio (Loladze and Elser, 2011). This suggests that there may be fundamental basis in the physiology of marine plankton for the Redfield ratio.

In the surface ocean, the ratio of inorganic N:P is on average 14.3:1, slightly lower than the Redfield ratio (Deutsch and Weber, 2012). As a result, N is widely believed to be the most important limiting nutrient in the oceans, a supposition which has been supported by various nutrient addition studies (Moore *et al.*, 2013). This may

seem surprising, given that there is an effectively inexhaustible supply of nitrogen in the atmosphere, which can be converted to a bioavailable form by nitrogen-fixing bacteria (Zehr and Kudela, 2011). By contrast, the only significant sources of new phosphorus to the oceans are from riverine transport and dust deposition (Schlesinger and Bernhardt, 2013), which are both primarily controlled by geological processes. To address the issue of why N, rather than P, appeared to limit productivity in much of the ocean, Tyrell (1999) proposed a distinction between the ultimate limiting nutrient – determining primary production over geological time scales – and the proximal one. Splitting the triple bond in dinitrogen to incorporate it into biomolecules is an energetically costly process. As a result, nitrogen-fixing bacteria are only competitive when N is deficient relative to other nutrients, such as P (Deutsch and Weber, 2012). If N-fixing activity resulted in a surplus of N, the N-fixers would be out-competed by other organisms. Complicating this picture, the nitrogenase enzyme, required for N-fixation, requires iron in its active site (Kustka *et al.*, 2003). As a result, N-fixation is often limited by iron availability (Chappell *et al.*, 2012). Support for the importance of Fe-limitation of N-fixation is found in the fact that the most P-depleted areas of the ocean are downwind of major deserts and thus receive substantial deposition of iron-rich dust (Jickells *et al.*, 2005). The subtropical north Atlantic, for example, receives substantial dust blowing from the Sahara which may explain the severe P-depletion observed there (Wu *et al.*, 2000) and the observation that N-fixers in the region are P-limited (Sañudo-Wilhelmy *et al.*, 2001).

The complex interplay of multiple nutrient cycles discussed above serves to illustrate the emerging concept that co-limitation by multiple nutrients may be the norm in the open ocean (Saito *et al.*, 2008; Moore *et al.*, 2013). Classical concepts of nutrient limitation had focused on the importance of a single limiting nutrient, typified by Liebig's 'Law of the Minimum', whereby the nutrient present in the least amount, relative to the needs of the consumer, would be limiting (De Baar, 1994). However, numerous studies have since found that the addition of multiple nutrients is required to stimulate the growth of natural communities (e.g. Thingstad *et al.*, 2005; Mills *et*

al., 2008). Co-limitation could emerge from a number of processes, acting both at the level of individual cells and at the community level. Saito *et al.* (2008) proposed three independent mechanisms by which an individual cell could experience nutrient co-limitation. First, independent co-limitation could arise if the stoichiometry of two elements in an organism was exactly equivalent to their stoichiometry in the environment. Biochemical substitution co-limitation could arise when two elements can perform the same role within a cell. This is most commonly observed for metal ions which can substitute for one another in the active site of an enzyme, or alternatively are present in two enzymes which carry out the same function. Perhaps the best-studied example of this form of co-limitation is zinc-cobalt co-limitation of carbonic anhydrase in phytoplankton (Saito and Goepfert, 2008). Finally, dependent co-limitation could arise where the acquisition of one nutrient element was dependent on another – for example, zinc-dependent alkaline phosphatase activity resulting in zinc-phosphorus co-limitation (Shaked *et al.*, 2006). In addition to these processes, community-level co-limitation could result from different members of the microbial community being limited by different nutrients. An obvious example of this is the limitation of N-fixers by P and Fe (Mills *et al.*, 2004) while other phytoplankton are limited by N availability. A recent study also demonstrated that different members of the bacterioplankton community responded to the addition of different nutrients (Sebastián and Gasol, 2013).

1.3.2 Responses to phosphorus limitation in bacteria

The archetypal regulator of the bacterial response to phosphate stress is the two-component PhoBR sensor kinase/response regulator from *E. coli* (Hsieh and Wanner, 2010). Other regulators have been described, which may act synergistically with, or independently of, PhoBR. For example, the phosphate-responsive PtrA, a cAMP receptor protein described in *Synechococcus,* is found in a number of strains, in some cases without an accompanying PhoBR (Ostrowski *et al.,* 2010). The Pho regulon, the suite of genes regulated according to phosphate availability by PhoBR, is, however, by far the best studied system for phosphate-responsive regulation in bacteria. In most

cases genes in the Pho regulon are preceded by a Pho box, two 11-nucleotide repeats to which PhoB can bind (Santos-Beneit, 2015). While at least 31 genes have been experimentally confirmed as being part of the Pho regulon in *E. coli* (Hsieh and Wanner, 2010), proteomics studies have suggested that hundreds more may be regulated by phosphate availability (Van Bogelen *et al.,* 1996). Due to the conserved nature of the Pho box, it can be computationally predicted in a range of bacteria (Yuan *et al.*, 2006). This enables a broader picture of the genes likely to be regulated by phosphate availability in these organisms. Pho regulon genes in a number of *Proteobacteria* include predicted components of the high-affinity phosphate uptake system, genes required for the metabolism of organophosphorus compounds, and genes for the biosynthesis of phosphorus-free lipids (Yuan, *et al.* 2006). These genes fit into two generally observed responses to nutrient limitation: an enhanced scavenging response, to increase the supply of nutrient to the cell; and a nutrient sparing response, to reduce the cell's requirement for that nutrient (Merchant, 2012). These complementary responses will be described in more detail in the following sections.

1.3.3 The enhanced scavenging response

Inorganic phosphate (PO_4^{3-}; Pi) represents the preferred source of P for most organisms. Hence one of the most common responses to P-limitation is the increased expression of components of the high-affinity Pi uptake transporter (Krol and Becker, 2004; Cox and Saito, 2013). This is particularly the case for the periplasmic phosphate-binding subunit, PstS (Scanlan *et al.*, 1993). This response enables microorganisms to more efficiently take up low concentrations of Pi in their environment. In the modern oceans, Pi is by far the dominant form of inorganic P (Karl, 2014). However, the ability to utilise the reduced form, phosphite (PO_3^{2-}), has been identified in some cyanobacteria from P-depleted regions (Feingersch *et al.*, 2012; Martínez *et al.*, 2012). Thus, these organisms appear able to expand the pool of P which is accessible to them, enabling them to better compete in these P-depleted environments.

The bulk of the P present in open ocean environments is present as dissolved organic P (DOP; Kolowith *et al.*, 2001). This may be present at nanomolar concentrations even when Pi is undetectable (Karl and Bjorkman, 2014). Therefore, the ability to access this P reservoir would seem to be an essential adaptation to life in some of the most P-depleted areas of the ocean. One of the most commonly detected genes which enable bacteria to access this reserve of P are alkaline phosphatases. Alkaline phosphatases are cytoplasmic, periplasmic or cell-surface associated proteins (Luo *et al.*, 2009) which show strong activity towards phosphate monoesters. Many alkaline phosphatases have quite loose substrate specificity and also show activity against phosphate diesters, such as is found in the DNA backbone and intact phospholipids (O'Brien and Herschlag, 2001). The most common genes exhibiting alkaline phosphatase activity in the surface ocean are PhoA and PhoX (Sebastian and Ammerman, 2009). Intriguingly, high alkaline phosphatase activity is also observed in the deep ocean, where phosphate concentrations are high (Koike and Nagata, 1997; Hoppe and Ullrich, 1999). This suggests that alkaline phosphatases may play roles beyond the provision of an additional source of P.

Although the majority of P in living organisms is present as a phosphate ester (C-O-P), many organisms are also capable of producing phosphonates (with a C-P bond) (Dyhrman *et al.*, 2009; McGrath *et al.*, 2013). These phosphonates seem to be highly enriched in marine dissolved organic matter, where they comprise 25% of total organic P (Clark *et al.*, 1998; Kolowith *et al.*, 2001). This is considerably more than in particulate organic matter sampled, suggesting that phosphate esters are preferentially degraded by marine bacteria (Kolowith *et al.*, 2001). Thus, phosphonates represent a major reservoir for P, and genes involved in phosphonate utilisation are common in P-depleted areas of the ocean (Coleman and Chisholm, 2010). In these environments, a wide variety of organisms seem able to degrade phosphonates, including cyanobacteria and diazotrophs (Dyhrman *et al.*, 2006; Feingersch *et al.*, 2012; Villarreal-Chiu *et al.*, 2012). Release of methane by SAR11 bacteria while metabolising methylphosphonate may contribute to the

supersaturation of waters in P-depleted oligotrophic regions with methane (Carini, White, *et al.*, 2014).

1.3.4 The phosphorus-sparing response

In addition to expressing a range of genes which enable them to access an expanded reservoir of P, bacteria respond to P-limitation by reducing the cellular requirement for phosphorus. Over the long term, bacteria may adapt to P-deficiency by reducing the size of their genome, as is predicted by genome streamlining theory (Giovannoni *et al.*, 2014). However, there are a number of shorter term changes by which bacteria can substantially reduce their P requirements. In fact bacteria from natural environments appear to have a remarkable plasticity in terms of their requirements for P. Bacterial communities from freshwater lakes were able to modify their N:P to match the stoichiometry of the supplied nutrients (Cotner *et al.*, 2010; Scott *et al.*, 2012), while further studies revealed a remarkable ability to reduce their phosphorus requirements to achieve N:P in excess of 1000:1 (Godwin and Cotner, 2015).

Table 1.1 Average stoichiometry and the percentage of total cell mass for selected cellular components.[a]

	% cell mass	**C:N**	**C:P**	**N:P**
Protein and free amino acids	30-65	3.3	-	-
RNA	3-15	2.2	3.7	1.7
DNA	0.5-3	2.3	3.8	1.7
Phospholipids	5-15	80	15	0.19
ATP	< 0.1	1.7	1.3	0.78

[a]Based on data collated from studies of a range of eukaryotic and cyanobacterial phytoplankton species by Geider and LaRoche (2002)

The majority of phosphorus in cells is typically present in RNA, DNA and phospholipids (Geider and La Roche, 2002; Sterner and Elser, 2002). Although small molecules such as adenosine triphosphate (ATP) are rich in P, their low abundance in the cell means that they represent only a small fraction of cellular P (Table 1.1.1). Polyphosphate represents an additional, but poorly characterised, cellular reservoir for P (Rao *et al.*, 2009). Under favourable growth conditions, in culture, ribosomal-P typically represents the largest fraction of P in cells (Geider and La Roche, 2002). Thus, one means of reducing cellular P requirements is to reduce the number of ribosomes, which would likely result in reduced growth rate. Studies have shown an increase in ribosomal-P with lower N:P nutrient supply (i.e. higher supply of P), while growth rate tends to be proportional to P-supply across a range of species (Elser *et al.*, 2003). In line with this hypothesis, in the extremely P-depleted Sargasso Sea, ribosomal-P was correlated with the rate of P supply (Zimmerman *et al.*, 2014).

Polyphosphate is commonly thought of as a form of P storage molecule. This is partly due to bacteria such as *E. coli* incorporating excess P into polyphosphate when grown with high P concentrations (Kornberg *et al.*, 1999). Therefore it would be expected that polyphosphate should be broken down upon P-limitation to liberate P for other purposes. However, polyphosphate has diverse roles in many cells beyond that of a storage molecule (Kornberg *et al.*, 1999). In fact high amounts of polyphosphate was observed in marine microbial communities in P-depleted waters, indicating that it must play an important role in the physiology of marine bacteria (Martin *et al.*, 2014). The precise role of polyphosphate in the cell is still poorly understood, but it seems to be more than a simple storage polymer and its relationship to P stress is not straightforward.

1.3.5 Phospholipids and lipid remodelling

Phospholipids represent a further reservoir of P within the cell (Table 1.1.1), which may account for 15-20% of total cellular phosphorus of bacteria in culture (Karl and Bjorkman, 2014). In natural environments with low P-concentrations, this proportion could well be higher. Marine phytoplankton have been observed to replace phospholipids with a range of phosphorus-free lipids, including sulfolipids and betaine lipids, in response to phosphorus limitation (Van Mooy *et al.*, 2006, 2009). This lipid remodelling process appears to be rapid, occurring within 24 hours of the onset of P limitation in one diatom strain, and reversible (Martin *et al.*, 2011). Similar responses have also been observed in land plants (Abida *et al.*, 2015), where it has been proposed that phospholipids could represent a means of storing phosphorus within the membrane (Frentzen, 2004). Cultured bacteria have also been observed to replace phospholipids with phosphorus-free lipids in response to P-limitation (Minnikin and Abdolrahimzadeh, 1974a; Minnikin *et al.*, 1974). However, combined culture and metagenomics studies failed to find evidence for heterotrophic bacterial genes known to be involved in the synthesis of phosphorus-free lipids in the marine environment (Van Mooy *et al.*, 2009). This suggested that heterotrophic bacteria may be at a competitive disadvantage against phytoplankton in P-depleted environments.

1.4 Membrane lipids in bacteria

1.4.1 Phospholipid biosynthesis and function

The classical view of the bacterial plasma membrane holds that it is composed primarily of the phospholipids phosphatidylethanolamine (PE) and phosphatidylglycerol (PG), with smaller amounts of cardiolipin (CL). This reflected the fact that knowledge of bacterial lipids is derived overwhelmingly from studies of *Escherichia coli* and *Bacillus subtilis* (Zhang and Rock, 2008). In this simplified view, the negatively charged PG and CL provided the net negative charge necessary for many essential processes in bacteria, such as cell division (Dowhan, 1997).

Meanwhile, PE, which is zwitterionic, diluted the negative charge conferred by the other two lipids. Due to its small head group, the three dimensional structure of PE is conical meaning that PE does not spontaneously assemble into a bilayer. A balance between bilayer and nonbilayer-forming lipids seems also to be important to bacterial fitness (Dowhan, 1997). Another phospholipid, phosphatidylcholine (PC), which is often considered more characteristic of eukaryotic membranes, is also fairly common in bacteria (Sohlenkamp *et al.*, 2003). A number of lipids are the result of modifications of phospholipids, such as the decoration of PG with lysine (Peschel *et al.*, 2001) to decrease the negative charge of the membrane as a protection against cationic antimicrobial peptides.

The biosynthesis of phospholipids has been reviewed in detail elsewhere (Zhang and Rock, 2008; Parsons and Rock, 2013) but key details will be presented below. An outline of the synthesis of the two major phospholipids in *E. coli* is presented in Figure 1.2. The key precursor for phospholipid synthesis is phosphatidic acid, which is itself formed by two sequential acylations of glycerol-3-phosphate (Kennedy, 1957; Lu *et al.*, 2006). Phosphatidic acid is primed for phospholipid synthesis by the addition of cytidine diphosphate (CDP) to form CDP-diacylglycerol (Ganong *et al.*, 1980), which serves as a substrate for the enzymes involved in the first committed steps of both PE and PG biosynthesis. To form PG, cytidine monophosphate (CMP) is replaced by glycerol-3-phosphate in a reaction mediated by PgsA (Miyazaki *et al.*, 1985). The resulting phosphatidylglycerol phosphate (PGP) is then dephosphorylated by one of three PGP phosphatases to form PG (Lu *et al.*, 2011). PG can further be used as a substrate for a number of cardiolipin synthases in the production of cardiolipin, which usually requires either the condensation of two PG molecules (Nishijima *et al.*, 1988; Guo and Tropp, 2000) or condensation of PG with PE (Tan *et al.*, 2012). Alternatively, CDP-DAG can serve as a substrate for phosphatidylserine synthase (PssA), which displaces CMP with L-serine to form phosphatidylserine (Ohta and Shibuya, 1977; DeChavigny *et al.*, 1991). Although an important lipid in eukaryotic membranes (Leventis and Grinstein, 2010), PS seems to be primarily a biosynthetic

intermediate in *E. coli* and is present at extremely low levels in the membrane (Ames, 1968). PE is then formed from PS by the action of PS decarboxylase (Kanfer and Kennedy, 1964; Hawrot and Kennedy, 1978). While most of the enzymes involved in the synthesis of PG and PE are integral membrane proteins, PssA is membrane-associated, with a high affinity for anionic regions (Louie *et al.*, 1986). Therefore, its affinity for the membrane, and consequently the rate of PE synthesis, is determined by the fraction of anionic lipids such as PG, thus maintaining the appropriate charge balance in the membrane (Zhang and Rock, 2008). In some bacteria that produce PC, CDP-DAG is also used as a substrate for PC synthase (Pcs), which condenses choline onto CDP-DAG with loss of CMP (De Rudder *et al.*, 1997, 1999). How this pathway competes with those for the formation of PE and PG has yet to be elucidated. An alternative path to PC formation in bacteria is the sequential *N*-methylation of PE three times, either by a single enzyme, such as PmtA in *Ensifer meliloti* (De Rudder *et al.*, 1997; Sohlenkamp *et al.*, 2003), or by multiple enzymes (Hacker *et al.*, 2008).

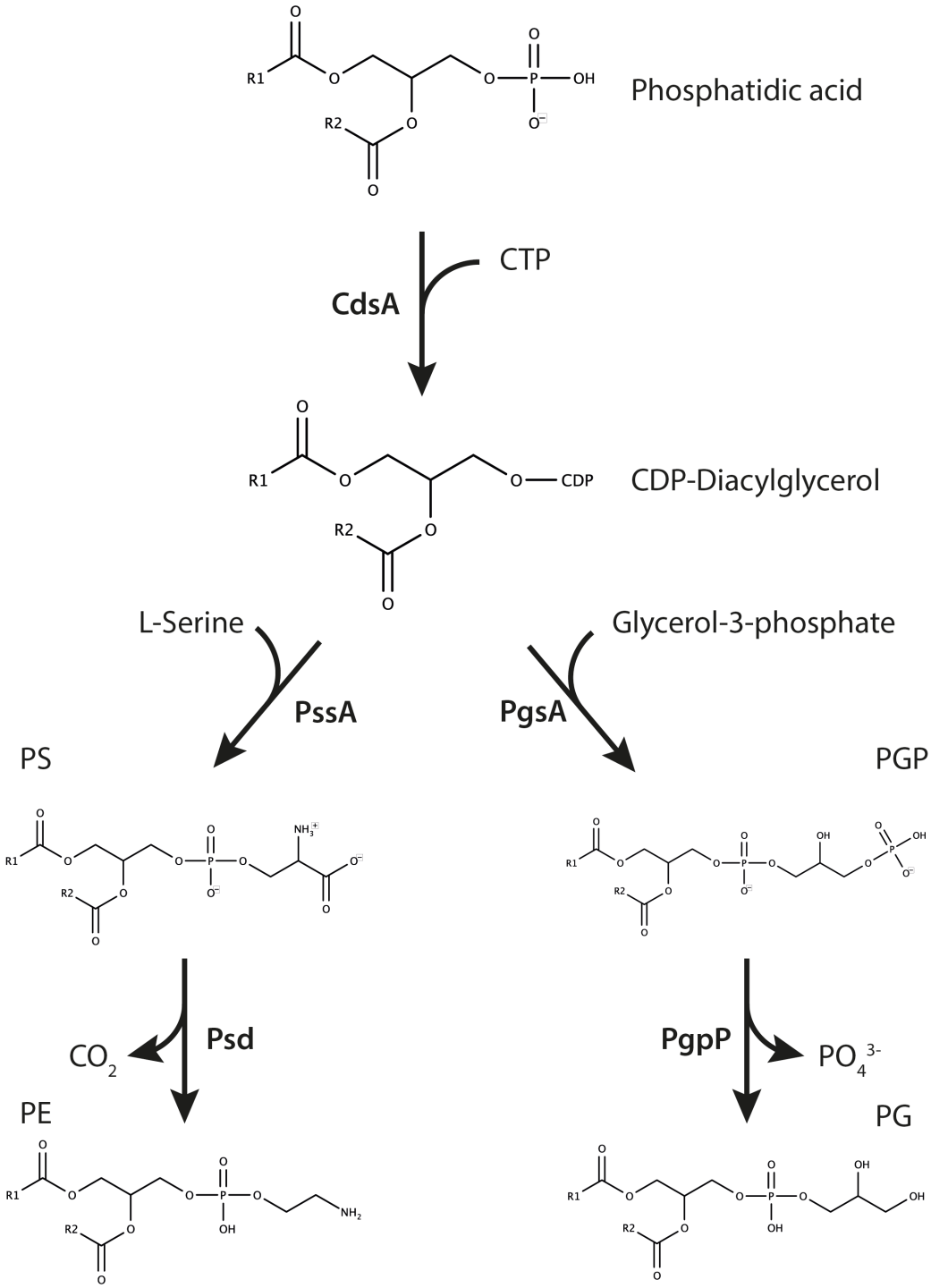
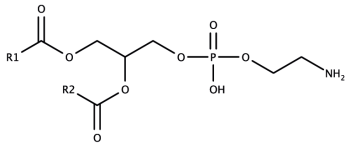


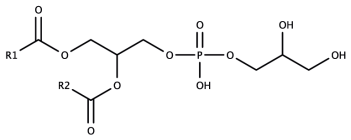
Figure 1.2 Biosynthesis of the two major phospholipids in *Escherichia coli*. Both phosphatidylethanolamine (PE) and phosphatidylglycerol (PG) are synthesised from phosphatidic acid and cytidine diphosphate–diacylglycerol (CDP-DAG) precursors. PE biosynthesis proceeds via phosphatidylserine (PS), while PG is synthesised via phosphatidylglycerol phosphate (PGP). Adapted from Zhang and Rock (2008).

1.4.2 Other glycerolipids in bacteria

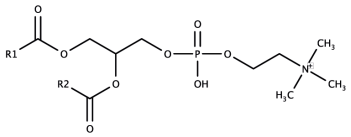
Increasingly, studies of a more diverse range of bacteria have revealed a huge variety of bacterial lipids (Sohlenkamp and Geiger, 2015). Many lipids share the hydrophobic diacylglycerol (DAG) backbone with phospholipids but lack phosphate (Fahy *et al.*, 2005). These include various glycolipids, which are also common in plant plastid membranes, with head groups typically comprising 1 to 3 glucose or galactose monomers (Hölzl and Dörmann, 2007). Indeed, a wide variety of glycosyltransferases involved in glycolipid synthesis have been characterised in bacteria, across a range of phyla (Hölzl and Dörmann, 2007). One interesting model organism for lipid bilayer studies, *Acholeplasma laidlawii*, has a membrane composed primarily of two glycolipids (Figure 1.3): monoglucosyldiacylglycerol (MGDG) and diglucosyldiacylglycerol (DGDG). It lacks a cell wall and is unable to synthesise its own fatty acids (Lindblom *et al.*, 1986). Thus, the fatty acid composition of its membrane can be set by supplying them exogenously. The bacterium therefore maintains membrane properties solely by controlling polar head group composition. Membrane curvature is maintained by a balance between the non-bilayer prone MGDG and the bilayer-forming DGDG (Lindblom *et al.*, 1986). MGDG thus plays a similar role in the membrane to PE and can in fact substitute for it in the membranes of *E. coli* mutants unable to synthesise PE, despite not being a native *E. coli* lipid (Wikström *et al.*, 2004).



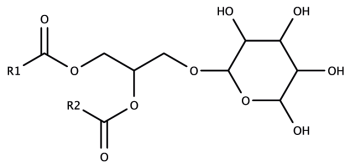
Phosphatidylethanolamine
(PE)



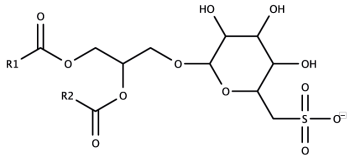
Phosphatidylglycerol
(PG)



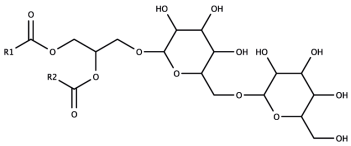
Phosphatidylcholine
(PC)



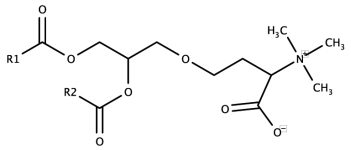
Monoglycosyl diacylglycerol
(MGDG)



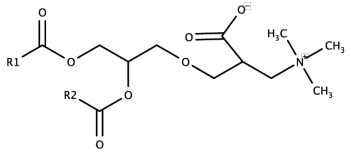
Sulfoquinovosyl diacylglycerol
(SQDG)



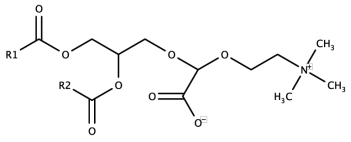
Diglycosyl diacylglycerol
(DGDG)



Diacylglyceryl trimethylhomoserine
(DGTS)



Diacylglyceryl hydroxymethyl-
trimethylalanine
(DGTA)



Diacylglyceryl carboxyhydroxy-
methylcholine
(DGCC)

Figure 1.3 Chemical structures of selected glycerolipids, included selected glycerophospholipids.

The sulfolipid sulfoquinovosyldiacylglycerol (SQDG; Figure 1.3) is one of the most abundant sulfur-containing organic molecules on earth as a result of its ubiquity in the photosynthetic membranes of plants and algae (Roy *et al.*, 2003). Although best-studied in the context of oxygenic photosynthesis, the biosynthesis of SQDG was first elucidated in the photoheterotrophic bacterium *Rhodobacter sphaeroides* (Benning *et al.*, 1995; Benning, 1998). Genes for the synthesis of SQDG have since been found in a number of heterotrophic species (Villanueva *et al.*, 2013), indicating that it is not unique to photosynthetic membranes. Since SQDG is negatively charged under physiological conditions, it may be able to substitute for PG under certain conditions, such as when P is limiting (Van Mooy *et al.*, 2006).

Another widespread group of phosphorus-free glycerolipids are the betaine lipids, which are common components of the membranes of algae, fungi and lower plants (Kato *et al.*, 1996; Kunzler and Eichenberger, 1997). There are three main classes of betaine lipid (Figure 1.3): diacylglyceryl trimethylhomoserine (DGTS), diacylglyceryl hydroxymethyl trimethyl-β-alanine (DGTA) and diacylgyceryl carboxy hydroxymethyl choline (DGCC; Dembitsky, 1996). One of these lipids, DGTS, has also been detected under P-limited growth conditions in bacterial membranes (Benning *et al.*, 1995). The biosynthesis of DGTS has been described in *R. sphaeroides* as a two-step process, mediated by BtaA and BtaB and requiring S-adenosylmethionine and diacylglycerol (Klug and Benning, 2001; Riekhof *et al.*, 2005). The chemical structure of DGTS is notably similar to that of PC, as both are zwitterionic, with a trimethylated quaternary nitrogen atom. Functional substitution of DGTS for PC has been demonstrated in a strain of the yeast *Saccharomyces cerevisiae* lacking the ability to synthesise PC (Riekhof *et al.*, 2014).

These diverse phosphorus-free glycerolipids, share a common requirement for DAG in their biosynthesis. However, DAG is not generated through the canonical phospholipid biosynthesis pathway, which instead proceeds via phosphatidic acid (Figure 1.2). DAG can be generated in *E. coli* by phospholipases which act on intact

phospholipids, such as the transfer of glycerol-1-phosphate from PG to periplasmic oligosaccharides, known as membrane-derived oligosaccharides (Goldberg *et al.*, 1981). The DAG by-product of this transfer can be re-phosphorylated by a diacylglycerol kinase to produce phosphatidic acid (Raetz and Newman, 1979). A similar method for producing DAG was described in *Ensifer meliloti* (formerly *Sinorhizobium meliloti*) where a phospholipase C, PlcP, degraded zwitterionic phospholipids in response to P stress (Zavaleta-Pastor *et al.*, 2010). This DAG could then serve as a substrate in the synthesis of multiple phosphorus-free lipids (Figure 1.4), while the liberated phosphoalcohol could presumably serve as a source of P for more essential cellular processes. In cyanobacteria and chloroplasts, the production of DAG directly from phosphatidic acid via phosphatidic acid phosphatases has also been described (Nakamura *et al.*, 2007). Cyanobacteria produce SQDG and glycolipids constitutively as part of their photosynthetic membranes. This contrasts with many heterotrophic bacteria in which they are produced in response to specific environmental stresses, and perhaps explains the apparently more streamlined synthesis of P-free glycerolipids in cyanobacteria.

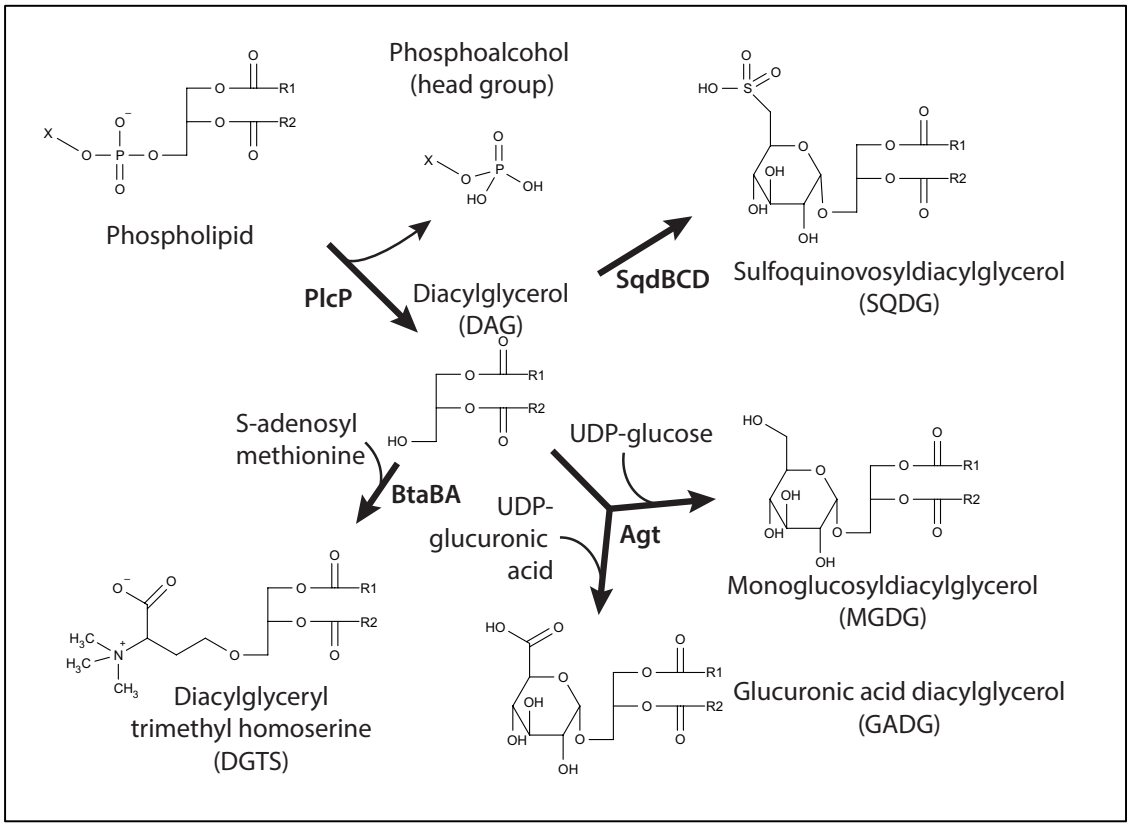


Figure 1.4. Proposed pathway of synthesis of non-phosphorus lipids through phospholipids and PlcP in marine heterotrophic bacteria. PlcP releases Diacylglycerol (DAG) from phospholipids which serves as a substrate by a variety of pathways. Three such pathways are shown here: the synthesis of the betaine lipid, diacylglyceryltrimethylhomoserine (DGTS) by BtaB and BtaA, the synthesis of the glycolipids monoglucosyldiacylglycerol (MGDG) and glucuronic acid diacylglycerol (GADG) by Agt, and the synthesis of sulfolipid sulfoquinovosyl diacylglycerol (SQDG) by SqdB, SqdC and SqdD.

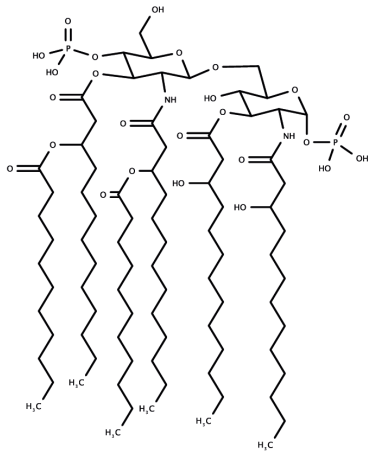
1.4.3 Lipid A and aminolipids

Lipids which lack the diacylglycerol backbone are widespread in bacteria, the best-studied of these being the lipid A component of lipopolysaccharide (LPS). LPS is located in the outer leaflet of the outer membrane of most Gram negative bacteria (Zhang *et al.*, 2013). The lipid A component of LPS is comprised of two glucosamine residues (Figure 1.5), each acylated on the amino and hydroxy groups of their 2nd and 3rd carbon atoms, respectively, with 3-hydroxy fatty acids (Whitfield and Trent, 2014). The fatty acids on one of the glucosamines are conjugated to additional "piggy back" fatty acids attached to the 3-hydroxy group (Figure 1.5), resulting in an overall hexa-acylated molecule (Whitfield and Trent, 2014). Unmodified lipid A is rarely found in the membranes of bacteria but rather serves as an anchor for the core oligosaccharide, while in some bacteria a further, highly variable, polysaccharide chain is added. The polar polysaccharide chains may help to defend gram negative cells from hydrophobic compounds, which otherwise would pass easily through the lipid membranes (Nikaido, 2003).

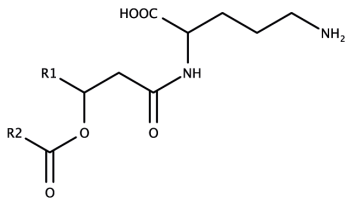
A related group of lipids, also common in bacteria, are the aminolipids such as ornithine lipid (Geiger *et al.*, 2010). Ornithine lipids are synthesised in a two step process initiated by N-acylation of the non-structural amino acid ornithine with a 3-hydroxy fatty acid by OlsB (Gao *et al.*, 2004). This is followed by the addition of a second, "piggy back", fatty acid to the 3-hydroxy group of the first, in a reaction mediated by OlsA (Weissenmayer *et al.*, 2002). The fatty acid structure of these lipids is similar to that of lipid A (Figure 1.5), and indeed OlsB is related to LpxD, which carries out the N-acylation of uridine diphosphate-*O*-3-acylglucosamine during lipid A biosynthesis (Gao *et al.*, 2004). An analogous biosynthesis can also be carried out by the bifunctional OlsF, which contains both N-acyltransferase and O-acyltransferase domains (Vences-Guzmán *et al.*, 2015). Given the structural similarities, ornithine lipid might exist alongside LPS in the outer leaflet of the outer membrane of gram negative bacteria (Nikaido, 2003). This hypothesis has been supported by studies showing an enrichment of ornithine lipid in the outer

membrane (Dees and Shively, 1982; Vences-Guzmán *et al.*, 2011). As a phosphorus-free lipid, ornithine lipid also appears to play a role in the response to P-limitation in a range of bacteria (Minnikin and Abdolrahimzadeh, 1974a; Geske *et al.*, 2013; Diercks *et al.*, 2015). A variety of hydroxylations of ornithine lipids may play a role in adaptation to acid stress (Vences-Guzmán *et al.*, 2012) by enabling a greater degree of hydrogen bonding between lipids, thus rendering the membrane less permeable. Several further studies have attempted to understand the roles of ornithine lipids in the cell. One study has observed a role in the maintenance of c-type cytochromes (Aygun-Sunar *et al.*, 2006), while another study observed impaired nodulation in an ornithine lipid-deficient mutant of *Rhizobium tropici* (Rojas-Jiménez *et al.*, 2005). However, a study in another *Rhizobium* species, *E. meliloti* found no evidence for a similar role in nodule formation (López-Lara *et al.*, 2005).

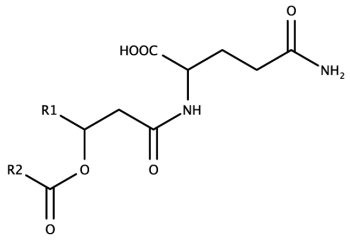
A number of other aminolipids, sharing a similar fatty acid structure to ornithine lipid, and with one or more amino acids in the polar head group, have also been described (Geiger *et al.*, 2010). Aminolipids with glutamine (Zhang *et al.*, 2009), lysine (Tahara, Yamada, *et al.*, 1976) and glycine (Kawazoe *et al.*, 1991) have been characterised in diverse bacteria (Figure 1.5). Although in none of these cases has the genes involved in their biosynthesis been described, several of the bacteria from which the lipids in question were extracted have homologs of OlsB or OlsF in their genomes. Thus, it is possible that various versions of these proteins exhibit different substrate specificities, synthesising different aminolipids (Geiger *et al.*, 2010). Further, more elaborate, aminolipids have also been described, with dual amino acid head groups. A tauro-ornithine lipid from *Gluconobacter cerinus* (named "cerilipin"; Tahara *et al.*, 1976) is synthesised by condensation of taurine onto ornithine lipid, although the enzymes involved are unknown (Tahara *et al.*, 1978). An analogous serine-glycine lipid was detected in a *Flavobacterium* and *Cytophaga marinus* (Batrakov *et al.*, 1998). How these diverse lipids fit into the scheme for lipid homeostasis and function developed in the simple *E. coli* model is not known. The fitness advantages conferred by these lipids in many cases remain mysterious.



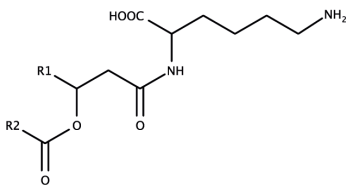
Lipid A



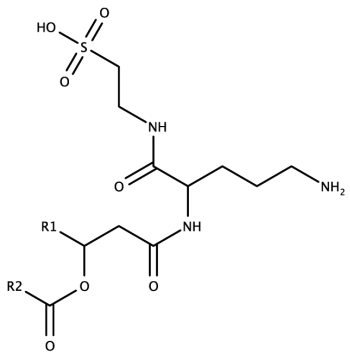
Ornithine lipid
(OL)



Glutamine lipid
(QL)



Lysine lipid
(KL)



Cerilipin
(Tauro-ornithine lipid)

Figure 1.5 Chemical structures of selected aminolipids and the lipid A component of lipopolysaccharide.

1.4.4 Lipids in the marine environment

Since they are presumed to be degraded rapidly, intact polar lipids (IPL) have been employed as tracers of living microbial biomass in the environment (Lipp *et al.*, 2008). Moreover, because certain lipid classes are restricted to particular groups of organisms, the study of the IPL profile of environmental samples has the potential to provide insights into the composition of the living microbial community. This is most evident when there is a well-established link between a lipid class and functional microbial group, for example between ladderane lipids and anammox bacteria (Jaeschke *et al.*, 2009). In contrast, a study conducted in marine surface waters found limited chemotaxonomic information in the polar lipid profiles (Brandsma *et al.*, 2012). Nonetheless, a number of lipid surveys have been carried out in marine surface and deep waters, which may provide information on the physiological state of the *in situ* community – for example, whether the community is experiencing P stress (Van Mooy *et al.*, 2006). One study along a transect in the Mediterranean found that the ratio of phospholipids to phosphorus-free lipids was correlated with phosphate concentration (Popendorf *et al.*, 2011). In particular, the sulfolipid SQDG and three betaine lipids were most abundant relative to phospholipids in the severely P-depleted Eastern Mediterranean. This was despite the community at the most P-depleted sites not being P-limited (Tanaka *et al.*, 2011). Another study in the similarly P-depleted Western North Atlantic found evidence for accumulation of P-free glycolipids in heterotrophic bacterial microcosms (Popendorf *et al.*, 2011). This would appear to contradict a previous assertion that major groups of heterotrophic marine bacteria were unable to produce P-free lipids (Van Mooy *et al.*, 2009). In the subtropical South Pacific, SQDG and glycolipids were suggested to primarily originate from phytoplankton, while heterotrophic bacteria were major contributors to pools of PG and PE (Van Mooy and Fredricks, 2010). The fatty acid profiles and depth distributions of betaine lipids, however, suggested a mixed source. Betaine lipids have traditionally been considered characteristic of eukaryotic algae (Kato *et al.*, 1996), despite more recent work demonstrating their occurrence in some bacteria (Benning *et al.*, 1995). Thus, recent surveys of polar lipids in the marine environment have

questioned common assumptions as to the predominant sources of some major classes of lipids.

Aminolipids have been studied much less frequently in environmental samples than glycerolipids. In lake waters, ornithine lipids were not detected at the surface but were present in the metalimnion and sediment (Ertefai *et al.*, 2008). Similarly, ornithine lipids were most abundant in anoxic waters of the Black Sea, and were not detected at the surface (Schubotz *et al.*, 2009). In both cases, sulfur oxidising bacteria were suggested as likely sources. This limited detection in environmental samples is surprising given how widespread the capacity to synthesise ornithine lipids is in bacteria (Vences-Guzmán *et al.*, 2015).

1.5 Recent advances in lipid analysis

Advances in mass spectrometry (MS) and liquid chromatography (LC) in recent decades have enabled the lipid content of organisms and environments to be studied in much greater detail than was previously possible. This has enabled the birth of a new field of lipidomics, which aims to study the entire set of lipids within a given sample (Shevchenko and Simons, 2010) following a similar philosophy to other 'omics disciplines such as genomics and metabolomics. Lipidomic studies have uncovered thousands of distinct molecular species of lipid even in relatively simple eukaryotic cells such as yeasts (Ejsing *et al.*, 2009). Increasingly, lipids have been recognised as dynamic components of the cell, with low abundance lipids playing an important role in cell signalling, including in some human diseases (Wymann and Schneiter, 2008). Moreover, lipids are not distributed uniformly across a cell: inhomogeneity arises through both active and passive processes. Lipids of similar thickness spontaneously sort themselves into clumps, resulting in regions with distinct chemical properties and lipid composition known as lipid rafts (Lingwood and Simons, 2010), which may play an important role in the spatial organisation of membrane proteins. In eukaryotes, different organelles tend to have different lipid

compositions (Van Meer *et al.*, 2008), which in part explains the enormous complexity of their lipidomes. In contrast, bacterial lipidomes seem to be substantially less complex (Gidden *et al.*, 2009), but evidence is accumulating for the presence of lipid aggregations similar to lipid rafts (López and Kolter, 2010; Bramkamp and Lopez, 2015). Negatively charged lipids may be localised to regions of high curvature in bacterial membranes (Renner and Weibel, 2011), helping to position the cell division apparatus.

1.5.1 Lipid extraction

From an analytical standpoint, lipids are usually functionally defined as the fraction of the cellular metabolites which are extracted in the hydrophobic phase of the two-phase Blight-Dyer (Bligh and Dyer, 1959) or Folch extraction techniques (Folch *et al.*, 1957). These methods involve partitioning cellular contents between an upper methanol-water phase and a lower chloroform phase, with proteins mostly precipitating at the interface. The material in the chloroform phase mostly consists of membrane lipids, although in some organisms a substantial fraction of the extracted material may be storage lipids, such as triacylglycerols, or polyhydroxybutanoate (Christie and Han, 2010). Some lipids with particularly large and hydrophilic head groups, including lipopolysaccharides, may be excluded. More recently, an alternative extraction method involving a single step extraction with methyl-*tert*-butylether has been developed which claims to be more amenable to high throughput studies (Matyash *et al.*, 2008).

1.5.2 Shotgun lipidomics

Mass spectrometers with high mass resolving power and accurate mass capabilities are theoretically capable of resolving complex lipid mixtures consisting of thousands or tens of thousands of lipid species (Russell and Edmondson, 1997; Han and Gross, 2005). This means that, in many cases, it is feasible to inject a lipid extract directly into a mass spectrometer without prior separation, a method referred to as 'shotgun'

lipidomics. If the mass accuracy of the instrument is sufficient, the molecular formula for the lipid may be calculated unambiguously from the mass of the intact ion. In some cases this may be sufficient to assign a structure. Alternatively, a number of mass spectrometric fragmentation strategies can also be used to determine the structure of lipids (Christie and Han, 2010). These rely on the commonalities in structure between all glycerolipids: characteristic fragment ions can usually be obtained for both fatty acids, and the head group can usually be determined by the formation of a characteristic fragment ion or its loss as a neutral fragment. In some cases the relative positions of fatty acid chains can also be determined based on the relative intensities of the ions corresponding to the two fatty acids (Hsu and Turk, 2000). The number of lipid species that can be detected by shotgun lipidomics can often be multiplied by changing the injection conditions so as to favour ionisation of different classes of lipids (Han *et al.*, 2006). One such strategy involves making the sample more basic by addition of lithium hydroxide, which results in better ionisation of PE relative to PG, and also results in the formation of lithium adducts of some lipids, which fragment in a more informative manner (Han and Gross, 2005).

1.5.3 Lipid analysis with chromatographic separation

In some samples, the lipid mixture may be too complex to resolve even using the most sophisticated mass spectrometers, or material may be present which would interfere with lipid ionisation, as is often the case with environmental samples. In these cases, and also when MS instruments with lower resolving power are to be used, chromatographic separation is typically used (Brouwers, 2011). Traditional lipid analysis has relied on thin layer chromatography (TLC) to separate lipids followed by staining with a range of dyes to determine lipid class (Fuchs *et al.*, 2011). Lipid classes can be quantified using this method, either by densitometric analysis of the spots following non-selective staining or by extracting the lipids from each spot and measuring fatty acids following formation of fatty acid methyl esters (Eder, 1995). While this TLC method is still widely used, it is relatively time-consuming and the

resolution attainable by a TLC plate is generally substantially worse than can be obtained by using a modern liquid chromatography (LC) column.

In modern high performance liquid chromatography (HPLC) applications, lipids are typically separated either using reversed phase chromatography, which separates lipids on the basis of their fatty acid chain length, or by normal phase or hydrophilic interaction chromatography (HILIC), which separate lipids according to their polar head groups (Sturt *et al.*, 2004; Brouwers, 2011; Sandra and Sandra, 2013). Reversed phase chromatography is perhaps the most widely applied and versatile format, with a hydrophobic stationary phase and a highly polar, aqueous, mobile phase. It allows for high resolution separation of lipids but also necessitates multiple standards for each lipid class to enable accurate quantification since species with different fatty acid compositions elute at different times (Ogiso *et al.*, 2008). Both HILIC and normal phase chromatrography employ a polar stationary phase with a less polar mobile phase (Strege, 1998). While normal phase chromatography is the traditional means to separate lipids (Sturt *et al.*, 2004; Christie and Han, 2010), the solvents used are toxic and poorly compatible with mass spectrometric detection (Nguyen and Schug, 2008; Cubbon *et al.*, 2010). As such, an increasing number of applications of HILIC separations of lipids have recently been described (Schwalbe-Herrmann *et al.*, 2010; Cífková *et al.*, 2012).

An HPLC system is most commonly interfaced to a mass spectrometer through an electrospray ionisation (ESI) interface. As the flow from the LC exits a needle tip into the ESI chamber, it is nebulised into a fine spray. A large voltage between the needle tip and the inlet to the MS results in molecules within the droplets becoming charged and attracted towards the inlet. Through an incompletely understood mechanism, the droplets decrease in size as they pass through the chamber until single, charged ions are left which can be introduced to the MS (Constantopoulos *et al.*, 2000). The ESI ensures that solvent from the LC system is removed prior to the introduction of ions into the high vacuum of the MS. Chemical conditions in the solvent, such as pH, can

have a major effect on the ionisation efficiency of different lipid classes (Brügger *et al.*, 1997). Once introduced to the MS, lipids can be identified using similar fragmentation strategies to those employed for shotgun lipidomics.

1.6 PhD project overview and aims

1.6.1 Strains used during this study

Members of the Roseobacter group were used to carry out the work described in this thesis. This group of marine bacteria make attractive model organisms since they are widespread and abundant in the marine environment (see for example Selje *et al.*, 2004) but also readily grown in culture and amenable to genetic manipulation (Piekarski *et al.*, 2009). I worked with two principal strains of Roseobacter: *Ruegeria pomeroyi* DSS-3 and *Phaeobacter* sp. MED193. Both strains have been genome sequenced; *R. pomeroyi* was the first Roseobacter to have its genome published (Moran *et al.*, 2004). Indeed *R. pomeroyi* has emerged as a prominent model Roseobacter, particularly for the study of sulfur flux from the ocean (Reisch *et al.*, 2013). It was isolated from seawater off the coast of Georgia, USA. In contrast, *Phaeobacter* sp. MED193 has been less well studied and no valid species description has been published. It was, however, isolated from Blanes Bay, a periodically phosphorus limited bay off the coast of Catalonia (Pinhassi *et al.*, 2006). Its genome includes homologs of the phospholipase C from *E. meliloti* that mediates lipid remodelling in that organism (Zavaleta-Pastor *et al.*, 2010) as well as of the betaine lipid synthesis genes BtaA and BtaB from *Rhodobacter sphaeroides* (Klug and Benning, 2001). These features made it a promising candidate to study lipid remodelling in marine bacteria.

1.6.2 Project aims

The aims of this project were as follows:

1. To determine whether model Roseobacter strains are capable of lipid remodelling in response to phosphorus limitation (Chapter 3). I hypothesised that *Phaeobacter* sp. MED193 was capable of carrying out lipid remodelling, with the betaine lipid DGTS produced as a replacement lipid.

2. To investigate the role of PlcP in the lipid remodelling response of marine bacteria to P limitation (Chapter 3). I hypothesised that PlcP was required for lipid remodelling in response to P limitation in *Phaeobacter* sp. MED193. Furthermore I suspected that the capacity to carry out lipid remodelling, mediated by PlcP, was more widespread than had previously been claimed and aimed to investigate this by analysing the distribution of PlcP homologs in the genomes of sequenced marine bacteria as well as in marine metagenomes.

3. To determine the function of a second OlsB homolog present in several Roseobacters (Chapter 4).. I hypothesised that this gene was required for the synthesis of a glutamine lipid that had recently been detected in *Rhodobacter sphaeroides* (Zhang *et al.*, 2009) and planned to test this by making a deletion mutant in *Ruegeria pomeroyi*, which also harbours the gene (Geiger *et al.*, 2010).

4. To investigate the potential role of aminolipids in the response of Roseobacter group bacteria to P limitation (Chapter 4). Ornithine lipids appear to substitute for phospholipids in the membranes of some bacteria during P limitation (Minnikin and Abdolrahimzadeh, 1974a; Geiger *et al.*, 1999) so I reasoned that this might provide an alternative means of lipid remodelling to PlcP-mediated remodelling. I tested the hypothesis that aminolipids (ornithine and glutamine lipids) would become more abundant in *R. pomeroyi* cultures grown in P-limited media and also that *R. pomeroyi* mutants unable to produce one or both of those aminolipids would grow less well in P-limited media.

5. To investgate the structure and biosynthesis of a seemingly novel aminolipid detected during preliminary lipid analysis of Roseobacter cultures (Chapter 5). I aimed to determine the structure of this lipid by high resolution accurate mass MS and to attempt to characterise the genes required for biosynthesis of this lipid in *R. pomeroyi*.

Chapter 2 Materials and Methods

2.1 Strains used and culture conditions

Two marine Roseobacter clade isolates, *Ruegeria pomeroyi* DSS-3 and *Phaeobacter* sp. MED193 (hereafter referred to as MED193) were used in this work. All growth of these strains was carried out at 30 ± 1 °C with shaking at 150 rpm. Stocks of both strains were stored at -80 °C in 20% (v/v) glycerol. These stocks were revived by inoculation in 5 mL marine broth (MB) medium, made using 30 g/litre Difco Marine Broth 2216 (Becton, Dickinson and Company, Sparks, MD, USA), followed by incubation at 30 °C for 48 – 72 hours with shaking at 150 rpm. Both strains were then routinely maintained on 1.5 % (w/v) agar plates (Agar No. 1, Oxoid, UK) stored at 4 °C. MB agar plates were using for maintenance of MED193 while *R. pomeroyi* was maintained on 1/2YTSS (2 g/litre yeast extract, 1.25 g/litre peptone, 20 g/litre sea salts, Sigma-Aldrich).

For growth in defined media, isolated colonies were picked initially into 5 mL rich medium (MB for MED193; ½ YTSS for *R. pomeroyi*). After 24 – 48 hours growth, when the culture was visibly cloudy, 1 mL of this initial culture was inoculated into 25 mL of defined medium and this starter culture was allowed to grow for a further 24 – 48 hours. Each of the Roseobacter isolates were grown in a different defined medium. MED193 was grown in the *Prochlorococcus* medium PCR-S11 (Rippka *et al.*, 2000), modified with a carbon source and vitamins, while *R. pomeroyi* was cultured in MAMS (see Table 2.1 for details). A 4% (v/v) inoculum was transferred from the starter culture into triplicate erlenmeyer flasks containing the appropriate defined medium. The volume of defined medium used was around 20 % of the total volume of the flask. Unless otherwise specified, growth was followed by measuring the optical density at 540 nm (OD_{540}) on a Shimadzu UV-1800 spectrophotometer.

Table 2.1 Composition of defined media.

Chemical	MAMS	Sea Salts	PCRS11
NaCl	30 g L^{-1}	-	-
Sea Salts (Sigma Aldrich)	-	30 g L^{-1}	-
Red Sea Salts (Red Sea Aquatics)	-	-	33.3 g L^{-1}
Glucose	10 mM	10 mM	10 mM
K_2HPO_4	1 mM	1 mM	-
Na_2HPO_4	-	-	5 - 50 µM
NH_4Cl	0.75 - 7.5 mM	0.75 - 7.5 mM	-
$(NH_4)_2SO_4$	-	-	400 µM
HEPES (pH 7.6)	10 mM	10 mM	-
HEPES (pH 7.5)	-	-	1 mM
$CaCl_2$	1.36 mM	-	-
$MgSO_4$	0.98 mM	-	-
$FeSO_4$	7.2 µM	-	-
Na_2MoO_4	83 µM	0.15 µM	-
$(NH_4)_6Mo_7O_{24}$	-	-	1.45 nM
$FeCl_3$	-	5.0 µM	8.0 µM
$FeCl_2$	7.5 µM	7.5 µM	-
$ZnCl_2$	370 nM	370 nM	-
$MnCl_2$	510 nM	510 nM	-
$MnSO_4$	-	-	30 nM
H_3BO_3	97 nM	97 nM	150 nM
$CoCl_2$	1.1 µM	1.1 µM	-
$Co(NO_3)_2$	-	-	1.5 nM
$CuCl_2$	12 nM	12 nM	-
$CuSO_4$	-	-	1.5 nM
$NiCl_2$	100 nM	100 nM	-
$Na_2MgEDTA$	-	-	8.0 µM
KBr	-	-	3.0 nM
KI	-	-	1.5 nM
$Cd(NO_3)_2$	-	-	1.5 nM
Na_2WO_4	-	-	0.3 nM
SeO_2	-	-	1.5 nM
$NiCl_2$	-	-	1.5 nM
$Cr(NO_3)_3$	-	-	0.3 nM
$VOSO_4$	-	-	0.3 nM
$KAl(SO_4)_2$	-	-	3.0 nM
Thiamine HCl	30 nM	30 nM	30 nM
Nicotinic acid	160 nM	160 nM	160 nM
Pyridoxine HCl	97 nM	97 nM	97 nM
p-Aminobenzoic acid	73 nM	73 nM	73 nM
Riboflavin	53 nM	53 nM	53 nM
Calcium Pantothenate	84 nM	84 nM	84 nM
Biotin	4.1 nM	4.1 nM	4.1 nM
Cyanocobalamin	1.5 nM	1.5 nM	1.5 nM
Folic acid	11 nM	11 nM	11 nM

2.1.1 Alkaline phosphatase assay

An assay for alkaline phosphatase activity was used to assess whether cultures were stressed by lack of phosphate availability. The assay follows the liberation of yellow-coloured *para*-nitrophenol (pNP) during incubation of *para*-nitrophenol phosphate (pNPP) with a bacterial culture. I prepared 10 mM pNPP (Sigma-Aldrich) in 10 mM pH 7 Tris solution. This was added to 900 µL aliquots of cell culture to obtain a final pNPP concentration of 1mM. Control incubations were set up in parallel with just the Tris buffer solution. These were incubated at room temperature in the dark until a visible yellow colour developed and the incubation time recorded. The time required varied widely depending on the density and P-stress of the culture (typically ranging from 10 minutes to several hours). The incubations were then centrifuged for 10 minutes at 14,000 x *g*. Triplicate 200 µL aliquots from the clarified supernatant were pipetted onto a 96-well plate and the absorbance at 405 nm measured using a BioRad iMark microplate reader. This absorbance was blank-corrected by subtracting the absorbance of the Tris-incubated control. The amount of pNP liberated was quantified by comparison of the blank-corrected A_{405} to a calibration curve constructed using pNP standards (Sigma-Aldrich) in the range 10 µM – 2 mM.

2.2 Cloning work

I used several *Escherichia coli* strains for cloning and protein expression work. All were grown routinely in Luria-Bertani (LB) medium at 37 °C and, for liquid cultures, shaking at 150 rpm. Stocks were kept in 20% (v/v) glycerol at -80 °C. Chemically competent JM109 cells (Promega) were used for routine cloning. Electrocompetent S17.1 cells were prepared in-house by washing mid-exponential phase cells 3 times in 10% v/v glycerol, then resuspending to achieve a 500-fold concentration of the initial culture. Aliquots (50 µL) were stored at -80 °C for subsequent use during electroporation.

2.2.1 Polymerase chain reaction (PCR)

PCR was performed routinely during cloning both to confirm the identity of constructed and preparatively. Unless otherwise specified, PCR was carried out using the non-proofreading *Taq* polymerase (Kapa Biosystems; Table 2.2). The high fidelity *Phusion* polymerase (Thermo Scientific) was used for the preparation of fragments for Gibson cloning (Section 2.2.3). Reagents for both methods are shown in Table 2.2. Primers were supplied by Sigma-Aldrich or by Integrated DNA Technologies (IDT) as a dried pellet which was resuspended in nuclease-free water to a concentration of 100 μM and stored at -20 °C. Working stock solutions (10 μM) of primers were prepared from this stock. A 10 mM solution of deoxy nucleoside triphosphates (dNTPs) was prepared from 100 mM stocks of dATP, dCTP, dGTP and dTTP (deoxy adenine, cytosine, guanine and thymidine triphosphates, respectively; all from Promega). The buffers used were provided by the polymerase manufacturers. For *Taq* polymerase, I used "Buffer A", which is a tris-ammonium sulfate buffer with 15 mM $MgCl_2$ while with *Phusion* polymerase I used HF buffer, which contains 7.5 mM $MgCl_2$.

The PCR mixtures in Table 2.2 were used to successfully amplify DNA from small amounts of cells picked from colonies growing on an agar plate (colony PCR) for both Roseobacter strains and *E. coli*. For PCR of purified DNA, the DNA concentration was quantified beforehand using a NanoDrop ND-1000 spectrophotometer. The DNA solution was then diluted to a concentration less than 1 nM and 1 μL added to the PCR mix. The results of the PCR were visualised by running on a 1% (w/v) agarose gel, stained with ethidium bromide, for 30 minutes at 80V with Tris/Borate/EDTA (TBE) buffer. The gel was viewed under UV light using a Gene Genius bio-imaging system (Syngene).

Table 2.2 Composition of reaction mixtures used for PCR with non-proofreading (*Taq*) polymerase and high-fidelity (*Phusion*) polymerase.[a]

	Taq	*Phusion*
MilliQ water	36.75	30.5
Buffer	5	10
DMSO	2	2
3.7% (w/v) BSA	1	1
10 mM dNTPs	1	1
10 µM forward primer	2	2.5
10 µM reverse primer	2	2.5
Polymerase	0.25	0.5
Total	**50**	**50**

[a] DMSO: dimethylsulfoxide, BSA: bovine serum albumin, dNTPs: deoxy nucleoside triphosphates.

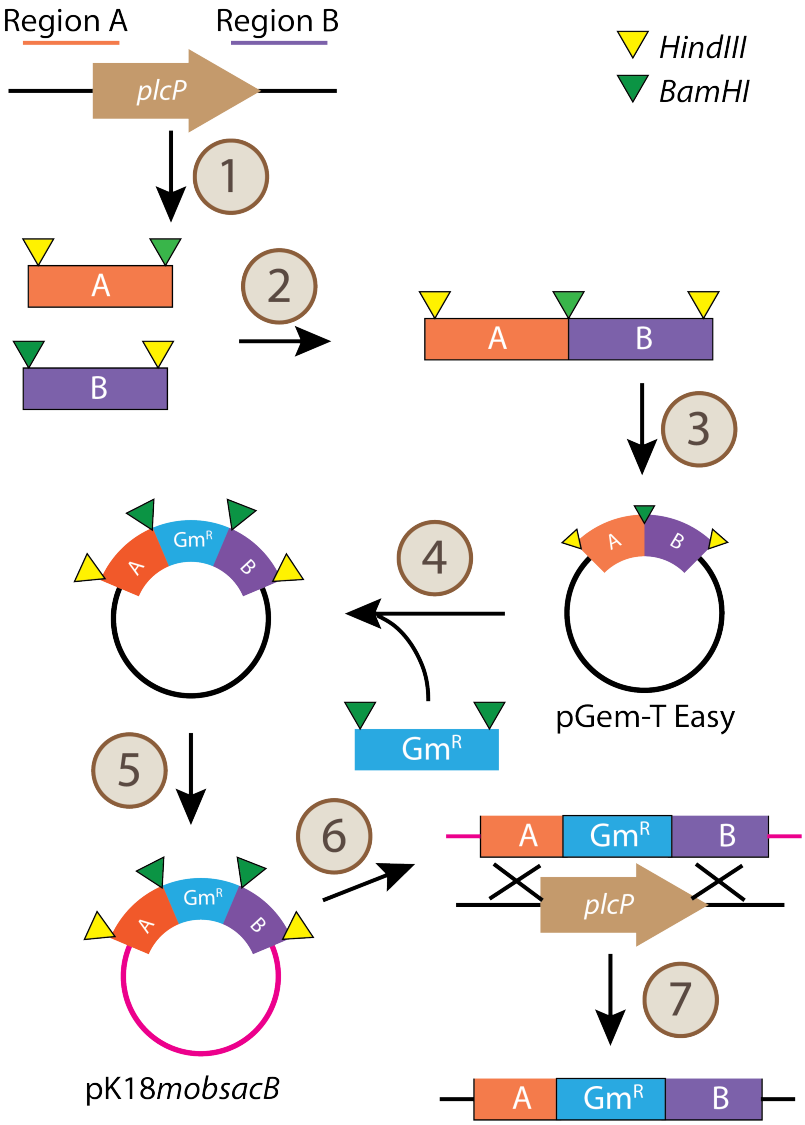


Figure 2.1 Cloning workflow schematic for the construction of *ΔplcP*. **1)** PCR amplification of regions upstream and downstream of *plcP* (regions A and B, respectively) from *Phaeobacter* sp. MED193 colonies. **2)** Restriction digest of purified A and B with *BamHI* and ligation to form AB. PCR amplification of ligation mix using A forward and B reverse primers. **3)** TA cloning of AB into the pGEM-T Easy vector. **4)** The gentamicin resistance cassette (Gm^R) from p34S-Gm is obtained by restriction digest with *BamHI*. This Gm^R is ligated into *BamHI*-digested pGEM-T::AB. **5)** A-Gm^R-B is excised from pGEM-T by *HindIII*, then ligated into *HindIII*-digested pK18*mobsacB*. **6)** pK18::A-Gm^R-B is introduced into *Phaeobacter* sp. MED193 cells by conjugation with *Escherichia coli* S17.1. Since pK18 does not replicate in MED193, only cells which undergo single or double recombination events will retain gentamicin resistance. **7)** If recombination occurs in both A and B regions, Gm^R is stably inserted into the chromosome in place of *plcP* while surrounding DNA should be unaffected.

2.2.2 Construction of deletion mutants in MED193 and R. pomeroyi

The same procedure, based upon the method described by Lidbury *et al* (2014), was used for the construction of deletion mutants for homologs of *plcP* in MED193 (MED193_17359) and of *olsB2* in *R. pomeroyi* (SPO2489). Primers were designed to amplify 500 – 700 base pair sequences to either side of the target gene. The upstream sequence I refer to as fragment A; the downstream sequence as fragment B (Figure 2.1). Both fragments were engineered to have two distinct restriction enzyme sites at either end. The restriction site at the internal end of the fragments was in both cases *BamHI* while at the external ends I engineered *HindIII* sites for MED193_17359 construct and *XbaI* for the SPO2489 construct (Table 2.3). I amplified the fragments by colony PCR using *Taq* polymerase. In order to minimise contamination by non-specific products and cellular DNA, I gel purified the PCR products by running the entire 50 µL reaction mix on a 1% (w/v) agarose gel, stained with ethidium bromide, for 30 minutes at 80V. Then I excised the strong band observed when the gel was placed on a UV visualiser into a 1.5 mL Eppendorf tube using a clean scalpel. I verified that the band was roughly the expected size by comparison to a 1 kilobase pair DNA ladder (Promega). To extract the DNA from the gel, I used a Macherey-Nagel PCR purification kit. Before ligating the fragments, I incubated each with FastDigest *BamHI* (Thermo Scientific) for 30 minutes at 37 °C. Ligation of A and B was carried out with T4 DNA ligase (Promega), incubated either for 1.5 hours at room temperature or overnight in the fridge (4 °C). To increase the yield of A-B ligation product, I performed another PCR using the A forward and B reverse primers and again gel purified the product. Since I used *Taq* polymerase (a non-proofreading polymerase) I was then able to TA clone this product in pGEM-T Easy (Promega), which I transformed into chemically competent JM109 cells as per the manufacturer's instructions. Transformants were selected by plating onto LB with 100 µg/mL ampicillin and spread with 25 µL of 0.5M isopropyl β-D-1-thiogalactopyranoside (IPTG) and X-gal. White colonies were screened for the presence of the expected A-B insert by PCR. Plasmids from colonies which tested positive for the presence of the A-B insert by PCR were extracted from overnight cultures using a GeneJET plasmid

miniprep kit (Thermo Scientific). The extracted plasmids were incubated with *BamHI* and *HindIII* (both FastDigest, Thermo Scientific) for 30 minutes at 30 °C to verify that they exhibited the expected restriction pattern. The AB inserts were then Sanger sequenced using the M13 primers (Table 2.3).

For a selectable marker, I chose the gentamicin resistance cassette (Gm^R) from p34S-Gm (Dennis and Zylstra, 1998). The cassette is flanked on either side by a number of sites for commonly used restriction enzymes, allowing it to be easily excised and ligated into a plasmid. I digested both p34S-Gm and pGEM-T::A-B with *BamHI*. Then I gel purified both the released Gm^R and the digested pGEM-T::A-B before ligating the resistance cassette into the plasmid. The ligation mixture was again transformed into JM109 and colonies recovered on LB-Gm plates (5 μg/mL gentamicin) were screened by PCR using A forward and Gm^R reverse primers. This allowed me to check both for the presence of Gm^R and its orientation relative to the surrounding A and B fragments.

To construct the deletion mutants by conjugation, I transferred the A-Gm^R-B insert into pK18*mobsacB* (Schäfer *et al.*, 1994), a mobilisable plasmid that does not replicate in Roseobacters. For restriction digest of pGEM-T::A-Gm^R-B and pK18 I used *HindIII* for the MED193_17359 construct and *XbaI* for the SPO2489 construct. I gel purified A-Gm^R-B and the linearised pK18 vector, then ligated the two together. The ligation mix was transformed into JM109 and colonies on LB – kanamycin plates (25 μg/mL) were screened by PCR to verify that they contained the expected insert. It was then necessary to introduce the plasmids into S17-1, an *E. coli* strain capable of high efficiency conjugation (Simon *et al.*, 1983). Plasmids were extracted from the JM109 colonies and checked by restriction digest. They were then electroporated into electrocompetent S17-1 cells at 18 kV/cm in 1 mm gap cuvettes (BTX, MA, USA). Cells were recovered in 1 mL SOC (Super Optimal Broth with Catabolite repression) medium for 1 hour at 37 °C before plating onto selective media.

For conjugation, I grew Roseobacter cells in defined medium to mid exponential phase (an OD_{540} of around 0.5). Around 5 mL of this culture was washed and resuspended in 0.3 mL sea salt medium (Table 2.1) prepared without added carbon or nitrogen source. Simultaneously, I had grown S17-1∷pK18::A-Gm^R-B overnight in LB-Gm. I harvested 0.5 mL of this culture and again washed and resuspended it in 0.3 mL of sea salts without carbon or nitrogen. The two cell suspensions were mixed, and the mixture spotted onto a 0.22 μm nitrocellulose membrane filter (Whatman) placed on top of a ½ YTSS agar plate. The conjugation plate was left at 30 °C for 24 hours before the filter was removed into a 50 mL Falcon tube, and the cells washed off into 1 mL sea salts medium (no carbon or nitrogen source). A 10-fold dilution series was prepared in sea salts medium (no carbon or nitrogen source) down to a 10^{-4} dilution. Aliquots (100 μL) from these dilutions were plated onto sea salts agar plates with 10 mM glucose, 2 mM glycine betaine (GBT) and 10 μg/mL gentamicin. Since *E. coli* is unable to use GBT as a sole nitrogen source, it is unable to grow on these plates. After incubation at 30 °C for 1-2 weeks, colonies were streaked onto two fresh agar plates, again made with the same sea salts medium. One of the plates was again spiked with gentamicin while the other was prepared with 50 μg/mL kanamycin. Colonies which grew on the gentamicin plate but not the plate with kanamycin were identified as potential double cross-over mutants (in which Gm^R inserted in place of the target gene, as desired). These colonies were screened by PCR using the Gm^R forward primers and the B reverse primers (Table 2.3). Finally, colonies that tested positive by both measures were sequenced using primers external to the A and B fragments on the bacterial chromosome (Table 2.3) to confirm that no mutations had been introduced outside of the target gene.

Table 2.3 Oligonucleotide primers used for molecular genetic work.[a]

Locus Tag	Use	Forward primer (5'-3')	Reverse primer (5'-3')
MED193_17359	Δ*plcP* construction (upstream region A).	GTCTAAGCTTTGAGGATGACGACGATGTTC	CTATGGATCCGGTGTCCGCTTCGTGACTAT
MED193_17359	Δ*plcP* construction (downstream region B).	CTATGGATCCTTGTCGAGCGAGACAATGG	ATCTAAGCTTCGCTCATATAGGGGGAGGTT
MED193_17359	Confirmation of Δ*plcP*.	AGCCATTTTTCACCACCAAG	CCCAGAACCCCGTAGTGATA
SPO2489	*ΔolsB2* construction (upstream region A).	CAATTCTAGAGGTGTTCCGTATTGTGGCAA	CAATGGATCCGCTCAACAGTATGGCATCCG
SPO2489	*ΔolsB2* construction (downstream region B).	CAATGGATCCGTTTACCGGGCTTGAGATCG	CAATTCTAGAGCCTCGATATCCAGATGGTTG
SPO2489	Confirmation of *ΔolsB2*	AGGTAACGCTGTTTTGTGGCAG	TCATCCTCCCAGATCTTGTCGC
Gm^R primers	Sequencing / confirmation	GCAGTCGCCCTAAAACAAAG	ATCTCGGCTTGAACGAATTG
M13 primers	Sequencing / confirmation	CACGACGTTGTAAAACGAC	GGATAACAATTTCACACAGG
MED193_17359 (plus promoter)	Δ*plcP*^{MED193} construction.	AGTCAAGCTTAACTGGTCAGCAAGCCAACT	AGTCGGATCCCATCGGGTAGATCCCCTATACA
PB7211_983	Confirmation of SAR11 *plcP*	GAGGTGTGCTCACCAGAAG	TCTGCTAAAGCAGTGCAGGA
PB7211_960	Confirmation of SAR11 *agt* (codon optimised for *E. coli*).	ATCCGCAAGTCAATGGTGTT	GTCACGTTTCACCGGATTTT
T7 primers	Sequencing *agt*	TAATACGACTCACTATAGGG	GCTAGTTATTGCTCAGCGG

[a] Restriction sites are underlined.

2.2.3 Gibson cloning for construction of deletion mutants in R. pomeroyi DSS-3

Some of the mutants used in this work were constructed using a streamlined protocol based on that described by Gibson *et al.* (2009). This protocol was substantially quicker than the one described in Section 2.2.2 and involved only a single cloning step. The end result of the cloning is a construct in pK18*mobsacB* analogous to that outlined in Figure 2.1, thus marker exchange mutagenesis was again achieved by obtained double-crossover mutants in which the gene of interest was replaced by Gm^R. Briefly, the method involves the generation of single stranded overhangs from double stranded DNA by a 5' exonuclease. Complementary overhangs are annealed by a DNA ligase and any single stranded gaps are filled by a proof-reading DNA polymerase, generating intact double stranded DNA. These processes are carried out simultaneously at a constant temperature of 50 °C in a thermocycler for 15 minutes to 1 hour. An assembly master mix, containing all the required enzymes and reagent, along with DH5α chemically competent cells, were obtained as part of the NEBuilder HiFi DNA Assembly Cloning kits (New England Biolabs).

This method was used to obtain mutants in *R. pomeroyi* DSS-3 deficient in a putative *olsA* (SPO1979) and a hypothesised cysteine sulfinic acid decarboxylase (CSAD; SPO3687). Prior to assembly, DNA fragments for A, B, the Gm^R cassette and pK18*mobsacB* were generated by PCR using the high-fidelity *Phusion* DNA polymerase (New England Biolabs). The primers used for the PCR were designed to have complementary overlap regions with adjacent primers of around 30 base pairs (Table 2.4). Regions A and B were amplified by colony PCR of *R. pomeroyi* DSS-3, while Gm^R and the linearised vector were prepared from purified stocks of p34S-Gm and pK18*mobsacB*, respectively. These PCR products were purified using a Macherey-Nagel PCR purification kit and the DNA concentration measured using a NanoDrop spectrophotometer. The purified DNA from each of the fragments was mixed together in the proportions recommended by the manufacturer and the

volume made up to 10 µL using MilliQ water prior to addition of 10 µL master mix. The reaction mix was then incubated for 1 hour at 50 °C before being placed on ice to stop the reaction. A 50 µL aliquot of the *E. coli* DH5α cells was chemically transformed with 1 µL of the reaction mix according to the manufacturer's instructions and the cells then plated on LB-Gm plates to select for lineages with correctly assembled constructs. After overnight growth at 37 °C, selected colonies were picked and checked for the presence of correctly assembled constructs by PCR using the A forward and B reverse primers. DH5α colonies testing positive for the correctly assembled construct by PCR were grown overnight in 5 mL of LB with 50 µg mL^{-1} kanamycin at 37 °C before the plasmids were extracted using plasmid miniprep kits.

Electrocompetent *R. pomeroyi* DSS-3 cells were prepared based on the method of Sebastian and Ammerman (2009). After growing cultures overnight in 50 mL ½YTSS, the culture was pelleted and washed 5 times with 10% v/v glycerol to remove salts before being resuspended in 0.5 mL for a 100-fold concentration. This was split into 50 µL aliquots, with unused aliquots being stored at -80 °C and remaining competent for at least two weeks. To one 50 µL aliquot was added 2 µL of the assembled and purified plasmid. This was electroporated into the cells at 12 kV/cm using 1 mm gap electroporation cuvettes and the cells allowed to recover in ½YTSS overnight prior to plating onto ½YTSS agar plates with 10 µg mL^{-1} gentamicin. Colonies growing on these plates were restreaked onto two fresh ½YTSS plates, one with gentamicin and one with kanamycin to check for double cross-over mutants, as described in Section 2.2.2. Colonies growing on gentamicin but not kanamycin were checked both by PCR and by sequencing using primers external to the A and B regions on the genome, plus the A and B reverse primers and Gm^R forward primer. This was compared to the sequence expected of a correctly assembled insert to check for any introduced point mutations.

Table 2.4 Primers used for Gibson cloning.[a]

Locus Tag	Use	Forward primer (5'-3')	Reverse primer (5'-3')
pK18mobsacB	Linearised vector	TGGCACTGGCCGTC GTTT	TACCGAGCTCGAATTC GTAATCATG
SPO1979	Δ*olsA* construction (upstream region A)	tatgacatgattacgaattcgagct cggtaGAATATGACCTT GGCGTGC	cgaacaggcttatgtACTCCGC GTGCCAGTACC
p34S-Gm	Δ*olsA* construction (Gm^R)	actggcacgcggagtACATA AGCCTGTTCGGTTCG	cagttcgacccgtccGCGGCGT TGTGACAATTT
SPO1979	Δ*olsA* construction (downstream region B)	ttgtcacaacgccgcGGACG GGTCGAACTGATC	cacgacgttgtaaaacgacggccagt gccaCGAACGGATCGTA CTGCTT
SPO1979	Confirmation of *ΔolsA*	GTGGGCGTCTATCGT CTGTT	AACGGTTCCTCGAGAC CTTC
SPO3687	Δ*csad* construction (upstream region A)	tatgacatgattacgaattcgagct cggtaATGTAGGTGTG CTGGGCTTC	cgaacaggcttatgtCCACGGT CAGGTGATAGTC
p34S-Gm	Δ*csad* construction (Gm^R)	atcacctgaccgtggACATA AGCCTGTTCGGTTCG	cggtgatcacgtcaaGCGGCGT TGTGACAATTT
SPO3687	Δ*csad* construction (upstream region B)	ttgtcacaacgccgcTTGAC GTGATCACCGAAATC	cacgacgttgtaaaacgacggccagt gccaGAGCAGATATCAG GCGGTTT
SPO3687	Confirmation of *Δcsad*	CCAATCTCTACCCGC CCTTCAA	GGCCGGGGTTGAACAG GTC

[a] Overlap sequences are in lower case and colour coded. Red: overlap with vector backbone; blue: junction between region A and Gm^R; green: junction between Gm^R and region B.

2.2.4 Construction of complemented ΔplcP strains

Primers were designed to amplify MED193_17359 plus a 400 base pair (bp) region upstream of the start codon, which was presumed to contain the promoter. Restriction sites at the 5' end of each primer were used to clone the product into the broad host plasmid pBBR1MCS-Km (Kovach *et al.*, 1995). This was electroporated into *E. coli* S17-1 before conjugation was carried out as described above. In this case, selection was on sea salt plates with 50 μg/mL kanamycin.

None of the isolated SAR11 clade strains are widely available in culture. Therefore, to complement the *ΔplcP* strain with the *plcP* homolog from SAR11, the gene from *Candidatus* Pelagibacter ubique sp. HTCC7211 (PB7211_983) was synthesised, fused to the 400 bp promoter from MED193 described above, by Genscript (Hong Kong). The construct was supplied in the pUC57 vector, flanked by *NdeI* and *BamHI* restriction sites. Before introducing into MED193 *ΔplcP*, the construct was cloned, using the restriction sites, into pBBR1MCS-Km. The plasmids were checked by restriction digests and by sequencing the insert region using M13 primers (Table 2.3). The resulting plasmid was introduced into MED193 *ΔplcP* by conjugation, as described previously.

2.2.5 Overexpression of the SAR11 agt glycosyltransferase homolog

The homolog of *agt* identified in SAR11 strain HTCC7211 (PB7211_960) was synthesised in pUC57 by Genscript (Hong Kong). Flanking *NdeI* and *BamHI* sites were used to clone the construct into pET28a, and this was transformed into *E. coli* strain BL21(DE3) for overexpression. As a control, I also constructed an empty vector strain by transforming BL21(DE3) with pET28a without any insert. Both strains were grown initially in LB with 25 μg/mL kanamycin to mid-exponential phase (an optical density at 600nm of around 0.5). A 1 mL aliquot from each culture was washed, then resuspended in M9 minimal medium (20 $g.L^{-1}$ glucose, 64 $g.L^{-1}$ $Na_2HPO_4.7H_2O$, 15

g.L^{-1} KH_2PO_4, 2.4 g.L^{-1} $MgSO_4$, 2.2 g.L^{-1} $CaCl_2$, 2.5 g.L^{-1} NaCl, 5 g.L^{-1} NH_4Cl) before being inoculated into 100 mL M9 medium. Expression was induced by the addition of 1 mM IPTG while a second culture was maintained without IPTG to act as a control. The cultures were maintained at 18 °C for 24 hours with shaking at 150 rpm. At this point the cells from around 5 mL of culture were pelleted by centrifugation at 9,000 x *g* for 10 minutes. Pellets were stored at -20 °C before lipid extraction within one week of collection.

The rest of the culture was pelleted separately, by centrifugation at 10,000 x *g* for 10 minutes and then resuspended in 2 mL of phosphate buffered saline (PBS) and transferred to a 15 mL Falcon tube (Sarstedt). These cells were lysed by passing through a French press at 1000 p.s.i four times. Incompletely lysed cells and large debris were then removed by centrifugation at 2,000 x *g* for 10 minutes. At this point a sample from the supernatant was collected to represent a 'whole cell' fraction before a further round of centrifugation, this time at 20,000 x *g* was performed to separate the lysate into soluble (the supernatant) and particulate (the pellet) fractions. The protein concentration in all three fractions was quantified using a Bradford assay kit (BioRad) using the manufacturer's method for microtiter plate readings. Briefly, this involved preparing a diluted dye reagent (1:5 v/v in distilled water) then filtering through a Whatman #1 filter to remove precipitate. Protein standards were prepared in the range 0.05 - 0.5 mg mL^{-1} using bovine serum albumin (BSA). Aliquots (10 µL) of each sample or standard were pipetted into the wells of a microplate and to these was added 200 µL of the diluted dye reagent. The microplate was the incubated at room temperature for 5 minutes to 1 hour to allow a stable colour to form before measuring the absorbance at 595 nm using a microplate reader. Aliquots containing around 10 µg protein were collected from each fraction, 5 µL of NuPAGE 4x lithium dodecylsulfate (LDS) loading buffer (Invitrogen) was added and the mixture made to 20 µL with distilled water. These were then boiled for 5 minutes to denature the proteins. The entire mixture was then loaded into a well in an 8-well 12% Bis-Tris polyacrylamide gel (Invitrogen) with MOPS running buffer (Invitrogen) and run at

200 V for 1 hour. PageRuler Plus prestained protein ladder (Thermo Scientific) was run alongside the samples. The gel was then stained by covering with Coomassie brilliant blue R-250 (BioRad) and shaking for around half an hour until clear bands were observed. The stained gel was imaged using the bio-imaging system with white light.

Fresh cells were grown for inclusion body purification and a lysate prepared as above. The lysate was centrifuged at 5,000 x *g* for 15 minutes to pellet cell debris and inclusion bodies. The pellet was resuspended and washed in 10 mL binding buffer (Novagen His Bind buffer diluted 8-fold in distilled water). This was repeated 3 times in order to remove trapped soluble proteins. After removing the supernatant, the washed pellet was resuspended in 2 mL of binding buffer with 6M urea and incubated on ice for 1 hour. Insoluble material was then removed by centrifugation at 16,000 x *g* for 30 minutes and the supernatant filtered through a 0.45 μm filter. To prepare resin for purification of the protein, 100 μL of Ni-NTA His Bind resin (Novagen) was washed twice with distilled water before being suspended in 100 μL of binding buffer. This was then added to each of the filtered supernatants and mixed at 4 °C on a rotary mixer for 1 hour. The resin was pelleted by centrifugation at 15,000 x *g* for 10 seconds and the supernatant discarded, save for 10 μL, which was reserved for protein quantification as described above. The resin was then washed with 100 μL of buffer and pelleted as above. This procedure was repeated until no blue colour developed following the Bradford assay of the supernatant. Once unbound proteins had been removed, bound proteins were eluted three times using 20 μL of His Bind elution buffer (Novagen). These elution fractions were run on a gel as described above. Selected bands at around 38 kDa were excised from the gel using a band cutter, and analysed by electrospray ionisation - mass spectrometry (ESI-MS) using an Orbitrap Fusion mass spectrometer (Thermo Scientific). The results were analysed using the Scaffold protein viewer (Searle, 2010).

2.3 Intact polar lipid extraction and analysis

The method described here differs from the one described in Sebastian *et al.* (2016). The differences reflect refinements that I have made over the course of my work. The results obtained in that paper used the method described there. However, subsequent results used the updated method I present here.

2.3.1 Lipid standards

All lipid standards were obtained as powders from Avanti Polar Lipids (Alabaster, AL.). An internal standard solution was prepared consisting of 0.25 μM diheptadecanoyl phosphatidylglycerol (17:0/17:0 PG), 0.2 μM diheptadecanoyl phosphatidylethanolamine (17:0/17:0 PE) and 0.25 μM diheptadecanoylphosphatidylcholine (17:0/17:0 PC) in 1:1 (v/v) chloroform : methanol. This was capped with argon, and stored at -80 °C until use. These standards were chosen because they represent the three main phospholipid species present in Roseobacter, and preliminary studies did not detect any lipids with the diheptadecanoyl fatty acid configuration.

2.3.2 Lipid extraction

Polar lipids were extracted by a modification of the Folch extraction method (Folch *et al.*, 1957). In order to increase sample throughput, I scaled the method down to a total volume of 2 mL, whilst approximately maintaining the original ratio of extraction solvent to sample wet weight (20:1). This has been found to be important for maximum extraction efficiency (Iverson *et al.*, 2001). The biomass collected was normalised to 1 mL of culture with an OD_{540} of 0.5, as follows:

$$Volume\ of\ culture\ collected\ (mL) = \frac{0.5}{OD_{540}}$$

The cells were washed with 1 mL of 20 g/L sea salts, then resuspended in 0.5 mL of LC-MS grade methanol (Sigma-Aldrich). At this stage the samples were spiked with 25 µL of the internal standard solution and the methanol suspension was transferred to a 2 mL glass Chromacol vial (Thermo Scientific). To this was added 1 mL of HPLC-grade chloroform (VWR) and 0.3 mL of Milli-Q water. The vials were capped with PTFE-lined lids (Thermo Scientific) and mixed on a rotary shaker at 4 °C for 1 hour. The vials were then centrifuged at 3,000 rpm for 10 minutes at 4 °C with the brake off to separate the solvent into two phases: an upper methanol-water phase, containing hydrophilic metabolites and a lower chloroform phase representing the lipid fraction. The chloroform phase was aspirated into a fresh vial using a glass Pasteur pipette. The remaining methanol-water phase was re-extracted by addition of a further 0.9 mL of chloroform and the chloroform phase was again aspirated into the same vial. The chloroform was removed under nitrogen at room temperature using a Techne (Staffordshire, UK) sample concentrator. The dried lipids were resuspended in 100 µL of 1:1 chloroform : methanol and 900 µL of acetonitrile added. These samples were either analysed immediately or stored at -80 °C for no more than 1 week before analysis.

2.3.3 Analysis of intact polar lipids

I analysed lipid samples by liquid chromatography-mass spectrometry (LC-MS), employing a Dionex 3400RS high performance liquid chromatography (HPLC) system coupled to an AmazonSL quadrupole ion trap (Bruker Scientific) via an electrospray ionisation interface. My main interest was to investigate variations in the abundance of different lipid classes. This dictated the use of either normal phase (Sturt *et al.*, 2004; Popendorf *et al.*, 2013) or hydrophilic interaction chromatography (HILIC; Cífková *et al.*, 2012), which separate lipids primarily according to their polar head group, rather than reverse phase chromatography, which separates lipids based on their fatty acid composition (Yamada *et al.*, 2013). Since the chromatography system was not compatible with normal phase solvents, I elected for the HILIC approach, using a BEH Amide XP column (Waters). The column was maintained at

30 °C, with a flow rate of 150 μL min^{-1}. Samples were run on a 15 minute gradient from 95% acetonitrile to 28% 10 mM of pH 9.2 ammonium acetate, with 10 minutes equilibration between samples.

I analysed each sample in both positive and negative ionisation modes. In both runs, I kept drying conditions the same (8 L min^{-1} drying gas at 250 °C; nebulising gas pressure of 1 psi). However, the end cap voltage was 4,500 V in positive mode and 3,500 V in negative mode, both with a 500 V offset.

2.3.4 Identification of lipid classes

Initial identification of the lipid classes made use of the MS^n capabilities of the ion trap mass spectrometer (i.e. its ability to subject target ions to multiple rounds of fragmentation). Various publications have identified mass transitions that are diagnostic for a particular lipid class (see Table 2.5): I used the presence of these transitions, in conjunction with the overall fragmentation spectrum, to assign observed mass peaks to a lipid class. Subsequently, assignment to lipid classes was based on a comparison of their intact mass and retention time to the identified peaks.

Table 2.5 Diagnostic ions used for the assignment of observed masses to a lipid class.[a]

Lipid Class	Neutral Loss	Product Ion	Reference
PC	189	-	(Yang *et al.*, 2009)
PE	-	196	(Han *et al.*, 2005)
PG	-	153	(Han *et al.*, 2004)
OL	-	115	(Zhang *et al.*, 2009)
QL	-	129	(Zhang *et al.*, 2009)
DGTS	-	236	(Popendorf *et al.*, 2013)

[a] All diagnostic ions refer to fragmentation of the $[M-H]^-$ ion in negative ion mode, with the exception of PC and DGTS, which were identified by their $[M+H]^+$ ion in positive mode.

2.3.5 *Structural assignment of an unidentified lipid class*

R. pomeroyi was grown in MAMS using a phosphates buffer system (27 mM total phosphate concentration) in place of HEPES. Lipid extracts were obtained as above, and analysed by direct infusion by nano-ESI onto a quadrupole – time of flight (Q-TOF) mass spectrometer. The most abundant peak in the negative ion spectrum corresponded to the *m/z* of one of the unidentified lipid peaks (656.4882). Therefore, this peak was selected for fragmentation. Spectra were obtained in profile mode and smoothed using a moving mean. Background correction using a linear baseline was applied with a 40% noise cutoff. For accurate mass determination, the centroid of each peak was used. No external lock mass was included in the infusion mix instead. The peak corresponding to $C_{17}H_{33}COO^-$ (*m/z* 281.2480; an 18:1 fatty acid carboxylate anion) was used as a lock mass. Calculation of candidate elemental formulae from the accurate mass considered formulae containing C0-100, H0-100, N0-10, O0-10, S0-4, P0-1. A conservative mass error of 100ppm was assumed.

2.3.6 *Automated lipid data processing and quantification*

I was primarily interested in investigating the relative abundance of lipid head group classes. Each lipid class is composed of a number of species which differ in their fatty acid composition. The abundance of these different species of the same lipid class can differ by several orders of magnitude. Moreover, multiple ions can appear in mass spectra due to the formation of multiple adducts. For the accurate and reproducible quantification of a class, I needed a method that could reliably detect each of the fatty acid variants and adducts present. Integrating the results by hand is possible but tedious and prone to error. For this reason I developed an automated pipeline for peak detection, identification and quantification.

The data output files were converted to the mzML format, then peak detection was carried out using the xcms package (Smith *et al.*, 2006) in the R statistical software environment. Peak lists were deisotoped and adducts detected using CAMERA (Kuhl

et al., 2012). Lipids were then identified using a combination of their neutral mass and retention time in comparison to the classes characterised as described above. The abundance of each lipid class was expressed relative to an internal standard. Where no internal standard of the same lipid class was available, the internal standard used was chosen based on its presumed chemical similarity to the analyte lipid class. The 17:0/17:0 PG internal standard was used for the lipids thought to be negatively charged under physiological pH (QL and the unidentified lipid class). PC was used as an internal standard for DGTS, since both share a quaternary ammonium group. Similarly, PE was thought suitable to serve as an internal standard for OL since both head groups are small and zwitterionic.

2.4 Growth experiments

I carried out a number of growth experiments during the course of this work. These used the strains and growth conditions described above (Section 2.1) and in most cases lipid extraction was carried out as described. There were, however, many small modifications between the growth experiments so I describe each below, including a reference to the figure in which the results are reported.

2.4.1 Characterisation of ΔplcP

Phaeobacter sp. MED193 wild type and *ΔplcP* were initially grown to late exponential phase in PCR-S11 (six cultures for each strain). Cells were then pelleted by centrifugation at 9,000 x *g* for 5 minutes. Cell pellets were resuspended in the same volume of either PCR-S11 medium with added P (50 μM) or with no added P (3 replicates per condition for both strains). The cultures were tracked for 4 days following resuspension. On each day, cell density was measured using the optical density at 540 nm (OD_{540}) and a 5 mL aliquot of culture was collected for lipid analysis. Alkaline phosphatase activity was also measured. Lipid extraction was performed as described above, however the method for LC-MS analysis differed. Samples were resuspended in 1:1 methanol to dichloromethane prior to separation

on a 150 mm Nucleosphere HILIC column (Macherey-Nagel) at 30 °C, with a flow rate of 150 μL min^{-1}. Samples were run on a gradient of 95% acetonitrile to 28% 10 mM ammonium acetate, with 2 and 5 minute holds at the start and end of each run, respectively. Following ionisation under conditions which were the same as those described previously, selected masses were targeted for MS^2 fragmentation, using the SmartFrag functionality of the Bruker TrapControl software to select an appropriate voltage. Masses selected for fragmentation were those identified as corresponding to DGTS (738.7 and 764.7) and the PC internal standard (762.7). The relative abundance of DGTS was expressed as the ratio of the sum of the peak areas for DGTS to the peak area for the PC internal standard, normalised to the OD_{540} of the culture. These results are reported in Figure 3.1.

This method for lipid analysis had a number of shortcomings, including the presence of a highly intense contaminant ion at *m/z* 590, which appeared to result from column bleed. This resulted in reduced sensitivity for analyte ions and was eliminated by replacing the column with the BEH-Amide column. Also, the biomass collected for lipid extraction was not standardised, which can result in variable amounts of ion suppression affecting the intensities of target ions in unpredictable ways. Therefore all subsequent lipid analyses were carried out as described in Section 2.3.

2.4.2 Growth of Phaeobacter sp. MED193 for lipidomics analysis

Wild type MED193 was grown as described in Section 2.4.1, with 6 biological replicates for each condition. Three technical replicate samples were collected for lipid analysis 2 days following resuspension in PCRS11 medium with or without added phosphate. The selection of this time point was based on the results of the previous resuspension experiment (Section 2.4.1; Figure 3.1) and supported by the results of alkaline phosphatase assays, which indicated that cultures in the medium lacking P were P-stressed. Lipids from these samples were extracted and analysed according to Section 2.3. Results reported in Figure 3.2 show the ratio of the mean

relative abundance of each major lipid class in cultures grown without added phosphate to that of cultures grown with added phosphate.

2.4.3 Identification of glutamine lipid and characterisation of ΔolsB2 and ΔolsA in Ruegeria pomeroyi DSS-3

R. pomeroyi wild type, *ΔolsB2* and *ΔolsA* were grown overnight in ½YTSS, at which point cells were harvested for lipid extraction and analysis, as described in Section 2.3. Three replicates were grown for each strain and results reported in Figures 4.2 and 4.4. The abundances of QL, OL and PE were reported relative to the abundance of the PE internal standard so that all values would be on the same scale.

2.4.4 Growth rate and lipid comparison of R. pomeroyi wild type and mutants

Three biological replicates of *R. pomeroyi* wild type, *ΔolsB2* and *ΔolsA* were grown in MAMS with either 0.5 mM or 5 mM phosphate. The lower phosphate concentration of 0.5 mM was chosen based on previous preliminary growth experiments in which *ΔolsA* had failed to grow at lower phosphate concentrations. Growth was tracked by measuring OD_{540} at regular intervals and alkaline phosphatase activity was measured prior to the collection of samples for lipid analysis, once the cultures were estimated to have entered late exponential phase. These samples, taken as three technical replicates from each culture, were extracted and analysed as described in Section 2.3. Growth rates were calculated by fitting a linear equation to the log transformed growth data from exponential phase cultures, using an implementation of the GrowthRates program (Hall *et al.*, 2014) in the R statistical software package (R Core Team, 2017). Pairwise comparisons of the growth rates of each strain grown at high and low P concentrations, as well as comparisons of the growth rates between strains grown at the same P concentration, were made using Student's t-test. Maximum yield was calculated as the mean of the maximum OD_{540} obtained for each strain-phosphate combination. Lipid class abundances were reported as the ratio of the abundance of in 0.5 mM P compared to 5 mM P. Since *ΔolsA* again failed to grow in 0.5 mM P, this strain was omitted from this comparison. These results are reported in Figure 4.5.

2.4.5 Growth of Δspo3687

R. pomeroyi DSS-3 *Δspo3687* was grown overnight in MAMS at which points samples were collected for lipid extraction and analysis as described in Section 2.3. Lipid extracts were analysed for the presence of the putative homotaurine lipid as shown in Figure 5.3 and compared to results for the wild type obtained previously.

2.5 Phylogenetic, comparative genomic and metagenome analyses

2.5.1 PlcP in the Global Ocean Sampling and Tara metagenomes

Phylogenetic analysis were performed using 225 PlcP-like sequences from marine bacterial isolates on IMG (Integrated Microbial Genomes) and 1129 sequences of environmental PlcP homologs obtained from the Global Ocean Sampling (GOS) database (Yooseph et al., 2008). These sequences were retrieved through BLASTP searches using *Phaeobacter* sp. MED193 PlcP (MED193_17359) as a query. An e-value $<10^{-40}$ was used as the cut off value. The metagenome sequences were organised into 243 clusters using the CD-HIT program (Li and Godzik, 2006), applying a 90% similarity threshold. The amino-acid sequences were aligned using MUSCLE (Edgar, 2004). A maximum likelihood tree was generated with the R software package phangorn using the JTT model. Metagenomic PlcP sequences were assigned to ecologically meaningful taxonomic groups based upon their proximity in the phylogeny to PlcP from sequenced isolates. In order compare the abundance of PlcP across different sampling sites, it is necessary to control for differences in sequencing depth at each site. This was done by normalising PlcP reads to the abundance of RecA, an essential single copy gene. A BLASTP search (e-value 10^{-40}) was performed using RecA from *Phaeobacter* sp. MED193 (MED193_04366). To correct for differences in sequence length of PlcP (265 a.a.) and RecA (354 a.a.), the number of PlcP hits was divided by the ratio of the lengths of PlcP and RecA.

For the analyses of PlcP in the Tara Oceans metagenomes I first performed a tBLASTn search using MED193_17359 as a query against the non-redundant Ocean Microbial

Reference Gene Catalog (OM-RGC; Sunagawa et al. 2015). These positive OM-RGC IDs were then cross-referenced to a gene profile table containing normalized counts of each non-redundant gene across the Tara Oceans stations (downloaded from http://ocean-microbiome.embl.de/companion.html). Each OM-RGC ID is accompanied by taxonomic information regarding the best blast hit for the sequence against the NCBI nr database. Using this information, PlcP homologues were grouped into ecologically relevant taxonomic groups and total counts of PlcP calculated for each taxonomic group, as well as the overall count at each station. In order to ensure that the results were comparable to those obtained using the GOS data, only surface water samples (5 m) were taken into account for the analyses. PlcP counts were normalized by RecA counts, obtained as described above using *Phaeobacter* sp. MED193 RecA as the query.

2.5.2 16S rRNA phylogeny of the Rhodobacteraceae and presence/absence of OlsB and OlsB2

A collection of 98 *Rhodobacteraceae* strains were identified for which whole genome sequences were available and so that at least one representative was included from each genus with a sequenced representative in the IMG database (as of August 2015). Additionally, a number of representatives from the pelagic Roseobacter cluster (Selje *et al.*, 2004) were included. A full list of strains included in the analysis is provided in Appendix 2. A BLASTP search using OlsB2 from *R. pomeroyi* DSS-3 (SPO2489) with an e-value cut-off of 10^{-5} was used to retrieve both OlsB and OlsB2 homologs from these genomes. These sequences were manually annotated as being either "OlsB", "OlsB2" or undetermined using the gene neighbourhood view in IMG. OlsB sequences were defined as being upstream of a likely OlsA sequence (annotated as a lyso-ornithine lipid acyltransferase in IMG). Sequences were annotated as OlsB2 if they were downstream of a BamE homolog while sequences matching neither of these criteria were classed as undetermined.

Phylogenetic analysis of the sequences was performed using an ete-build pipeline (Huerta-Cepas, Serra, *et al.*, 2016). Alignment was performed using the metaligner functionality of that pipeline, which creates a consensus alignment from Muscle, Mafft and ClustalOmega alignments using M-coffee to scan the alignments, retaining only consistent columns (Edgar, 2004; Wallace *et al.*, 2006; Sievers *et al.*, 2011; Katoh and Standley, 2013). All alignment programs were run with default parameters and MAFFT was run with the "--auto" on, ensuring that it would automatically determine the most appropriate alignment algorithm to use for the supplied sequences. The consensus alignment was trimmed to remove columns with more than 50% gaps using TrimAl (Capella-Gutiérrez *et al.*, 2009) and this trimmed alignment was manually checked using Jalview (Waterhouse *et al.*, 2009). The pmodeltest functionality of ete3-build, based on ProtTest (Abascal *et al.*, 2005), was then used to identify the most appropriate model of protein evolution for this collection of sequences. This was done by comparing empirical substitution matrix (from JTT, WAG, VT, LG and MtREV), while also testing whether substitution rates should be gamma distributed (+Γ), whether to allow for a proportion of invariable sites (+I) and whether the equilibrium frequencies of amino acids should be estimated from the alignment (+F; Reeves, 1992; Yang, 1993; Cao *et al.*, 1994). Models were compared using Akaike's information criterion which is defined as AIC = 2*(number of parameters)-2*(log likelihood). Lower scores indicate a better model, thus it balances improvements in the likelihood with also penalising additional parameters (Akaike, 1974). The most appropriate model was thus identified as WAG+I+Γ+F. Using this substitution model, a maximum likelihood phylogeny was then constructed with RaxML with 100 bootstrapped replicates (Stamatakis, 2014). An unrooted phylogenetic tree was then drawn using the ete3 tree visualisation software (Huerta-Cepas, Serra, *et al.*, 2016), with each edge coloured according to whether all of its descendant terminal nodes (leaves) were annotated as being OlsB, OlsB2 or undetermined (Figure 4.8). Edges were coloured blue or red if all descendant leaves were OlsB or OlsB2, respectively. If descendant leaves had mixed annotations or were all undetermined, the edges were coloured black.

To investigate the evolutionary history of the OlsB and OlsB2 in the *Rhodobacteraceae*, 16S rRNA sequences from the aforementioned 98 strains, plus those for two *Hyphomonadaceae* (*Hyphomonas neptunium* and *Hirschia baltica*), which were used as an outgroup, were downloaded from the SILVA database (Quast *et al.*, 2013). A maximum likelihood phylogeny was constructed for these sequences using the same pipeline as described above, with the exception that no model testing was performed and a generalised time reversible (GTR) gamma model for nucleic acids was used. The presence or absence of OlsB and OlsB2 (as determined above) was indicated by blue or red boxes, respectively, plotted alongside the phylogeny (Figure 4.7).

2.5.3 Detection of aminolipid synthesis genes in the Tara metagenomes

Instead of performing multiple BLAST searches as described, for example, in Howard *et al.* (2008), I queried the Tara metagenomes using a single Hidden Markov model profile constructed from a nucleotide alignment of reference OlsB, OlsB2 and OlsF sequences using Hmmer3 with an e-value cutoff of 10^{-5} (Eddy, 2011; Wheeler and Eddy, 2013). Profile-based homology searches such as this can be significantly more sensitive than pairwise comparison tools such as BLAST (Eddy, 1996; Park *et al.*, 1998). The reference sequences were chosen to represent OlsB, OlsB2 and OlsF sequences whose functions had been validated experimentally. Ideally this validation would involve direct testing of gene function, as has been carried out for OlsB in *Ensifer meliloti* and for the OlsF of *Serratia proteamaculans* (Gao *et al.*, 2004; Vences-Guzmán *et al.*, 2015). However, given the lack of sequences satisfying this criterion, I also included sequences from bacteria that have been shown to produce ornithine lipid, or in the case of OlsB2, glutamine lipid. A reference alignment of these sequences was obtained as described in Section 2.5.2.

This reference alignment was also used to construct a maximum likelihood phylogenetic tree using the GTRGamma model of RaxML (Stamatakis, 2014) with

100 bootstrap replicates. In order to classify the sequences retrieved from the metagenomes by the Hmm search, their maximum likelihood placement onto this reference phylogeny was determined using pplacer (Matsen *et al.*, 2010). Each metagenome sequence was thus placed onto a single best-scoring edge of the phylogenetic tree, and was classified as either OlsB, OlsB2 or OlsF based on the identity of the reference sequences on the leaves descending from this edge. Sequences placed onto edges with descendant leaves having more than one identity were left unclassified. Taxonomic information for each metagenomics sequence was obtained by performing best hit searches using BLASTx against the NCBI non-redundant protein database (nr). The abundance of each gene in the Tara metagenomes was standardised to RecA abundance as described in Section 2.5.1, with RecA sequences retrieved using an Hmm search (e-value cutoff 10^{-5}) with the high quality reference profile from the Pfam database (Punta *et al.*, 2012).

2.5.4 Presence of homotaurolipids in cultured Roseobacters

To assess the distribution of homotaurolipids (HTLs) in the Roseobacter group, I analysed lipid extracts from a number of readily cultivable strains. The strains analysed were mostly from clade 1 as defined by Simon *et al* (2017) although *Dinoroseobacter shibae* is more basal within the *Rhodobacteraceae* and recent studies indicate that *Stappia stellulata* may not belong to the same Family, despite currently being classed within the *Rhodobacteraceae* (Pujalte *et al.*, 2013). All strains were grown overnight in marine broth before samples for lipid analysis were collected, extracted and analysed as described in Section 2.3. A 16S rRNA phylogeny was constructed using sequences downloaded from the SILVA database and in the same way as described for the 16S phylogeny described in Section 2.5.2. This phylogeny was then plotted alongside homotaurolipid abundance using ete3 (Huerta-Cepas, Serra, *et al.*, 2016). Since there are no authentic standards available to quantify HTL abundance, the absolute abundance estimates (as a ratio to the abundance of the PG internal standard) do not have a straightforward interpretation like a concentration or molar amount would. However, comparisons of the abundance of the same lipid

class between closely related species should be approximately valid so I have presented HTL abundances across strains as a proportion of the maximum abundance (which was detected in *Roseovarius nubinhibens* ISM). These proportions were log-transformed for plotting purposes since in most of the strains the abundance of HTL was less than 10% of that detected in *R. nubinhibens*.

2.5.5 Comparative genomics to identify potential genes involved in HTL synthesis

The genomes of the strains grown in Section 2.5.4 were downloaded from the NCBI database. Not all of the strains used had been genome sequenced - of the 16 strains initially grown, genome sequences were available for 11. These consisted of 9 strains that had been found to produce HTL and 2 strains (*S. stellulata* and *D. shibae*) that did not. In order to identify genes potentially involved in HTL synthesis, I assigned each gene from the 11 genomes to an orthologous group using eggNOG mapper (Huerta-Cepas, Forslund, *et al.*, 2016). This program conducts a BLAST search of each sequence against eggNOG database (Huerta-Cepas, Szklarczyk, *et al.*, 2016) of orthologous genes, with the query sequence being annotated with the same orthologous group as the best BLAST hit. Orthologous groups that were present in the genomes of all HTL-producing strains but absent from the genomes of *S. stellulata* and *D. shibae* were considered to be potentially involved in HTL synthesis.

2.6 Statistical analysis

Linear regression was carried out to investigate potential relationships between the abundance of aminolipid synthesis genes, identified as described in Section 2.5.3, and a measure of relative phosphate availability. While the ratio of nitrogen to phosphorus is traditionally used for this purpose in microbial ecology (Redfield, 1958), the use of a ratio has issues when values for the denominator are close to or at zero, as is often the case for phosphate measurements in the ocean. For this reason I chose to use an alternative measure for the relative abundance of these two nutrients which also has an intuitive interpretation: the excess of nitrogen relative to the phosphorus for an

'average' Redfield ratio phytoplankton (Michaels *et al.*, 1996; Deutsch and Weber, 2012):

$$N^* = [NO_3^-] - 16.[PO_4^{3-}]$$

The data for the abundance of OlsB, OlsB2 and OlsF takes the form of count data and initial visual inspection of these counts suggested that a normal approximation was not appropriate for modelling them. Therefore it was natural to choose a generalised linear model from the Poisson family, which are explicitly formulated for count data (O'Hara and Kotze, 2010). The Poisson distribution assumes that the variance is proportional to the mean, an assumption that is often violated with ecological data. An initial investigation into this issue revealed that data for aminolipid synthesis genes were overdispersed (a variance larger than the mean), so I instead adopted a negative binomial model, which is often more appropriate for such data (Ver Hoef and Boveng, 2007). Since the response variable of a Poisson regression has to be a non-negative integer, a typical strategy to model rate data is to include the normalising parameter as an offset in the regression, with its coefficient set to 1 (Zuur *et al.*, 2013). As such I modelled the raw count data for the genes as the response variable with RecA abundance, corrected for the relative lengths of the Hmm profiles used, included as an offset. This is the equivalent of modelling the ratio of each aminolipid synthesis gene to RecA. A log link was used in all cases, which ensures that predicted count values cannot be negative.

In order to assess whether N* was a significant predictor for the abundance of each aminolipid synthesis gene, I compared two models for each gene. A base model was constructed in which the abundances of each of the microbial groups to which sequences for that gene were assigned as per Section 2.5.3 were included as covariates (Table 2.6). This was compared to second model that was identical to the base model but with the addition of a term for N* using likelihood ratio tests. The abundance values for each microbial group were calculated from metagenomic 16S Illumina tag

(miTAG) data (Logares *et al.*, 2014), which is available from http://ocean-microbiome.embl.de/companion.html.

Table 2.6 Microbial groups used as covariates in models of aminolipid synthesis gene abundance.

Microbial group	OlsB	OlsB2	OlsF
Rhodobacteraceae	Yes	Yes	
Rhodospirillales	Yes		
SAR11	Yes		
Rhizobiales	Yes		
Other *Alphaproteobacteria*	Yes		
Gammaproteobacteria	Yes	Yes	Yes
Bacteroidetes			Yes
Other phyla	Yes		

To investigate possible differences in the relationship between aminolipid synthesis gene abundance and N* between microbial groups, I constructed negative binomial GLMs with the group-specific counts of each gene as the response variable. Microbial group was included as a categorical factor, alongside group abundance and N*. The interaction between microbial group and N* was also modelled. ANOVA tests indicated that the interaction term was statistically significant for OlsB, indicating differences in the slope of the relationship between N* and OlsB abundance differed between microbial groups. Inspection of z-scores revealed this interaction term was statistically significant ($p < 0.05$) for two groups: *Rhizobiales* and Other phyla. Negative binomial regressions for each of these groups individually, including the group abundance and N* as covariates was used to test whether the slope of the regression was significantly greater than zero. All models were fitted using the R statistical software package.

Chapter 3 Lipid remodelling is a widespread strategy in marine heterotrophic bacteria upon phosphorus deficiency

3.1 Introduction

Bacteria and phytoplankton compete for scarce nutrients, such as phosphorus, in the ocean (Zubkov *et al.*, 2007). The relative competitiveness of these two groups in environments where nutrients are scarce has important implications for global biogeochemical cycles (Havskum *et al.*, 2003). Heterotrophic bacteria have long been considered to thrive in the oligotrophic ocean, where respiration often exceeds primary production (Del Giorgio *et al.*, 1997) and this success was theorised to be in part a result of successful competition for nutrients by dominant heterotrophic bacteria in these environments (Cotner and Biddanda, 2002). Recently, however, contradictory evidence regarding the ability of marine heterotrophic bacteria to reduce the P content of their membrane has emerged. Membrane phospholipids have been estimated to comprise around 15–20 % of total cellular P in bacteria (Karl and Bjorkman, 2014). Thus, bacteria would be able to significantly reduce their cellular P quota by replacing these lipids with P-free lipids, as has been documented in laboratory studies (Minnikin and Abdolrahimzadeh, 1974; Minnikin *et al.*, 1974). In environments where P is a limiting nutrient, a reduced P quota should increase competitiveness (Droop, 1974; Thingstad and Rassoulzadegan, 1999). A series of studies documented such lipid remodelling by both eukaryotic phytoplankton and cyanobacteria in response to P limitation (Van Mooy *et al.*, 2006, 2009). Surprisingly, no evidence was found for lipid remodelling by the most abundant groups of marine heterotrophic bacteria (Van Mooy *et al.*, 2009) and studies in the oligotrophic Pacific Ocean found that heterotrophic bacteria were the major producers of phospholipids (Van Mooy *et al.*, 2006, 2008). This apparently obligate requirement for phospholipids would appear to place heterotrophic bacteria at a competitive disadvantage to phytoplankton in P-depleted environments. However, in one study,

heterotrophic bacteria from the P-depleted Sargasso Sea accumulated glycolipids when incubated with glucose in microcosms (Popendorf *et al.*, 2011), suggesting that at least some marine bacterioplankton are capable of lipid remodelling. In fact, analysis of metagenomics data has revealed that genes involved in adaptation to P-deficiency seem to be specifically acquired as a response to low P availability in the environment (Coleman and Chisholm, 2010). Therefore, it is possible that only heterotrophic bacteria specifically adapted to P-depleted environments in the ocean have acquired the ability to remodel their lipid membrane in response to P-stress.

In this study, I investigated the ability of bacteria from the abundant marine Roseobacter group to carry out lipid remodelling in response to P-limitation, employing *Phaeobacter* sp. MED193, isolated from the P-depleted Mediterranean Sea, as a model system. Using a marker-exchange mutagenesis strategy, I demonstrate that a phospholipase C (PlcP), first characterised in the soil bacterium *Ensifer meliloti* (Zavaleta-Pastor *et al.*, 2010), is central to lipid remodelling in a model Roseobacter strain. Furthermore, I show that this PlcP is abundant in marine metagenomes and distributed across most major lineages of marine bacteria.

3.2 Results

3.2.1 plcP is expressed under P-deficiency conditions in Phaeobacter *sp. MED193 and is essential for lipid remodelling*

To test the hypothesis that at least some marine heterotrophic bacteria synthesize non-P lipids in response to P deficiency I worked with an isolate from the marine Roseobacter group, *Phaeobacter* sp. MED193. The genome of this strain contains a homolog sharing 57% similarity with the newly discovered *plcP* from a terrestrial rhizobium (Zavaleta-Pastor *et al.*, 2010). I subjected the strain to abrupt P starvation, following initial growth in P-replete medium (phosphate concentration 50 μM), and studied the P starvation response over time (Section 2.4.1). Alkaline phosphatase activity was used as a diagnostic for P deficiency (Thingstad and Mantoura, 2005) and activity was indeed induced in the P-starved cultures (Figure 3.1 a).

Phaeobacter sp. MED193 synthesised a mixture of phospholipids and non-P lipids under both P-replete and P-depleted conditions (Figure 3.2 a, b). Phospholipids detected were phosphatidylethanolamine (PE), phosphatidylglycerol (PG) and small amounts of phosphatidylcholine (PC). Non-P lipids consisted of a number of amino acid-containing lipids: ornithine lipid (OL), glutamine lipid (QL), and two further aminolipids (AAL1 and AAL2). In addition to these, the P-free betaine lipid diacylglyceryl-*N,N,N*-trimethyl homoserine (DGTS) was abundant in P-starved cultures. There were also a number of lipids present whose structures could not be confidently identified. The proportion of the phospholipids PE and PG decreased following P starvation, whereas the non-P lipid DGTS became more abundant, apparently becoming the dominant lipid in P-starved cultures (Figure 3.2). The abundance of DGTS increased rapidly in cultures which were suspended in defined medium lacking any added source of P, reaching a maximum around 2 days following re-suspension (Figure 3.1 b).

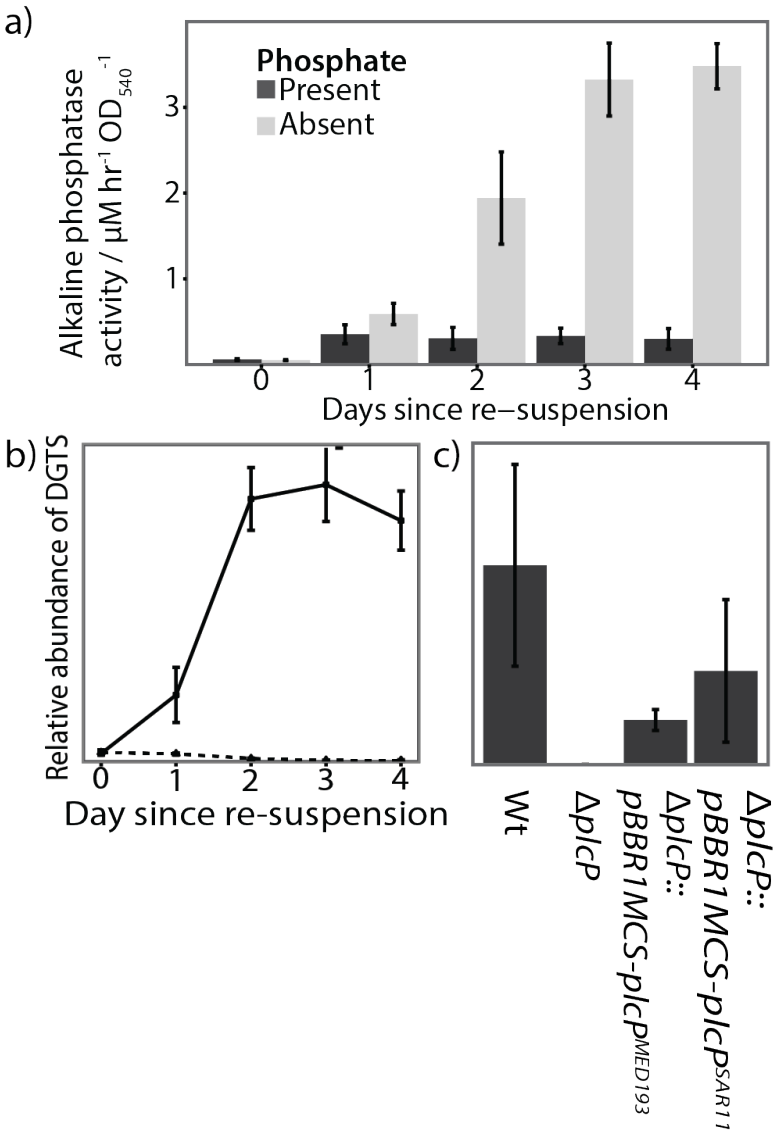


Figure 3.1 Phosphorus starvation induces alkaline phosphatase activity **(a)** and lipid remodelling **(b, c)** in *Phaeobacter* sp. MED193. (a) Bar chart showing induction of alkaline phosphatase activity in cultures incubated without added phosphate. Alkaline phosphatase activity was measured as the rate of formation of coloured *para*-nitrophenol following *para*-nitrophenol phosphate addition. Line graph (b) shows the accumulation of diacylglyceryl-*N,N,N*-trimethylhomoserine (DGTS) in *Phaeobacter* sp. MED193 over time in P-deplete (solid line) but not in P-replete (50 µM PO_4^{3-}; dashed line) media. DGTS synthesis is observed in wild type (Wt) but not in a *plcP* deletion mutant (*ΔplcP*) (c). Complementation with *plcP* from MED193 (*ΔplcP*::pBBR1MCS-$plcP^{MED193}$) or from SAR11 (*ΔplcP*::pBBR1MCS-$plcP^{SAR11}$) is sufficient to restore DGTS synthesis. Error bars represent the standard deviation of three independent replicates. OD_{540}: optical density at 540 nm.

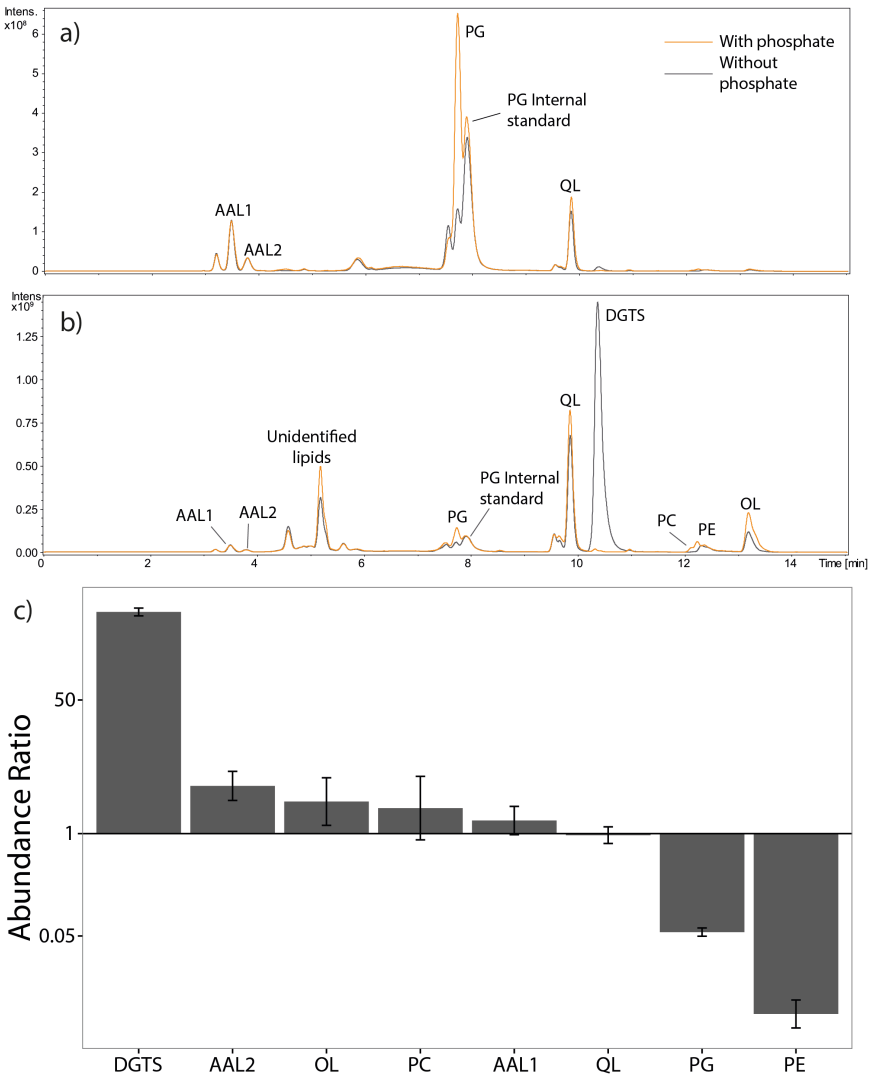


Figure 3.2 Lipid composition of *Phaeobacter* sp. MED193 (wild type) cultures grown with 50 µM phosphate and without added phosphate. Phospholipid abundance significantly decreased in P-starved cultures while betaine lipid abundance increased. Chromatograms a) and b) show, respectively, representative negative and positive chromatograms of whole lipid extracts from MED193 cultures. c) Ratio of the abundance of lipid classes in MED193 cultures grown with and without phosphate. Ratios were calculated by dividing the mean abundance of each lipid class in extracts from cultures grown without phosphate by that of cultures grown with phosphate. Abundance values for each lipid class were calculated by normalising peak areas to phospholipid internal standards. Phosphatidylglycerol (PG) internal standard was used to normalise PG, ornithine lipid (OL), glutamine lipid (QL), and two uncharacterised aminolipids (AAL1 and AAL2). Phosphatidylethanolamine (PE) internal standard was used to normalise PE. Phosphatidylcholine (PC) internal standard was used to normalise PC and diacylglyceryl-*N,N,N*-trimethylhomoserine (DGTS). Error bars show ± standard deviation of three biological replicates.

To determine whether PlcP is directly involved in lipid remodelling in *Phaeobacter* sp. MED193, I constructed a deletion mutant, Δ*plcP*, in which the *plcP* homolog was replaced by a gentamicin resistance cassette following homologous recombination between two flanking regions up- and downstream of the gene to be deleted (see Section 2.2.2 for details, Figure 2.1). Intermediate constructs were validated using a combination of PCR and restriction digests (Figure 3.3 a-e) while the final mutant strain was validated by PCR using primers targeting the gentamicin resistance cassette and the downstream homologous sequence (Table 2.3, Figure 3.3 f). This PCR product was also sequenced to check for the introduction of any point mutations. When grown under P-limiting conditions, the betaine lipid DGTS, was not detectable in the mutant (Figure 3.1 c). To eliminate the possibility of polar effects during mutant construction, I also verified that DGTS synthesis was restored by complementation of the Δ*plcP* mutant (Section 2.2.4) with both the native *plcP* of *Phaeobacter* sp. MED193 (Δ*plcP*::pBBR1MCS-$plcP^{MED193}$) and the *plcP* of the SAR11 strain HTCC7211 (Δ*plcP*::pBBR1MCS-$plcP^{SAR11}$). The genes were introduced to *ΔplcP* as inserts on the broad host-range plasmid, pBBR1MCS with the size of the insert checked by restriction digest of the flanking *BamHI* and *HindIII* sites (Figure 3.4). In both cases, complementation restored the synthesis of DGTS (Figure 3.1c). The ability of $plcP^{SAR11}$ to complement *ΔplcP* indicates that both enzymes have a similar functional behaviour.

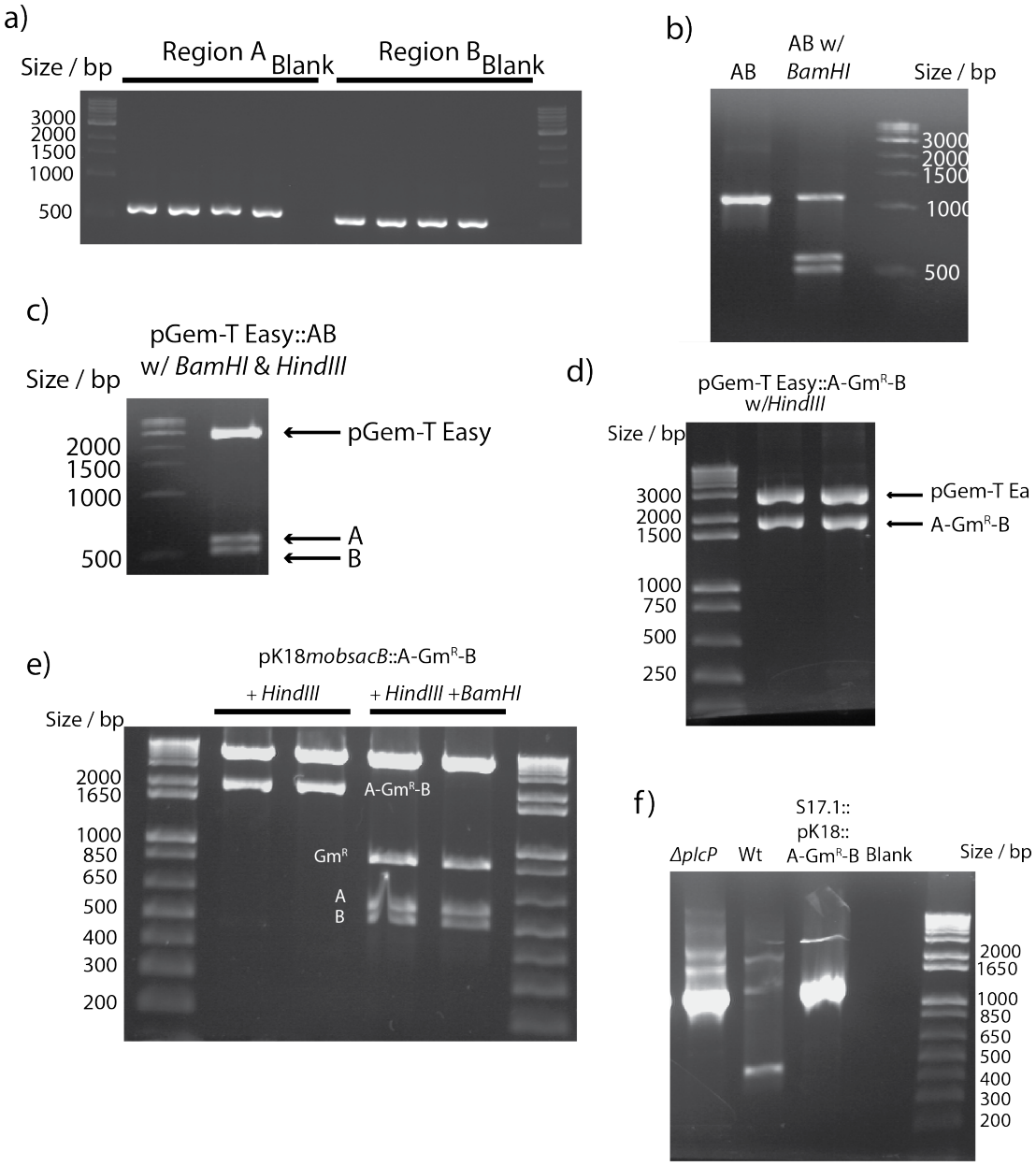


Figure 3.3 Validation of intermediate constructs during *ΔplcP* construction (**a-e**) and final *ΔplcP* mutant (**f**). **a)** PCR of *Phaeobacter* sp. MED193 colonies using primers for region A (591 base pairs) or region B (528 base pairs). **b)** Restriction digest of AB using *BamHI*. **c)** Restriction digest of pGEM-T::AB with *BamHI* and *HindIII*. **d)** Restriction digest of pGEM-T::A-Gm^R-B with *HindIII* yielding A-Gm^R-B insert (1956 base pairs). **e)** Restriction digest of pK18*mobsacB*::A-Gm^R-B with either *HindIII* alone or *HindIII* plus *BamHI*, yielding either A-Gm^R-B or the individual components, regions A and B plus Gm^R (865 base pairs). **f)** *ΔplcP* in *Phaeobacter* sp. MED193 using gentamicin resistance cassette (Gm^R) forward primer and B reverse. Run alongside wild type (Wt) strain, *Escherichia coli* S17.1 carrying the pK18 construct used to make the deletion mutant and a distilled water blank.

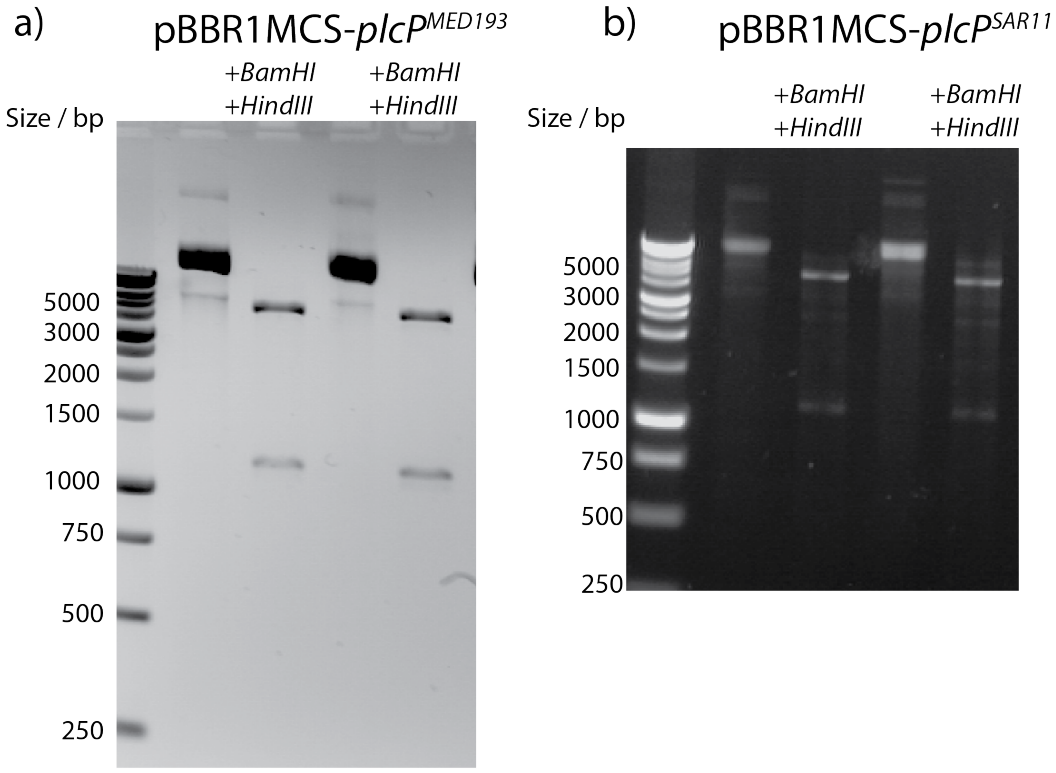


Figure 3.4 Restriction digests to confirm plasmids used for complement mutant construction in *Phaeobacter* sp. MED193. **a)** Undigested and restricted pBBR1MCS-Km with native MED193 *plcP* plus promoter region. **b)** Undigested and restricted pBBR1MCS-Km with *plcP* from SAR11 fused to the MED193 promoter region. In both cases restriction with *BamHI* and *HindIII* results in a fragment between 1000 and 1500 base pairs (bp), corresponding to the expected sizes of 1193 bp and 1250 bp for the MED193 and SAR11 *plcP* plus promoter fragments, respectively.

3.2.2 *PlcP is abundant and widely distributed among diverse phyla in the marine environment*

Having established that PlcP is central to the lipid remodelling response to P deficiency, I proceeded to investigate the abundance and distribution of PlcP in the marine environment. I mined the genomes of marine isolates in the Integrated Microbial Genomes (IMG) database as well as the Global Ocean Sampling (GOS) (Rusch *et al.*, 2007) and Tara Oceans metagenomes (Sunagawa *et al.*, 2015) for PlcP homologs (Section 2.5.1). BLAST searches retrieved 216 PlcP homologs in marine isolates (Appendix 1), and more than 1,100 unique environmental PlcP homologs in the GOS metagenomes. PlcP was also found to be ubiquitous in the Tara Oceans metagenomes. Phylogenetic analysis of these sequences provided evidence that PlcP is widespread among diverse bacterial phyla, including *Alphaproteobacteria*, *Gammaproteobacteria* and *Bacteroidetes* (Figure 3.5 a). Notably, PlcP is present in the two tropical oceanic representatives of the ubiquitous SAR11 clade ('*Candidatus* Pelagibacter ubique' sp. HTCC7211 and HTCC7217), which was also concurrently shown by Carini *et al.* (2015). *Alphaproteobacteria* appeared to dominate in the GOS and Tara Oceans dataset, accounting for around 75% of the hits (Figure 3.5 b, c). Close relatives of the SAR11 clade PlcP, represented by strain HTCC7211, accounted for 54% and 30% of the hits in the GOS and Tara Oceans datasets, respectively. I also found PlcP homologs in single amplified genomes from a number of other abundant lineages of marine bacteria. These included the clade SAR116 of *Alphaproteobacteria*, the gammaproteobacterial SAR86, a flavobacterium and a verrucomicrobium (Dupont *et al.*, 2012; Swan *et al.*, 2013). As an adaptation to P deficiency, PlcP is expected to be more abundant in areas of the oceans where low P availability exerts a strong selective pressure. Analyses of the relative abundance of PlcP in the metagenomes from the GOS and the Tara Oceans expedition show that PlcP is indeed more abundant in the Mediterranean Sea and the North Atlantic (Figure 3.6), which are characterized by low phosphate concentrations and high N/P ratios compared to other areas of the oceans (Fanning, 1992; Wu *et al.*, 2000; Krom *et al.*, 2010).

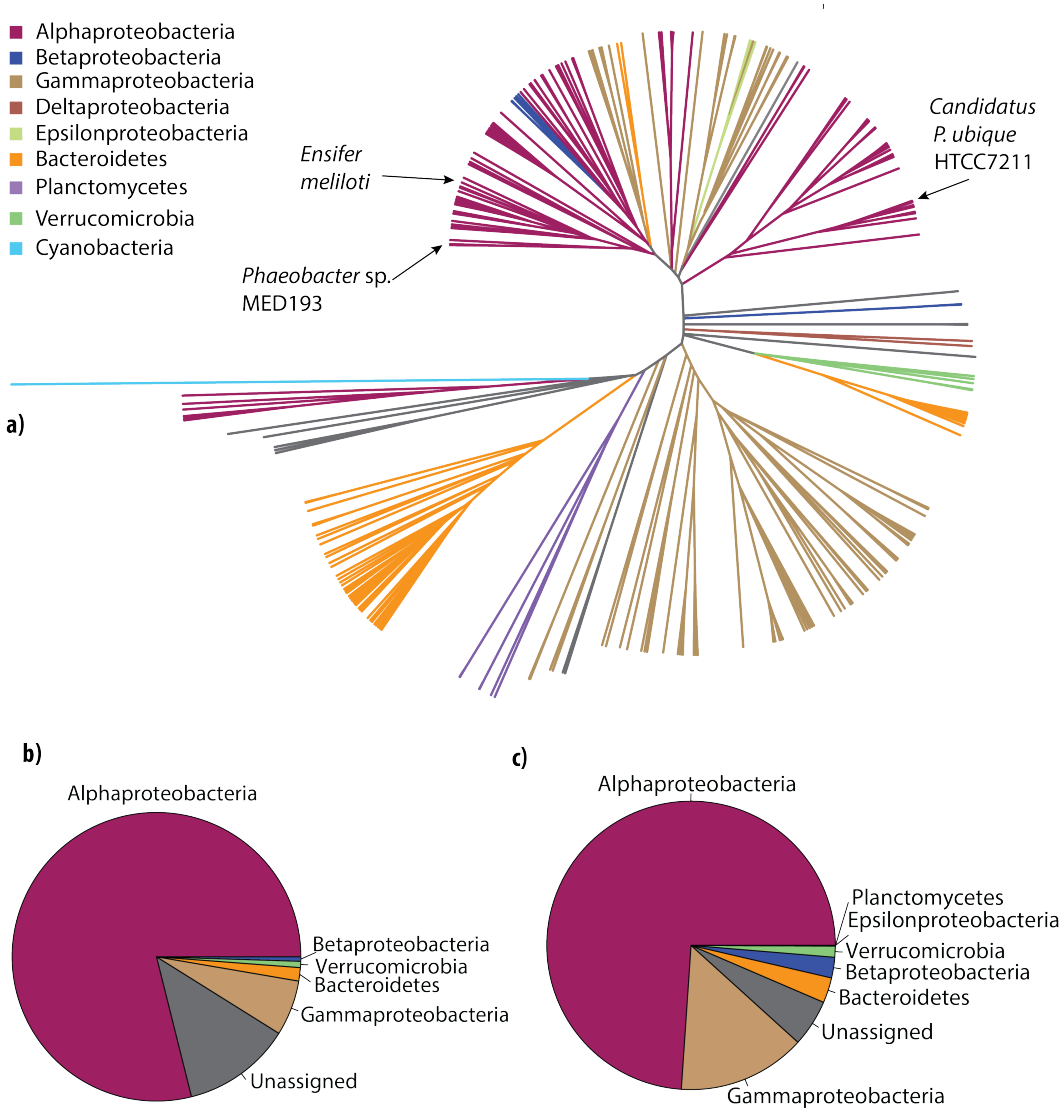


Figure 3.5 Distribution of PlcP in marine bacteria. a) Phylogenetic relationships of PlcP homologs in marine bacterial isolates and all non-redundant PlcP sequences in the Global Ocean Sampling dataset. Sequences were retrieved by BLASTP using *Phaeobacter* sp. MED193 PlcP (MED193_17359) as query (e-value < 10^{-40}). Uncoloured sequences are metagenome sequences that did not cluster alongside any isolate sequences. b,c) Percent distribution of metagenome hits within each phylogenetic group in the Global Ocean Sampling and Tara Oceans datasets, respectively.

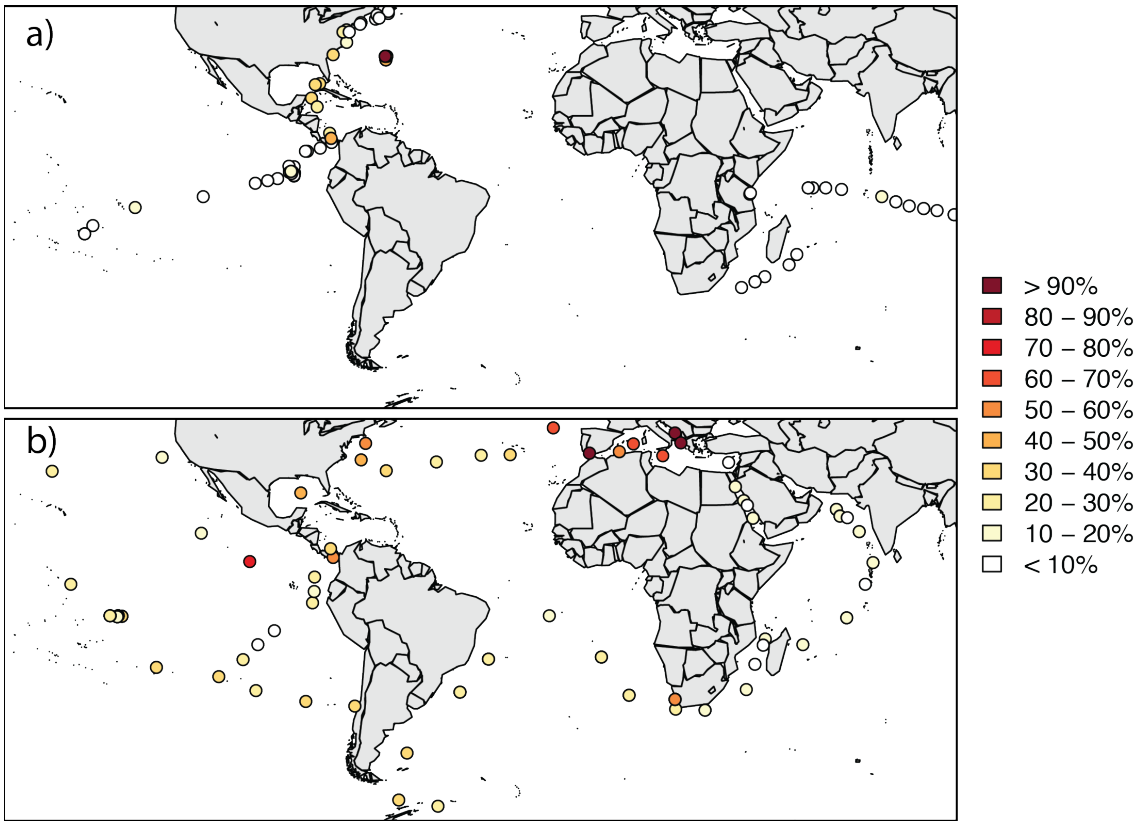


Figure 3.6 Distribution of PlcP homologs in marine metagenomes. Sampling sites are coloured according to the estimated percentage of cells with a PlcP homolog. a) Distribution in the Global Ocean Sampling database. BLASTP searches were performed using *Phaeobacter* sp. MED193 PlcP (MED193_17359) as query (e-value < 10^{-40}). The number of reads at each site was normalized using BLASTP hits to RecA from MED193 (MED193_04366; e-value < 10^{-40}). b) Distribution of PlcP in the Tara dataset.

3.2.3 Genomic context of plcP

In order to shed further light on the role of PlcP I analysed its genomic context in marine isolates in the IMG dataset. Based on the genomic context of *plcP*, there are broadly 6 types of gene arrangements, and a gene putatively involved in non-P lipid synthesis was detected in the neighbourhood of *plcP* in around two thirds of the strains investigated (Appendix 1).

Type 1 is present in several strains, including *Phaeobacter* sp. MED193, in which the two genes necessary for the synthesis of the betaine lipid DGTS, *btaA* and *btaB* (Riekhof *et al.*, 2005), are found downstream of *plcP* (Figure 3.7 a). This is consistent with the lipid results obtained in strain MED193 cultures and its mutants (Figure 3.1 b, c). The most commonly observed gene in the vicinity of *plcP* (type 2, present in over 25% of the bacterial isolates analysed, Appendix 1) was a putative glycosyltransferase, homologous to the promiscuous glycosyltransferase *agt* recently identified in *Agrobacterium tumefaciens* (Semeniuk *et al.*, 2014). Strains with agt homologs downstream of *plcP* include the environmentally relevant '*Candidatus* Pelagibacter ubique' sp. HTCC7211 and HTCC7217 (Figure 3.7 a). In *A. tumefaciens*, this gene was found to be required for the synthesis of GADG and MGDG under P deficiency (Semeniuk *et al.*, 2014), which is consistent with the recent findings of Carini *et al.* (2015).

To confirm the role of the glycosyltransferase *agt* in the synthesis of these lipids, the gene (PB7211_960) from '*Candidatus* Pelagibacter ubique' sp. HTCC7211 was codon-optimised for *E. coli*, chemically synthesised and inserted into a pET28a expression vector (see Section 2.2.5). The pET28a-agt^{SAR11} construct was checked by both sequencing and restriction digests of flanking *NdeI* and *BamHI* sites (Figure 3.8a). Protein extracts from *E. coli* BL21(DE3)::pET28a-agt^{SAR11} cultures, both with and without added IPTG (induced and uninduced, respectively), were run on an SDS-PAGE gel. A dense band running just above the 35 kDa size marker appeared to be

enriched in the whole cell and pellet fractions of the IPTG-induced culture (Figure 3.8 b). This corresponds to the theoretical weight of the PB7211_960 protein product of 37.9 kDa, suggesting that the overexpressed protein might have been forming inclusion bodies. Purification of inclusion bodies from the cell pellet fraction, and subsequent His-tag purification resulted in the elution of a clear band on an SDS-PAGE gel at around 38 kDa (Figure 3.8 c). The bands for the second elution fractions of both the induced and uninduced cultures were excised and sequenced using an Orbitrap mass spectrometer. Peptides retrieved from the two bands were 55 and 57 amino acids in length (for the induced and uninduced cultures respectively), with 94 and 97% coverage of the PB7211_960 protein and were 100% identical to it. These results indicated the successful overexpression of agt^{SAR11} in *E. coli* BL21(DE3). Analysis of lipid extracts from these cultures indicated the accumulation of MGDG and GADG (Figure 3.7 b-d), which were not observed in lipid extracts from the same strain transformed with an empty vector (Figure 3.7 b), confirming that this *agt* homolog from HTCC7211 is sufficient and responsible for the synthesis of the two glycolipids, MGDG and GADG.

Types 3 and 5 (Figure 3.7 a) include a number of marine heterotrophic bacteria where *plcP* is located immediately upstream of some putative glycosyltransferases, one of which has been confirmed to be responsible for the synthesis of di- and tri-glycosyldiacylglycerols (Devers *et al.*, 2011). However, in type 4 and type 6, *plcP* does not seem to form an operon with its neighbouring genes although genes predicted to be involved in the synthesis of non-P lipids (e.g. *btaBA*) are, in many cases, found elsewhere in their genomes (Appendix 1). Together, these analyses of the genomic context of *plcP* support its role in the synthesis of non-P glycolipids and betaine lipids in marine heterotrophic bacteria (Figure 1.5).

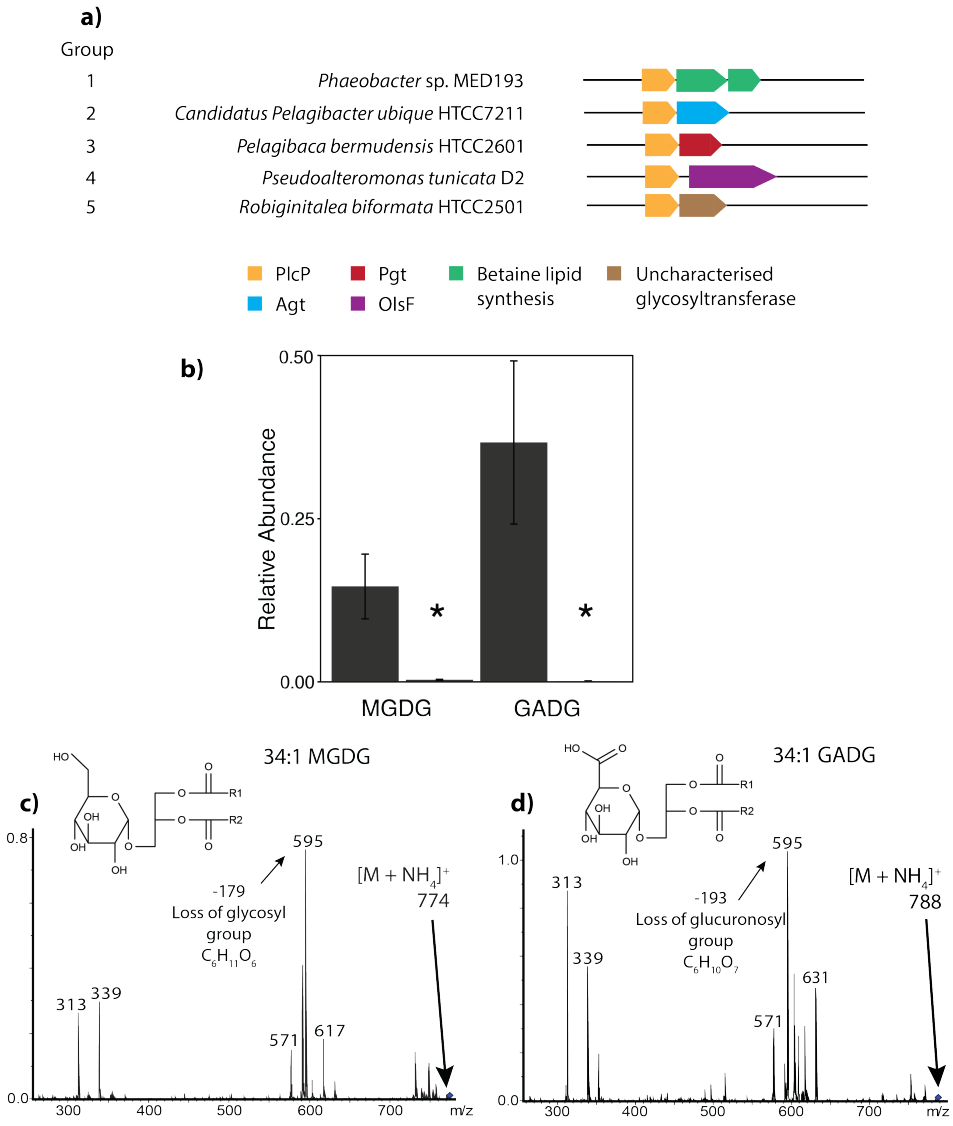


Figure 3.7 PlcP often appears in an operon with a glycosyltransferase. **a)** Genomic context of *plcP*. Representatives are shown of each of the major type of *plcP* genomic context, based on an analysis of sequence marine bacteria (Appendix 1). **b)** Heterologous expression of the putative glycosyltransferase *agt* from SAR11 HTCC7211 results in the accumulation of two glycolipids: monoglycosyl diacylglycerol (MGDG) and glucuronic acid diacylglycerol (GADG) (dark grey bars), but not in the control harbouring an empty expression plasmid (light grey bars, also marked with asterisks). Abundance is expressed as the peak area of glycolipid relative to that of a phosphatidylcholine internal standard. **c)** Fragmentation spectra for representative MGDG and **d)** GADG species obtained from recombinant *E. coli* harbouring the HTCC7211 *agt* homolog. The two species differ in the neutral loss corresponding to the polar head group (179 and 193 *m/z* for the loss of hexosyl and hexuronic acid groups, respectively). In each case this loss yields diacylglycerol (595.6) and a further two peaks are present corresponding to monoacylglycerol with 18:1 (339) and 16:0 (313) fatty acids.

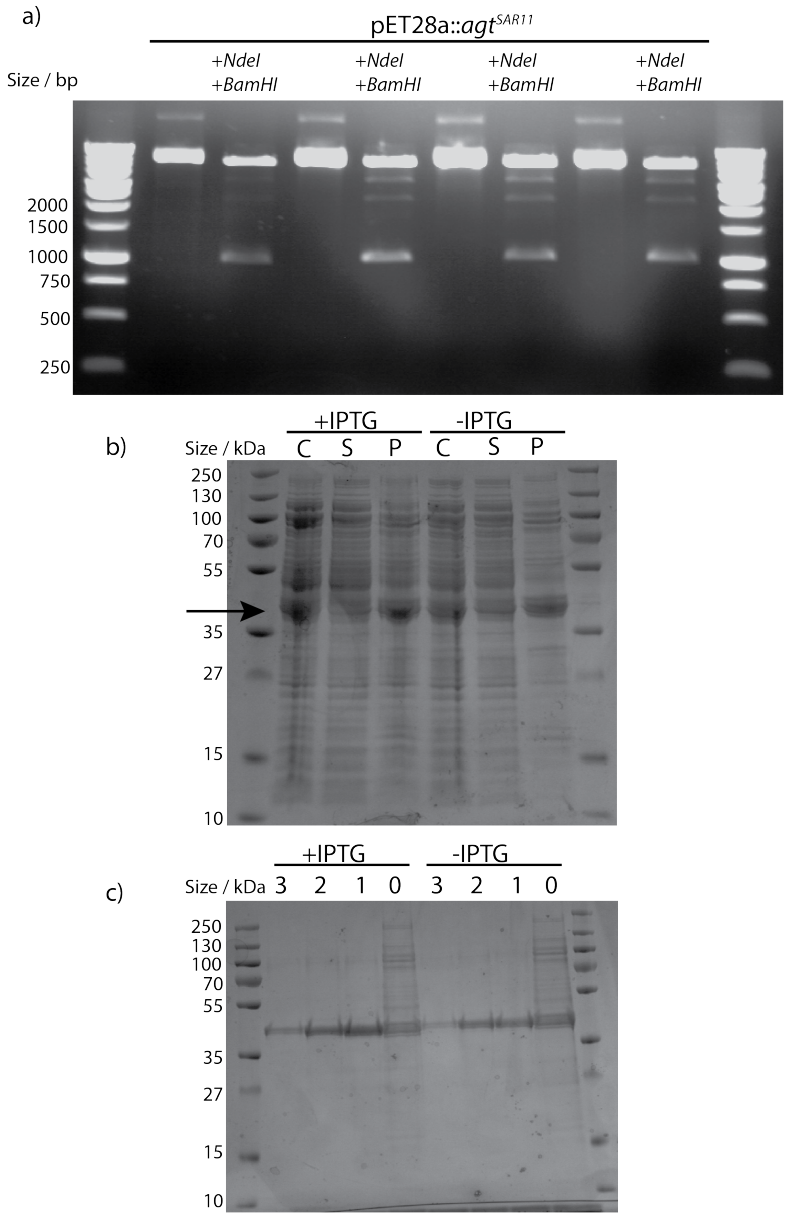


Figure 3.8 Confirmation of Agt-overexpressing *E. coli* BL21(DE3). **a)** Restriction digest of the pET28a construct into which *agt* from SAR11 was inserted. Restriction digest with *NdeI* and *HindIII* resulted in the formation of a fragment around 1000 base pairs (bp), consistent with the expected size of 1010 bp. **b)** SDS - polyacrylamide gel electrophoresis (SDS-PAGE) of extracts from induced (+IPTG) and uninduced (-IPTG) BL21(DE3) harvested after growth in M9 medium. The whole cell profile (C) was run alongside separate supernatant (S) and pellet (P) fractions. An arrow points to a potentially enriched band in the C and P fractions. **c)** SDS-PAGE of purified inclusion following 0 to 3 rounds of His tag purification (denoted along the top). The band from the induced culture following 2 rounds of His tag purification was cut out of the gel and sequenced.

3.3 Discussion

3.3.1 Lipid remodelling is a widespread strategy in marine heterotrophic bacteria under low-P conditions

The ability to synthesize non-P lipids seems to be widespread amongst marine heterotrophic bacteria adapted to low P environments. I found that in these P-depleted regions, such as the North Atlantic Subtropical Gyre and the Mediterranean Sea, PlcP may be present in the majority of bacterial cells (Figure 3.6). While alphaproteobacterial PlcP sequences are the most abundant in marine metagenomes, PlcP appears to be broadly distributed amongst bacteria with distinct ecologies. Whereas the planktonic oligotrophic SAR11 clade are specialised for the high affinity uptake of small compounds (Giovannoni *et al.*, 2005), PlcP was also identified in the SAR86 clade, believed to dominate the uptake of high molecular weight organic matter (Dupont *et al.*, 2012). Members of groups with a particle-associated lifestyle, such as the *Roseobacter* clade and *Bacteroidetes* (Wagner-Döbler and Biebl, 2006; Fernández-Gómez *et al.*, 2013), also harboured PlcP homologs. Lipid remodelling mediated by PlcP therefore does not appear to be restricted to bacteria with particular lifestyles but rather constitutes part of a generic response to low P among marine bacteria from P-depleted environments.

One previous explanation for the accumulation of P-free lipids in marine heterotrophic bacteria is that they are used to increase cell size without changing the cellular P quota (Thingstad *et al.*, 2005; Van Mooy *et al.*, 2009). However, I observed a pronounced reduction in the abundance of phospholipids in P-starved cultures of wild-type *Phaeobacter* sp. MED193 (Figure 3.2), indicating a *bona fide* P-sparing response. In line with this observation, my collaborators observed reductions in the amount of phospholipids per cell without obvious changes in cell size or in estimated cell volume (Sebastián *et al.*, 2016). Together with the very low membrane P content found in environmental heterotrophic bacteria (Sebastián *et al.*, 2016), these data suggest an alternative strategy whereby lipid remodelling is used to reduce the cellular P quota, in line with what was recently observed in a SAR11 isolate (Carini et al. 2015).

Enzymes involved in the synthesis of these P-free glycerolipids, including DGTS, SQDG, MGDG and GADG, require DAG as a substrate (Klug and Benning, 2001; Hölzl *et al.*, 2005; Semeniuk *et al.*, 2014). Bacterial lipid synthesis proceeds via a shared phosphatidic acid (phosphorylated DAG) intermediate (Parsons and Rock, 2013). The generation of DAG requires either the de-phosphorylation of phosphatidic acid or the removal of the polar head group from a phospholipid by the action of a phospholipase C (e.g. PlcP). *Phaeobacter* sp. MED193 as well as the SAR11 representative, HTCC7211, appear to lack homologs of bacterial phosphatidic acid phosphatases (Icho and Raetz, 1983). Thus, it is likely that the major means of generating DAG in marine heterotrophic bacteria is through the action of PlcP (Figure 1.5). This is consistent with the abolition of DGTS synthesis in the *Phaeobacter* sp. MED193*ΔplcP* deletion mutant (Figure 3.1) and the complementation of the *Phaeobacter* sp. MED193 *Δ plcP* deletion mutant with the PlcP of SAR11 strain HTCC7211. On the other hand, in *E. coli,* which is not able to produce PlcP, DAG can be generated through the degradation of PG by MdoB (Parsons and Rock, 2013), which explains why glycolipids could accumulate in *E. coli* following over-expression of *agt* from '*Candidatus* Pelagibacter ubique' sp. HTCC7211 without an exogenous supply of DAG. In contrast to the widespread distribution of PlcP homologs across several phyla of marine heterotrophic bacteria, they are largely absent from the genomes of cyanobacteria (Figure 3.5), which are also capable of lipid remodelling (Van Mooy *et al.*, 2006, 2009). It is likely that cyanobacteria instead generate the required DAG through the action of phosphatidic acid phosphatases which are found in their genomes (Nakamura *et al.*, 2007).

The ability to remodel lipids in response to P scarcity appears to be ubiquitous in phytoplankton (Van Mooy *et al.*, 2009), whereas in heterotrophic bacteria it seems to be more restricted to those adapted to low-P conditions (Figure 3.6). An explanation for this may be that the typical replacement lipids in cyanobacteria, SQDG and MGDG, are also required for the proper function of photosynthetic membranes, for example in association with photosystem II (Umena *et al.*, 2011). Heterotrophic

bacteria, by contrast, appear to require primarily PE and PG as bulk membrane lipids (Parsons and Rock, 2013). There is a selective pressure against the accumulation of genes that confer an insufficient advantage in bacterial genomes (Mira *et al.*, 2001) and this deletional bias has been invoked as an explanation for the patchy distribution of genes conferring an adaptation to P deficiency in the marine environment (Coleman and Chisholm, 2010). An example of this is found among the SAR11 bacteria. '*Candidatus* Pelagibacter ubique' sp. HTCC7211, isolated from the P-deplete Sargasso Sea contains a higher number and diversity of P acquisition genes than other SAR11 strains (e.g. *P. ubiqu*e HTCC1062), thereby increasing its ability to use alternative sources of P (Carini, White, *et al.*, 2014), and is capable of lipid remodelling (Carini *et al.* 2015). In fact, HTCC7211 becomes dominant during the summer months in the Sargasso Sea (Vergin *et al.*, 2013), when this region becomes highly stratified and P-depleted. In contrast, *P. ubiqu*e HTCC1062, isolated from P-rich waters of the coastal North East Pacific, can only grow on phosphate (Carini, White, *et al.*, 2014), and the lack of *plcP* in its genome suggests that it is not capable of P-lipid remodelling. Thus, it is likely that genes involved in the synthesis of non-P lipids are specifically acquired by marine heterotrophic bacteria as an adaptation to low P conditions. This may explain why non-P lipids in marine heterotrophic bacteria seem to be much more diverse than those in their photosynthetic counterparts.

Whereas sulfolipids are the preferred substitution lipid for cyanobacteria, and betaine lipids for eukaryotic phytoplankton (Van Mooy *et al.*, 2009), heterotrophic bacteria synthesize a broad spectrum of lipids upon P deficiency, including betaine lipids (DGTS and DGTA), and a variety of glycolipids and aminolipids (Figure 3.1, 3.2, 3.5; Sebastián *et al.*, 2016). Sulfolipids (SQDG), the glycolipid DGDG and betaine lipids (DGTS, DGTA) have been traditionally ascribed to phytoplankton (Van Mooy and Fredricks, 2010), although production of DGTA in dark seawater incubations using glucose as a precursor has been previously reported (Popendorf *et al.*, 2011). Furthermore, the gene *sqdB*, involved in SQDG biosynthesis, has recently been detected in environmental heterotrophic bacteria (Villanueva *et al.*, 2013). However,

sequences affiliated to heterotrophic bacteria appear to be rare in pelagic marine environments (Van Mooy *et al.*, 2006). Similarly, homologs of glycosyltransferases known to be involved in DGDG synthesis like *pgt* (Hölzl *et al.*, 2005; Devers *et al.*, 2011), are rare in the marine environment although found in some marine heterotrophic bacteria (e.g. *Pelagibaca bermudensis*; Appendix 1). Despite the lack of the relevant genes characterised as being involved in the synthesis of these lipids, a recent survey found substantial quantities of DGTA, SQDG and DGDG of likely heterotrophic origin in Mediterranean sea water samples (Sebastián *et al.*, 2016). How marine heterotrophic bacteria synthesize these lipids awaits further characterisation.

3.3.2 *The glycosyltransferase Agt is a key enzyme in the synthesis of glycolipids in marine heterotrophic bacteria*

My findings also indicate that the glycosyltransferase Agt is likely to be the enzyme mediating the synthesis of MGDG and GADG in SAR11 bacteria since it resulted in the formation of these lipids when overexpressed in *E. coli* (Figure 3.7). My collaborators also reported that these lipids are abundant in the membrane lipids of natural bacterial communities in P-deplete waters of the Mediterranean Sea (Sebastián *et al.*, 2016). Substitution of the acidic phospholipid PG by glucuronic acid-containing glycolipids has also been documented in P-limited cultures of *Brevundimonas diminuta* (formerly *Pseudomonas diminuta*) (Minnikin *et al.*, 1974) and recently in the Sargasso sea SAR11 isolate HTCC7211 (Carini *et al.* 2015). Therefore it seems that despite the metabolic diversity of the SAR11 clade (Schwalbach et al. 2010, Carini *et al.* 2014), SAR11 isolates from two separate P-depleted systems behave similarly. Both the phospholipid PG and GADG are anionic under physiological conditions so it may be that they could be interchanged while maintaining the biophysical properties of the membrane. A similar substitution of PG for the anionic SQDG has already been proposed for phytoplankton (Van Mooy *et al.*, 2006). Interestingly, the PG content of *Phaeobacter* sp. MED193 membranes declined upon P-starvation (Figure 3.2) despite the lack of any concomitant increase in the abundance of alternative negatively charged lipids to replace it. PlcP in

Sinorhizobium meliloti only displayed activity towards the zwitterionic phospholipids PE and PC (Zavaleta-Pastor *et al.*, 2010) but the substrate specificity of SAR11 PlcP, as well as environmental PlcP homologs warrants further investigation. In Eastern Mediterranean sites, PlcP was more abundant than the ubiquitous, single-copy gene RecA, which may indicate that it is present in two or more copies per genome in some organisms. This is also observed in some sequenced isolates, such as *Labrenzia alexandrii* and *Rhodopirellula baltica*: the function of these secondary PlcP copies remains unresolved, and it is unclear how prevalent they are in the environment. One possibility is that in bacteria with multiple PlcP copies these copies may be complementary, with differing substrate specificities.

3.3.3 Physiological consequences of lipid remodelling

Despite the variety of surrogate non-P lipids, marine heterotrophic bacteria nonetheless seem to revert to phospholipid-dominated membranes under P sufficiency, similar to what has been observed in phytoplankton (Martin *et al.*, 2011). Yet a study in yeast has shown that DGTS can functionally substitute for the phospholipid phosphatidylcholine (PC) without an apparent phenotype in culture (Riekhof *et al.*, 2014). It is therefore puzzling why heterotrophic bacteria do not simply maintain a constant low P membrane composition. In plants, phospholipids have been proposed as P storage molecules which can be mobilised upon P deficiency (Tjellström *et al.*, 2008). Indeed, heterotrophic bacterial communities in the Mediterranean Sea have been seen to accumulate phospholipids upon relief of P limitation (Sebastián *et al.*, 2016), while there was no major change in community composition, suggesting excess P is stored in phospholipids. This is consistent with the use of phospholipids as a P reservoir: on return to P limitation, PlcP mediated degradation of membrane phospholipids might result in a net release of P for diversion to other cellular uses. However, heterotrophic bacteria in regions of the ocean not characterised by P-depletion also appear to have phospholipid-rich membranes (Van Mooy *et al.*, 2006, 2008), suggesting that phospholipids constitute the 'default' membrane lipids for heterotrophic bacteria. An additional explanation

for the prevalence of phospholipids under P sufficiency is that membrane proteins appear to have evolved in a phospholipid-dominated environment. For example, the activity of certain membrane transporters can be enhanced by specific interactions with phospholipids (Laganowsky *et al.*, 2014). This implies that membrane functions might be compromised to some degree in a remodelled membrane. One means by which organisms have been assumed to limit the effects of remodelling on membrane function, have been by maintaining bulk biophysical properties, such as membrane charge. This is exemplified by the presumed substitutions in phytoplankton of the anionic PG for SQDG and zwitterionic PC by betaine lipids (Van Mooy *et al.*, 2006). However, my results appear to show a replacement of both anionic and zwitterionic phospholipids by just the zwitterionic DGTS (Figure 3.2) in *Phaeobacter* sp. MED193, indicating that net charge may not be maintained following remodelling by at least some marine bacteria.

In summary, my results, in concert with other recent findings (Carini *et al.*, 2015; Sebastián *et al.*, 2016), demonstrate that the ability to substitute phospholipids for non-P lipids under P deprivation is a common strategy in marine heterotrophic bacterial communities adapted to low-P environments. Central to the remodelling process is a phospholipase, PlcP, which is widespread in marine surface waters where phosphorus is scarce. My data points to lipid remodelling as an important ecological adaptation enabling natural heterotrophic bacteria to thrive in low P marine environments.

Chapter 4 Glutamine lipids are widespread in Roseobacter group bacteria

4.1 Introduction

Aminolipids are common components of the bacterial membrane that have an amino acid head group, and share a common fatty acid structure (Geiger *et al.*, 2010). The best characterised aminolipid, ornithine lipid (OL), is estimated to be present in more than 50% of sequenced bacteria (Vences-Guzmán *et al.*, 2015). As a phosphorus-free lipid, OL is often produced alongside other such lipids in response to P-limitation (Minnikin and Abdolrahimzadeh, 1974; Geiger *et al.*, 1999) in order to reduce the P cost of the membrane. In contrast to P-free glycerolipids, OL lacks a diacylglycerol backbone so phospholipase C-mediated degradation of phospholipids is not required for OL synthesis (Sebastián *et al.*, 2016). While in some bacteria, including SAR11 strain HTCC7211, OL has only been detected in the membranes of P-limited cultures (Geiger *et al.*, 1999; Carini *et al.*, 2015), other bacteria appear to produce OL constitutively (Rojas-Jiménez *et al.*, 2005; Comerci *et al.*, 2006). This implies a role for OL in these organisms beyond simply serving as replacements for phospholipids. Although present in both membranes, OL appears to be localised to the outer membrane in many bacteria (Dees and Shively, 1982; Rojas-Jiménez *et al.*, 2005). Accordingly, OL may be involved in changing outer membrane permeability (Dees and Shively, 1982; Nikaido, 2003). Indeed, hydroxylated ornithine lipids do seem to protect against acid and temperature stress in *Rhizobium tropici,* which may influence the ability of this nodule forming bacterium to establish functional symbioses (Vences-Guzmán *et al.*, 2011). Similarly, several *Desulfovibrio* strains produced increased amounts of OL when grown at higher temperatures (Seidel *et al.*, 2013). Ornithine lipid is also required for the production of a number of c-type cytochromes in the membrane of the purple non-sulfur bacterium *Rhodobacter capsulatus* (Aygun-Sunar *et al.*, 2006).

Ornithine lipid synthesis is a two-step process mediated either by two enzymes, OlsB and OlsA, which seem to be found primarily in *Alphaproteobacteria* (Geiger *et al.*, 2010), or by a single multidomain protein, OlsF, which is prevalent in *Gammaproteobacteria* and *Bacteroidetes* (Vences-Guzmán *et al.*, 2015). First, a 3-hydroxy fatty acid is condensed onto the α-amino group of ornithine by either OlsB or the *N*-acyltransferase domain of OlsF (Gao *et al.*, 2004; Vences-Guzmán *et al.*, 2015), resulting in the formation of lyso-ornithine lipid (Figure 4.1). This is followed by the acylation of the 3-hydroxy group of the first fatty acid, catalysed by OlsA or the *O*-acyltransferase domain of OlsF (Weissenmayer *et al.*, 2002; Vences-Guzmán *et al.*, 2015). In bacteria employing the OlsBA pathway, *olsB* and *olsA* are present adjacent to one another in the genome apparently in an operon (Geiger *et al.*, 2010). Some bacteria from the *Rhodobacteraceae* harbour a second *olsB* homolog (*olsB2*) which is separated from *olsBA* in the genome (Geiger *et al.*, 2010). Gene duplication is often followed by diversification in the function of the duplicate gene (Zhang, 2003; Taylor and Raes, 2004). Thus, OlsB2 might have distinct substrate specificity to OlsB and mediate the synthesis of a different aminolipid. In fact, one *olsB2*-containing bacterium, *Rhodobacter sphaeroides*, produces glutamine lipid (QL) in addition to OL (Zhang *et al.*, 2009; Geiger *et al.*, 2010), suggesting that OlsB2 may be involved in the synthesis of this lipid.

A number of aminolipids other than OL have been described in bacteria, including the aforementioned glutamine lipids (Zhang *et al.*, 2009), lysine lipids (Tahara, Yamada, *et al.*, 1976) and glycine lipids (Batrakov *et al.*, 1999). More complex aminolipids containing multiple amino acids have also been detected (Tahara, Kameda, *et al.*, 1976). Very little is known about how these aminolipids are synthesised, with the enzymes involved largely unknown. The bifunctional acyltransferase, OlsF, may play a role in lysine lipid synthesis (Moore *et al.*, 2015; Vences-Guzmán *et al.*, 2015). However, no studies have directly confirmed this activity. Without knowledge of the genes involved it is difficult to estimate how widely distributed the ability to produce these lipids is in bacteria.

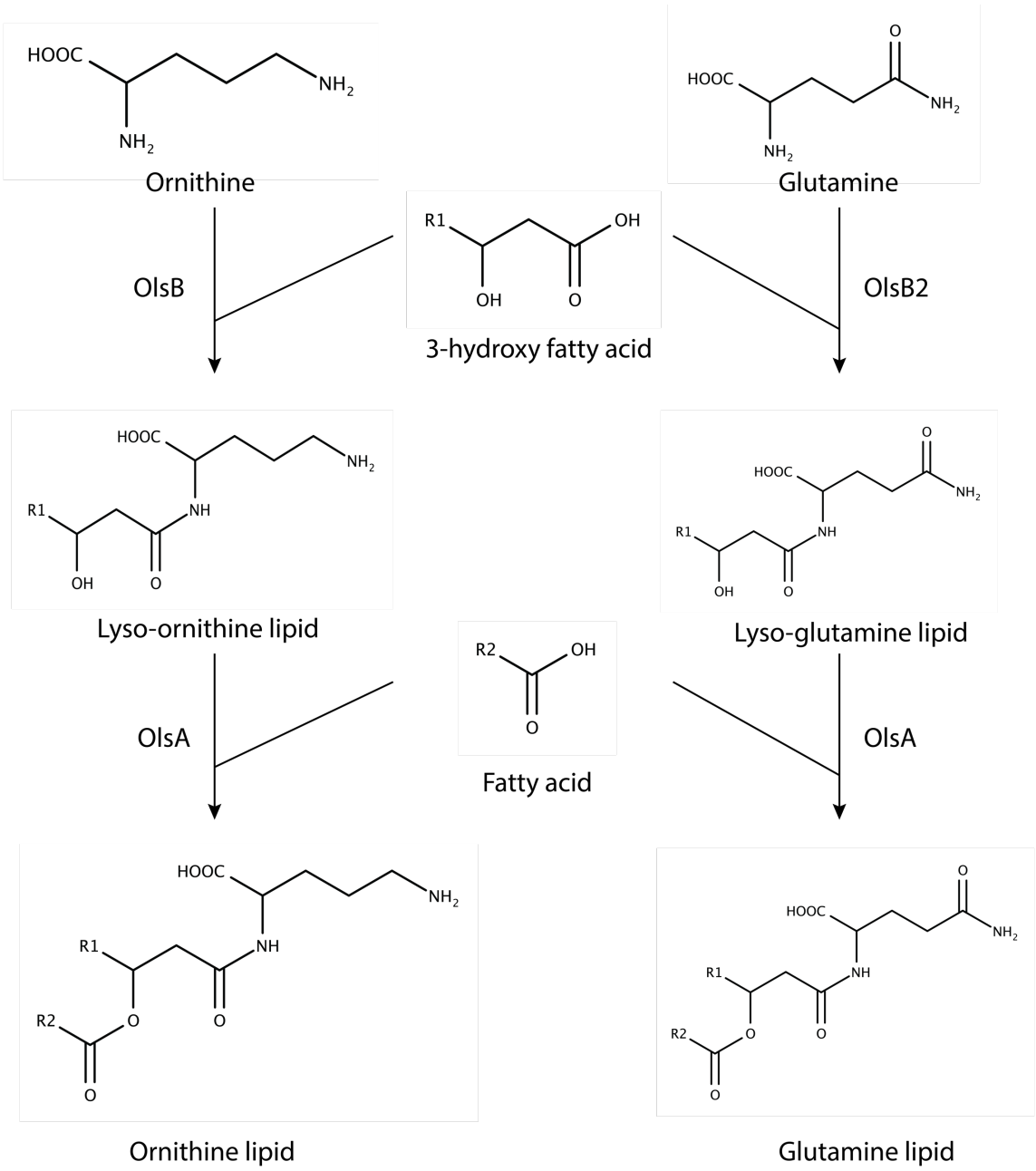


Figure 4.1 Overview of the pathway for the biosynthesis of ornithine lipids (OL) alongside the hypothesised pathway for glutamine lipid (QL) biosynthesis. Biosynthesis of OL proceeds via a two-step pathway.(Weissenmayer *et al.*, 2002; Gao *et al.*, 2004). Lyso-OL is formed by the condensation of a 3-hydroxy fatty acid with ornithine, mediated by the *N*-acyltransferase OlsB. This is converted into OL by the esterification of a second fatty acid to the hydroxyl group of the first by OlsA. Geiger (2010) hypothesised that QL might be synthesised in a manner analogous to OL, using a second OlsB homologue (OlsB2) to carry out the initial *N*-acylation of glutamine. The subsequent *O*-acylation of lyso-QL to form QL could then be mediated by OlsA.

A few studies have reported detecting aminolipids in the marine environment. One such study, conducted in the Black Sea, detected OL in deeper, anoxic, waters but not at the surface (Schubotz *et al.*, 2009). Similar results were reported in a study of lake water (Ertefai *et al.*, 2008). The failure to detect OL in surface waters is puzzling, given the widespread distribution of *olsB* and *olsF* in the genomes of sequenced bacteria, and its presence in some strains of the widespread SAR11 clade (Carini *et al.*, 2015; Vences-Guzmán *et al.*, 2015). To my knowledge, no aminolipids other than OL have been reported in aquatic environments. At present it is unclear whether the lack of reported aminolipids in marine surface waters reflects shortcomings in the analytical techniques used to detect lipids in these environments, or a genuine lack of these lipids. One way to begin addressing this issue is to investigate the abundance and distribution of genes involved in the synthesis of aminolipids in environmental metagenomes. However, this approach has been hampered by a lack of knowledge of how aminolipids other than OL are synthesised. I tackle this issue by testing the hypothesis that the *olsB2*, present in the genomes of some *Rhodobacteraceae,* mediates glutamine lipid (QL) synthesis (Geiger *et al.*, 2010). Glutamine lipids may also play a role in lipid membrane remodelling in response to P limitation. It has been proposed that QL carry a net negative charge and could thus substitute for PG in the membrane (Zhang *et al.*, 2011). I previously found no evidence for such a response in *Phaeobacter* sp. MED193 (Figure 3.2), which is capable of PlcP-mediated remodelling (Figure 3.1). It is possible ornithine and glutamine lipids play a role in lipid remodelling in bacteria unable to produce alternative P-free lipids. One such Roseobacter bacterium is *Ruegeria pomeroyi* DSS-3, which also harbours two *olsB* genes (Geiger *et al.*, 2010), but does not have PlcP or any characterised genes for the synthesis of P-free glycerolipids. *R. pomeroyi* DSS-3 is also one of the best-characterised members of the Roseobacter group (Moran *et al.*, 2004). Using *R. pomeroyi* DSS-3 as a model organism, I show that OlsB2 is required for QL synthesis. With this information, I proceed to investigate the likely distribution of this capability to synthesise glutamine lipids both phylogenetically and by examining marine

metagenomes. Furthermore, I test the hypothesis that QL can substitute for phospholipids in this Roseobacter in response to P-limitation.

4.2 Results

4.2.1 Glutamine lipids are present in cultures of Ruegeria pomeroyi DSS-3

Although the presence of glutamine lipids in *Ruegeria pomeroyi* DSS-3 has been hypothesised (Geiger *et al.*, 2010), it has never been demonstrated. To test this hypothesis, I obtained lipid extracts from *R. pomeroyi* after growth overnight in the rich sea salts medium ½ YTSS (Section 2.4.3). When analysed by LC-MS in negative ionisation mode, these revealed a similar lipid profile to what I had previously observed for *Phaeobacter* sp. MED193 (Figure 3.2) including prominent peaks corresponding to PG, OL, PE and a number of uncharacterised lipids (Figure 4.2 a). Of particular interest were a number of peaks with mass to charge ratios of 717, corresponding to the most abundant species observed in *R. sphaeroides* (Figure 4.2; Zhang *et al.*, 2009). In order to investigate the structure of these ions, I performed multiple rounds of fragmentation (MS^n), using a quadrupole ion trap MS. This analysis revealed that the peak with a retention time of 9.8 minutes fragmented in a manner consistent with that previously reported for QL (Zhang *et al.*, 2009). Initial fragmentation yielded a major product ion with *m/z* 435, resulting from the neutral loss of a mass of 282, which is indicative of an 18:1 fatty acid (Figure 4.2 c). This pattern of initial fatty acid loss is typical for all amino acid lipids (Moore *et al.*, 2016). Further fragmentation of the *m/z* 435 ion resulted in the formation of a characteristic ion with *m/z* 145, corresponding to glutamine after the loss of a proton (Figure 4.2 d). The neutral loss of 290 is consistent with the loss of a 20:3 ketene, which was reported to result from a 3-OH 20:1 fatty acid in the R1 position (Figure 4.1; Zhang *et al.*, 2009). An additional ion in the MS^3 spectrum, with mass 127, was also observed by Zhang *et al.* (2009) who proposed that it resulted from the cyclisation of glutamate following loss of water (Figure 4.2 d). Thus, these results are consistent with those for the QL previously reported by Zhang *et al.* (2009) and demonstrate the presence of 3-OH

20:1/18:1 glutamine lipid in *R. pomeroyi*. An additional lower abundance ion of mass 692, eluting at the same time as the QL 3-OH 20:1/18:1 showed a loss of 256 on initial fragmentation, yielding an MS^2 spectrum with the same major ions present as for the *m/z* 717 ion (Figure 4.2 c). This neutral loss is consistent with the loss of a 16:0 fatty acid. The abundance of the resulting fragment ions was too low to obtain an MS^3 spectrum. However, it is likely that this species is a QL 3-OH 20:1/16:0.

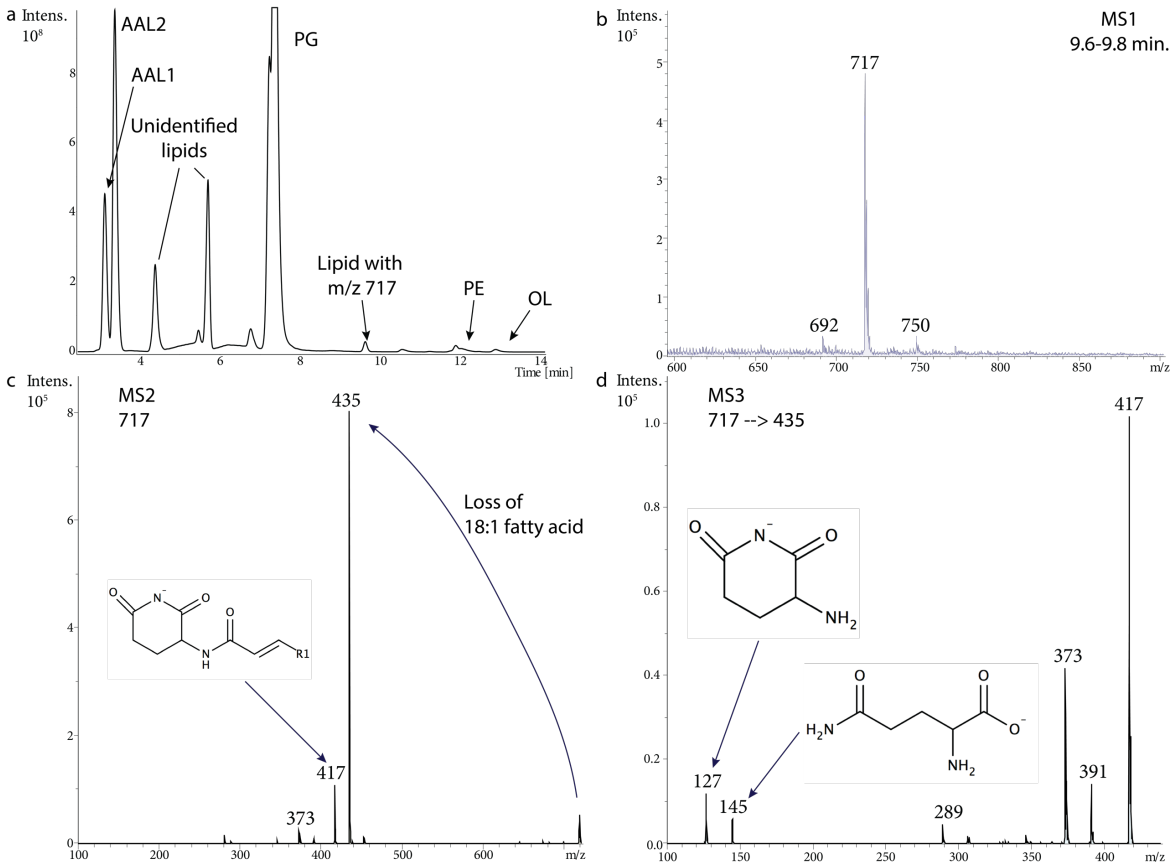


Figure 4.2 Glutamine lipids are present in lipid extracts from *Ruegeria pomeroyi* DSS-3 cultures grown in ½YTSS . **a)** Representative base peak chromatogram obtained after running lipid extracts from *R. pomeroyi* cultures on a liquid chromatography mass spectrometer in negative ion mode. The y-axis has been expanded to allow better visualisation of the peak with *m/z* 717 **b)** Mass spectrum showing molecular ions detected in the peak eluting between 9.6 and 9.8 minutes. The most abundant ion, with *m/z* 717, was selected for further fragmentation (**c,d**). After a first round of fragmentation (**c),** the major ion with *m/z* 435 was consistent with a loss of an 18:1 fatty acid. This ion was selected for MS^3 fragmentation (**d**), which yielded diagnostic ions with *m/z* 145 and 127, corresponding to glutamate and to an ion resulting from the cyclisation of glutamate following loss of water, respectively.

4.2.2 A second homolog of olsB is required for glutamine lipid synthesis in R. pomeroyi DSS-3

Having confirmed the presence of QL in *R. pomeroyi*, I set about testing the hypothesis of Geiger *et al.* (2010) that one of the two homologs of *olsB* present in the genomes of some *Rhodobacteraceae* was involved in QL biosynthesis. I investigated the genomes of 5 *Rhodobacteraceae* for *olsB* homologs by conducting a nucleotide BLAST search using OlsB from *Ensifer meliloti* (Figure 4.3). Of the 5 strains investigated, 3 (*R. pomeroyi* DSS-3, *Phaeobacter* sp. MED193 and *Rhodobacter sphaeoides* 4.2.1) have now been shown to produce both OL and QL (Figure 4.2; Zhang *et al.*, 2009; Sebastián *et al.*, 2016). Of the remaining two strains, *Dinoroseobacter shibae* DFL-12 is a member of the Roseobacter group but is phylogenetically distant from most other Roseobacter isolates (Simon *et al.*, 2017) while *Stappia stellulata* DSM 5886 is part of a cluster that is phylogenetically distinct from the rest of the *Rhodobacteraceae* (Pujalte *et al.*, 2013). These five strains therefore cover the breadth of phylogenetic diversity within the Family. Using an E-value cut-off of 10^{-5}, 2 matches were retrieved from BLAST searches of 4 out of the 5 strains, while a single match was obtained for *S. stellulata*. In the four strains with two matches, the best BLAST hit was immediately upstream of a gene annotated as OlsA (lyso-ornithine lipid acyltransferase; Figure 4.3 a) on the IMG website. The second retrieved match was downstream of a gene predicted to encode BamE, a component of the outer membrane β-barrel assembly machinery (Rigel *et al.*, 2012), in three of the strains, but not in the genome of *D. shibae* (Figure 4.3 b). The neighbourhood of the *olsB* homolog from *S. stellulata* was not syntenic with any of the other strains investigated (Figure 4.3 a).

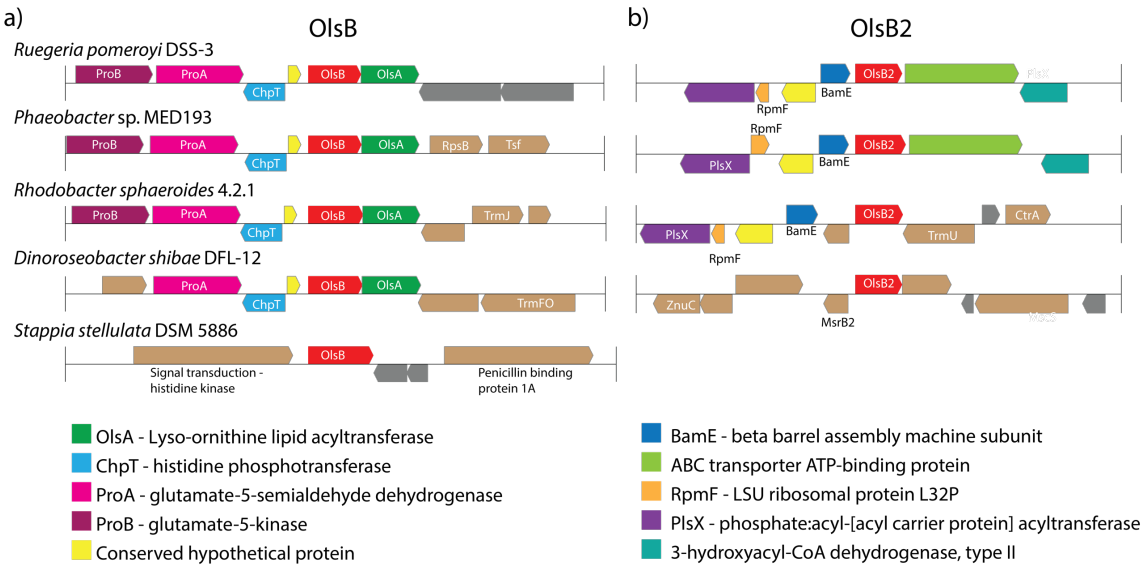


Figure 4.3 Gene neighbourhoods of the OlsB homologs (shown in red) in selected *Rhodobacteraceae*. The neighbourhoods are aligned to the location of the genes highlighted in red, which were retrieved by a nucleotide BLAST search using OlsB from *Ensifer meliloti* (SM11_chr0026) using an e-value cut-off of 10^{-5}. Genes within 3000 base pairs of the retrieved OlsB homologs are shown. **a)** An alignment of best BLAST hits to SM11_chr0026 (putative OlsB orthologs) in each genome. **b)** An alignment of further genes retrieved by the BLAST search (putative OlsB2 orthologs). Genes with the same predicted protein product (or which are reciprocal best BLAST hits, in the case of hypothetical proteins) in at least two of the neighbourhoods are coloured, and their predicted functions listed below the alignments. Genes shown in brown have a predicted protein product but were not shared by any of the other neighbourhoods. Grey genes encode hypothetical proteins and were not conserved between neighbourhoods.

The *olsB* from *E. meliloti* is present immediately upstream of *olsA*; the same ordering as was found for the best hit in four of the *Rhodobacteraceae* genomes (Figure 4.3 a). Genes with similar synteny between organisms often share the same function (Koonin, 2005). Thus, these *olsB* homologs are likely true orthologs of *olsB* from *E. meliloti*. In contrast, paralogs – genes which share an evolutionary history but which have different functions – can often be distinguished by their differing synteny. The genome neighbourhoods of the *olsB* homologs shown in Figure 4.3 b are quite different to that of *olsB*, suggesting that these might in fact be paralogs. In line with

the nomenclature of Geiger *et al.* (2010), I refer to these genes as *olsB2*. In order to test whether *olsB2* in *R. pomeroyi* was involved in QL synthesis, I constructed a deletion mutant, in which the gene (*spo2489*) was disrupted by marker exchange with a gentamicin resistance cassette. Colony PCR of this strain using primers targeting Gm^R and the downstream homologous region resulted in a product around 1500 base pairs long, which is consistent with results on a purified plasmid containing the construct (Figure 4.4). The sequence of the mutated region was checked by sequencing, which confirmed that no mutations had been introduced into the homologous flanking regions. Glutamine lipid was not detectable in this *ΔolsB2* strain, whereas OL levels were not decreased (Figure 4.5). Therefore, OlsB2 is required for QL synthesis in *R. pomeroyi* but does not appear to be involved in the synthesis of OL.

The second step of ornithine lipid biosynthesis requires acylation of the 3-hydroxy group of the lyso-ornithine lipid produced by OlsA (Figure 4.1; Gao *et al.*, 2004), which is located immediately downstream of OlsB in the genome of *E. meliloti*. There is, however, no gene with predicted *O*-acyltransferase activity in the vicinity of *olsB2* that might carry out this second step (Figure 4.3 b). As such, I hypothesised that OlsA might exhibit flexible substrate specificity, with activity towards lyso-glutamine lipid in addition to its role in OL biosynthesis. Such loose substrate specificity has previously been suggested by the finding that OlsA from *Rhodobacter capsulatus* is able to acylate phosphatidic acid in addition to lyso-ornithine lipid (Aygun-Sunar *et al.*, 2007). A deletion mutant was constructed in *R. pomeroyi* for *olsA* (*spo1979; ΔolsA*) and validated by sequencing of the mutated region, as well as PCR using primers targeting the inserted gentamicin resistance cassette. This PCR resulted in a product around 1500 base pairs long, consistent with an expected size of 1516 base pairs in a correctly constructed mutant (Figure 4.4). As predicted by the model, *ΔolsA* had significantly lower quantities of both QL and OL compared to the wild type (Figure 4.5). While QL was undetectable in the *ΔolsA* mutant, a small peak with an elution

time and mass corresponding to OL was still present (Figure 4.5), indicating a possible alternate pathway for lyso-OL acylation in *R. pomeroyi*.

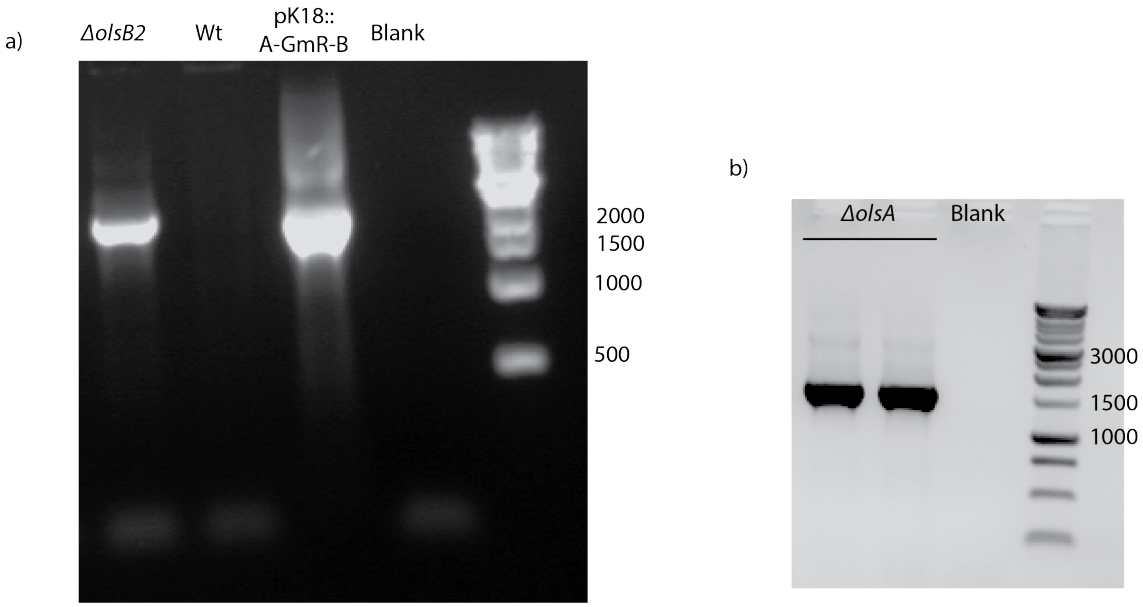


Figure 4.4 Confirmation of mutants in *Ruegeria pomeroyi* DSS-3 by colony PCR. **a)** *ΔolsB2* in *R. pomeroyi* DSS-3 again obtained using Gm^R forward and B reverse primers. Run alongside wild type (Wt) *R. pomeroyi* and the purified pK18 construct. **b)** *ΔolsA* in *R. pomeroyi* using the Gm^R forward primer paired with a primer targeting chromosomal DNA downstream of region B. The size of the product observed is comparable to the expected product size for a correctly assembled mutant (1516 base pairs).

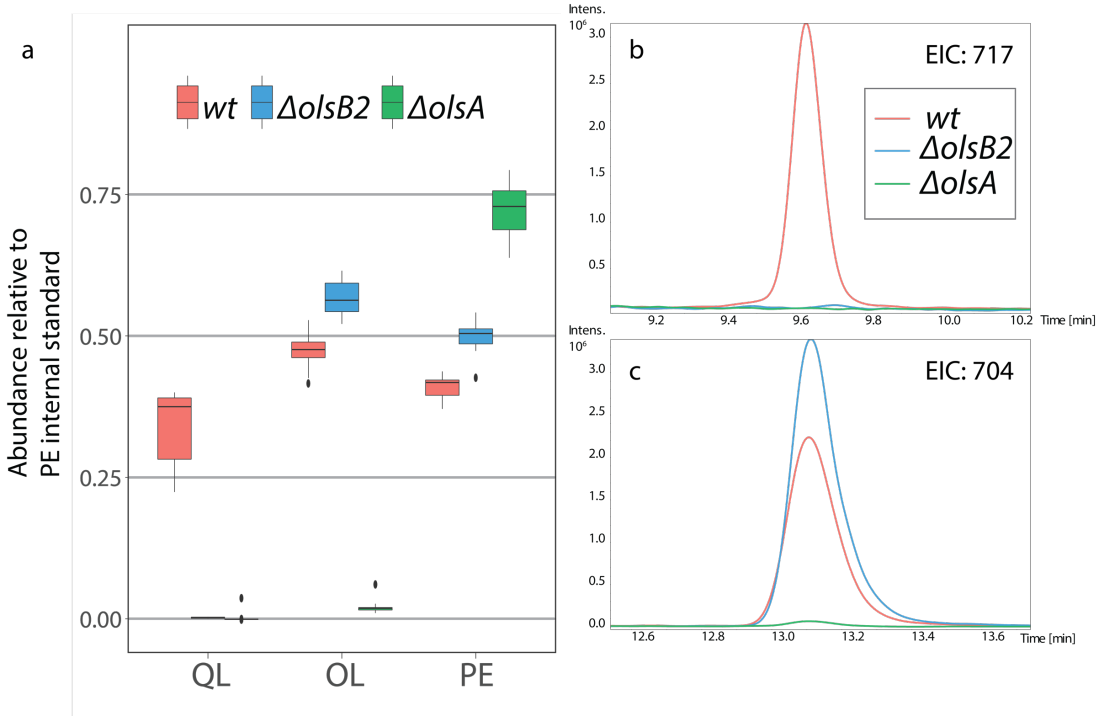


Figure 4.5 Boxplots (**a**) show that glutamine lipid (QL) was not detected in both *ΔolsB2* and *ΔolsA* mutants relative to the wild type (wt). Ornithine lipid (OL) levels were unaffected in *ΔolsB2* but OL was below the detection limit in *ΔolsA*. Upper and lower bounds of the coloured boxes represent the 75th and 25th percentiles, respectively of the data, while the thicker central line is the median. Upper and lower whiskers extend from the 2nd to the 98th percentiles. Outliers are represented by circular points. Based on results obtained from six biological replicates in each case. (**b**, **c**) Representative extracted ion chromatograms (EIC) obtained after analysing lipid extracts from the three *R. pomeroyi* strains by liquid chromatography – mass spectrometry in negative ion mode. Ions with mass 717 (**b**) and 704 (**c**) correspond to the intact mass of the major species of QL and OL, respectively. The time windows plotted corresponds to the known retention times of these species.

4.2.3 *The growth rates of* ΔolsB2 *and* ΔolsA *are impaired in phosphate-limited medium*

To investigate the consequences of the loss of function of OlsB2 and OlsA, I grew the two mutant strains alongside the wild type in defined MAMS medium with either 5 mM or 0.5 mM phosphate ("high Pi" and "low Pi", respectively; see Section 2.4.4). In high Pi medium the maximum growth rates of the three strains did not differ significantly (Figure 4.6), ranging from 0.110 ± 0.008 OD hr^{-1} (change in optical density at 540 nm per hour, with 95% confidence interval) in the wild type to 0.096 ± 0.012 OD hr^{-1} in *ΔolsA*. The final yield of *ΔolsA* (1.629 ± 0.376 OD units, 95% CI) was, however, significantly lower than for the other two strains (2.765 ± 0.073 and 2.762 ± 0.155 OD units for wild type and *ΔolsB2*, respectively). When the strains were grown in the low Pi medium, however, *ΔolsA* failed to grow, or grew extremely slowly (Figure 4.6 a). The growth rate of *ΔolsB2* was also reduced relative to high Pi medium (t-test, t = -6.349, $p = 0.003$). The growth rate of the wild type was lower than in high Pi medium (0.096 ± 0.008 OD hr^{-1}), though this difference was not significant. In low Pi medium, *ΔolsB2* (maximum growth rate: 0.077 ± 0.012 OD hr^{-1}) grew significantly slower than the wild type (t-test, t = -2.28, $p = 0.042$). The yields of both wild type and *ΔolsB2* were also lower in low Pi medium compared to high Pi medium (Figure 4.6 a), but did not differ significantly from each other.

The lower growth rates and yields of these cultures in low Pi medium indicated that they were P-limited. This assessment was supported by higher rates of alkaline phosphatase activity in the low Pi cultures, with activity of 6.25 ± 0.97 μM pNP hr^{-1} OD^{-1} in low Pi grown wild type cultures, compared to 0.86 ± 0.092 μM pNP hr^{-1} OD^{-1} in the high Pi cultures (t-test; t = -12.3, $p < 0.001$). Lipid analysis of samples collected from these cultures in late exponential phase did not, however, indicate substantial remodelling of the lipid membrane (Figure 4.6 b) in the manner observed for marine strains possessing PlcP (Sebastián *et al.*, 2016). Lipid abundances relative to a PE internal standard were systematically lower in samples from cultures grown with 0.5 mM phosphate (Figure 4.5 b). As lipid concentrations are typically proportional to

biomass (Christie and Han, 2010), this suggests that the amount of biomass from which lipids were extracted was less for the low Pi cultures than those grown at high Pi. The OD_{540} was used to standardise the amount of biomass collected, and internal standards were added prior to extraction (see Section 2.3): these results suggest a different relationship between OD and biomass at the two different phosphate concentrations. Despite this complication, the results of the lipid analysis show that in the wild type strain the ratio of QL abundance in low Pi compared to high Pi cultures was similar to that of the phospholipids, while the ratio of OL was lower (Figure 4.6). If substantial lipid remodelling had occurred, the ratios for these aminolipids would have been higher than for the phospholipids. Similarly, there was little evidence for lipid remodelling in *ΔolsB2* as the low-Pi-to-high-Pi ratios for OL and two putative aminolipids (AAL1 and AAL2) were lower than for the phospholipids PG and PC (Figure 4.6 b). Curiously, the ratio for PE in *ΔolsB2* was substantially lower than for either of the other phospholipids or the ratio observed in the wild type. This does suggest a shift away from PE in the membranes of P-limited *ΔolsB2* cells. However, it is unclear what lipids might replace it.

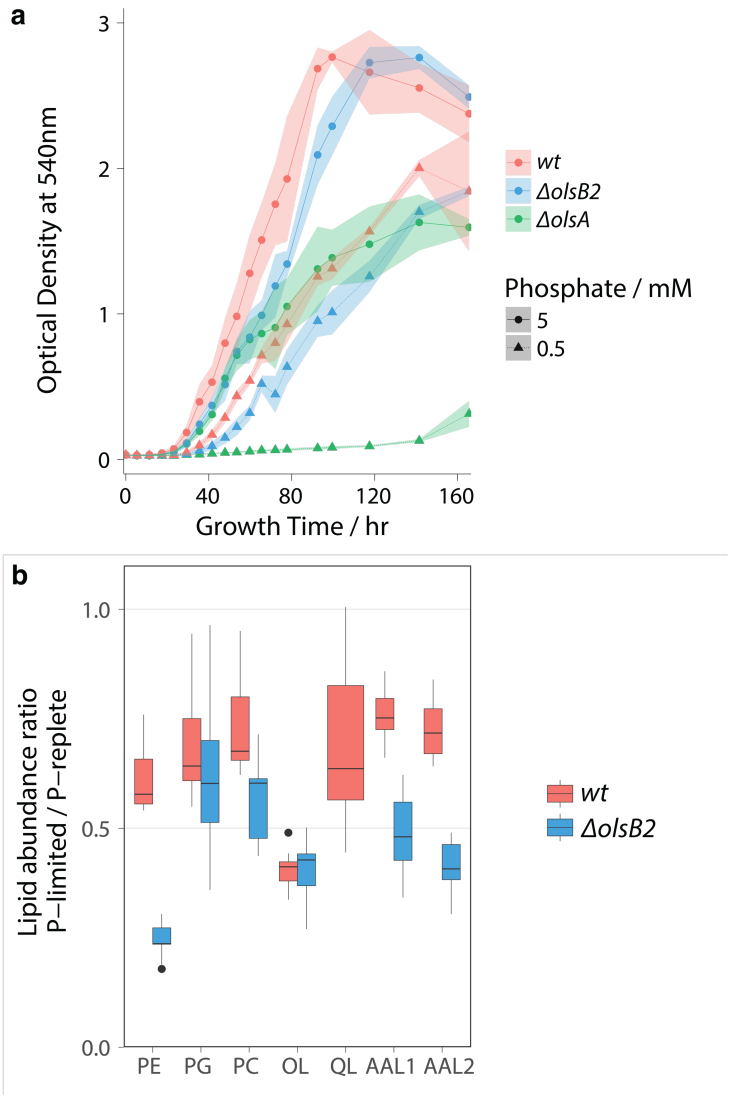


Figure 4.6 Growth of wild type and mutant *Ruegeria pomeroyi* DSS-3 in medium with high and low phosphate. **a)** Wild type (*wt*) *R. pomeroyi* and strains deficient in *olsB2* and *olsA* were grown in defined MAMS medium with either 5 mM (circles) or 0.5 mM (triangles) of added phosphate. Points represent the mean of 3 biological replicates while the shaded ribbons cover the region within one standard deviation of the mean. **b)** Boxplots showing the ratio of the abundance of major lipid classes in wild type and *ΔolsB2 R. pomeroyi* after growth with 0.5 mM phosphate relative to 5 mM phosphate. Cell samples for lipid analysis were collected late-exponential phase cultures. No ratio is reported for the glutamine lipid (QL) in *ΔolsB2* because the low abundance values recorded resulted in extremely high variability. Boxes represent the 25th, 50th and 75th percentiles of the data with their lower, central and upper lines, respectively. Whiskers extend from the 5th to the 95th percentiles. PE: phosphatidylethanolamine; PG: Phosphatidylglycerol; PC: phosphatidylcholine; OL: ornithine lipid; AAL1 and AAL2: unidentified, putative aminolipids.

4.2.4 OlsB2 appears to be ubiquitous within a subclade of the Rhodobacteraceae

Although a previous study identified OlsB2 in several *Rhodobacteraceae* (Geiger *et al.*, 2010), no attempt has been made to investigate the distribution of the gene more broadly. To this end, I constructed a 16S rRNA phylogeny of the *Rhodobacteraceae*, including at least one representative from each genus for which a fully or partially sequenced genome was available (as of August 2015). Many of the internal nodes of the phylogeny are poorly supported, likely due to a lack of sequence variation in the 16S rRNA. However the phylogeny does confidently separate the *Stappia/Labrenzia* cluster from the rest of the Family. Alongside this phylogeny, I indicated the presence of a predicted *olsB* and *olsB2* in the genome (Figure 4.7). This analysis indicated that *olsB* is conserved throughout the *Rhodobacteraceae,* with the exception of *Ketogulonicigenium vulgare*, but was absent from the *Hyphomodaceae* strains included as outgroups. In contrast, *olsB2* is less widely distributed, being absent from the deeply-branching *Stappia/Labrenzia* cluster. These organisms are currently assigned as *Rhodobacteraceae* but recent phylogenetic analyses place them within the *Rhizobiales* (Simon *et al.*, 2017). Aside from this cluster, virtually all *Rhodobacteraceae* genomes contained *olsB2.* Exceptions lacking *olsB2* included two pelagic Roseobacters, *Roseobacter* sp. CHAB-I-5 and *Rhodobacterales* sp. HTCC2255. These strains are both members of the pelagic Roseobacter cluster (Billerbeck *et al.*, 2016), a grouping defined based on similarities in gene content and which is amongst the most widely distributed and abundant subclades within the Roseobacter group. This group does not, however, appear to be defined by a lack of *olsB2* since other members, including *Planktomarina temperata* RCA23 and *Rhodobacterales* sp. HTCC2150 do possess copies of *olsB2*. Since the two pelagic Roseobacter cluster genomes lacking *olsB2* are not fully closed, it is possible that *olsB2* is present in these genomes was missed during sequencing.

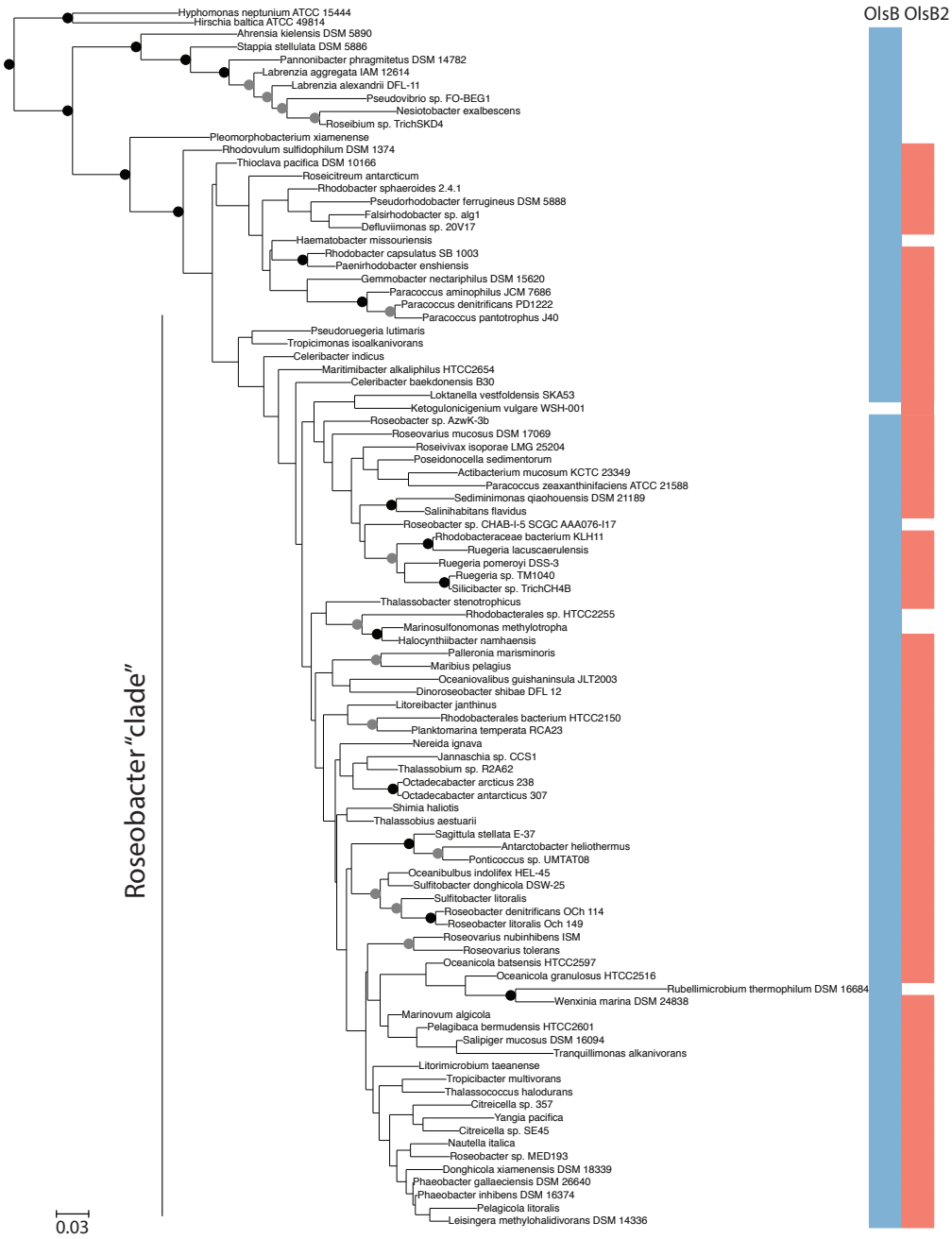


Figure 4.7 Maximum likelihood 16S rRNA phylogeny of *Rhodobacteraceae* with full or partially sequenced genomes. Sequences were downloaded from the SILVA database, aligned using Metaligner (Section 2.5.2) and a maximum likelihood phylogeny constructed using RaxML. Bootstrap support for nodes is indicated by filled circles. A black circle indicates support >0.70 while grey indicates support 0.50 – 0.70. Two sequences from the *Hyphomonadaceae* served as outgroups. Plotted alongside the phylogeny, the coloured boxes indicate the presence of one or two *olsB* homologues in the genome, as detected by BLAST searches using an e-value cut-off of 10^{-5}.

4.2.5 Aminolipid synthesis genes in ocean metagenomes

Growth experiments with both *Phaeobacter* sp. MED193 and *R. pomeroyi* DSS-3 have failed to uncover any significant substitution of phospholipids by either OL or QL under P-limited growth conditions (Figure 3.2, Figure 4.6). Moreover, the conserved nature of both OlsB and OlsB2 within the *Rhodobacteraceae* (Figure 4.7) is inconsistent with the patchier distribution of genes whose primary role seems to be in the acclimation to P-depletion. One such gene, the phospholipase PlcP (see Chapter 3), is present in 31 of the 98 *Rhodobacteraceae* genomes presented in Figure 4.7. This contrasts with organisms in which OL does appear to serve primarily to substitute for phospholipids under P-stress (López-Lara *et al.*, 2005; Carini *et al.*, 2015). It seems likely that aminolipids may have different roles in different organisms. Genes involved specifically in P-stress responses tend to have distinct biogeographies in the ocean, being markedly more abundant in regions that tend to be depleted in P (Coleman and Chisholm, 2010; Sebastián *et al.*, 2016). This appears to be the case for OlsB in SAR11 as it is found in a strain, HTCC7211, from the commonly P-depleted Sargasso Sea but not in most other strains. I reasoned that in microbial groups that use aminolipids primarily to substitute for phospholipids when P is scarce, the abundance of aminolipid synthesis genes should be related to P availability. Other groups might not show this relationship since, for these groups, the presence of aminolipids is not primarily an adaptation to periodic P-depletion.

To begin disentangling these factors, I set about determining the distribution of aminolipid synthesis genes in the *Tara* metagenome data set (Pesant *et al.*, 2015). Similarity searches, such as BLAST, retrieve a mixture of OlsB and OlsB2 sequences so an alternative means of reliably classifying these sequences was required. Although my analysis of gene neighbourhoods (Figure 4.3) had indicated that the two genes could often be distinguished by the presence of adjacent *olsA* or *bamE*, the contigs in the *Tara* metagenomes are too short for such synteny analysis. I therefore verified whether OlsB and OlsB2 sequences from the *Rhodobacteraceae* genomes in Figure 4.7 could be reliably separated phylogenetically (Figure 4.8). This analysis showed that

these sequences were consistently separated into two major clades (100% bootstrap support), and that these two clades were congruent with a classification based on synteny (the red and blue edges in Figure 4.8). A number of sequences could not be classified using the synteny-based approach (black edges in Figure 4.8). The largest group of these are from the *Stappia*/*Labrenzia* cluster, where the neighbourhood of the *olsB* homologue is mostly similar to that observed for *S. stellulata* (Figure 4.3). Two sequences in the blue (OlsB) cluster were at the end of a genomic scaffold so synteny could not be determined.

To classify environmental sequences using this phylogenetic approach, I created an alignment of reference sequences of OlsB, OlsB2 and the *N*-acyltransferase domain of OlsF. The functions of these genes had either been verified experimentally or, in the case of the marine strains, the strain had been shown to produce OL or QL. This alignment was used to build a reference phylogeny (Figure 4.8 a) and the alignment was also used to build an Hmmer profile with which to query the *Tara* dataset (e-value cut-off 10^{-5}). Sequences retrieved by this query were then placed onto the reference tree using the phylogenetic placement algorithm, pplacer (Matsen *et al.*, 2010). This algorithm determines the maximum likelihood placement onto a reference phylogeny of each metagenomics sequence in turn and has been developed with short metagenomics reads in mind. The environmental sequences were then classified as either OlsB, OlsB2 or OlsF based on the reference sequence that they were placed closest to. Sequences placed on internal nodes of the tree, with descendants of multiple reference types, could not be unambiguously classified so are referred to as "Other" in Figure 4.8 b.

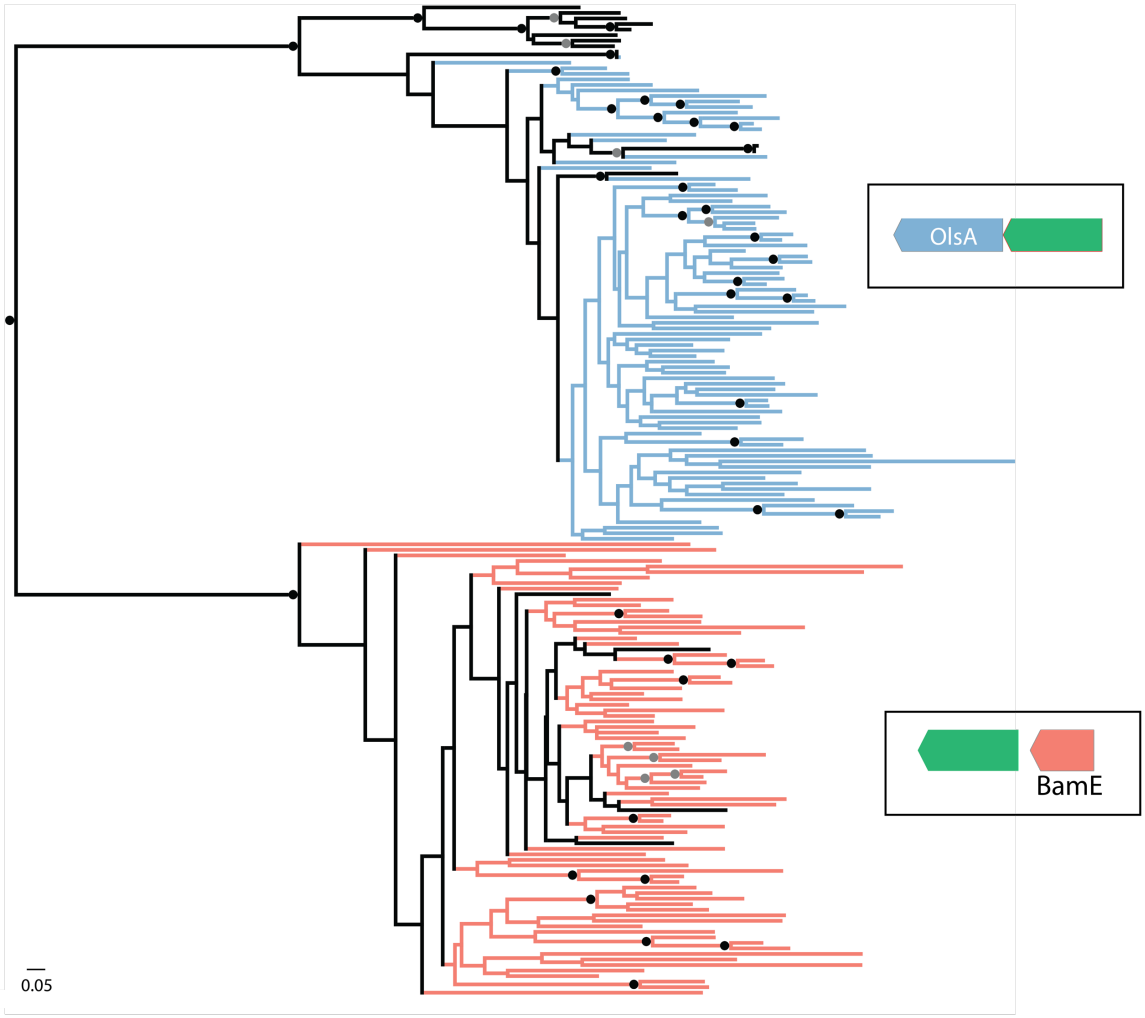


Figure 4.8 Phylogeny of OlsB homologues from the genomes of the strains used to construct the phylogeny in Figure 4.7. Protein sequences were aligned using Metaligner (Section 2.5.2) and a maximum likelihood phylogenetic tree inferred using RaxML with 100 bootstrap replicates. Nodes with >49% bootstrap support are indicated with circles (grey: 50 – 69%; black: >70%). Edges are coloured according to the synteny of the genes on their descendant leaves, as indicated by the inset schematics. Genes in red are downstream of a nearby predicted BamE gene while blue indicates that the genes are immediately upstream of a predicted OlsA. Leaves in black matched neither of these criteria. Two sequences, Ga0056889_10411 from *Falsirhodobacter* sp. alg1 and Rhd02DRAFT_00243 from *Marinosulfonomonas* sp. PRT002, both placed within the blue "OlsB" cluster, could not be classified based on synteny since they were located at the end of a contig with no genes downstream of them.

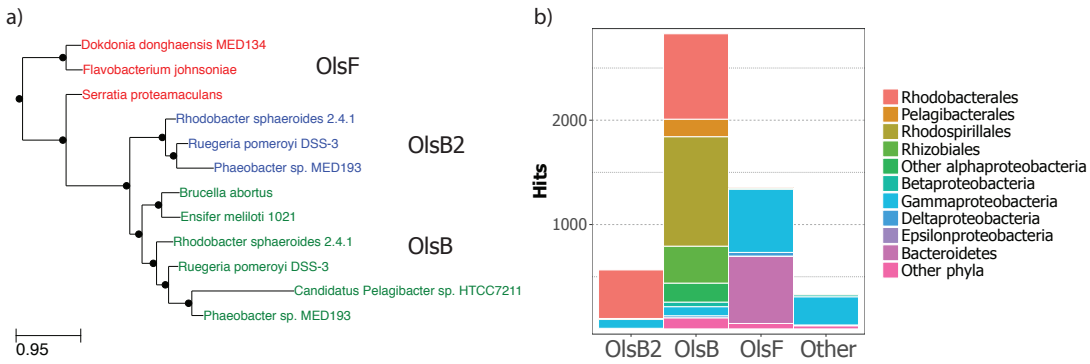


Figure 4.9 Maximum likelihood phylogeny showing the evolutionary relationship between the N-acyltransferase domains of selected OlsF, OlsB and OlsB2 proteins. The functions of the selected sequences have either been experimentally validated or the appropriate lipid (ornithine lipid for OlsF or OlsB, and glutamine lipid for OlsB2) has been detected in lipid extracts from pure cultures of the organism.

Out of a total of 5097 sequences retrieved, more than half, 2837, were classified as OlsB. The next more numerous was OlsF (1352 sequences) and then OlsB2 (567 sequences); 341 sequences were unassigned (Figure 4.8). In order to examine the taxonomic composition of these environmental sequences, I retrieved the taxonomy of the best BLAST hit for each sequences in the NCBI nr database. This showed that the majority of OlsB2 sequences were from the *Rhodobacteraceae* (Figure 4.8 b), although a sizeable minority (15 %) were gammaproteobacterial. The majority of OlsB sequences (91%) originated from *Alphaproteobacteria* with the *Rhodospirillales* and *Rhodobacteraceae* comprising 37% and 29% of the total, respectively. *Gammaproteobacteria* and *Bacteroidetes* contributed to OlsF sequences in roughly equal proportions, aligning with the reported distribution of OlsF amongst these two groups (Vences-Guzmán *et al.*, 2015).

I proceeded to look at the distribution of OlsB, OlsB2 and OlsF in data derived from 0.22 µm-filtered surface water samples in the *Tara* dataset (Figure 4.10). The abundance of OlsB, relative to the abundance of the conserved, single copy, gene

RecA, was comparatively high in the western north Atlantic and in the Mediterranean (Figure 4.10 b). These are areas characterised by P-scarcity (Wu *et al.*, 2000; Thingstad *et al.*, 2005) and where there was a high abundance of the central phospholipase C, PlcP, involved in the lipid remodelling response to P depletion (Sebastián *et al.*, 2016). Several sites outside these two regions also exhibited high abundances of OlsB, for example in the South Pacific and Southern Ocean. Phosphorus is not typically considered to be especially scarce in these regions so the high abundance of OlsB observed suggests that multiple factors may be determining the distribution of this gene in surface ocean microbial communities. No obvious patterns were observed in the distributions of OlsB2 or OlsF (Figure 4.10 a, c).

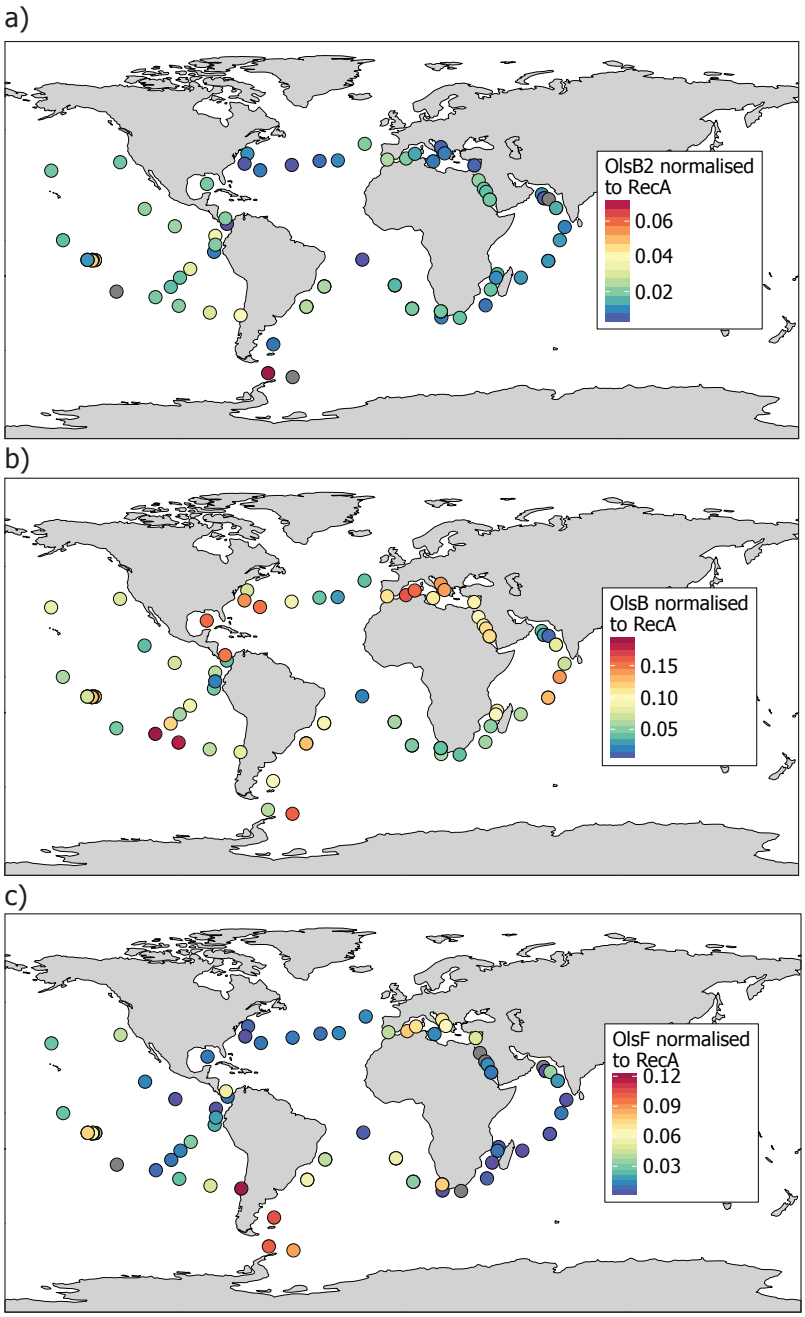


Figure 4.10 Global maps of the abundance of a) OlsB2, b) OlsB and c) OlsF, normalised to the abundance of the conserved single copy gene RecA, in the Tara metagenome dataset. Only surface seawater samples (collected at 5m depth) are shown, and viral metagenomes (collected with a lower filter size less than 0.1 μm) were excluded. Sequences were retrieved from the metagenomes using a Hmmer search and then classified as either OlsB, OlsB2 or OlsF using the phylogenetic placement algorithm pplacer, as described in more detail elsewhere in the text. Grey circles indicate no sequences corresponding to that gene were detected in the sample.

In order to disentangle the multiple sources of variation likely to be driving the distribution of OlsB, I used linear regression analysis to test the hypothesis that the aminolipid synthesis genes are more abundant in P-depleted areas of the ocean. In order to model this relationship, I used a measure of the relative concentrations of inorganic N and P called N* (Deutsch and Weber, 2012). This statistic is defined as:

$$N^* = [NO_3^-] - 16.[PO_4^{3-}]$$

A value of N* = 0 is equivalent to an N:P of 16:1. As a linear combination of the concentrations of N and P, it has an advantage over the more commonly used ratio N:P in that it does not inflate to very high values when the concentration of P is close to zero. The mean value for N* in the *Tara* samples was -3.8, which is comparable to a global average N* of -3.5 (Deutsch and Weber, 2012), and indicates that the ocean as a whole is generally deficient in nitrate relative to phosphate. A test for the correlation between the abundance of OlsB and N* revealed a significant positive correlation (Pearson $r = 0.33$, $n = 186$, $p < 0.001$), whereas correlations between N* and the other two genes were not significant. I asked whether OlsB abundance could be described simply as a function of the abundance of the major microbial groups found to harbour OlsB in the metagenome (Figure 4.9 b). I fitted a generalised linear model (GLM) with a negative binomial distribution to the abundance data for each of the aminolipid synthesis genes and compared a model only incorporating terms for the abundance of each of the major microbial groups found to harbour the gene (Figure 4.9 b) with one also including N* as an independent variable. The abundance of each microbial group was estimated using miTAG data (Logares *et al.*, 2014), which are derived from metagenomic 16S rDNA sequences. The slope coefficients for N* were significantly different from zero ($p < 0.05$; Table 4.1, Figure 4.11 a, c) for both OlsB and OlsB2, albeit only marginally so in the case of OlsB2. N* did not significantly predict OlsF abundance (Table 4.1; Figure 4.11 d). In order to test whether the inclusion of the N* term significantly improved the fit of the models, I performed a likelihood ratio test (LRT) comparing the model for each gene with microbial

abundance data alone against the model also including N*. This comparison showed that the models for OlsB and OlsB2 were significantly improved by the inclusion of N*, while the model for OlsF was not (Table 4.1).

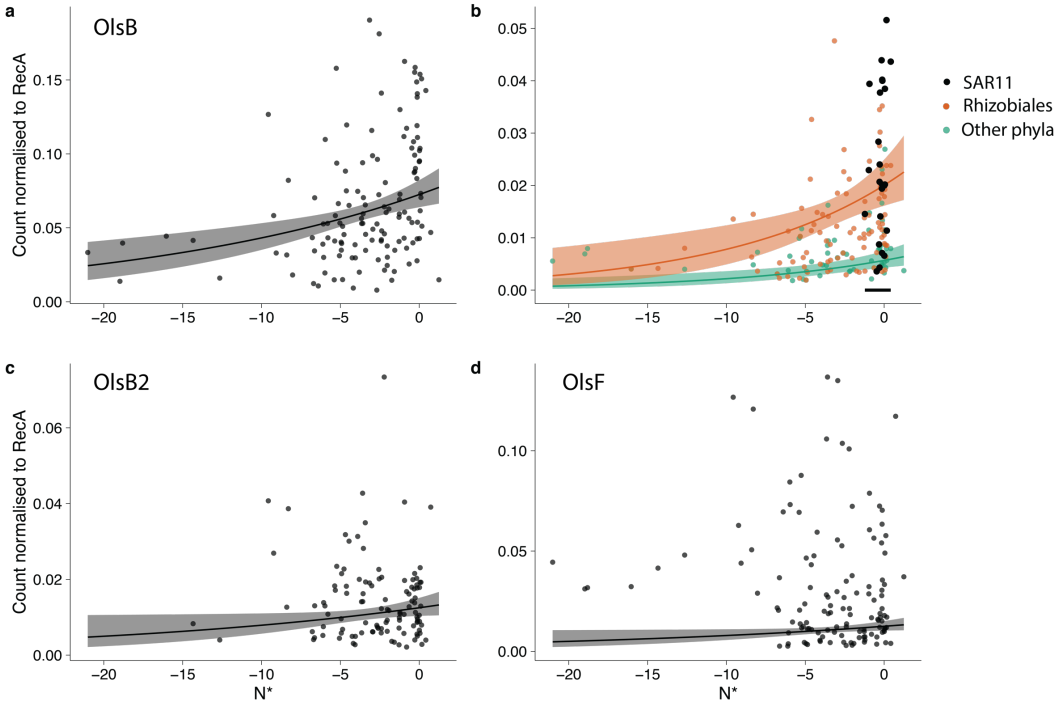


Figure 4.11 Scatterplots and generalised linear model (GLM) fits showing the relationship between aminolipid synthesis gene counts and N* in the *Tara* metagenomes. **a, c** and **d** show GLM fits for the relationship between N* and OlsB, OlsB2 and OlsF, respectively. All other model parameters were held constant at their mean values. N* significantly predicted the abundances of OlsB and OlsB2 but not OlsF. Scatterplots show the original data points. **b**) Lines showing the fitted relationships between N* and OlsB counts for the microbial groups for which the regression slope was significantly greater than zero. Points show the original data values. Also shown are the original values for SAR11 OlsB. The horizontal bar below the data shows the range of N* values at which SAR11 OlsB was found. Counts for aminolipid synthesis genes were divided by counts for RecA and multiplied by a correction factor to correct for differences in sequence length.

Table 4.1 Likelihood ratio test (LRT) comparisons of generalised linear models for aminolipid synthesis gene abundance with and without the inclusion of N* as an independent variable.[a]

N*					**Model comparison**	
Gene	Slope Coefficient	Std. Error	*z*-value	*p*	LRT statistic	LRT *p*
OlsB	0.201	0.056	3.62	<0.001	13.2	<0.001
OlsB2	0.179	0.087	2.06	0.040	4.35	0.037
OlsF	- 0.078	0.077	-1.03	0.304	1.22	0.269

[a] Coefficients are given for the N* term along with z-scores and associated p-values for the inclusion of N* as a parameter.

These GLMs thus suggest that OlsB and OlsB2 are found more commonly in bacteria present in waters that are relatively deficient in P. Based on the widespread occurrence of OlsB and OlsB2 in the *Rhodobacteraceae* and the lack of apparent response to P-limitation in tested strains, I hypothesised that the relationship between N* and OlsB or OlsB2 abundance differed between microbial groups. Using the data for aminolipid synthesis gene abundance within each microbial group (Figure 4.9 b), I modelled the relationship between the abundance of each gene in each group and N*. This analysis showed that there were significant differences between microbial groups in this relationship between the group-specific OlsB abundance and N* (χ^2 = 22.29, df = 8, p = 0.004) but this was not the case for either OlsB2 or OlsF. The slope of this relationship was significant for two groups: the *Rhizobiales* (β = 0.44 ± 0.22 (95% CI), z = 3.85, p < 0.001) and for "Other phyla" (β = 0.45 ± 0.32 (95% CI), z = 2.76, p = 0.017), indicating that OlsB is more common within these groups when P is scarcer relative to N (Figure 4.11 b). The heterogeneous "Other phyla" group is comprised of bacteria from several phyla, the most numerous being the *Verrucomicrobia* (41 sequences) while more than 20 sequences from each of the *Acidobacteria, Fibrobacteres, Balneolaeota* and *Spirochaetes* were also present. The small number of sequences from each group meant that models built using only sequences attributed to these phyla lacked sufficient statistical power to discern any significant relationship. Although the GLM did not detect a significant relationship

between SAR11 OlsB and N*, visual analysis of scatterplots (Figure 4.11 b) indicated OlsB sequences attributed to SAR11 were only found at relatively high values of N*. The minimum N* value at which SAR11 OlsB detected was -1.22, which is greater than the mean value for N* in the dataset as a whole, which is consistent with the presence of OlsB only in SAR11 strains adapted for relatively P-depleted waters.

4.3 Discussion

Aminolipids are a poorly studied class of lipids, which seem to be found exclusively in bacteria. Only the biosynthesis of ornithine lipid had been characterised previously (Weissenmayer *et al.*, 2002; Gao *et al.*, 2004). In this study I used marker exchange mutagenesis to show that OlsB2 in *R. pomeroyi* DSS-3 is required for the production of glutamine lipid. A second gene, OlsA, which had previously been shown to convert lyso-OL to OL (Weissenmayer *et al.*, 2002), was also required for synthesis of QL (Figure 4.5). These findings indicate that glutamine lipid biosynthesis likely proceeds via a two-step process, analogous to the synthesis of ornithine lipid (Figure 4.1). The first step in QL biosynthesis, the *N*-acylation of glutamine with a 3-hydroxy fatty acid, is mediated by OlsB2. The second step, the *O*-acylation of the hydroxyl group of the first fatty acid, appears to be catalysed by the acyltransferase OlsA. OlsA can also acylate glycerol-3-phosphate to form phosphatidic acid, an intermediate in phospholipid biosynthesis (Aygun-Sunar *et al.*, 2007), indicating that it has relatively broad substrate specificity. Disruption of OlsA did not result in an accumulation of lyso-OL or lyso-QL, which might be expected to accumulate based on the proposed biosynthesis pathway (Figure 4.1). However, this lack of detectable lyso-OL is consistent with prior studies in *E. meliloti ΔolsA* strains (López-Lara *et al.*, 2005). It would appear that these lysolipids are under tight control in the cell and rapidly degraded if they are not acylated by OlsA to form the intact aminolipid. Lysophospholipids can have quite different effects on membrane curvature compared to diacylated lipids with the same head group (Fuller and Rand, 2001) so it seems likely that an accumulation of lysoaminolipids could have a deleterious effect on

membrane integrity. *R. pomeroyi* DSS-3 *ΔolsA* did not completely lack OL (Figure 4.5), suggesting that alternative pathways may exist for the acylation of lyso-OL. Enzymes which might partially substitute for OlsA are unclear, although numerous *O*-acyltransferases are present in the genome.

The role for QL in Roseobacters remains, to a large extent, mysterious. A slight growth defect was observed for *ΔolsB2,* deficient in QL but not OL, in P-depleted media relative to the wild type (Figure 4.6). This could have been due to the increased requirements for phosphate of this strain as a result of the lack of a phosphorus-free lipid. However, lipid analysis found no evidence for substantial remodelling in either strain following growth in low-P medium (Figure 4.6 b) and the growth yield was similar to that of the wild type. Thus, evidence for an involvement of QL in a response to P-limitation is inconclusive from these data. The strain lacking OlsA, and thus deficient in both QL and OL, exhibited a more severe growth phenotype, reaching a lower final yield in high-P medium and showing almost no growth in low-P medium (Figure 4.6). It is likely that OL and QL functionally substitute for one another to some extent, resulting in the more severe phenotype when both are depleted. Repeated attempts to grow *ΔolsA* in a range of phosphate concentrations below 0.5 mM (50 μM – 0.25 mM) reproduced this lack of growth. The reason for this lack of viability in low-P medium is unexplained: one explanation could be that a sufficient concentration of phosphate ions is required to stabilise the membrane in the absence of either aminolipid. This would be analogous to the phenotype of *E. coli* mutants lacking PE, which require divalent cations (such as Ca^{2+}) for viability (DeChavigny *et al.*, 1991).

OlsB2 is widespread within the *Rhodobacteraceae*, although it is absent from a clade including the genera *Stappia* and *Labrenzia* (Figure 4.7). Notably, this "*Stappia* group", although taxonomically part of the *Rhodobacteraceae*, appears to be phylogenetically distinct, possibly affiliated to the *Rhizobiales* or forming its own Family (Pujalte *et al.*, 2013; Simon *et al.*, 2017). It appears that OlsB2 was acquired

early during the emergence of the *Rhodobacteraceae* and was subsequently lost from the small number of strains in which it is absent, presumably as they adapted to niches in which QL no longer conferred a selective advantage. A possible exception is *Rhodobacteraceae* bacterium HTCC2255 which, based on a whole genome phylogeny, appears to have diverged from the rest of the *Rhodobacteraceae* early during their evolution (Billerbeck *et al.*, 2016). It is therefore possible that this divergence pre-dated the emergence of OlsB2.

The capability to synthesise OL, through OlsB, seems to be even more highly conserved in the *Rhodobacteraceae* (Figure 4.7). This observation, combined with the growth phenotype observed for the *ΔolsA* strain lacking both aminolipids and the constitutive production of OL and QL along with the lack of an obvious response of these lipids to P-limitation in either *R. pomeroyi* or *Phaeobacter* sp. MED193 (Figures 3.3, 4.5), point to these aminolipids playing an important role in the biology of Roseobacters even when P is plentiful. The previous finding of a role for OL in maintaining optimal amounts of *c*-type cytochromes in *Rhodobacter capsulatus* supports this view of aminolipids playing an integral role in *Rhodobacteraceae* biology (Aygun-Sunar *et al.*, 2006). This contrasts to previous findings in *E. meliloti* that the lack of OL had a minimal impact on fitness except when P was limiting (López-Lara *et al.*, 2005). Several strains also appear to only synthesise OL when grown in P-depleted medium (Minnikin and Abdolrahimzadeh, 1974; Carini, Campbell, *et al.*, 2014). These observations suggest a model whereby aminolipids play different roles in different bacteria: in some strains the capacity to produce aminolipids has largely been acquired as an adaptation to P scarcity whilst in other bacteria they play a more integral role in the cell. Analysis of aminolipid synthesis genes in the *Tara* metagenomes data set provided some support for this hypothesis (Figure 4.11), at least in the case of OlsB. The abundance of this gene showed an overall positive relationship with N*, indicating that it may provide a selective advantage in P-depleted conditions. When looking at OlsB abundance broken down into broad microbial groups, however, this relationship was not universal. Indeed, the two

groups contributing the most to overall OlsB abundance, the *Rhodobacteraceae* and the *Rhodospirillales* (Figure 4.9), showed no significant relationship with N*. Of the groups contributing most to the overall positive relationship between OlsB abundance and N*, the *Rhizobiales* includes cultured representatives such as *E. meliloti* in which OL production seems primarily to occur as a response P-limitation (Geiger *et al.*, 1999). There was also a significantly positive overall relationship between OlsB2 abundance and N* in the *Tara* dataset but in this case, neither of the group-wise abundances of *Rhodobacteraceae* or *Gammaproteobacteria*, the two major sources of OlsB2 in the metagenomes (Figure 4.9), showed a significant relationship to N*. The regression using the complete data for each gene are inevitably more powerful due to the larger number of counts, which may explain the ability to detect an effect with the overall data but not with group-specific abundances. There was also no evidence that the abundance of OlsF was influenced by N* (Figure 4.11). Unlike OlsB and OlsB2, which are primarily found in *Alphaproteobacteria,* OlsF is the aminolipid synthesis gene most commonly found in *Gammaproteobacteria* and *Bacteroidetes* (Figure 4.9). A recent lipidomic analysis of one marine *Bacteroidetes* strain, *Dokdonia* sp. MED134, showed the presence of several aminolipid classes which comprised a substantial proportion of the lipidome even in P-replete conditions (Sebastián *et al.*, 2016). Little is known about the role of aminolipids in *Bacteroidetes* or whether OlsF is responsible for synthesis of all these lipids.

The likelihood that the primary role of aminolipids in *Rhodobacteraceae* is not as an adaptation to P-limitation raises the question of what role these lipids have in these organisms. Ornithine lipid is often found in the outer membrane of gram negative bacteria (Dees and Shively, 1982; Rojas-Jiménez *et al.*, 2005). An analysis of the synteny of OlsB2 revealed that it most commonly occurs immediately downstream of BamE, a component of the β-barrel outer membrane protein assembly machinery (Knowles *et al.*, 2009). This proximity hints that QL may be primarily involved in the outer membrane, since genes with similar functions are often located close to one another in bacterial genomes (Overbeek *et al.*, 1999), although the localisation of QL

remains undetermined. In *R. pomeroyi* DSS-3 both OL and QL had similar fatty acid profiles, suggesting a similar localisation within the membrane. For both lipids, a major species with a 3-hydroxy 20:1 and 18:1 fatty acids was present, while a second, less abundant, species with a 3-hydroxy 20:1 and 16:0 fatty acid was also observed (Figure 4.2). Notably, this fatty acid composition is quite distinct from that of major glycerolipids in the cell, in which C20 fatty acids are rare. Although they detected a greater diversity of QL and OL species, Zhang *et al.* (2009) similarly found long chain fatty acids in the R1 position of these lipids in *R. sphaeroides*. Lipids of similar fatty acyl chain length can tend to spontaneously cluster together in the membrane, in order to minimise the hydrophobic mismatch between lipids of different lengths (Lingwood and Simons, 2010). This might indicate these aminolipids have a different distribution in the membrane to phospholipids, which would be consistent with the observation that OL is enriched in the outer membrane. As a result of the similarities between the acyl-oxyacyl fatty acid structure of aminolipids and lipid A, it has been proposed that OL might partially substitute for lipopolysaccharide in the outer leaflet of the outer membrane (Nikaido, 2003). Like aminolipids, the fatty acid composition of lipid A is often quite distinct from that of phospholipids and can vary greatly, from C10 up to C22 (Steimle *et al.*, 2016). In two *Rhodobacteraceae*, *R. capsulatus* and *R. sphaeroides*, the lipid A fatty acids appear to be relatively short (C10 – C14; Wilkinson, 1997), which would result in a substantial hydrophobic mismatch if they occurred together with the long chain aminolipids. Unfortunately, my attempts to separate the inner and outer membranes of *R. pomeroyi* DSS-3 were unsuccessful so I was unable to determine the localisation of QL.

If QL is found to be enriched in the outer membrane, it may play a role in modulating outer membrane permeability, as has been postulated for OL (Nikaido, 2003). Several studies have found that OL become more abundant in the membranes of bacteria grown at higher temperatures (Taylor *et al.*, 1998; Seidel *et al.*, 2013). Some rhizobial species respond to acid stress by increasing the abundance of hydroxylated OL (Vences-Guzmán *et al.*, 2011). Hydroxylation increases the capacity for hydrogen

bonding between adjacent lipids, which is likely to reduce membrane fluidity and permeability (Nikaido, 2003). Glutamine lipids appear to have a higher capacity for lateral hydrogen bonding than unmodified OL due to the presence of an additional carbonyl group to act as an H-bond acceptor. Thus they may confer greater resistance against environmental stressors compared to OL alone. It is noteworthy that many Roseobacter strains thrive under phytoplankton bloom conditions (Teeling *et al.*, 2012) and even in direct association with phytoplankton (Grossart *et al.*, 2005). Such environments are highly enriched in rival microorganisms relative to bulk seawater so a capacity to resist bactericidal compounds is likely crucial to success (Strom, 2008). Some Roseobacters are able to produce their own allelopathic agents, such as the 'roseobacticide' produced by *Phaeobacter gallaeciensis* when grown alongside aging cultures of *Emiliana huxleyi* (Seyedsayamdost, Case, *et al.*, 2011). Presumably these Roseobacters are resistant to their own toxins and a component of this resistance is likely to be the relative impermeability of their membrane.

In conclusion, my results have demonstrated the role of an N-acyltransferase, OlsB2, in the synthesis of glutamine lipids in *Rhodobacteraceae*. This capacity is shown to be widespread within this Family, including being present in most of the Roseobacter group of abundant marine bacteria. I found no evidence that QL is involved in a response to phosphorus limitation and, indeed, the presence of QL under P-replete conditions argues for an important role in the normal functioning of the cell. Future work will hopefully help to establish what the likely role of this lipid is and how it has contributed to the distinctive biology of these important marine organisms.

Chapter 5 A novel class of sulfonolipid is widespread in marine Roseobacters

5.1 Introduction

Bacterial lipids are highly diverse yet there is currently little understanding of the benefits that this diversity provides (Sohlenkamp and Geiger, 2015). The study of bacterial lipid diversity is currently hampered by a lack of knowledge of both the chemical structures of many of these lipids, and the identity of genes involved in their synthesis. Once the chemical structure of a lipid is known, analytical strategies can be devised to detect the lipid in both the natural environment and cultures (Moore *et al.*, 2016). This can also help to direct studies into the lipid's biosynthesis, knowledge of which can provide a clearer idea of the likely distribution of a lipid in bacteria. One of the best studied aminolipids, ornithine lipid had been detected in diverse cultured bacteria since the 1960s (Shively and Knoche, 1969). However, it was not until the genes involved in its biosynthesis were elucidated that it became clear how widespread the capacity to produce OL was (Gao *et al.*, 2004; Vences-Guzmán *et al.*, 2015).

Recent studies, including my own work, have revealed that marine bacteria produce a wide variety of poorly characterised aminolipids. Sebastian *et al.* (2016) found a *Bacteroidetes* strain, *Dokdonia* sp. MED134, produced 4 uncharacterised aminolipids, in addition to ornithine lipid and a hydroxylated ornithine lipid. Combined, these lipids comprised more than half the total polar lipids of that strain, and this proportion increased to almost three quarters under P limitation. Similarly, I have found a number of uncharacterised lipids in two Roseobacter strains (Figure 3.2, Figure 4.2). Of particular interest are two lipid peaks of comparable intensity to the major phospholipids in both *Phaeobacter* sp. MED193 and *Ruegeria pomeroyi* DSS-3 (AAL1 and AAL2 in Figure 3.2 and Figure 4.2). The relatively low molecular weights of both lipids suggested that they were not phospholipids but rather, most probably,

aminolipids. As these lipids were seemingly such a major component of the lipidome of multiple Roseobacter strains, I characterised the chemical structure of one of them (AAL1) using high resolution accurate mass spectrometry. This identified the lipid as a seemingly novel class of sulfonolipid. Since the genes involved in the synthesis of this lipid are unknown, I investigated its distribution within the Roseobacter group by analysing lipid extracts from a battery of Roseobacter strains. Finding that some, but not all, of these strains produced the sulfonolipid, I conducted a comparative genomics investigation to identify candidate genes potentially involved in its biosynthesis.

5.2 Results

5.2.1 A previously uncharacterised lipid in lipid extracts from Ruegeria pomeroyi *DSS-3 is a homotaurine-containing lipid*

During LC-MS analysis of lipid extracts from *Phaeobacter* sp. MED193 and *R. pomeroyi* DSS-3, two prominent peaks in the base peak chromatogram eluted between 4 and 5 minutes. The most prominent ions in the earlier eluting peak had *m/z* values of 630.7, 656.7 and 670.7; in the second peak there was a similar pattern in the relative intensities of the masses but these were 16 *m/z* units higher. To elucidate the structure of these lipids, the most intense species, at 656.7 *m/z* (AAL1-656) was selected for high resolution MS/MS analysis on a quadrupole – time of flight (QTOF) mass spectrometer (Figure 5.1a). I scanned a range of collision energies in order to obtain maximum structural information. At low collision energy (40 eV) the major species formed corresponded to a neutral loss of 282 mass units (Figure 5.1b). This is consistent with the neutral loss of an 18:1 fatty acid. A second peak at *m/z* 281 is likely the carboxylate anion of an 18:1 fatty acid. Further fragmentation, at higher collision energies (up to 90 eV), yielded a major ion at *m/z* 237 (Figure 5.1c). This ion likely corresponds to a 16:0 fatty acid present as a ketene, which would be consistent with the fragmentation scheme proposed for ornithine lipids (Zhang *et al.*, 2009). These results, therefore suggest a lipid class with a similar fatty acyl backbone structure to the aminolipids, such as ornithine and glutamine lipid.

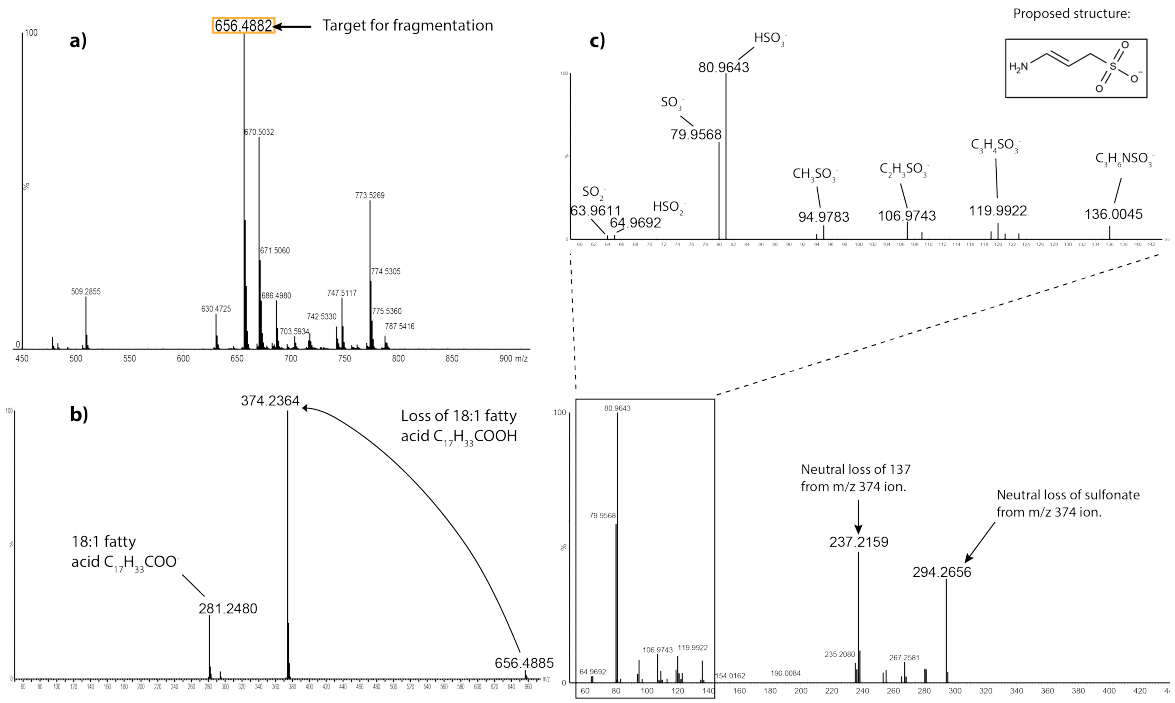


Figure 5.1 A previously unidentified class of homotaurine lipids is present in *Ruegeria pomeroyi* DSS-3. a) Intact masses of *R. pomeroyi* lipids, measured using a high resolution, accurate mass quadrupole-time of flight mass spectrometer in negative ion mode. The identity of the most abundant ion, highlighted, was unknown, so this ion was fragmented in order to elucidate its structure. b) Fragmentation spectrum at 40 eV collision energy. The most abundant species corresponds to the loss of a 18:1 fatty acid. c) Fragmentation spectrum at 90 eV collision energy. The lower spectrum shows a peak at 237.2159, consistent with the presence of a 16:0 fatty acid. The upper spectrum shows an expanded view of the spectrum in the *m/z* range below 140. Ions are annotated with their predicted elemental composition. Inset is the proposed structure of the 136 *m/z* fragment (3-aminopropene sulfonic acid).

Prominent peaks at 80 and 81 *m/z*, respectively, were apparent in the fragmentation spectrum obtained at 90 eV collision energy (Figure 5.1c). The accurate masses for these ions were 79.9568 and 80.9643. Of the candidate formulae within 100 ppm of the measured mass, SO_3^- and HSO_3^- appear most plausible, with mass errors of 0.182 ppm and 4.194 ppm, respectively. A smaller peak doublet at *m/z* 63.9611 and 64.9692 was also present in the 90 eV spectrum. These masses are unambiguously assigned to SO_2^- (12.506 ppm) and HSO_2^- (8.08 ppm). Taken together, these results demonstrate the presence of a sulfonate group in the lipid.

An ion at 136 *m/z* corresponded to the deprotonated head group. The mass determined here is larger than that of deprotonated taurine (*m/z* 124). Since the head group includes a sulfonate (SO_3^-) group, the plausible formula most closely corresponding to the accurate mass is $C_3H_6NSO_3$ (Table 5.1). This is consistent with the structure being 3-amino propene sulfonate (Figure 5.1).

Table 5.1 Nominal and accurate masses of proposed head group fragments.

Nominal *m/z*	Accurate *m/z*	Formula	Mass Error / ppm
136	136.0045	$C_3H_6NSO_3$	12.784
120	119.9922	$C_3H_4SO_3$	34.048
107	106.9743	$C_2H_3SO_3$	51.315
95	94.9783	CH_3SO_3	16.737

Charge remote fragmentation is a gas-phase process, which can occur at both low and high collision energies, whereby fragmentation takes place without involving charge transfer reactions (Cheng and Gross, 2000). It is favoured when the charge is fixed in one location by a strongly ionised group and can result in very informative fragmentation patterns, for example along the alkyl chain of a fatty acid. Sulfonate is a strongly acidic group and consequently, when analysed in negative ionisation mode, negative charge tends to be fixed to the deprotonated sulfonic acid. This property has been used to characterise the structure of the sulfo-glycolipid SQDG (Kim *et al.*, 1997) and also of taurine- and amino propane sulfonate-conjugated bile acids (Stroobant *et*

al., 1995). Charge-remote fragmentation would appear to explain the series of fragment peaks observed for AAL1-656 at 90eV in the *m/z* range between the intact head group at 136 and the sulfonate peaks at 80 and 81 (Table 5.1). Of formulae containing at least one sulfur and three oxygen atoms, only one lay within 100 ppm of 119.9922 and 94.9783. Two matches were obtained for *m/z* 106.9743 ($C_2H_3SO_3$ and $CHNO_3S$). Of these, $C_2H_3SO_3$ had the lower mass error (55.988 ppm compared to 61.574 ppm) and appeared more consistent with the formulae assigned to the adjacent peaks. The formulae assigned to these masses support the structure proposed in Figure 5.1. However, the mass error is somewhat high, possibly due to the low number of ions at high collision energy. To account for these results, I outline a proposed fragmentation scheme in Figure 5.2. The unsaturation in the head group arises during fragmentation, implying that the intact head group is 3-aminopropane sulfonate (homotaurine).

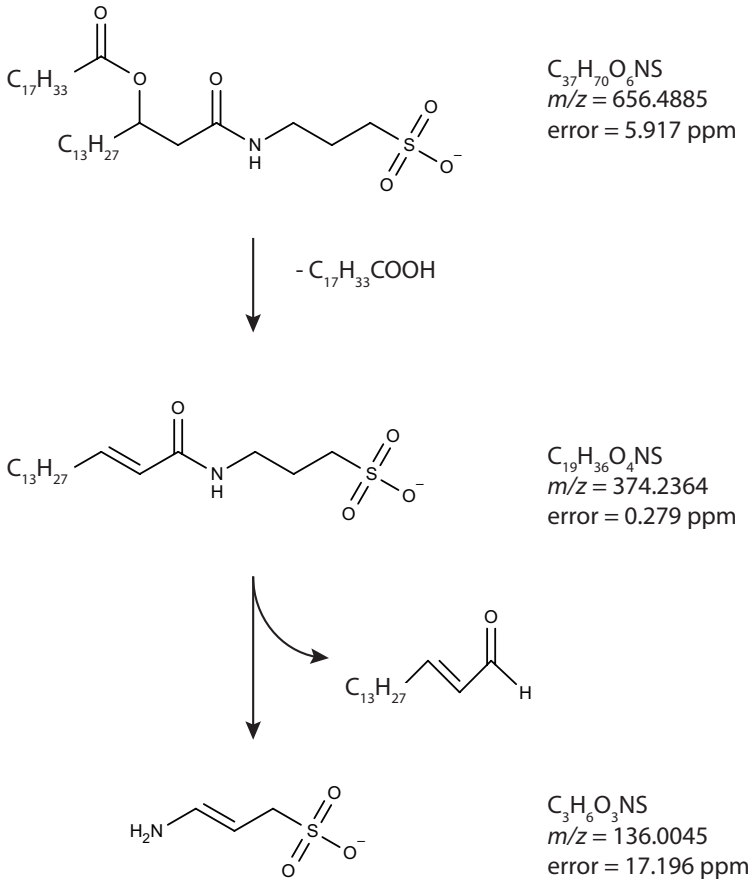


Figure 5.2 Proposed fragmentation scheme for the homotaurine lipid with *m/z* 656. Mass errors indicate the difference between the measured and theoretical masses of the proposed species.

5.2.2 The homotaurine component of homotaurine lipids is not derived from sulfonate-containing buffer molecules

Sulfonate compounds, such as homotaurine, are relatively rare in bacteria. The production of similar compounds has not, to my knowledge, been reported in Roseobacters, although a recent paper reported taurine biosynthesis in some cyanobacteria and *Bacteroidetes* (Agnello *et al.*, 2013). One potential source for this sulfonate group could be molecules of the buffer in the medium used to conduct the aforementioned studies, HEPES (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid), which contains an ethanesulfonic acid group. To ensure that HEPES molecules were not being used for the production of homotaurine lipids (HTLs), I grew *R. pomeroyi* DSS-3 in the defined MAMS medium containing only phosphates buffer. In the phosphates buffer medium, sulfate was the only supplied sulfur source so any sulfonate-containing compounds must be have been produced *de novo* by the bacteria. Since HTLs were detected in cultures grown with phosphates buffer (Figure 5.3 b), the homotaurine component of these lipids must have been produced by *R. pomeroyi*.

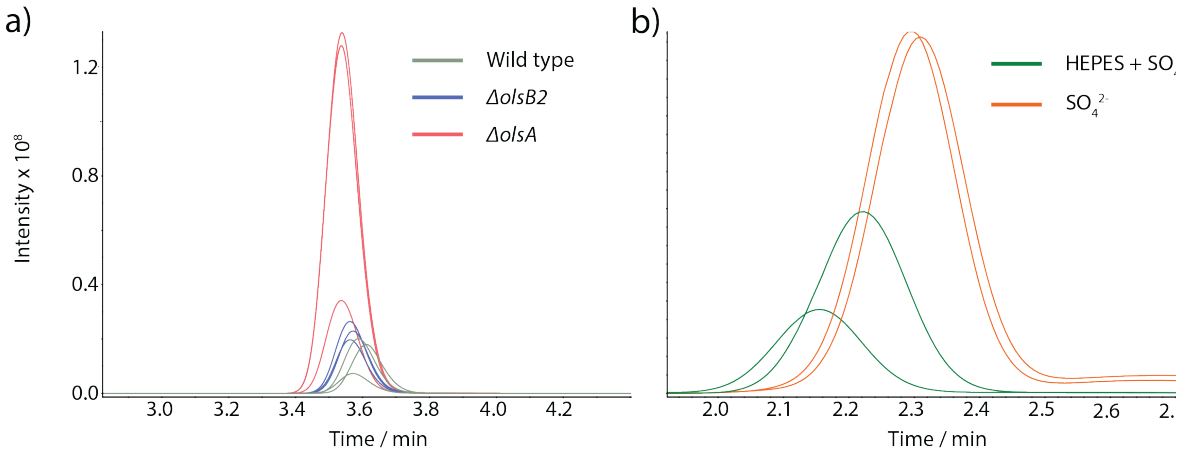


Figure 5.3 Extracted ion chromatograms showing ions of *m/z* 656, the major homotaurine lipid (HTL) species, in lipid samples from *Ruegeria pomeroyi* DSS-3 analysed in negative ion mode. **a)** After overnight growth in ½YTSS, the aminolipid mutants constructed in Chapter 4 (*ΔolsA* and *ΔolsB2*) produced HTL at levels comparable to, or in excess of, the wild type. **b)** HTL abundance after growth of *R. pomeroyi* in MAMS defined medium either with (green line) or without (orange line) HEPES buffer. The medium without HEPES was instead buffered to the same pH using phosphates buffer. The results in **b** were obtained using a different chromatographic method to those presented in **a**, including a different column (described in more detail in Section 2.4.1). This explains the different retention times between the two panels.

5.2.3 A cysteine sulfinic acid decarboxylase homolog is not required for homotaurolipid synthesis

Although HTLs are aminolipids, similar in structure to ornithine and glutamine lipids, I found that neither OlsB2 nor OlsA are required for HTL synthesis (Figure 5.3 a). This suggested that synthesis of HTL proceeds via a different pathway to the other aminolipids. Lacking suitable candidate enzymes that might be involved in this process, I instead attempted to disrupt homotaurine synthesis as a means to investigate the phenotype resulting from a lack of HTL. The biosynthesis of homotaurine has not, to my knowledge, been described for any organism. However, the synthesis of taurine, a structural analogue of homotaurine, occurs via the oxygenation of cysteine to cysteine sulfinic acid by a cysteine dioxygenase (CDO). This is followed by the decarboxylation of cysteine sulfinic acid to yield hypotaurine

(Agnello *et al.*, 2013), which is then thought to undergo spontaneous oxygenation to form taurine. The enzymes involved in this process, CDO and cysteine sulfinic acid decarboxylase (CSAD) are present in a number of bacteria, and those from *Synechococcus* have been shown to produce taurine (Agnello *et al.*, 2013). Homologs of both CSAD and CDO are present in the genomes of several Roseobacters, including *R. pomeroyi* DSS-3 and *Phaeobacter* sp. MED193, where they occur adjacent to one another, in an apparent operon (Figure 5.4). I hypothesised that homotaurine might be synthesised by a mechanism analogous to that of taurine. Homocysteine, an intermediate in the methionine and cysteine biochemical cycles (Ferla and Patrick, 2014), might serve as a precursor, being dioxygenated to form homocysteine sulfinic acid before being decarboxylated to form hypohomotaurine, which could then be converted to homotaurine (Figure 5.4). To test whether this was the case, I constructed a mutant in *R. pomeroyi* DSS-3 in which the CSAD homolog, *spo3687,* was disrupted by marker exchange mutagenesis with a gentamicin resistance cassette. Colony PCR of this mutant using primers targeting the start and end of the homology regions flanking the resistance cassette, yielded a product of around 2600 base pairs (Figure 5.5). This is consistent with the expected size for this product in the mutant but smaller than would be expected for the wild type (3102 base pairs). The identity of the mutant was confirmed by sequencing of these PCR products. Following growth of this mutant, *Δspo3687*, in ½ YTSS medium, I analysed lipid extracts from the cultures using LC-MS. Homotaurolipids were still present and their abundance was not significantly different compared to wild type cultures grown alongside the mutant (Figure 5.4). Given that HTL synthesis was not abolished in the mutant, I ceased my investigations into the function of this gene and did not attempt to further assess any putative role it might have in taurine or homotaurine biosynthesis.

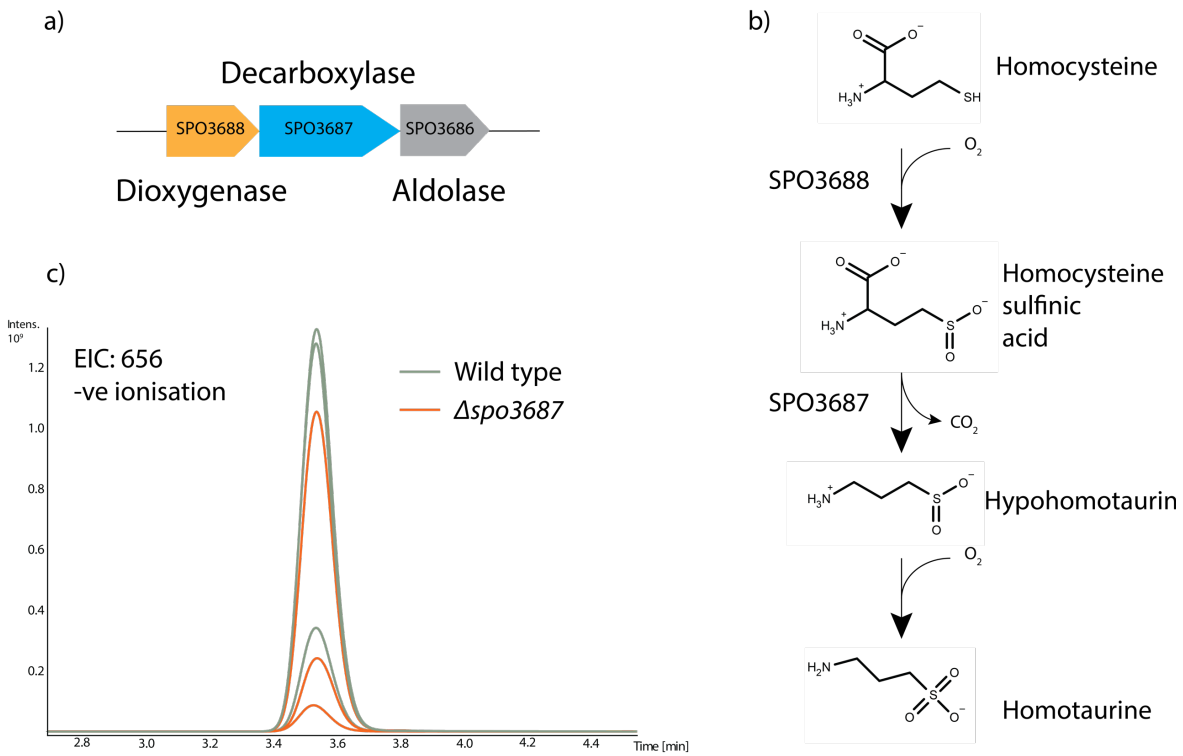


Figure 5.4 Investigation of the putative role of a cysteine sulfinic acid decarboxylase (CSAD) homolog in homotaurolipid biosynthesis. **a)** Gene organisation of the CSAD homolog (*spo3687*) in the *Ruegeria pomeroyi* DSS-3 genome. Immediately upstream is a homolog of cysteine dioxygenase (SPO3688). **b)** The hypothesised pathway for homotaurine biosynthesis by analogy to the synthesis of taurine. **c)** Extracted ion chromatogram for the major species of homotaurolipid (*m/z* 656), collected in negative ionisation mode. Chromatograms are shown for 3 biological replicates of both wild type *R. pomeroyi* and a deletion mutant (*Δspo3687*) in which the putative CSAD, *spo3687*, has been inactivated.

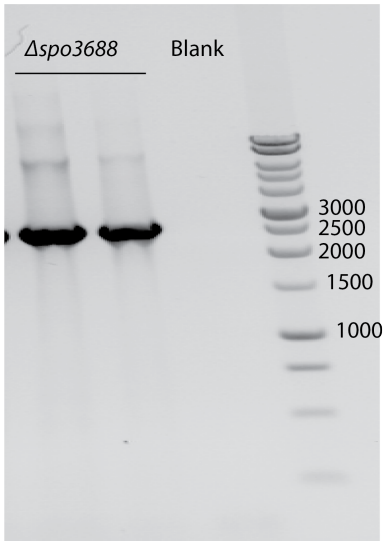


Figure 5.5 Confirmation of *Δspo3688* in *R. pomeroyi* using the A forward and B reverse primers. The size of the observed product is similar to the expected product size for a correctly assembled mutant (2617 base pairs) and smaller than would be expected for the wild type (3102 base pairs).

5.2.4 A subset of the Roseobacter clade produces homotaurine lipids when grown in marine broth medium

Faced with the absence of a suitable genetic marker for the capacity to synthesise HTL, I investigated the presence of HTL in a collection of Roseobacter strains. I selected 16 strains, in addition to *R. pomeroyi* DSS-3, to obtain as much coverage as possible of the Roseobacter group as a whole (Figure 5.6). The selected strains included *Stappia stellulata*, which recent phylogenetic studies indicate is not a member of the *Rhodobacteraceae* (Pujalte *et al.*, 2013), which served as an outgroup. These strains were each grown in the complex seawater-mimic medium, marine broth, overnight, before cells were harvested for lipid analysis. Homotaurolipid was detected in all the strains tested apart from *S. stellulata* and *Dinoroseobacter shibae* (Figure 5.6). The separation of these two strains from the remaining Roseobacter sequences is in line with previous results showing *D. shibae* branching deeply within the *Rhodobacteraceae* phylogeny (Simon *et al.*, 2017).

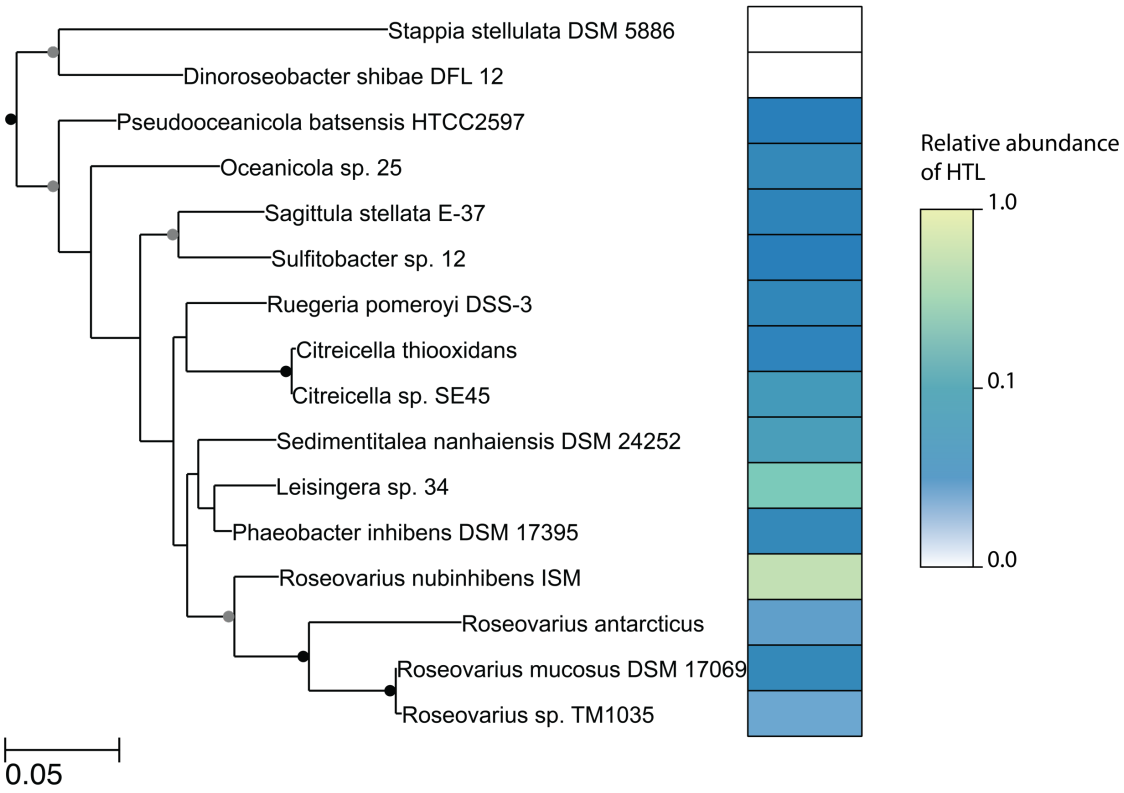


Figure 5.6 Phylogeny of 16S rRNA sequences from selected *Rhodobacteraceae* plotted alongside information on the abundance of homotaurolipid (HTL). Sequences for the phylogeny were downloaded from the SILVA database and aligned using MAFFT. A maximum likelihood phylogeny was then constructed using RaxML and 100 bootstrap replicates used to assess the consistency of the tree. Nodes with more than 50% bootstrap support are indicated with circles (grey: 50 – 69 %; black: > 70%). *Stappia stellulata* 16S rRNA was used as an outgroup. The abundance of HTL measured in each strain, as a proportion of the maximum abundance measured, is plotted alongside the phylogeny. Abundances were calculated as the peak area of HTL species relative to a phosphatidylglycerol internal standard.

Attempting to generate hypotheses as to genes potentially involved in the synthesis of HTL, I conducted a comparative genomics investigation into the Roseobacter strains whose lipid profiles had been analysed. Several of the strains analysed had not been genome sequenced so these were, of necessity, excluded from the analysis. I reasoned that the synthesis of HTL would require an *N*-acyltransferase activity to acylate homotaurine, analogous to that mediated by OlsB in the synthesis of ornithine lipid (Gao *et al.*, 2004). Therefore, I looked for predicted *N*-acyltransferases that were present in all of the strains that produced HTL in marine broth (the 'producers') while being absent from the strains that did not produce HTL (the 'non-producers'). It should be noted that this analysis is confounded to some extent by the phylogenetic structure of the strains being analysed: the producer strains are more closely related to each other than to the two non-producer strains (Figure 5.6). Therefore I expected that some proportion of the candidate genes identified could be false positives. I assigned all the genomic sequences from the 9 sequenced producer strains and 2 non-producer strains to orthologous groups (OGs) using the eggNOG-mapper software (Huerta-Cepas, Forslund, *et al.*, 2016), which provides a consistent pipeline for sequence annotation and OG assignment by comparison to the eggNOG database (Huerta-Cepas, Szklarczyk, *et al.*, 2016). I identified a group of 1417 'core' OGs which were present in the genomes of all producer strains of which 1060 were also present in the two non-producers (Figure 5.7 a). A further 285 of these core OGs were present in *D. shibae* while 35 were present in *S. stellulata,* leaving 37 candidate genes present in all producer strains but not in the genomes of the non-producers (Figure 5.7 a). Two of these candidate genes were annotated as being potential acyltransferases (Table 5.2) so I investigated the neighbourhood of these genes in more detail. One of these genes, a predicted phospholipid glycerol acyltransferase (05CDD) occurs downstream of the ChvIG response regulator system (Figure 5.7 b), which regulates the response to acid stress in *Agrobacterium tumefaciens* (Li *et al.*, 2002) and is crucial for host colonisation by *Ensifer meliloti* (Bélanger *et al.*, 2009; Wang *et al.*, 2010). In all strains other than *Pseudooceanicola batsensis* and *Roseovarius nubinhibens*, a predicted 3-hydroxyacyl-CoA dehydrogenase, homologous to FadB, is present

immediately downstream of 05CDD on the opposite strand (Figure 5.7 b). FadB is a multifunctional enzyme involved in the β-oxidation of fatty acids (Heath *et al.*, 2002). It catalyses the conversion of enoyl-CoA to 3-hydroxyacyl-CoA as well as the subsequent dehydrogenation of 3-hydroxyacyl-CoA to 3-ketoacyl-CoA. It also has further enoyl-CoA isomerase and 3-hydroxyacyl-CoA epimerase activities (Heath *et al.*, 2002). A second 05CDD ortholog was present in the genome of *Sagittula stellata* (SSE37_RS24755). However, the genomic neighbourhood surrounding this gene was not syntenic with any of the other strains (Figure 5.7 b), suggesting that it may not be a true ortholog. The gene neighbourhood of the second candidate gene (08UX5), a predicted GcN5-related *N*-acetyltransferase (GNAT), appeared slightly less conserved (Figure 5.7 c). In all cases, there was a predicted inosine-uridine-preferring nucleoside hydrolase immediately upstream of 08UX5. Such nucleoside hydrolases hydrolyse ribonucleosides to release ribose sugar and the nucleic acid base (Petersen and Møller, 2001; Versées and Steyaert, 2003), allowing the host organism to salvage the base from the environment. Another conserved gene, present in 5 of the 9 strains, was annotated as YdiU, a conserved hypothetical selenoprotein with a possible protein kinase domain (Dudkiewicz *et al.*, 2012).

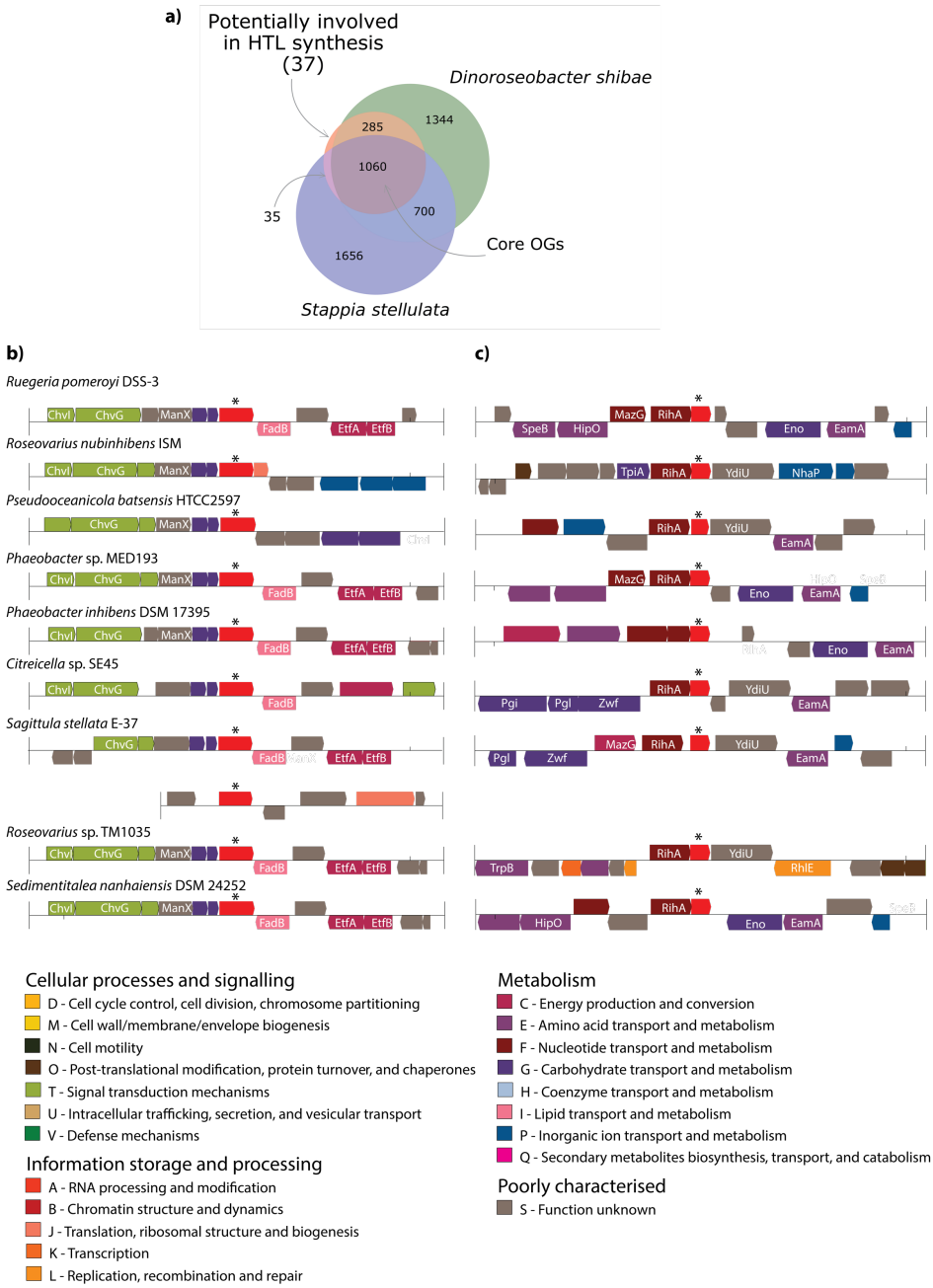


Figure 5.7 Identification of two candidate acyltransferases to be investigated for involvement in homotaurolipid (HTL) biosynthesis. **a)** Venn diagram illustrating the logic used to identify the candidate genes. The central circle represents the 1417 orthologous groups common to all HTL-producing strains. Below are genomic neighbourhoods for genes in eggNOG OGs **b)** 05CDD and **c)** 08UX5. The candidate gene is highlighted in red, while other neighbouring genes are coloured according to COG category.

Table 5.2 Candidate eggNOG orthologous groups (OGs).[a]

OG	Locus tag (*R. pomeroyi)*	COG category	eggNOG annotation
07TNN	SPO0365	S	Domain of unknown function (DUF1989)
05CDD	**SPO0716**	**S**	**Phospholipid glycerol acyltransferase**
01VSN	SPO0838	M	Bacterial sugar transferase
05MEE	SPO0968	L	Nudix Hydrolase
05K39	SPO1109	S	Endonuclease Exonuclease phosphatase
05NU0	SPO1287	S	Glyoxalase bleomycin resistance protein dioxygenase
08EFD	SPO1424	P	Sodium hydrogen exchanger
08K1F	SPO1437	I, Q	Dehydrogenase
05D82	SPO1743	E	Glutamate dehydrogenase
05S99	SPO1919	P	Tellurite resistance protein
090P8	SPO1963	G	phosphoglycerate mutase
05N59	SPO2284	M	Transglycosylase
01QFF	SPO2332	P	Acriflavin resistance protein
08YCB	SPO2344	E	Sarcosine oxidase, gamma subunit
08UX5	**SPO2471**	**S**	**Acetyltransferase, (GNAT) family**
01QH6	SPO2697	C	CoA-binding domain protein
07QRG	SPO2741	L	DNA polymerase iii
05X3J	SPO2754	S	
060TV	SPO2757	S	EF hand domain protein
01TDR	SPO2815	P	ABC transporter permease protein
08TCB	SPO3054	S	
0808N	SPO3431	S	Domain of unknown function (DUF3576)
08RQB	SPO3439	I	Enoyl-coA hydratase isomerase family protein
08R76	SPO3444	M	3-deoxy-d-manno-octulosonic-acid transferase
07RFB	SPO3740	O	Catalyzes the reversible oxidation-reduction of methionine sulfoxide in proteins to methionine
6405	SPO0725	S	SH3 type 3
01RKE	SPO2816	P	ABC transporter permease protein
05DM9	SPO0035	S	Core-2 I-branching enzyme family protein
0627S	SPO_RS14095	S	
07SF7	SPO2814	E	Peptide opine nickel uptake family ABC transporter periplasmic substrate-binding protein
05SH6	SPO3502	S	Pfam:DUF3118
08U9J	SPO3839	S	Type I secretion target repeat protein
05WTG	SPO2776	S	
05Q6E	SPO3716	S	
08T58	SPO3836	S	Mg^{2+} transport protein CorA
064YK	SPO1241	S	
05U51	SPO0776	S	Thioesterase

[a]Groups in bold have a putative acyltransferase function. Annotations are based on those provided in the eggnog database. Where there is no annotation, the function is unknown.

5.3 Discussion

5.3.1 A novel class of sulfonolipids in marine Roseobacters

In this work I have shown that a previously uncharacterised lipid, which appears to be widespread amongst marine Roseobacters, is a novel aminolipid with a probable homotaurine head group (Figure 5.1,Figure 5.2). While homotaurine is the most likely head group, based on the MS fragmentation data alone it is not possible to rule out the sulfonate group being located on the first or second carbon, with reference to the amino group. A study using a complementary technique, such as nuclear magnetic resonance (NMR) could resolve this uncertainty. The presence of a sulfonate group means that these lipids also fall into the broad category of sulfonolipids. The most abundant, and best studied, lipid of this type is the plant sulfolipid, SQDG (Figure 5.8), which is present in the membranes of most oxygenic phototrophs (Frentzen, 2004), as well as in some heterotrophic bacteria (Villanueva *et al.*, 2013). The plant sulfolipid is an abundant and probably structural component of photosynthetic membranes, although the recovery of SQDG molecules in crystal structures of photosystem proteins suggests specific binding of these lipids and a likely essential role in photosynthesis (Umena *et al.*, 2011). A number of other sulfolipids appear to elicit potent responses when certain organisms are exposed to them. A sulfo-aminolipid (copepodamide, Figure 5.8), produced by a number species of copepods, a major zooplankton group, was found to induce toxin production in the dinoflagellate, *Alexandrium minutum* (Selander *et al.*, 2015), likely as a defence against predation. In a separate example, a sulfonolipid (rosette-inducing factor, Figure 5.8) produced by the *Bacteroidetes* bacterium *Algoriphagus machipongonensis* induced the development of multicellular rosettes in a choanoflagellate (Alegado *et al.*, 2012). It seems in both of these examples that sulfonolipids are used by the sensing organism as a marker for the presence of another organism with which it interacts (either as a predator or as a symbiont). The fact that sulfo-aminolipids appear to be relatively rare across the tree of life likely makes them well suited to mediate such chemical interactions where a high degree of specificity is required.

Lipids similar to those produced by *A. machipongonensis* have been described in a number of *Bacteroidetes*, particularly amongst the *Cytophaga* (Godchaux and Leadbetter, 1980, 1983, 1984). They tend to be localised to the outer membrane, and seem to play a role in the gliding motility of these organisms (Abbanat *et al.*, 1986; Godchaux and Leadbetter, 1988). The sulfonolipids of the *Bacteroidetes* differ from those that I have described in Roseobacters in that they are composed of a base, termed capnine (Figure 5.8), similar to the sphingoid bases of sphingolipids, which may be *N*-acylated to form the full sulfonolipid (Godchaux and Leadbetter, 1980). In this way they are similar structurally to sphingolipids whereas the Roseobacter homotaurolipids are more similar to aminolipids such as ornithine lipid.

The most abundant species of OL and QL in *R. pomeroyi* DSS-3 had a 20:0/18:1 fatty acid composition under the growth conditions I used (Figure 4.2). These fatty acids are relatively unusual amongst other lipid classes in *R. pomeroyi*. By contrast, HTL had a fatty acid composition more akin to the phospholipids, with 16:0 and 18:1 being the most abundant groups (Figure 5.1). This might indicate that HTLs have a similar localisation in the cell membrane to phospholipids, unlike OL which is thought to be localised to the outer membrane (Dees and Shively, 1982; Wee and Wilkinson, 1988). However, the distribution of lipids between the membranes in members of the Roseobacter group has yet to be established.

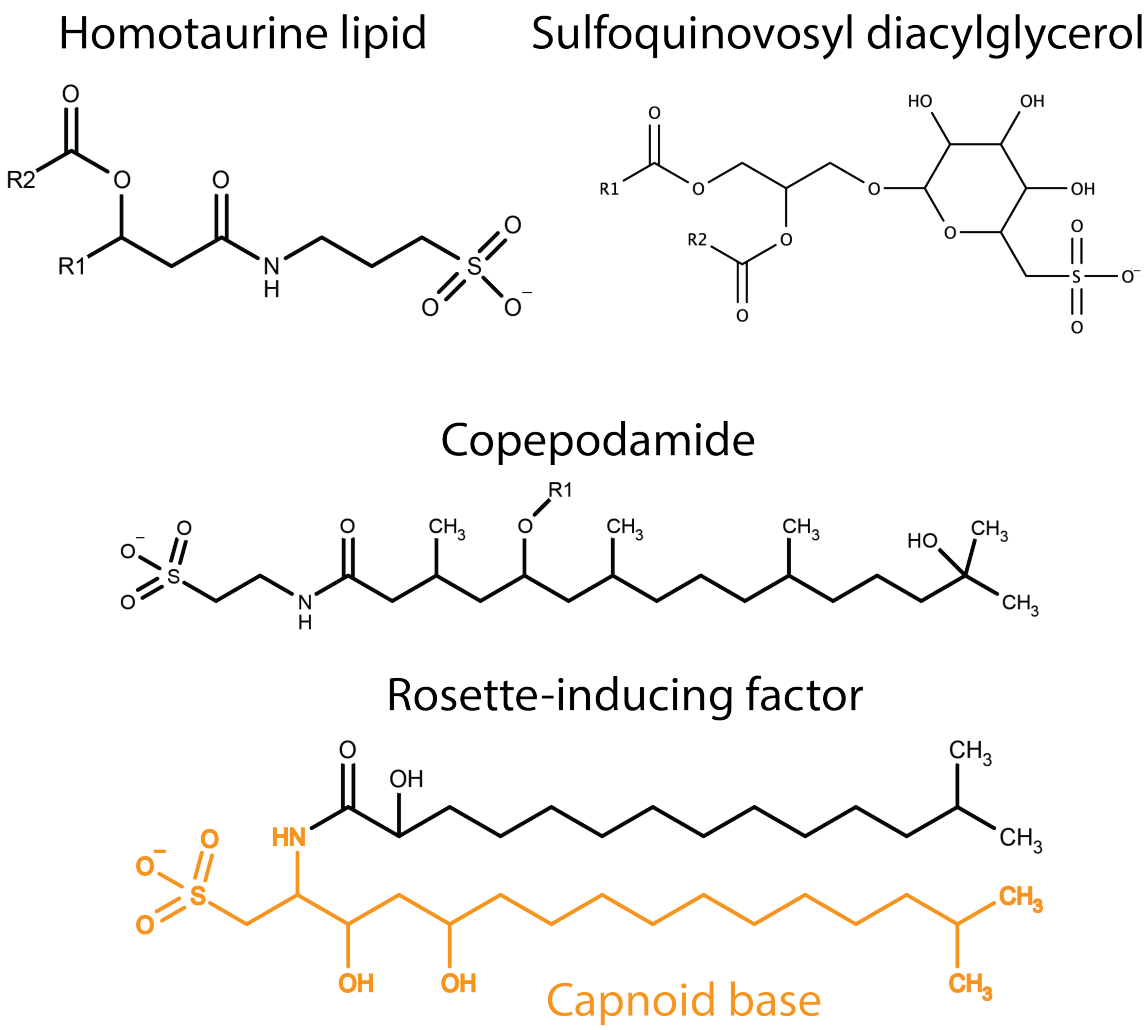


Figure 5.8 Structures of the sulfonolipids mentioned in the Discussion. The R1 group of copepodamide can be docosahexaenoic acid, eicosahexaenoic acid or stearidonic acid (Selander *et al.*, 2015).

A survey of the distribution of HTL among isolates from the Roseobacter group indicated that the ability to produce this lipid is widely distributed within the group (Figure 5.6). One strain, *D. shibae*, the most basal of the strains examined, lacked any HTL under the conditions assessed, as did an outgroup strain, *S. stellulata*. The lack of HTL likely indicates that these strains lack the capacity to produce the lipid, as other Roseobacter strains examined seemed to produce HTL constitutively. However, it is possible that these strains might have the capacity to produce HTL but only do so under certain conditions. This pattern is observed for ornithine lipid, which is produced constitutively in some bacteria, such as *R. pomeroyi* DSS-3 (Chapter 4), but in others is only produced as a response to P-depletion (Minnikin and Abdolrahimzadeh, 1974; Carini *et al.*, 2015). The sparse sampling of the more basal Roseobacters in this study means that it is not possible to ascertain whether an absence of HTL is typical for these organisms. In particular, no representatives from the clades of pelagic Roseobacters, which are amongst the most abundant in the marine environment (Selje *et al.*, 2004; Billerbeck *et al.*, 2016), were analysed in this work. In addition to their abundance, these organisms have a notably different ecology to many of the other Roseobacters, which often live in close association with particles and phytoplankton or other larger organisms (Wagner-Döbler and Biebl, 2006). Thus, it would be interesting to establish whether HTL was also present in these pelagic organisms as this would give further insights into its role in the ecology of marine Roseobacters.

One important step to developing a better understanding of the role of HTLs will be to discover how they are synthesised. The two candidate acyltransferases, identified through comparative genomics of a library of Roseobacter strains, may serve as a starting point for future investigations into how this lipid is synthesised (Figure 5.7; Table 5.2). The genomic neighbourhood of 05CDD in particular appears to hold promise, as several neighbouring genes could plausibly be functionally related to HTL synthesis (Figure 5.7 b), although it is important to stress that there is no direct evidence for this. These include a homolog of FadB, a 3-hydroxyacyl-CoA

dehydrogenase. If HTL synthesis occurs in a manner analogous to that of OL, a 3-hydroxy fatty acid would be required (Weissenmayer *et al.*, 2002) as a substrate for the first step in HTL synthesis. One of the activities of FadB involves the conversion of enoyl-CoA to 3-hydroxyacyl-CoA during the degradation of fatty acids (Heath *et al.*, 2002). It is therefore possible that FadB is involved in the production of 3-hydroxy fatty acids for the synthesis of aminolipids, such as HTL. An alternative source of such 3-hydroxy fatty acids is during fatty acid biosynthesis, when 3-hydroxyacyl-acyl carrier protein (ACP) is produced. During lipid A biosynthesis in *E. coli* the enzymes LpxA and LpxD are responsible for the acylation of *N*-acetylglucosamine with 3-hydroxy fatty acids (Whitfield and Trent, 2014). The competition for 3-hydroxyacyl-ACP between these enzymes on the one hand and the dehydratases FabA and FabZ on the other is a potentially rate determining step in lipid A biosynthesis (Heath *et al.*, 2002). Thus the presence of alternative sources of 3-hydroxyacyl fatty acids could enable finer control over the rate of synthesis of aminolipids such as HTLs.

5.3.2 The presence of homotaurine in Roseobacters

The identification of a homotaurine moiety in aminolipids from Roseobacter group bacteria raises the question of how this moiety is synthesised. I was unable to determine the biosynthesis of HTL, however, previous work on the synthesis of the aminolipids OL and QL, as well as the synthesis of the related N-acyl amino acids, suggests that synthesis is likely to proceed by the N-acylation of homotaurine (Chapter 4; Brady *et al.*, 2004; Gao *et al.*, 2004). Since no exogenous supply of homotaurine was provided, this would imply that homotaurine must be produced by the cell. The presence of homotaurine has been documented in some red algae (Miyasawa *et al.*, 1970; Ito *et al.*, 1977) but never previously in bacteria, as far as I am aware. As a structural analogue of γ-aminobutyric acid (GABA), homotaurine appears to activate GABA-receptors in the mammalian brain, a property which has led to its proposal as a potential therapeutic for Alzheimer's disease (Aisen *et al.*, 2011). Enzymes involved in the catabolism of GABA also show similar levels of activity against homotaurine as they do with GABA as a substrate (De Gracia and

Jollès-Bergeret, 1973; Mayer and Cook, 2009). Indeed, several Roseobacters, including *R. pomeroyi* DSS-3, are able to use homotaurine as a sole nitrogen source (Mayer and Cook, 2009). The relative availability of GABA and homotaurine in environments such as seawater has not been studied so it remains unclear which of these compounds is the primary substrate of the enzymes capable of their catabolism in the natural environment.

Multicellular eukaryotes, such as zooplankton, have traditionally been considered as the major producers of sulfonated compounds, particularly taurine, in oceans (Webb and Johannes, 1967). Taurine likely acts as an osmolyte in these organisms (Yancey, 2005). Recent metaproteomics studies of ocean surface water have consistently uncovered a high abundance of ABC transporters involved in taurine uptake (Sowell *et al.*, 2009, 2011; Williams *et al.*, 2012). This suggests a high and consistent flux of taurine, or structurally similar compounds, in the surface ocean: it is unclear whether production by metazoans is sufficient to satisfy this implied demand. Genes required for the production of taurine are present in many *Synechococcus* strains, and also some strains of marine heterotrophic bacteria (Agnello *et al.*, 2013), suggesting that bacteria may represent an additional source of taurine in the surface ocean. Sulfonated amino acid compounds appear to play a major but poorly understood role in the marine carbon, nitrogen and sulfur cycles. It will be interesting to investigate how homotaurine production by marine Roseobacters fits in to this broader picture.

5.3.3 Conclusions

This study has demonstrated that a newly described class of lipids, the homotaurolipids, are a major component of the membranes of a number of Roseobacter strains. I was unable to identify the enzymes involved in the production of this lipid but comparative genomics of HTL-producing strains has identified several candidate acyltransferases which could be the subject of future efforts to identify genes involved in the production of this lipid.

Chapter 6 Conclusions and future perspectives

6.1 Lipid remodelling in response to phosphorus scarcity

This work, combined with that of several others (Carini *et al.*, 2015; Sebastián *et al.*, 2016), has demonstrated that lipid remodelling is an important mechanism by which marine heterotrophic bacteria respond to P scarcity. While such lipid remodelling has been known about for more than 40 years (Minnikin and Abdolrahimzadeh, 1974; Minnikin *et al.*, 1974) and has been described in a wide range of organisms, from bacteria to plants and algae (Tjellström *et al.*, 2008; Van Mooy *et al.*, 2009), little was known about how widely distributed it was amongst bacteria in natural environments. In fact, a previous study had claimed that marine heterotrophic bacteria were likely unable to carry out such remodelling (Van Mooy *et al.*, 2009). This confusion may partly be explained by the patchy distribution of the genes involved – closely related SAR11 strains differ in their ability to remodel their lipids, for example (Carini *et al.*, 2015). The distribution of the lipid remodelling capability, as denoted by the presence of PlcP, indicates that horizontal gene transfer between bacteria adapted to P-depleted marine environments has been a major factor in its proliferation, as seems to be typical for genes related to dealing with P scarcity (Coleman and Chisholm, 2010). It will be interesting to determine whether a similar pattern emerges in other environments, such as fresh water and soil, which are also frequently P-depleted. There is reason to think that this may be the case, since PlcP was first characterised in a soil bacterium, *E. meliloti* (Zavaleta-Pastor *et al.*, 2010).

In this work, I did not attempt to study the enzymology of PlcP, though this will be an important topic for future research. In their original paper describing PlcP in *E. meliloti,* Zavaleta-Pastor *et al.* (Zavaleta-Pastor *et al.*, 2010) reported that PlcP was active against the zwitterionic phospholipids PE and PC but not PG. This is in line with my observations that PE in *Phaeobacter* sp. MED193 was much more greatly reduced in P-starved cells than PG (Chapter 3). Both *E. meliloti* and *Phaeobacter* sp.

MED193 produce DGTS as the major P-free lipid when P is scarce (Geiger *et al.*, 1999; Sebastián *et al.*, 2016), which presumably enables the cells to maintain the overall charge balance of their membranes since both it and PE are zwitterionic. Interestingly, the equivalent gene adjacent to PlcP in SAR11 and many other marine bacteria is a homolog of the bifunctional glycosyltransferase Agt from *Agrobacterium tumefaciens*, which can synthesise both the neutral MGDG and the negatively charged GADG (Semeniuk *et al.*, 2014). This raises the question of whether PlcP from these organisms is also able to degrade PG, or whether they possess some alternative means to degrade PG.

6.1.1 Implications for ocean biogeochemistry

This finding of an extensive capacity for lipid remodelling in certain regions of the ocean potentially has important implications for the biogeochemistry of these regions. The plasticity of P in phytoplankton relative to C and N has been shown to have an important effect on atmospheric CO_2 as predicted by oceanic box models (Galbraith and Martiny, 2015). Models of heterotrophic bacteria have often been assumed them to have relatively inflexible stoichiometry relative to phytoplankton (Wang *et al.*, 2012), studies of lake heterotrophic bacterial communities have shown variations in cellular C:P over three orders of magnitude (Godwin and Cotner, 2015). In light of the accumulating evidence that the stoichiometric flexibility of heterotrophic bacteria has been underestimated, the potential biogeochemical implications of this should be investigated. The inclusion of heterotrophic bacteria in such large-scale ocean-atmosphere model is, however, still in its infancy (Follows and Dutkiewicz, 2011).

It remains unclear how significant a saving of cellular P lipid remodelling represents. Around 20% of cellular P in bacteria is typically bound up within the membrane (Karl and Bjorkman, 2014). Since cells often scale back rRNA production when P is scarce (Elser *et al.*, 2003; Makino *et al.*, 2003), without a means to replace phospholipids it

seems likely that the membrane would consume an even larger proportion of cellular resources when P is scarce. The potential for a substantial reduction in the cellular P quota should thus provide bacteria capable of lipid remodelling a competitive advantage when P is limiting (Thingstad and Rassoulzadegan, 1999). So far, however, there has been no quantitative assessment of the amount of P that can be spared by lipid remodelling. Indeed, there are remarkably few estimates of lipid P as a proportion of total cellular P. Estimates of the amount of P that can be spared in this way will be important if the kind of stoichiometric flexibility implied by lipid remodelling is to be incorporated into ecological models.

Bacteria that are capable of lipid remodelling often produce substantial quantities of phospholipids when P-replete (Geiger *et al.*, 1999; Carini *et al.*, 2015; Sebastián *et al.*, 2016). This raises the question of why these bacteria do not simply retain membranes comprised primarily of P-free lipids. One suggestion, based on work done on lipid remodelling in plants, is that phospholipids might act as a storage compound for phosphorus, to be mobilised when P becomes scarce (Tjellström *et al.*, 2008). This was based on the observation that lipid remodelling in *Arabidopsis thaliana* was reversible, with phospholipids rapidly replacing the substitute glycolipids once the P supply was restored. Similar results have also been observed in experiments with diatoms (Martin *et al.*, 2011) although such remodelling has yet to be reported in bacteria. An alternative, or complementary, explanation for these observations could be that a higher proportion of membrane phospholipids is advantageous *per se*. Strains closely related to those which can remodel in response to P-scarcity but which lack PlcP themselves seem to constitutively produce large amounts of phospholipids (Chapter 4; Carini *et al.*, 2015). Moreover, phospholipids, particularly PE and PG, appear to be more widely distributed through Bacteria than any other major class of membrane lipid (Sohlenkamp and Geiger, 2015). Membrane proteins have evolved alongside their lipid partners (Lee, 2011) and through much of evolutionary history, it appears that the lipid portion of the bacterial membrane has included a substantial fraction of phospholipids. Therefore, it is possible that there is a fitness trade off

involved in replacing phospholipids with P-free lipids. From studies in *E. coli* we know that mutants unable to produce a class of phospholipids are viable but exhibit severe phenotypes. Lack of PE, for example, confers a requirement for divalent cations (DeChavigny *et al.*, 1991) and results in the mis-assembly of some proteins such as LacY, which is inverted relative to its conformation in wild type cells (Dowhan and Bogdanov, 2009). There are, however, a number of key differences between cells that have remodelled their lipids in response to P stress and mutants unable to synthesise a phospholipid. Firstly, most bacteria possessing PlcP appear to produce one or more substitute P-free lipids, such as DGTS or MGDG, which can presumably serve as a partial substitute for the phospholipids it replaces. Furthermore, in the strains studied so far, phospholipids are not completely removed after lipid remodelling (See for example Chapter 3; Carini *et al.*, 2015). Thus, there may still be sufficient amounts of these phospholipids remaining to engage in highly specific interactions with membrane proteins (Contreras *et al.*, 2011). These differences explain why the phenotype of cells which have undergone lipid remodelling is much less severe than for mutants entirely lacking a lipid class. However, quantitative differences in the rates of cellular processes, such as nutrient uptake or of respiration, could nonetheless have important implications for the biogeochemistry of regions in the ocean where P is scarce and warrants further investigation.

6.2 Aminolipids in marine bacteria

In addition to describing the lipid remodelling response to P scarcity in marine bacteria, this work also investigated the role of a particular group of P-free lipids, the aminolipids. I described the genes involved in the biosynthesis of glutamine lipid (Chapter 4) and also identified a novel, putative homotaurine-containing lipid that appears to be prevalent in Roseobacters (Chapter 5). The study of these lipids was initially motivated by the possibility that they might be involved in the response to P depletion. However, I found no strong evidence that these lipids are involved in the response to P scarcity in the Roseobacters despite studies documenting the role of OL

as a substitute for phospholipids in other organisms (Geiger *et al.*, 1999; Carini *et al.*, 2015). Thus, although aminolipids, or at least OL, may represent an adaptation to P scarcity in some bacteria, this does not appear to be universal. In the Roseobacter strains I studied, aminolipids were a major component of the membrane under all conditions studied. This begs the question of what role these lipids play in the biology of Roseobacters and other microbial groups in which they are present.

A previous finding that OLs are required for the maintenance of optimal levels of c-type cytochromes in *Rhodobacter capsulatus* (Aygun-Sunar *et al.*, 2006) hinted that these lipids may be involved in the core biology of the *Rhodobacteraceae*. I found some evidence to support this in that an *R. pomeroyi* mutant lacking OlsA, and hence unable to synthesise OL or QL, exhibited growth defects, particularly in media with lower concentrations of phosphate. However, there is currently little understanding as to why aminolipids are seemingly so important to the biology of *Rhodobacteraceae* while other organisms have evolved to be completely independent of them. Bacteria such as the SAR11 strain HTCC1062 are able to successfully compete in marine surface waters while bearing a membrane comprised of only two polar lipids: PE and PG (Van Mooy *et al.*, 2009). This is a strict subset of the lipids produced by Roseobacters such as *R. pomeroyi*, so what benefit do these bacteria get from their more complex membrane? It seems likely that the greater number of lipid classes facilitates more complex membrane functions in some way, although what this might mean and how it might be achieved is not clear. In eukaryotes, sphingolipid-enriched lipid rafts are thought to help enable lateral spatial organisation of the membrane by concentrating functionally related proteins into lipid patches (Lingwood and Simons, 2010). There is also some evidence for the existence of structures similar to lipid rafts in bacteria (Bramkamp and Lopez, 2015) although it remains unclear to what extent these are similar to those described in eukaryotes. Lateral membrane organisation can also arise through other means, for example cardiolipin has been proposed to spontaneously localise to regions of high negative membrane curvature (Renner and Weibel, 2011) such as the septa of dividing rod-shaped bacteria, which could in turn

help to localise the cell division apparatus. Almost nothing is currently known about the physicochemical properties of aminolipids so it is difficult to predict their behaviour within the bacterial membrane. Such studies will be required in order to gain a better understanding of how these lipids might behave *in vivo*.

My characterisation of the biosynthetic pathway for QL made it only the second aminolipid for which a synthetic pathway has been described, after OL (Weissenmayer *et al.*, 2002; Gao *et al.*, 2004). This is despite the substantial diversity of lipids with similar structures that have been detected in bacteria (Geiger *et al.*, 2010). One possibility is that some homologs of OlsB/OlsF exhibit different substrate specificity, as I have described for OlsB2 (Chapter 4). There may also be unrelated pathways for aminolipid synthesis, however, as appears to be the case for the HTL I detected in several Roseobacters (Chapter 5). Mutants of *R. pomeroyi* deficient in OlsB2 or OlsA were unaffected in their ability to produce HTL, although it would be informative to test whether a mutant unable to produce OlsB was still able to produce HTL. However, these results indicate that HTL synthesis is likely to proceed by a different pathway to the synthesis of OL or QL, despite the similarities in structure. I identified some candidate genes which may be involved in this process, and it would be interesting to investigate whether these candidates do indeed play a role in HTL synthesis. Identification of the genes involved in producing these unusual lipids would facilitate studies of their role in the Roseobacters, such as whether they may be sensed by algal partners, as is the case for some *Bacteroidetes* sulfonolipids (Alegado *et al.*, 2012). Given the wealth of genomic data now available, once the genes involved in producing these sulfonolipids are known it will be possible to gain a much better understanding of how widespread the ability to produce these lipids is within bacteria.

6.3 Summing up

This work has highlighted the importance of lipid remodelling as a response to P scarcity in marine bacteria. A diverse array of phosphorus-free lipids are used to

substitute for phospholipids and the consequences of this substitution will be an important area for further work. Not all phosphorus-free lipids are primarily deployed as substitutes, however. Studies of a number of aminolipids that are present in the Roseobacters group indicated that these did not seem to be involved in the response to P-limitation. Characterisation of the enzymes involved in glutamine lipid synthesis, only the second aminolipid for which a biosynthetic pathway has been described, revealed that the ability to produce this lipid is highly conserved within the *Rhodobacteraceae*. A further class of structurally related aminolipids with a novel homotaurine-containing head group was also described in a number of Roseobacters. Comparative genomics of HTL-producing strains has yielded candidates for future investigations into the synthesis of this novel class of lipids. Taken together, these results provide fresh insights into the importance of lipid diversity for bacterial ecology. However, an understanding of this topic is still in its infancy and many fundamental questions remain to be answered.

Chapter 7 References

Abascal, F., Zardoya, R., and Posada, D. (2005) ProtTest: selection of best-fit models of protein evolution. *Bioinformatics* **21**: 2104–2105.

Abbanat, D.R., Leadbetter, E.R., Godchaux, W., and Escher, A. (1986) Sulphonolipids are molecular determinants of gliding motility. *Nature* **324**: 367–369.

Abida, H., Dolch, L.-J., Meï, C., Villanova, V., Conte, M., Block, M.A., et al. (2015) Membrane Glycerolipid Remodeling Triggered by Nitrogen and Phosphorus Starvation in *Phaeodactylum tricornutum*. *Plant Physiol.* **167**: 118–36.

Agnello, G., Chang, L.L., Lamb, C.M., Georgiou, G., and Stone, E.M. (2013) Discovery of a substrate selectivity motif in amino acid decarboxylases unveils a taurine biosynthesis pathway in prokaryotes. *ACS Chem. Biol.* **8**: 2264–2271.

Aisen, P.S., Gauthier, S., Ferris, S.H., Saumier, D., Haine, D., Garceau, D., et al. (2011) Tramiprosate in mild-to-moderate Alzheimer's disease - A randomized, double-blind, placebo-controlled, multi-centre study (the alphase study). *Arch. Med. Sci.* **7**: 102–111.

Akaike, H. (1974) A New Look at the Statistical Model Identification. *IEEE Trans. Automat. Contr.* **19**: 716–723.

Alegado, R.A., Brown, L.W., Cao, S., Dermenjian, R.K., Zuzow, R., Fairclough, S.R., et al. (2012) A bacterial sulfonolipid triggers multicellular development in the closest living relatives of animals. *Elife* **1**: e00013.

Alonso, C., Warnecke, F., Amann, R., and Pernthaler, J. (2007) High local and global diversity of *Flavobacteria* in marine plankton. *Environ. Microbiol.* **9**: 1253–1266.

Ames, G.F. (1968) Lipids of *Salmonella typhimurium* and *Escherichia coli*: stucture and metabolism. *J. Bacteriol.* **95**: 833–843.

Amin, S.A., Hmelo, L.R., van Tol, H.M., Durham, B.P., Carlson, L.T., Heal, K.R., et al. (2015) Interaction and signalling between a cosmopolitan phytoplankton and associated bacteria. *Nature* **522**: 98-101.

Armbrust, E.V. (2009) The life of diatoms in the world's oceans. *Nature* **459**: 185–192.

Arrigo, K. (2004) Marine microorganisms and global nutrient cycles. *Nature* **437**: 349–356.

Arrigo, K.R. (1999) Phytoplankton Community Structure and the Drawdown of Nutrients and CO_2 in the Southern Ocean. *Science* **283**: 365–367.

Aygun-Sunar, S., Bilaloglu, R., Goldfine, H., and Daldal, F. (2007) *Rhodobacter capsulatus* OlsA is a bifunctional enyzme active in both ornithine lipid and phosphatidic acid biosynthesis. *J. Bacteriol.* **189**: 8564–8574.

Aygun-Sunar, S., Mandaci, S., Koch, H.G., Murray, I.V.J., Goldfine, H., and Daldal, F. (2006) Ornithine lipid is required for optimal steady-state amounts of *c*-type cytochromes in *Rhodobacter capsulatus*. *Mol. Microbiol.* **61**: 418–435.

Azam, F. (1998) Microbial control of the oceanic carbon flux: The plot thickens. *Science* **280**: 694–696.

Azam, F., Fenchel, T., and Field, J. (1983) The Ecological Role of Water-Column Microbes in the Sea. *Mar. Ecol. Prog. Ser.* **10**: 257–263.

Azam, F., Smith, D.C., Steward, G.F., and Hagström, Å. (1994) Bacteria-organic matter coupling and its significance for oceanic carbon cycling. *Microb. Ecol.* **28**: 167–179.

De Baar, H.J.W. (1994) von Liebig's Law of the Minimum and Plankton Ecology. *Prog. Oceanogr.* **33**: 347–386.

Batrakov, S.G., Nikitin, D.I., Mosezhnyi, A.E., and Ruzhitsky, A.O. (1999) A glycine-containing phosphorus-free lipoaminoacid from the gram-negative marine bacterium *Cyclobacterium marinus* WH. *Chem. Phys. Lipids* **99**: 139–143.

Batrakov, S.G., Nikitin, D.I., Sheichenko, V.I., and Ruzhitsky, a O. (1998) A novel sulfonic-acid analogue of ceramide is the major extractable lipid of the gram-negative marine bacterium *Cyclobacterium marinus* WH. *Biochim. Biophys. Acta* **1391**: 79–91.

Behrenfeld, M.J. and Boss, E.S. (2014) Resurrecting the ecological underpinnings of ocean plankton blooms. *Ann. Rev. Mar. Sci.* **6**: 167–94.

Behrenfeld, M.J. and Falkowski, P.G. (1997) Photosynthetic rates derived from satellite-based chlorophyll concentration. *Limnol. Oceanogr.* **42**: 1–20.

Behrenfeld, M.J., O'Malley, R.T., Siegel, D.A., McClain, C.R., Sarmiento, J.L., Feldman, G.C., et al. (2006) Climate-driven trends in contemporary ocean productivity. *Nature* **444**: 752–5.

Behringer, D.C. and Butler IV, M.J. (2006) Stable isotope analysis of production and trophic

relationships in a tropical marine hard-bottom community. *Oecologia* **148**: 334–341.

Bélanger, L., Dimmick, K.A., Fleming, J.S., and Charles, T.C. (2009) Null mutations in *Sinorhizobium meliloti exoS* and *chvI* demonstrate the importance of this two-component regulatory system for symbiosis. *Mol. Microbiol.* **74**: 1223–1237.

Benning, C. (1998) Biosynthesis and Function of the Sulfolipid Sulfoquinovosyl Diacylglycerol. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **49**: 53–75.

Benning, C., Huang, Z., and Gage, D. (1995) Accumulation of novel glycolipids and a betaine lipid in cells of *Rhodobacter sphaeroides* grown under phosphate limitation. *Arch. Biochem. Biophys.* **317**: 103–111.

Billerbeck, S., Wemheuer, B., Voget, S., Poehlein, A., Giebel, H.-A., Brinkhoff, T., et al. (2016) Biogeography and environmental genomics of CHAB-I-5, a pelagic lineage of the marine Roseobacter clade. *Nat. Microbiol.* **1**: 16063.

Bligh, E.G. and Dyer, W.J. (1959) A rapid method of total lipid extraction and purification. *Can. J. Biochem. Physiol.* **37**: 911–917.

Brady, S.F., Chao, C.J., and Clardy, J. (2004) Long-Chain *N*-Acyltyrosine Synthases from Environmental DNA. *Appl. Environ. Microbiol.* **70**: 6865–6870.

Bramkamp, M. and Lopez, D. (2015) Exploring the Existence of Lipid Rafts in Bacteria. *Microbiol. Mol. Biol. Rev.* **79**: 81–100.

Brandsma, J., Hopmans, E.C., Philippart, C.J.M., Veldhuis, M.J.W., Schouten, S., and Sinninghe Damsté, J.S. (2012) Low temporal variation in the intact polar lipid composition of North Sea coastal marine water reveals limited chemotaxonomic value. *Biogeosciences* **9**: 1073–1084.

Breitbart, M. (2012) Marine Viruses: Truth or Dare. *Ann. Rev. Mar. Sci.* **4**: 425–448.

Brouwers, J.F. (2011) Liquid chromatographic-mass spectrometric analysis of phospholipids. Chromatography, ionization and quantification. *Biochim. Biophys. Acta* **1811**: 763–75.

Brügger, B., Erben, G., Sandhoff, R., Wieland, F.T., and Lehmann, W.D. (1997) Quantitative analysis of biological membrane lipids at the low picomole level by nano-electrospray ionization tandem mass spectrometry. *Proc. Natl. Acad. Sci. U. S. A.* **94**: 2339–2344.

Buchan, A., González, J., and Moran, M. (2005) Overview of the Marine Roseobacter Lineage. *Appl. Environ. Microbiol.* **71**: 5665–5677.

Buchan, A., LeCleir, G.R., Gulvik, C.A., and González, J.M. (2014) Master recyclers: features and functions of bacteria associated with phytoplankton blooms. *Nat. Rev. Microbiol.* **12**: 686–698.

Cao, Y., Adachi, J., Janke, A., Paabo, S., and Hasegawa, M. (1994) Phylogenetic-Relationships among Eutherian Orders Estimated from Inferred Sequences of Mitochondrial Proteins - Instability of a Tree-Based on a Single-Gene. *J. Mol. Evol.* **39**: 519–527.

Capella-Gutiérrez, S., Silla-Martínez, J.M., and Gabaldón, T. (2009) trimAl: A tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* **25**: 1972–1973.

Carini, P., Campbell, E.O., Morré, J., Sañudo-Wilhelmy, S.A., Cameron Thrash, J., Bennett, S.E., et al. (2014) Discovery of a SAR11 growth requirement for thiamin's pyrimidine precursor and its distribution in the Sargasso Sea. *ISME J.* **8**: 1727-1738.

Carini, P., Mooy, B.A.S. Van, Thrash, J.C., White, A., Zhao, Y., and Campbell, E.O. (2015) SAR11 lipid renovation in response to phosphate starvation. *Proc. Natl. Acad. Sci.* **112:** 7767-7772

Carini, P., White, A.E., Campbell, E.O., and Giovannoni, S.J. (2014) Methane production by phosphate-starved SAR11 chemoheterotrophic marine bacteria. *Nat. Commun.* **5**: 4346.

Carlson, C.A. and Hansell, D.A. (2015) DOM Sources, Sinks, Reactivity and Budgets. In, Carlson,C.A. and Hansell,D.A. (eds), *Biogeochemistry of Marine Dissolved Organic Matter*. Academic Press, pp. 65–126.

Carlson, C.A., Morris, R., Parsons, R., Treusch, A.H., Giovannoni, S.J., and Vergin, K. (2009) Seasonal dynamics of SAR11 populations in the euphotic and mesopelagic zones of the northwestern Sargasso Sea. *ISME J.* **3**: 283–295.

Chappell, P.D., Moffett, J.W., Hynes, A.M., and Webb, E.A. (2012) Molecular evidence of iron limitation and availability in the global diazotroph *Trichodesmium*. *ISME J.* **6**: 1728–39.

Cheng, C. and Gross, M. (2000) Applications and mechanisms of charge-remote fragmentation. *Mass Spectrom. Rev.* **19**: 398–420.

Christie, W.W. and Han, X. (2010) Lipid Analysis Fourth Edition. Woodhead Publishing, Cambridge, UK.

Church, M.J. (2008) Resource Control of Bacterial Dynamics in the Sea. In, Kirchman,D.L. (ed), *Microbial Ecology of the Oceans*. John Wiley and Sons, Inc., Hoboken, New Jersey, pp. 335–382.

Cífková, E., Holčapek, M., Lísa, M., Ovčačíková, M., Lyčka, A., Lynen, F., and Sandra, P. (2012) Nontargeted quantitation of lipid classes using hydrophilic interaction liquid chromatography-electrospray ionization mass spectrometry with single internal standard and response factor approach. *Anal. Chem.* **84**: 10064–70.

Clark, L.L., Ingall, E.D., and Benner, R. (1998) Marine phosphorus is selectively remineralized. *Nature* **393**: 426.

Coleman, M.L. and Chisholm, S.W. (2010) Ecosystem-specific selection pressures revealed through comparative population genomics. *Proc. Natl. Acad. Sci.* **107**: 18634–18639.

Comerci, D.J., Altabe, S., De Mendoza, D., and Ugalde, R.A. (2006) *Brucella abortus* synthesizes phosphatidylcholine from choline provided by the host. *J. Bacteriol.* **188**: 1929–1934.

Constantopoulos, T.L., Jackson, G.S., and Enke, C.G. (2000) Challenges in achieving a fundamental model for ESI. *Anal. Chim. Acta* **406**: 37–52.

Contreras, F., Ernst, A.M., and Wieland, F. (2011) Specificity of Intramembrane Protein – Lipid Interactions. *Cold Spring Harb. Perspect. Biol.* **3**: a004705.

Cotner, J.B. and Biddanda, B.A. (2002) Small players, large role: Microbial influence on biogeochemical processes in pelagic aquatic ecosystems. *Ecosystems* **5**: 105–121.

Cotner, J.B., Hall, E.K., Scott, J.T., and Heldal, M. (2010) Freshwater Bacteria are Stoichiometrically Flexible with a Nutrient Composition Similar to Seston. *Front. Microbiol.* **1**: 132.

Cottrell, M.T. and Kirchman, D.L. (2000) Community Composition of Marine Bacterioplankton Determined by 16S rRNA Gene Clone Libraries and Fluorescence In Situ Hybridization. *Appl. Environ. Microbiol.* **66**: 5116–5122.

Cottrell, M.T. and Kirchman, D.L. (2000) Natural assemblages of marine proteobacteria adn members of the *Cytophaga-Flavobacter* clustler consuming low- and high-molecular-weight dissolved organic matter. *Appl. Environ. Microbiol.* **66**: 1692–1697.

Cox, A.D. and Saito, M.A. (2013) Proteomic responses of oceanic *Synechococcus* WH8102 to phosphate and zinc scarcity and cadmium additions. *Front. Microbiol.* **4**: 387.

Cubbon, Si., Antonio, C., Wilson, J., and Thomas-Oates, J. (2010) Metabolomic applications of HILIC-LC-MS. *Mass Spectrom. Rev.* **29**: 671–684.

DeChavigny, A., Heacock, P.N., and Dowhan, W. (1991) Sequence and inactivation of the pss gene of *Escherichia coli*: Phosphatidylethanolamine may not be essential for cell viability. *J. Biol. Chem.* **266**: 5323–5332.

Dees, C. and Shively, J.M. (1982) Localization and quantitation of the ornithine lipid of *Thiobacillus thiooxidans*. *J. Bacteriol.* **149**: 798–799.

Dembitsky, V. (1996) Betaine ether-linked glycerolipids: Chemistry and biology. *Prog. Lipid Res.* **35**: 1–51.

Dennis, J.J. and Zylstra, G.J. (1998) Plasposons: modular self-cloning minitransposon derivatives for rapid genetic analysis of gram-negative bacterial genomes. *Appl. Environ. Microbiol.* **64**: 2710–5.

Deutsch, C. and Weber, T. (2012) Nutrient Ratios as a Tracer and Driver of Ocean Biogeochemistry. *Ann. Rev. Mar. Sci.* **4**: 113–141.

Devers, E.A., Wewer, V., Dombrink, I., Dörmann, P., and Hölzl, G. (2011) A processive glycosyltransferase involved in glycolipid synthesis during phosphate deprivation in *Mesorhizobium loti*. *J. Bacteriol.* **193**: 1377–84.

Diercks, H., Semeniuk, A., Gisch, N., Moll, H., Duda, K. a, and Hölzl, G. (2015) Accumulation of Novel Glycolipids and Ornithine Lipids in *Mesorhizobium loti* under Phosphate Deprivation. *J. Bacteriol.* **197**: 497–509.

Dittmar, T. (2015) Reasons Behind the Long-Term Stability of Dissolved Organic Matter. In, Carlson,C.A. and Hansell,D.A. (eds), *Biogeochemistry of Marine Dissolved Organic Matter*. Academic Press, pp. 369–388.

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

Dowhan, W. and Bogdanov, M. (2009) Lipid-dependent membrane protein topogenesis. *Annu. Rev. Biochem.* **78**: 515–40.

Droop, M.R. (1974) The nutrient status of algal cells in continuous culture. *J. Mar. Biol.*

Assoc. United Kingdom **9**: 825–855.

Dudkiewicz, M., Szczepinska, T., Grynberg, M., and Pawlowski, K. (2012) A novel protein kinase-like domain in a selenoprotein, widespread in the tree of life. *PLoS One* **7**: e32138.

Dunne, J.P., Sarmiento, J.L., and Gnanadesikan, A. (2007) A synthesis of global particle export from the surface ocean and cycling through the ocean interior and on the seafloor. *Global Biogeochem. Cycles* **21**: 1–16.

Dupont, C.L., Rusch, D.B., Yooseph, S., Lombardo, M.-J., Richter, R.A., Valas, R., et al. (2012) Genomic insights to SAR86, an abundant and uncultivated marine bacterial lineage. *ISME J.* **6**: 1186–99.

Durham, B.P., Sharma, S., Luo, H., Smith, C.B., Amin, S.A., Bender, S.J., et al. (2014) Cryptic carbon and sulfur cycling between surface ocean plankton. *Proc. Natl. Acad. Sci.* **112**: 453-457

Dyhrman, S.T., Benitez-nelson, C.R., Orchard, E.D., Haley, S.T., and Pellechia, P.J. (2009) A microbial source of phosphonates in oligotrophic marine systems. *Nat. Geosci.* **2**: 696–699.

Dyhrman, S.T., Chappell, P.D., Haley, S.T., Moffett, J.W., Orchard, E.D., Waterbury, J.B., and Webb, E.A. (2006) Phosphonate utilization by the globally important marine diazotroph *Trichodesmium*. *Nature* **439**: 68–71.

Eddy, S.R. (2011) Accelerated profile HMM searches. *PLoS Comput. Biol.* **7**: e1002195.

Eddy, S.R. (1996) Hidden Markov models. *Curr. Opin. Struct. Biol.* **6**: 361–365.

Eder, K. (1995) Gas chromatographic analysis of fatty acid methyl esters. *J. Chromatogr. B* **671**: 113–131.

Edgar, R.C. (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32**: 1792–7.

Ejsing, C.S., Sampaio, J.L., Surendranath, V., Duchoslav, E., Ekroos, K., Klemm, R.W., et al. (2009) Global analysis of the yeast lipidome by quantitative shotgun mass spectrometry. *Proc. Natl. Acad. Sci. U. S. A.* **106**: 2136–41.

Elser, J.J., Acharya, K., Kyle, M., Cotner, J., Makino, W., Markow, T., et al. (2003) Growth rate-stoichiometry couplings in diverse biota. *Ecol. Lett.* **6**: 936–943.

Elser, J.J., Dobberfubl, D.R., Mackay, N.A., and Schampel, J.H. (1995) Organism Size, Life History, and N:P Stoichiometry Toward a unified view of cellular and ecosystem processes. *Bioscience* **46**: 674–684.

Ertefai, T.F., Fisher, M.C., Fredricks, H.F., Lipp, J.S., Pearson, A., Birgel, D., et al. (2008) Vertical distribution of microbial lipids and functional genes in chemically distinct layers of a highly polluted meromictic lake. *Org. Geochem.* **39**: 1572–1588.

Fahy, E., Subramaniam, S., Brown, H.A., Glass, C.K., Merrill, A.H., Murphy, R.C., et al. (2005) A comprehensive classification system for lipids. *J. Lipid Res.* **46**: 839–61.

Falkowski, P.G., Barber, R.T., and Smetacek, V. (1998) Biogeochemical controls and feedbacks on ocean primary production. *Science* **281**: 200–206.

Falkowski, P.G., Fenchel, T., and Delong, E.F. (2008) The microbial engines that drive Earth's biogeochemical cycles. *Science* **320**: 1034–9.

Fanning, K.A. (1992) Nutrient provinces in the sea: Concentration ratios, reaction rate ratios, and ideal covariation. *J. Geophys. Res.* **97**: 5693.

Feingersch, R., Philosof, A., Mejuch, T., Glaser, F., Alalouf, O., Shoham, Y., and Béjà, O. (2012) Potential for phosphite and phosphonate utilization by *Prochlorococcus*. *ISME J.* **6**: 827–34.

Ferla, M.P. and Patrick, W.M. (2014) Bacterial methionine biosynthesis. *Microbiology* **160**: 1571–1584.

Fernández-Gómez, B., Richter, M., Schüler, M., Pinhassi, J., Acinas, S.G., González, J.M., and Pedrós-Alió, C. (2013) Ecology of marine *Bacteroidetes*: a comparative genomics approach. *ISME J.* **7**: 1026–37.

Field, C.B. (1998) Primary Production of the Biosphere: Integrating Terrestrial and Oceanic Components. *Science* **281**: 237–240.

Flombaum, P., Gallegos, J.L., Gordillo, R.A., Rincon, J., Zabala, L.L., Jiao, N., et al. (2013) Present and future global distributions of the marine *Prochlorococcus* and *Synechococcus*. *Proc. Natl. Acad. Sci.* **110**: 9824–9829.

Folch, J., Lees, M., and Stanley, G.H.S. (1957) A simple method for the isolation and purification of total lipides from animal tissues. *J. Biol. Chem.* **226**: 497–509.

Follows, M.J. and Dutkiewicz, S. (2011) Modeling diverse communities of marine microbes.

Ann. Rev. Mar. Sci. **3**: 427–451.

Follows, M.J., Dutkiewicz, S., Grant, S., and Chisholm, S.W. (2007) Emergent Biogeography of Microbial Communities in a Model Ocean. *Science* **315**: 1843–1846.

Freitas, S., Hatosy, S., Fuhrman, J. a, Huse, S.M., Welch, D.B.M., Sogin, M.L., and Martiny, A.C. (2012) Global distribution and diversity of marine *Verrucomicrobia*. *ISME J.* **6**: 1499–505.

Frentzen, M. (2004) Phosphatidylglycerol and sulfoquinovosyldiacylglycerol: anionic membrane lipids and phosphate regulation. *Curr. Opin. Plant Biol.* **7**: 270–6.

Fuchs, B., Süss, R., Teuber, K., Eibisch, M., and Schiller, J. (2011) Lipid analysis by thin-layer chromatography--a review of the current state. *J. Chromatogr. A* **1218**: 2754–74.

Fuchs, B.M., Spring, S., Teeling, H., Quast, C., Wulf, J., Schattenhofer, M., et al. (2007) Characterization of a marine gammaproteobacterium capable of aerobic anoxygenic photosynthesis. *Proc. Natl. Acad. Sci.* **104**: 2891–2896.

Fuhrman, J. a (2009) Microbial community structure and its functional implications. *Nature* **459**: 193–9.

Fuller, N. and Rand, R.P. (2001) The influence of lysolipids on the spontaneous curvature and bending elasticity of phospholipid membranes. *Biophys. J.* **81**: 243–54.

Galbraith, E.D. and Martiny, A.C. (2015) A simple nutrient-dependence mechanism for predicting the stoichiometry of marine ecosystems. *Proc. Natl. Acad. Sci.* **112**: 8199-8204.

Ganesh, S., Parris, D.J., DeLong, E.F., and Stewart, F.J. (2014) Metagenomic analysis of size-fractionated picoplankton in a marine oxygen minimum zone. *ISME J.* **8**: 187–211.

Ganong, B.R., Leonard, J.M., and Raetz, C.R. (1980) Phosphatidic acid accumulation in the membranes of *Escherichia coli* mutants defective in CDP-diglyceride synthetase. *J. Biol. Chem.* **255**: 1623–1629.

Gao, J.-L., Weissenmayer, B., Taylor, A.M., Thomas-Oates, J., López-Lara, I.M., and Geiger, O. (2004) Identification of a gene required for the formation of lyso-ornithine lipid, an intermediate in the biosynthesis of ornithine-containing lipids. *Mol. Microbiol.* **53**: 1757–70.

Geider, R. and La Roche, J. (2002) Redfield revisited: variability of C:N:P in marine

microalgae and its biochemical basis. *Eur. J. Phycol.* **37**: 1–17.

Geiger, O., González-Silva, N., López-Lara, I.M., and Sohlenkamp, C. (2010) Amino acid-containing membrane lipids in bacteria. *Prog. Lipid Res.* **49**: 46–60.

Geiger, O., Röhrs, V., Weissenmayer, B., Finan, T.M., and Thomas-Oates, J.E. (1999) The regulator gene phoB mediates phosphate stress-controlled synthesis of the membrane lipid diacylglyceryl-*N,N,N*-trimethylhomoserine in *Rhizobium* (*Sinorhizobium*) *meliloti*. *Mol. Microbiol.* **32**: 63–73.

Geske, T., Vom Dorp, K., Dörmann, P., and Hölzl, G. (2013) Accumulation of glycolipids and other non-phosphorous lipids in *Agrobacterium tumefaciens* grown under phosphate deprivation. *Glycobiology* **23**: 69–80.

Ghai, R., Mizuno, C.M., Picazo, A., Camacho, A., and Rodriguez-Valera, F. (2013) Metagenomics uncovers a new group of low GC and ultra-small marine *Actinobacteria*. *Sci. Rep.* **3**: 2471.

Ghiglione, J.-F., Galand, P.E., Pommier, T., Pedrós-Alió, C., Maas, E.W., Bakker, K., et al. (2012) Pole-to-pole biogeography of surface and deep marine bacterial communities. *Proc. Natl. Acad. Sci.* **109**: 17633–8.

Gibson, D.G., Young, L., Chuang, R.-Y., Venter, J.C., Hutchison, C. a, and Smith, H.O. (2009) Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* **6**: 343–345.

Gidden, J., Denson, J., Liyanage, R., Ivey, D.M., and Lay, J.O. (2009) Lipid compositions in *Escherichia coli* and *Bacillus subtilis* during growth as determined by MALDI-TOF and TOF/TOF mass spectrometry. *Int. J. Mass Spectrom.* **283**: 178–184.

Del Giorgio, P.A., Cole, J.J., and Cimbleris, A. (1997) Respiration rates in bacteria exceed phytoplankton production in unproductive aquatic systems. *Nature* **385**: 148–151.

Giovannoni, S.J., Cameron Thrash, J., and Temperton, B. (2014) Implications of streamlining theory for microbial ecology. *ISME J.* **8**: 1553–1565.

Giovannoni, S.J. and Stingl, U. (2005) Molecular diversity and ecology of microbial plankton. *Nature* **437**: 343–8.

Giovannoni, S.J., Tripp, H.J., Givan, S., Podar, M., Vergin, K.L., Baptista, D., et al. (2005) Genome streamlining in a cosmopolitan oceanic bacterium. *Science* **309**: 1242–5.

Giovannoni, S.J. and Vergin, K.L. (2012) Seasonality in ocean microbial communities. *Science* **335**: 671–6.

Godchaux, W. and Leadbetter, E.R. (1980) *Capnocytophaga* spp. contain sulfonolipids that are novel in procaryotes. *J. Bacteriol.* **144**: 592–602.

Godchaux, W. and Leadbetter, E.R. (1988) Sulfonolipids are localized in the outer membrane of the gliding bacterium *Cytophaga johnsonae*. *Arch. Microbiol.* **150**: 42–47.

Godchaux, W. and Leadbetter, E.R. (1984) Sulfonolipids of Gliding Bacteria. *J. Biol. Chem.* **259**: 2982–2990.

Godchaux, W. and Leadbetter, E.R. (1983) Unusual sulfonolipids are characteristic of the *Cytophaga-Flexibacter* group. *J. Bacteriol.* **153**: 1238–1246.

Godwin, C.M. and Cotner, J.B. (2015) Aquatic heterotrophic bacteria have highly flexible phosphorus content and biomass stoichiometry. *ISME J.* **9**: 2324-2327

Goldberg, D.E., Rumley, M.K., and Kennedy, E.P. (1981) Biosynthesis of membrane-derived oligosaccharides: a periplasmic phosphoglyceroltransferase. *Proc. Natl. Acad. Sci.* **78**: 5513–7.

De Gracia, D.G. and Jollès-Bergeret, B. (1973) Sulfinic and sulfonic analogs of γ-aminobutyric acid and succinate semialdehyde, new substrates for the amino-butyrate aminotransferase and the succinate semialdehyde dehydrogenase of *Pseudomonas fluorescens*. *Biochim. Biophys. Acta - Enzymol.* **315**: 49–60.

Grossart, H.P., Levold, F., Allgaier, M., Simon, M., and Brinkhoff, T. (2005) Marine diatom species harbour distinct bacterial communities. *Environ. Microbiol.* **7**: 860–873.

Guo, D. and Tropp, B.E. (2000) A second Escherichia coli protein with CL synthase activity. *Biochim. Biophys. Acta - Mol. Cell Biol. Lipids* **1483**: 263–274.

Hacker, S., Sohlenkamp, C., Aktas, M., Geiger, O., and Narberhaus, F. (2008) Multiple phospholipid *N*-methyltransferases with distinct substrate specificities are encoded in *Bradyrhizobium japonicum*. *J. Bacteriol.* **190**: 571–580.

Hall, B.G., Acar, H., Nandipati, A., and Barlow, M. (2014) Growth rates made easy. *Mol. Biol. Evol.* **31**: 232–238.

Han, X. and Gross, R.W. (2005) Shotgun lipidomics: electrospray ionization mass spectrometric analysis and quantitation of cellular lipidomes directly from crude

extracts of biological samples. *Mass Spectrom. Rev.* **24**: 367–412.

Han, X., Yang, J., Cheng, H., Ye, H., and Gross, R.W. (2004) Toward fingerprinting cellular lipidomes directly from biological samples by two-dimensional electrospray ionization mass spectrometry. *Anal. Biochem.* **330**: 317–31.

Han, X., Yang, K., Cheng, H., Fikes, K.N., and Gross, R.W. (2005) Shotgun lipidomics of phosphoethanolamine-containing lipids in biological samples after one-step *in situ* derivatization. *J. Lipid Res.* **46**: 1548–60.

Han, X., Yang, K., Yang, J., Fikes, K.N., Cheng, H., and Gross, R.W. (2006) Factors influencing the electrospray intrasource separation and selective ionization of glycerophospholipids. *J. Am. Soc. Mass Spectrom.* **17**: 264–74.

Hansell, D.A., Carlson, C.A., Repeta, D.J., and Schlitzer, R. (2009) Dissolved Organic Matter in the Ocean A Controversy Stimulates New Insights. *Oceanography* **22**: 202–211.

Havskum, H., Thingstad, T.F., Scharek, R., Peters, F., Berdalet, E., Sala, M.M., et al. (2003) Silicate and labile DOC interfere in structuring the microbial food web via algal-bacterial competition for mineral nutrients: Results of a mesocosm experiment. *Limnol. Oceanogr.* **48**: 129–140.

Hawrot, E. and Kennedy, E.P. (1978) Phospholipid composition and membrane function in phosphatidylserine decarboxylase mutants of *Escherichia coli*. *J. Biol. Chem.* **253**: 8213–8220.

Heath, R.J., Jackowski, S., and Rock, C.O. (2002) Fatty acid and phospholipid metabolism in prokaryotes. In, Vance,D.E. and Vance,J.E. (eds), *Biochemistry of Lipids, Lipoproteins and Membranes.*, pp. 55–92.

Ver Hoef, J.M. and Boveng, P.L. (2007) Quasi-poisson vs. negative binomial regression: How should we model overdispersed count data? *Ecology* **88**: 2766–2772.

Hölzl, G. and Dörmann, P. (2007) Structure and function of glycoglycerolipids in plants and bacteria. *Prog. Lipid Res.* **46**: 225–43.

Hölzl, G., Leipelt, M., Ott, C., Zähringer, U., Lindner, B., Warnecke, D., and Heinz, E. (2005) Processive lipid galactosyl/glucosyltransferases from *Agrobacterium tumefaciens* and *Mesorhizobium loti* display multiple specificities. *Glycobiology* **15**: 874–886.

Hoppe, H. and Ullrich, S. (1999) Profiles of ectoenzymes in the Indian Ocean: phenomena of phosphatase activity in the mesopelagic zone. *Aquat. Microb. Ecol.* **19**: 139–148.

Howard, E.C., Henriksen, J.R., Buchan, A., Reisch, C.R., Bürgmann, H., Welsh, R., et al. (2006) Bacterial taxa that limit sulfur flux from the ocean. *Science* **314**: 649–52.

Howard, E.C., Sun, S., Biers, E.J., and Moran, M.A. (2008) Abundant and diverse bacteria involved in DMSP degradation in marine surface waters. *Environ. Microbiol.* **10**: 2397–410.

Hsu, F.F. and Turk, J. (2000) Characterization of phosphatidylethanolamine as a lithiated adduct by triple quadrupole tandem mass spectrometry with electrospray ionization. *J. Mass Spectrom.* **35**: 595–606.

Huerta-Cepas, J., Forslund, K., Szklarczyk, D., Jensen, L.J., von Mering, C., and Bork, P. (2016) Fast genome-wide functional annotation through orthology assignment by eggNOG-mapper. *Mol. Biol. Evol.* doi: 10.1093/molbev/msx148

Huerta-Cepas, J., Serra, F., and Bork, P. (2016) ETE 3: Reconstruction, Analysis, and Visualization of Phylogenomic Data. *Mol. Biol. Evol.* **33**: 1635–1638.

Huerta-Cepas, J., Szklarczyk, D., Forslund, K., Cook, H., Heller, D., Walter, M.C., et al. (2016) EGGNOG 4.5: A hierarchical orthology framework with improved functional annotations for eukaryotic, prokaryotic and viral sequences. *Nucleic Acids Res.* **44**: D286–D293.

Icho, T. and Raetz, C.R.H. (1983) Multiple genes for membrane-bound phosphatases in *Escherichia coli* and their action on phospholipid precursors. *J. Bacteriol.* **153**: 722–730.

Ito, K., Miyazawa, K., and Matsumoto, F. (1977) Amino acid composition of the ethanolic extractives from 31 species of marine red algae. *Hiroshima Daigaku Sui-Chikusangakubu Kiyo* **16**: 77–90.

Iverson, S.J., Lang, S.L., and Cooper, M.H. (2001) Comparison of the Bligh and Dyer and Folch methods for total lipid determination in a broad range of marine tissue. *Lipids* **36**: 1283–7.

Jaeschke, A., Rooks, C., Trimmer, M., Nicholls, J.C., Hopmans, E.C., Schouten, S., and Sinninghe Damsté, J.S. (2009) Comparison of ladderane phospholipid and core lipids as indicators for anaerobic ammonium oxidation (anammox) in marine sediments. *Geochim. Cosmochim. Acta* **73**: 2077–2088.

Jensen, P.R. and Lauro, F.M. (2008) An assessment of actinobacterial diversity in the marine environment. *Antonie Van Leeuwenhoek* **94**: 51–62.

Jiao, N., Herndl, G.J., Hansell, D. a, Benner, R., Kattner, G., Wilhelm, S.W., et al. (2010) Microbial production of recalcitrant dissolved organic matter: long-term carbon storage in the global ocean. *Nat. Rev. Microbiol.* **8**: 593–9.

Jiao, N. and Zheng, Q. (2011) The microbial carbon pump: from genes to ecosystems. *Appl. Environ. Microbiol.* **77**: 7439–44.

Jickells, T.D., An, Z.S., Andersen, K.K., Baker, A.R., Bergametti, G., Brooks, N., et al. (2005) Global Iron Connections Between Desert Dust, Ocean Biogeochemistry, and Climate. *Science* **308**: 67–72.

Johnson, Z.I., Zinser, E.R., Coe, A., McNulty, N.P., Woodward, E.M.S., and Chisholm, S.W. (2006) Niche partitioning among *Prochlorococcus* ecotypes along ocean-scale environmental gradients. *Science* **311**: 1737–40.

Kabisch, A., Otto, A., König, S., Becher, D., Albrecht, D., Schüler, M., et al. (2014) Functional characterization of polysaccharide utilization loci in the marine *Bacteroidetes* "*Gramella forsetii*" KT0803. *ISME J.* **8**: 1492-1502

Kanfer, J. and Kennedy, E.P. (1964) Metabolism and Function of Bacterial Lipids II. Biosynthesis of Phospholipids in *Escherichia coli*. *J. Biol. Chem.* **239**: 1720–1726.

Karl, D.M. and Bjorkman, K.M. (2014) Dynamics of Dissolved Organic Phosphorus. In, Hansell,D.A. and Carlson,C.A. (eds), *Biogeochemistry of Marine Dissolved Organic Matter*. Academic Press, pp. 233–334.

Kato, M., Sakai, M., Adachi, K., Ikemoto, H., and Sano, H. (1996) Distribution of betaine lipids in marine algae. *Phytochemistry* **42**: 1341–1345.

Katoh, K. and Standley, D.M. (2013) MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Mol. Biol. Evol.* **30**: 772–780.

Kawazoe, R., Okuyama, H., Reichardt, W., and Sasaki, S. (1991) Phospholipids and a novel glycine-containing lipoamino acid in *Cytophaga johnsonae* Stanier strain C21. *J. Bacteriol.* **173**: 5470–5475.

Kennedy, E.P. (1957) Metabolism of lipides. *Annu. Rev. Biochem.* **26**: 119–148.

Kettle, A.J. and Andreae, M.O. (2000) Flux of dimethylsulfide from the oceans : A comparison of updated data sets and flux models. *J. Geophys. Res.* **105**: 26,793-26,808.

Kim, Y.H., Yoo, J.S., and Kim, M.S. (1997) Structural Characterization of Sulfoquinovosyl , Monogalactosyl and Digalactosyl Diacylglycerols by FAB-CID-MS / MS. *J. Mass Spectrom.* **32**: 968–977.

Klausmeier, C.A., Litchman, E., Daufresne, T., and Levin, S.A. (2004) Optimal nitrogen-to-phosphorus stoichiometry of phytoplankton. *Nature* **429**: 171–174.

Klug, R.M. and Benning, C. (2001) Two enzymes of diacylglyceryl-*O*-4'-(*N,N,N*,-trimethyl)homoserine biosynthesis are encoded by *btaA* and *btaB* in the purple bacterium *Rhodobacter sphaeroides*. *Proc. Natl. Acad. Sci.* **98**: 5910–5.

Knowles, T.J., Scott-Tucker, A., Overduin, M., and Henderson, I.R. (2009) Membrane protein architects: the role of the BAM complex in outer membrane protein assembly. *Nat. Rev. Microbiol.* **7**: 206–214.

Koebnik, R., Locher, K.P., and Van Gelder, P. (2000) Structure and function of bacterial outer membrane proteins: barrels in a nutshell. *Mol. Microbiol.* **37**: 239–253.

Koike, I. and Nagata, T. (1997) High potential activity of extracellular alkaline phosphatase in deep waters of the central Pacific. *Deep. Res. Part II Top. Stud. Oceanogr.* **44**: 2283–2294.

Kolowith, L., Ingall, E., and Benner, R. (2001) Composition and cycling of marine organic phosphorus. *Limnol. Oceanogr.* **46**: 309–320.

Koonin, E. V (2005) Orthologs, paralogs, and evolutionary genomics. *Annu. Rev. Genet.* **39**: 309–338.

Kornberg, A., Rao, N.N., and Ault-Riché, D. (1999) Inorganic polyphosphate: A molecule of many functions. *Annu. Rev. Biochem.* **68**: 89–125.

Kovach, M.E., Elzer, P.H., Hill, D.S., Robertson, G.T., Farris, M.A., Roop, R.M., and Peterson, K.M. (1995) Four new derivatives of the broad-host-range cloning vector pBBR1MCS, carrying different antibiotic-resistance cassettes. *Gene* **166**: 175–6.

Krol, E. and Becker, A. (2004) Global transcriptional analysis of the phosphate starvation response in *Sinorhizobium meliloti* strains 1021 and 2011. *Mol. Genet. Genomics* **272**: 1–17.

Krom, M.D., Emeis, K.C., and Van Cappellen, P. (2010) Why is the Eastern Mediterranean phosphorus limited? *Prog. Oceanogr.* **85**: 236–244.

Kuhl, C., Tautenhahn, R., Böttcher, C., Larson, T.R., and Neumann, S. (2012) CAMERA: An integrated strategy for compound spectra extraction and annotation of liquid chromatography/mass spectrometry data sets. *Anal. Chem.* **84**: 283–289.

Kunzler, K. and Eichenberger, W. (1997) Betaine lipids and zwitterionic phospholipids in plants and fungi. *Phytochemistry* **46**: 883–892.

Kustka, A., Saudo-Wilhelmy, S., Carpenter, E.J., Capone, D.G., and Raven, J.A. (2003) A revised estimate of the iron use efficiency of nitrogen fixation with special reference to the marine cyanobacterium *Trichodesmium* spp. (*Cyanophyta*). *J. Phycol.* **39**: 12–25.

Laganowsky, A., Reading, E., Allison, T.M., Ulmschneider, M.B., Degiacomi, M.T., Baldwin, A.J., and Robinson, C. V. (2014) Membrane proteins bind lipids selectively to modulate their structure and function. *Nature* **510**: 172–175.

Lauro, F.M., McDougald, D., Thomas, T., Williams, T.J., Egan, S., Rice, S., et al. (2009) The genomic basis of trophic strategy in marine bacteria. *Proc. Natl. Acad. Sci.* **106**: 15527–15533.

Lee, A.G. (2011) Lipid-protein interactions. *Biochem. Soc. Trans.* **39**: 761–766.

Leventis, P.A. and Grinstein, S. (2010) The Distribution and Function of Phosphatidylserine in Cellular Membranes. *Annu. Rev. Biophys.* **39**: 407–427.

Li, L., Jia, Y., Hou, Q., Charles, T.C., Nester, E.W., and Pan, S.Q. (2002) A global pH sensor: *Agrobacterium* sensor protein ChvG regulates acid-inducible genes on its two chromosomes and Ti plasmid. *Proc. Natl. Acad. Sci.* **99**: 12369–74.

Li, W. and Godzik, A. (2006) Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* **22**: 1658–9.

Lidbury, I., Murrell, J.C., and Chen, Y. (2014) Trimethylamine *N*-oxide metabolism by abundant marine heterotrophic bacteria. *Proc. Natl. Acad. Sci.* **111**: 2710–2715.

Lindblom, G., Brentel, I., Sjölund, M., Wikander, G., and Wieslander, A. (1986) Phase equilibria of membrane lipids from *Acholeplasma laidlawii*: importance of a single lipid forming nonlamellar phases. *Biochemistry* **25**: 7502–7510.

Lingwood, D. and Simons, K. (2010) Lipid rafts as a membrane-organizing principle.

Science **327**: 46–50.

Lipp, J.S., Morono, Y., Inagaki, F., Hinrichs, K., Group, O.G., and Group, G. (2008) Significant contribution of *Archaea* to extant biomass in marine subsurface sediments. *Nature* **454**: 991–4.

Logares, R., Sunagawa, S., Salazar, G., Cornejo-Castillo, F.M., Ferrera, I., Sarmento, H., et al. (2014) Metagenomic 16S rDNA Illumina tags are a powerful alternative to amplicon sequencing to explore diversity and structure of microbial communities. *Environ. Microbiol.* **16**: 2659–2671.

Loladze, I. and Elser, J.J. (2011) The origins of the Redfield nitrogen-to-phosphorus ratio are in a homoeostatic protein-to-rRNA ratio. *Ecol. Lett.* **14**: 244–50.

López-Lara, I.M., Gao, J.-L., Soto, M.J., Solares-Pérez, A., Weissenmayer, B., Sohlenkamp, C., et al. (2005) Phosphorus-free membrane lipids of *Sinorhizobium meliloti* are not required for the symbiosis with alfalfa but contribute to increased cell yields under phosphorus-limiting conditions of growth. *Mol. Plant. Microbe. Interact.* **18**: 973–982.

López, D. and Kolter, R. (2010) Functional microdomains in bacterial membranes. *Genes Dev.* **24**: 1893–902.

Louie, K., Chen, Y., and Dowhan, W. (1986) Substrate-Induced Membrane Association of Phosphatidylserine Synthase from *Escherichia coli*. *J. Bacteriol.* **165**: 805–812.

Lozier, M.S. (2010) Deconstructing the Conveyor Belt. *Science* **328**: 1507–1511.

Lu, Y.-H.H., Guan, Z., Zhao, J., and Raetz, C.R. (2011) Three phosphatidylglycerol-phosphate phosphatases in the inner membrane of *Escherichia coli*. *J. Biol. Chem.* **286**: 5506–5518.

Lu, Y.J., Zhang, Y.M., Grimes, K.D., Qi, J., Lee, R.E., and Rock, C.O. (2006) Acyl-Phosphates Initiate Membrane Phospholipid Synthesis in Gram-Positive Pathogens. *Mol. Cell* **23**: 765–772.

Luo, H. (2014) Evolutionary origin of a streamlined marine bacterioplankton lineage. *ISME J.* **9**: 1423-1433.

Luo, H., Benner, R., Long, R. a, and Hu, J. (2009) Subcellular localization of marine bacterial alkaline phosphatases. *Proc. Natl. Acad. Sci.* **106**: 21219–23.

Makino, W., Cotner, J.B., Sterner, R.W., and Elser, J.J. (2003) Are bacteria more like plants

or animals? Growth rate and resource dependence of bacterial C : N : P stoichiometry. *Funct. Ecol.* **17**: 121–130.

Martin, P., Dyhrman, S.T., Lomas, M.W., Poulton, N.J., and Van Mooy, B. a S. (2014) Accumulation and enhanced cycling of polyphosphate by Sargasso Sea plankton in response to low phosphorus. *Proc. Natl. Acad. Sci.* **111**: 8089-8094.

Martin, P., Van Mooy, B.A.S., Heithoff, A., and Dyhrman, S.T. (2011) Phosphorus supply drives rapid turnover of membrane phospholipids in the diatom *Thalassiosira pseudonana*. *ISME J.* **5**: 1057–60.

Martínez, A., Osburne, M.S., Sharma, A.K., Delong, E.F., and Chisholm, S.W. (2012) Phosphite utilization by the marine picocyanobacterium *Prochlorococcus* MIT9301. *Environ. Microbiol.* **14**: 1363–1377.

Martiny, A.C., Pham, C.T.A., Primeau, F.W., Vrugt, J.A., Moore, J.K., Levin, S.A., and Lomas, M.W. (2013) Strong latitudinal patterns in the elemental ratios of marine plankton and organic matter. *Nat. Geosci.* **6**: 279–283.

Matsen, F.A., Kodner, R.B., and Armbrust, E.V. (2010) pplacer: linear time maximum-likelihood and Bayesian phylogenetic placement of sequences onto a fixed reference tree. *BMC Bioinformatics* **11**: 538.

Matyash, V., Liebisch, G., Kurzchalia, T. V, Shevchenko, A., and Schwudke, D. (2008) Lipid extraction by methyl-*tert*-butyl ether for high-throughput lipidomics. *J. Lipid Res.* **49**: 1137–46.

Mayer, J. and Cook, A.M. (2009) Homotaurine metabolized to 3-sulfopropanoate in *Cupriavidus necator* H16: Enzymes and genes in a patchwork pathway. *J. Bacteriol.* **191**: 6052–6058.

McGrath, J.W., Chin, J.P., and Quinn, J.P. (2013) Organophosphonates revealed: new insights into the microbial metabolism of ancient molecules. *Nat. Rev. Microbiol.* **11**: 412–9.

van Meer, G., Voelker, D.R., and Feigenson, G.W. (2008) Membrane lipids: where they are and how they behave. *Nat. Rev. Mol. Cell Biol.* **9**: 112–24.

Merchant, S.S. and Helmann, J.D. (2012) Elemental economy: microbial strategies for optimizing growth in the face of nutrient limitation. In, Poole,R.K. (ed), *Advances in microbial physiology*. Academic Press, pp. 91–210.

Michaels, A.F., Olson, D., Sarmiento, J.L., Ammerman, J.W., Fanning, K., Jahnke, R., et al. (1996) Inputs, losses and transformations of nitrogen and phosphorus in the pelagic North Atlantic Ocean. *Biogeochemistry* **35**: 181–226.

Mills, M., Moore, C., Langlois, R., Milne, A., Achterberg, E.P., Nachtigall, K., et al. (2008) Nitrogen and phosphorus co-limitation of bacterial productivity and growth in the Oligotrophic Sub-Tropical North Atlantic. *Limnol. Oceanogr.* **53**: 824–834.

Mills, M.M., Ridame, C., Davey, M., La Roche, J., and Geider, R.J. (2004) Iron and phosphorus co-limit nitrogen fixation in the eastern tropical North Atlantic. *Nature* **429**: 292–4.

Minnikin, D.E. and Abdolrahimzadeh, H. (1974) The replacement of phosphatidylethanolamine and acidic phospholipids by an ornithine-amide lipid and a minor phosphorus-free lipid in *Pseudomonas fluorescens* NCMB 129. *FEBS Lett.* **43**: 257–260.

Minnikin, D.E., Abdolrahimzadeh, H., and Baddiley, J. (1974) Replacement of acidic phosphates by acidic glycolipids in *Pseudomonas diminuta*. *Nature* **249**: 268–269.

Mira, A., Ochman, H., and Moran, N. a. (2001) Deletional bias and the evolution of bacterial genomes. *Trends Genet.* **17**: 589–596.

Miyasawa, K., Ito, K., and Matsumoto, F. (1970) Occurrence of (+)-2-hydroxy-3-aminopropanesulfonic acid and 3-aminopropanesulfonic acid in a red alga, *Grateloupia livida*. *Nippon Suisan Gakkaishi* **36**: 109–114.

Miyazaki, C., Kuroda, M., Ohta, A., and Shibuya, I. (1985) Genetic manipulation of membrane phospholipid composition in *Escherichia coli*: *pgsA* mutants defective in phosphatidylglycerol synthesis. *Proc. Natl. Acad. Sci.* **82**: 7530–4.

Møller, E.F. (2005) Sloppy feeding in marine copepods: Prey-size-dependent production of dissolved organic carbon. *J. Plankton Res.* **27**: 27–35.

Moore, C.M., Mills, M.M., Arrigo, K.R., Berman-Frank, I., Bopp, L., Boyd, P.W., et al. (2013) Processes and patterns of oceanic nutrient limitation. *Nat. Geosci.* **6**: 701–710.

Moore, E.K., Hopmans, E.C., Rijpstra, W.I.C., Andrea, I.S., Villanueva, L., Wienk, H., et al. (2015) Lysine and novel hydroxylysine lipids in soil bacteria: amino acid membrane lipid response to temperature and pH in *Pseudopedobacter saltans*. *Front. Microbiol.* **6**: 637.

Moore, E.K., Hopmans, E.C., Rijpstra, W.I.C., Villanueva, L., and Sinninghe Damsté, J.S. (2016) Elucidation and identification of amino acid containing membrane lipids using liquid chromatography/high-resolution mass spectrometry. *Rapid Commun. Mass Spectrom.* **30**: 739–750.

Van Mooy, B.A.S. and Fredricks, H.F. (2010) Bacterial and eukaryotic intact polar lipids in the eastern subtropical South Pacific: Water-column distribution, planktonic sources, and fatty acid composition. *Geochim. Cosmochim. Acta* **74**: 6499–6516.

Van Mooy, B.A.S., Fredricks, H.F., Pedler, B.E., Dyhrman, S.T., Karl, D.M., Koblízek, M., et al. (2009) Phytoplankton in the ocean use non-phosphorus lipids in response to phosphorus scarcity. *Nature* **458**: 69–72.

Van Mooy, B.A.S., Moutin, T., Duhamel, S., Rimmelin, P., and Van Wambeke, F. (2008) Phospholipid synthesis rates in the eastern subtropical South Pacific Ocean. *Biogeosciences* **5**: 133–139.

Van Mooy, B.A.S., Rocap, G., Fredricks, H.F., Evans, C.T., and Devol, A.H. (2006) Sulfolipids dramatically decrease phosphorus demand by picocyanobacteria in oligotrophic marine environments. *Proc. Natl. Acad. Sci.* **103**: 8607–12.

Moran, M.A., Buchan, A., González, J.M., Heidelberg, J.F., Whitman, W.B., Kiene, R.P., et al. (2004) Genome sequence of *Silicibacter pomeroyi* reveals adaptations to the marine environment. *Nature* **432**: 910–3.

Morris, J.J., Lenski, R.E., and Zinser, E.R. (2012) The Black Queen Hypothesis: Evolution of Dependencies through Adaptive Gene Loss. *MBio* **3**: 1–7.

Morris, R.M., Frazar, C.D., and Carlson, C.A. (2012) Basin-scale patterns in the abundance of SAR11 subclades, marine *Actinobacteria* (OM1), members of the Roseobacter clade and OCS116 in the. *Environ. Microbiol.* **14**: 1133–1144.

Morris, R.M., Rappe, M.S., Connon, S.A., Vergin, K.L., Siebold, W.A., Carlson, C.A., and Giovannoni, S.J. (2002) SAR11 clade dominates ocean surface bacterioplankton communities. *Nature* **420**: 806–810.

Nakamura, Y., Tsuchiya, M., and Ohta, H. (2007) Plastidic phosphatidic acid phosphatases identified in a distinct subfamily of lipid phosphate phosphatases with prokaryotic origin. *J. Biol. Chem.* **282**: 29013–21.

Newton, R.J., Griffin, L.E., Bowles, K.M., Meile, C., Gifford, S., Givens, C.E., et al. (2010) Genome characteristics of a generalist marine bacterial lineage. *ISME J.* **4**: 784–98.

Nguyen, H.P. and Schug, K. a. (2008) The advantages of ESI-MS detection in conjunction with HILIC mode separations: Fundamentals and applications. *J. Sep. Sci.* **31**: 1465–1480.

Nikaido, H. (2003) Molecular basis of bacterial outer membrane permeability revisited. *Microbiol. Mol. Biol. Rev.* **67**: 593–656.

Nikrad, M.P., Cottrell, M.T., and Kirchman, D.L. (2014) Growth activity of gammaproteobacterial subgroups in waters off the west Antarctic Peninsula in summer and fall. *Environ. Microbiol.* **16**: 1513–23.

Nishijima, S., Asami, Y., Uetake, N., Yamagoe, S., Ohta, A., and Shibuya, I. (1988) Disruption of the *Escherichia coli* cls gene responsible for cardiolipin synthesis. *J. Bacteriol.* **170**: 775–780.

O'Brien, P.J. and Herschlag, D. (2001) Functional interrelationships in the alkaline phosphatase superfamily: Phosphodiesterase activity of *Escherichia coli* alkaline phosphatase. *Biochemistry* **40**: 5691–5699.

O'Hara, R.B. and Kotze, D.J. (2010) Do not log-transform count data. *Methods Ecol. Evol.* **1**: 118–122.

Ogiso, H., Suzuki, T., and Taguchi, R. (2008) Development of a reverse-phase liquid chromatography electrospray ionization mass spectrometry method for lipidomics, improving detection of phosphatidic acid and phosphatidylserine. *Anal. Biochem.* **375**: 124–31.

Ohta, A. and Shibuya, I. (1977) Membrane phospholipid synthesis and phenotypic correlation of an *Escherichia coli* pss mutant. *J. Bacteriol.* **132**: 434–443.

Orsi, W.D., Smith, J.M., Liu, S., Liu, Z., Sakamoto, C.M., Wilken, S., et al. (2016) Diverse, uncultivated bacteria and archaea underlying the cycling of dissolved protein in the ocean. *ISME J.* **10:** 2158-2173.

Overbeek, R., Fonstein, M., D'Souza, M., Pusch, G.D., and Maltsev, N. (1999) The use of gene clusters to infer functional coupling. *Proc. Natl. Acad. Sci.* **96**: 2896–2901.

Park, J., Karplus, K., Barrett, C., Hughey, R., Haussler, D., Hubbard, T., and Chothia, C. (1998) Sequence comparisons using multiple sequences detect three times as many remote homologues as pairwise methods. *J. Mol. Biol.* **284**: 1201–1210.

Parsons, J.B. and Rock, C.O. (2013) Bacterial lipids: metabolism and membrane

homeostasis. *Prog. Lipid Res.* **52**: 249–76.

Partensky, F. and Garczarek, L. (2010) *Prochlorococcus*: Advantages and Limits of Minimalism. *Ann. Rev. Mar. Sci.* **2**: 305–331.

Paulsen, M.L., Doré, H., Garczarek, L., Seuthe, L., Müller, O., Sandaa, R.-A., et al. (2016) *Synechococcus* in the Atlantic Gateway to the Arctic Ocean. *Front. Mar. Sci.* **3**: 191.

Pesant, S., Not, F., Picheral, M., Kandels-Lewis, S., Le Bescot, N., Gorsky, G., et al. (2015) Open science resources for the discovery and analysis of Tara Oceans data. *Sci. Data* **2**: 150023.

Peschel, a, Jack, R.W., Otto, M., Collins, L. V, Staubitz, P., Nicholson, G., et al. (2001) *Staphylococcus aureus* resistance to human defensins and evasion of neutrophil killing via the novel virulence factor MprF is based on modification of membrane lipids with L-lysine. *J. Exp. Med.* **193**: 1067–1076.

Petersen, C. and Møller, L.B. (2001) The RihA, RihB, and RihC ribonucleoside hydrolases of *Escherichia coli*. Substrate specificity, gene expression, and regulation. *J. Biol. Chem.* **276**: 884–894.

Piekarski, T., Buchholz, I., Drepper, T., Schobert, M., Wagner-Doebler, I., Tielen, P., and Jahn, D. (2009) Genetic tools for the investigation of Roseobacter clade bacteria. *BMC Microbiol.* **9**: 265.

Pinhassi, J., Gómez-Consarnau, L., Alonso-Sáez, L., Sala, M., Vidal, M., Pedrós-Alió, C., and Gasol, J. (2006) Seasonal changes in bacterioplankton nutrient limitation and their effects on bacterial community composition in the NW Mediterranean Sea. *Aquat. Microb. Ecol.* **44**: 241–252.

Pomeroy, L.R., Williams, P.J. le B., Azam, F., and Hobbie, J.E. (2007) The microbial loop in the planktonic communities in lakes with various trophic status. *Oceanography* **20**: 28–33.

Pommier, T., Canback, B., Riemann, L., Bostrom, K.H., Simu, K., Lundberg, P., et al. (2007) Global patterns of diversity and community structure in marine bacterioplankton. *Mol. Ecol.* **16**: 867–880.

Popendorf, K.J., Fredricks, H.F., and Van Mooy, B.A.S. (2013) Molecular ion-independent quantification of polar glycerolipid classes in marine plankton using triple quadrupole MS. *Lipids* **48**: 185–95.

Popendorf, K.J., Lomas, M.W., and Van Mooy, B.A.S. (2011) Microbial sources of intact polar diacylglycerolipids in the Western North Atlantic Ocean. *Org. Geochem.* **42**: 803–811.

Popendorf, K.J., Tanaka, T., Pujo-Pay, M., Lagaria, A., Courties, C., Conan, P., et al. (2011) Gradients in intact polar diacylglycerolipids across the Mediterranean Sea are related to phosphate availability. *Biogeosciences* **8**: 3733–3745.

Pujalte, M.J., Lucena, T., Ruvira, M.A., Arahal, D.R., and Macian, M.C. (2013) The Family *Rhodobacteraceae*. In, Rosenberg,E. (ed), *The Prokaryotes: Alphaproteobacteria and Betaproteobacteria*. Springer-Verlag, pp. 440–512.

Punta, M., Coggill, P., Eberhardt, R., Mistry, J., Tate, J., Boursnell, C., et al. (2012) The Pfam protein families databases. *Nucleic Acids Res.* **38**: D211-D222.

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., et al. (2013) The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. *Nucleic Acids Res.* **41**: 590–596.

Raetz, C.R.H. and Newman, K.F. (1979) Diglyceride kinase mutants of *Escherichia coli*: Inner membrane association of 1,2-diglyceride and its relation to synthesis of membrane-derived oligosaccharides. *J. Bacteriol.* **137**: 860–868.

Rao, N.N., Gómez-García, M.R., and Kornberg, A. (2009) Inorganic polyphosphate: essential for growth and survival. *Annu. Rev. Biochem.* **78**: 605–47.

Rappé, M.S. and Giovannoni, S.J. (2003) The uncultured microbial majority. *Annu. Rev. Microbiol.* **57**: 369–394.

Redfield, A. (1958) The biological control of chemical factors in the environment. *Am. Sci* **46**: 205–221.

Reeves, J.H. (1992) Heterogeneity in the substitution process of amino acid sites of proteins coded for by mitochondrial DNA. *J. Mol. Evol.* **35**: 17–31.

Reisch, C.R., Crabb, W.M., Gifford, S.M., Teng, Q., Stoudemayer, M.J., Moran, M.A., and Whitman, W.B. (2013) Metabolism of dimethylsulphoniopropionate by *Ruegeria pomeroyi* DSS-3. *Mol. Microbiol.* **89**: 774–91.

Renner, L.D. and Weibel, D.B. (2011) Cardiolipin microdomains localize to negatively curved regions of *Escherichia coli* membranes. *Proc. Natl. Acad. Sci.* **108**: 6264–6269.

Ridgewell, A. and Arndt, S. (2015) Why Dissolved Organics Matter: DOC in Ancient Oceans and Past Climate Change. In, Carlson,C.A. and Hansell,D.A. (eds), *Biogeochemistry of Marine Dissolved Organic Matter*. Academic Press, pp. 1–20.

Riekhof, W.R., Andre, C., and Benning, C. (2005) Two enzymes, BtaA and BtaB, are sufficient for betaine lipid biosynthesis in bacteria. *Arch. Biochem. Biophys.* **441**: 96–105.

Riekhof, W.R., Naik, S., Bertrand, H., Benning, C., and Voelker, D.R. (2014) Phosphate starvation in fungi induces the replacement of phosphatidylcholine with the phosphorus-free betaine lipid diacylglyceryl-*N,N,N*-trimethylhomoserine. *Eukaryot. Cell* **13**: 749–57.

Rigel, N.W., Schwalm, J., Ricci, D.P., and Silhavy, T.J. (2012) BamE modulates the Escherichia coli beta-barrel assembly machine component BamA. *J. Bacteriol.* **194**: 1002–8.

Rippka, R., Coursin, T., Hess, W., Lichtle, C., Scanlan, D.J., Palinska, K. a., et al. (2000) *Prochlorococcus marinus* Chisholm *et al.* 1992 subsp. *pastoris* subsp. nov. strain PCC 9511, the first axenic chlorophyll a_2/b_2-containing cyanobacterium (*Oxyphotobacteria*). *Int. J. Syst. Evol. Microbiol.* **50**: 1833–1847.

Rojas-Jiménez, K., Sohlenkamp, C., Geiger, O., Martínez-Romero, E., Werner, D., and Vinuesa, P. (2005) A ClC chloride channel homolog and ornithine-containing membrane lipids of *Rhizobium tropici* CIAT899 are involved in symbiotic efficiency and acid tolerance. *Mol. Plant. Microbe. Interact.* **18**: 1175–1185.

Roy, A.B., Hewlins, M.J.E., Ellis, A.J., Harwood, J.L., and White, G.F. (2003) Glycolytic Breakdown of Sulfoquinovose in Bacteria: A Missing Link in the Sulfur Cycle. *Appl. Environ. Microbiol.* **69**: 6434–6441.

De Rudder, K.E., Thomas-Oates, J.E., and Geiger, O. (1997) *Rhizobium meliloti* mutants deficient in phospholipid *N*-methyltransferase still contain phosphatidylcholine. *J. Bacteriol.* **179**: 6921–6928.

De Rudder, K.E.E., Sohlenkamp, C., and Geiger, O. (1999) Plant-exuded choline is used for rhizobial membrane lipid biosynthesis by phosphatidylcholine synthase. *J. Biol. Chem.* **274**: 20011–20016.

Rusch, D.B., Halpern, A.L., Sutton, G., Heidelberg, K.B., Williamson, S., Yooseph, S., et al. (2007) The Sorcerer II Global Ocean Sampling expedition: northwest Atlantic through eastern tropical Pacific. *PLoS Biol.* **5**: e77.

Russell, D. and Edmondson, R. (1997) High-resolution Mass Spectrometry and Accurate Mass Measurements with Emphasis on the Characterization of Peptides and Proteins by Matrix-assisted Laser Desorption / Ionization Time-of-Flight Mass Spectrometry. *J. Mass Spectrom.* **32**: 263–276.

Saito, M.A. and Goepfert, T.J. (2008) Zinc-cobalt colimitation of *Phaeocystis antarctica*. *Limnol. Oceanogr.* **53**: 266–275.

Saito, M. a., Goepfert, T.J., and Ritt, J.T. (2008) Some thoughts on the concept of colimitation: Three definitions and the importance of bioavailability. *Limnol. Oceanogr.* **53**: 276–290.

Sandra, K. and Sandra, P. (2013) Lipidomics from an analytical perspective. *Curr. Opin. Chem. Biol.* **17**: 847–53.

Sañudo-Wilhelmy, S.A., Kustka, a B., Gobler, C.J., Hutchins, D. a, Yang, M., Lwiza, K., et al. (2001) Phosphorus limitation of nitrogen fixation by *Trichodesmium* in the central Atlantic Ocean. *Nature* **411**: 66–9.

Sañudo-Wilhelmy, S. a, Gómez-Consarnau, L., Suffridge, C., and Webb, E. a (2014) The role of B vitamins in marine biogeochemistry. *Ann. Rev. Mar. Sci.* **6**: 339–67.

Scanlan, D.J., Mann, N.H., and Carr, N.G. (1993) The response of the picoplanktonic marine cyanobacterium *Synechococcus* species WH7803 to phosphate starvation involves a protein homologous to the periplasmic phosphate-binding protein of *Escherichia coli*. *Mol. Microbiol.* **10**: 181–191.

Scanlan, D.J., Ostrowski, M., Mazard, S., Dufresne, A., Garczarek, L., Hess, W.R., et al. (2009) Ecological genomics of marine picocyanobacteria. *Microbiol. Mol. Biol. Rev.* **73**: 249–299.

Schäfer, A., Tauch, A., Jäger, W., Kalinowski, J., Thierbach, G., and Pühler, A. (1994) Small mobilizable multi-purpose cloning vectors derived from the *Escherichia coli* plasmids pK18 and pK19: selection of defined deletions in the chromosome of *Corynebacterium glutamicum*. *Gene* **145**: 69–73.

Schattenhofer, M., Fuchs, B.M., Amann, R., Zubkov, M. V, Tarran, G.A., and Pernthaler, J. (2009) Latitudinal distribution of prokaryotic picoplankton populations in the Atlantic Ocean. *Environ. Microbiol.* **11**: 2078–2093.

Schlesinger, W.H. and Bernhardt, E.S. (2013) Biogeochemistry: An analysis of global change 3rd ed.

Schubotz, F., Wakeham, S.G., Lipp, J.S., Fredricks, H.F., and Hinrichs, K.-U. (2009) Detection of microbial biomass by intact polar membrane lipid analysis in the water column and surface sediments of the Black Sea. *Environ. Microbiol.* **11**: 2720–34.

Schwalbe-Herrmann, M., Willmann, J., and Leibfritz, D. (2010) Separation of phospholipid classes by hydrophilic interaction chromatography detected by electrospray ionization mass spectrometry. *J. Chromatogr. A* **1217**: 5179–83.

Scott, J.T., Cotner, J.B., and Lapara, T.M. (2012) Variable stoichiometry and homeostatic regulation of bacterial biomass elemental composition. *Front. Microbiol.* **3**: 42.

Searle, B.C. (2010) Scaffold: A bioinformatic tool for validating MS/MS-based proteomic studies. *Proteomics* **10**: 1265–1269.

Sebastian, M. and Ammerman, J.W. (2009) The alkaline phosphatase PhoX is more widely distributed in marine bacteria than the classical PhoA. *ISME J.* **3**: 563–72.

Sebastián, M. and Gasol, J.M. (2013) Heterogeneity in the nutrient limitation of different bacterioplankton groups in the Eastern Mediterranean Sea. *ISME J.* **7**: 1665–8.

Sebastián, M., Smith, A.F., González, J.M., Fredricks, H.F., Mooy, B. Van, Koblízek, M., et al. (2016) Lipid remodelling is a widespread strategy in marine heterotrophic bacteria upon phosphorus deficiency. *ISME J.* **10**: 968–978.

Seidel, M., Rütters, H., Rullkötter, J., and Sass, H. (2013) Phosphate-free ornithine lipid contents in *Desulfovibrio* spp. respond to growth temperature. *Org. Geochem.* **59**: 133–142.

Selander, E., Kubanek, J., Hamberg, M., Andersson, M.X., Cervin, G., and Pavia, H. (2015) Predator lipids induce paralytic shellfish toxins in bloom-forming algae. *Proc. Natl. Acad. Sci.* **112**: 6395–6400.

Selje, N., Simon, M., and Brinkhoff, T. (2004) A newly discovered Roseobacter cluster in temperate and polar oceans. *Nature* **427**: 445–8.

Semeniuk, A., Sohlenkamp, C., Duda, K., and Hölzl, G. (2014) A bifunctional glycosyltransferase from *Agrobacterium tumefaciens* synthesizes monoglucosyl and glucuronosyl diacylglycerol under phosphate deprivation. *J. Biol. Chem.* **289**: 10104–14.

Seyedsayamdost, M.R., Carr, G., Kolter, R., and Clardy, J. (2011) Roseobacticides: Small molecule modulators of an algal-bacterial symbiosis. *J. Am. Chem. Soc.* **133**: 18343–

18349.

Seyedsayamdost, M.R., Case, R.J., Kolter, R., and Clardy, J. (2011) The Jekyll-and-Hyde chemistry of *Phaeobacter gallaeciensis*. *Nat. Chem.* **3**: 331–335.

Shaked, Y., Xu, Y., Leblanc, K., and Morel, F.M.M. (2006) Zinc availability and alkaline phosphatase activity in *Emiliania huxleyi*: Implications for Zn-P co-limitation in the ocean. *Limnol. Oceanogr.* **51**: 299–309.

Shevchenko, A. and Simons, K. (2010) Lipidomics: coming to grips with lipid diversity. *Nat. Rev. Mol. Cell Biol.* **11**: 593-598.

Shively, J.M. and Knoche, H.W. (1969) Isolation of an ornithine-containing lipid from *Thiobacillus thiooxidans*. *J. Bacteriol.* **98**: 829–830.

Sievers, F., Wilm, A., Dineen, D., Gibson, T.J., Karplus, K., Li, W., et al. (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **7**: 539.

Simon, M., Scheuner, C., Meier-Kolthoff, J.P., Brinkhoff, T., Wagner-Döbler, I., Ulbrich, M., et al. (2017) Phylogenomics of *Rhodobacteraceae* reveals evolutionary adaptation to marine and non-marine habitats. *ISME J.* **11**: 1483-1499.

Simon, R., Priefer, U., and Pühler, A. (1983) A broad host range mobilization system for *in vivo* genetic engineering: Transposon mutagenesis in gram negative bacteria. *Bio/Technology* **1**: 784–791.

Slightom, R.N. and Buchan, A. (2009) Surface colonization by marine roseobacters: integrating genotype and phenotype. *Appl. Environ. Microbiol.* **75**: 6027–37.

Smith, C.A., Want, E.J., O'Maille, G., Abagyan, R., and Siuzdak, G. (2006) XCMS: processing mass spectrometry data for metabolite profiling using nonlinear peak alignment, matching, and identification. *Anal. Chem.* **78**: 779–787.

Smith, D.C., Simon, M., Alldredge, A.L., and Azam, F. (1992) Intense hydrolytic enzyme activity on marine aggregates and implications for rapid particle dissolution. *Nature* **39**: 139–142.

Sohlenkamp, C. and Geiger, O. (2015) Bacterial membrane lipids: diversity in structures and pathways. *FEMS Microbiol. Rev.* **40**: 133-159.

Sohlenkamp, C., López-Lara, I.M., and Geiger, O. (2003) Biosynthesis of

phosphatidylcholine in bacteria. *Prog. Lipid Res.* **42**: 115–62.

Sowell, S.M., Abraham, P.E., Shah, M., Verberkmoes, N.C., Smith, D.P., Barofsky, D.F., and Giovannoni, S.J. (2011) Environmental proteomics of microbial plankton in a highly productive coastal upwelling system. *ISME J.* **5**: 856–65.

Sowell, S.M., Wilhelm, L.J., Norbeck, A.D., Lipton, M.S., Nicora, C.D., Barofsky, D.F., et al. (2009) Transport functions dominate the SAR11 metaproteome at low-nutrient extremes in the Sargasso Sea. *ISME J.* **3**: 93–105.

Stamatakis, A. (2014) RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**: 1312–1313.

Steimle, A., Autenrieth, I.B., and Frick, J.S. (2016) Structure and function: Lipid A modifications in commensals and pathogens. *Int. J. Med. Microbiol.* **306**: 290–301.

Sterner, R.W. and Elser, J.J. (2002) Ecological Stoichiometry. Princeton University Press, Princeton.

Strege, M. a (1998) Hydrophilic interaction chromatography-electrospray mass spectrometry analysis of polar compounds for natural product drug discovery. *Anal. Chem.* **70**: 2439–2445.

Strom, S.L. (2008) Microbial ecology of ocean biogeochemistry: a community perspective. *Science* **320**: 1043–5.

Stroobant, V., De Hoffman, E., Libert, R., and Van Hoof, F. (1995) Fast-Atom Bombardment Mass Spectrometry and Low Energy Collision-Induced Tandem Mass Spectrometry of Tauroconjugated Bile Acid Anions. *J. Am. Soc. Mass Spectrom.* **6**: 588–596.

Sturt, H.F., Summons, R.E., Smith, K., Elvert, M., and Hinrichs, K.-U. (2004) Intact polar membrane lipids in prokaryotes and sediments deciphered by high-performance liquid chromatography/electrospray ionization multistage mass spectrometry--new biomarkers for biogeochemistry and microbial ecology. *Rapid Commun. Mass Spectrom.* **18**: 617–28.

Sunagawa, S., Coelho, L.P., Chaffron, S., Kultima, J.R., Labadie, K., Salazar, G., et al. (2015) Structure and function of the global ocean microbiome. *Science* **348**: 1–10.

Suttle, C.A. (2005) Viruses in the sea. *Nature* **437**: 356–361.

Swan, B.K., Tupper, B., Sczyrba, A., Lauro, F.M., Martinez-Garcia, M., González, J.M., et al. (2013) Prevalent genome streamlining and latitudinal divergence of planktonic bacteria in the surface ocean. *Proc. Natl. Acad. Sci.* **110**: 11463–8.

Tahara, Y., Kameda, M., Yamada, Y., and Kondo, K. (1976) A new lipid; the ornithine and taurine-containing "Cerilipin". *Agric. Biol. Chem.* **40**: 243–244.

Tahara, Y., Shinmoto, K., Yamada, Y., and Kondo, K. (1978) Enzymatic Synthesis of Tauroornithine Lipid in *Gluconobacter cerinus*. *Agric. Biol. Chem.* **42**: 205–206.

Tahara, Y., Yamada, Y., and Kondo, K. (1976) A new lysine-containing lipid isolated from *Agrobacterium tumefaciens*. *Agric. Biol. Chem.* **40**: 1449–1450.

Tan, B.K., Bogdanov, M., Zhao, J., Dowhan, W., Raetz, C.R., and Guan, Z. (2012) Discovery of a cardiolipin synthase utilizing phosphatidylethanolamine and phosphatidylglycerol as substrates. *Proc. Natl. Acad. Sci.* **109**: 16504–16509.

Tanaka, T., Thingstad, T.F., Christaki, U., Colombet, J., Cornet-Barthaux, V., Courties, C., et al. (2011) Lack of P-limitation of phytoplankton and heterotrophic prokaryotes in surface waters of three anticyclonic eddies in the stratified Mediterranean Sea. *Biogeosciences* **8**: 525–538.

Taylor, C.J., Anderson, A.J., and Wilkinson, S.G. (1998) Phenotypic variation of lipid composition in *Burkholderia cepacia*: A response to increased growth temperature is a greater content of 2-hydroxy acids in phosphatidylethanolamine and ornithine amide lipid. *Microbiology* **144**: 1737–1745.

Taylor, J.S. and Raes, J. (2004) Duplication and divergence: the evolution of new genes and old ideas. *Annu. Rev. Genet.* **38**: 615–43.

R Core Team (2017) R: A language and environment for statistical computing.

Teeling, H., Fuchs, B.M., Becher, D., Klockow, C., Gardebrecht, A., Bennke, C.M., et al. (2012) Substrate-controlled succession of marine bacterioplankton populations induced by a phytoplankton bloom. *Science* **336**: 608–11.

Thingstad, T.F., Krom, M.D., Mantoura, R.F.C., Flaten, G. a F., Groom, S., Herut, B., et al. (2005) Nature of phosphorus limitation in the ultraoligotrophic eastern Mediterranean. *Science* **309**: 1068–71.

Thingstad, T.F., Øvreås, L., Egge, J.K., Løvdal, T., and Heldal, M. (2005) Use of non-limiting substrates to increase size; a generic strategy to simultaneously optimize

uptake and minimize predation in pelagic osmotrophs? *Ecol. Lett.* **8**: 675–682.

Thingstad, T.F. and Rassoulzadegan, F. (1999) Conceptual models for the biogeochemical role of the photic zone microbial food web, with particular reference to the Mediterranean Sea. *Prog. Oceanogr.* **44**: 271–286.

Thingstad, T.F.R. and Mantoura, R.F.C. (2005) Titrating excess nitrogen content of phosphorous-deficient eastern Mediterranean surface water using alkaline phosphatase activity as a bio-indicator. *Limnol. Oceanogr. Methods* **3**: 94–100.

Tjellström, H., Andersson, M.X., Larsson, K.E., and Sandelius, A.S. (2008) Membrane phospholipids as a phosphate reserve: the dynamic nature of phospholipid-to-digalactosyl diacylglycerol exchange in higher plants. *Plant. Cell Environ.* **31**: 1388–98.

Todd, J.D., Curson, A.R.J., Dupont, C.L., Nicholson, P., and Johnston, A.W.B. (2009) The dddP gene, encoding a novel enzyme that converts dimethylsulfoniopropionate into dimethyl sulfide, is widespread in ocean metagenomes and marine bacteria and also occurs in some *Ascomycete* fungi. *Environ. Microbiol.* **11**: 1624–1625.

Todd, J.D., Kirkwood, M., Newton-Payne, S., and Johnston, A.W.B. (2012) DddW, a third DMSP lyase in a model Roseobacter marine bacterium, *Ruegeria pomeroyi* DSS-3. *ISME J.* **6**: 223–6.

Todd, J.D., Rogers, R., Li, Y.G., Wexler, M., Bond, P.L., Sun, L., et al. (2007) Structural and regulatory genes required to make the gas dimethyl sulfide in bacteria. *Science* **315**: 666–9.

Tripp, H.J., Kitner, J.B., Schwalbach, M.S., Dacey, J.W.H., Wilhelm, L.J., and Giovannoni, S.J. (2008) SAR11 marine bacteria require exogenous reduced sulphur for growth. *Nature* **452**: 741–4.

Tyrrell, T. (1999) The relative influences of nitrogen and phosphorus on oceanic primary production. *Nature* **400**: 525–31.

Umena, Y., Kawakami, K., Shen, J.-R., and Kamiya, N. (2011) Crystal structure of oxygen-evolving photosystem II at a resolution of 1.9 Å. *Nature* **473**: 55–60.

Vences-Guzmán, M.Á., Geiger, O., and Sohlenkamp, C. (2012) Ornithine lipids and their structural modifications: from A to E and beyond. *FEMS Microbiol. Lett.* **335**: 1–10.

Vences-Guzmán, M.Á., Guan, Z., Escobedo-Hinojosa, W.I., Bermúdez-Barrientos, J.R., Geiger, O., and Sohlenkamp, C. (2015) Discovery of a bifunctional acyltransferase

responsible for ornithine lipid synthesis in *Serratia proteamaculans*. *Environ. Microbiol.* **17**: 1487–1496.

Vences-Guzmán, M.Á., Guan, Z., Ormeño-Orrillo, E., González-Silva, N., López-Lara, I.M., Martínez-Romero, E., et al. (2011) Hydroxylated ornithine lipids increase stress tolerance in *Rhizobium tropici* CIAT899. *Mol. Microbiol.* **79**: 1496–514.

Vergin, K.L., Beszteri, B., Monier, A., Thrash, J.C., Temperton, B., Treusch, A.H., et al. (2013) High-resolution SAR11 ecotype dynamics at the Bermuda Atlantic Time-series Study site by phylogenetic placement of pyrosequences. *ISME J.* **7**: 1322–32.

Versées, W. and Steyaert, J. (2003) Catalysis by nucleoside hydrolases. *Curr. Opin. Struct. Biol.* **13**: 731–738.

Villanueva, L., Bale, N., Hopmans, E.C., Schouten, S., and Damsté, J.S.S. (2013) Diversity and distribution of a key sulpholipid biosynthetic gene in marine microbial assemblages. *Environ. Microbiol.* **16**: 774–787.

Villarreal-Chiu, J.F., Quinn, J.P., and McGrath, J.W. (2012) The genes and enzymes of phosphonate metabolism by bacteria, and their distribution in the marine environment. *Front. Microbiol.* **3**: 19.

Voget, S., Wemheuer, B., Brinkhoff, T., Vollmers, J., Dietrich, S., Giebel, H.-A., et al. (2014) Adaptation of an abundant Roseobacter RCA organism to pelagic systems revealed by genomic and transcriptomic analyses. *ISME J.* **9**: 1–14.

Wagner-Döbler, I. and Biebl, H. (2006) Environmental biology of the marine Roseobacter lineage. *Annu. Rev. Microbiol.* **60**: 255–80.

Wallace, I.M., O'Sullivan, O., Higgins, D.G., and Notredame, C. (2006) M-Coffee: Combining multiple sequence alignment methods with T-Coffee. *Nucleic Acids Res.* **34**: 1692–1699.

Wang, C., Kemp, J., Da Fonseca, I.O., Equi, R.C., Sheng, X., Charles, T.C., and Sobral, B.W.S. (2010) *Sinorhizobium meliloti* 1021 loss-of-function deletion mutation in *chvI* and its phenotypic characteristics. *Mol. Plant. Microbe. Interact.* **23**: 153–60.

Wang, H., Sterner, R.W., and Elser, J.J. (2012) On the "strict homeostasis" assumption in ecological stoichiometry. *Ecol. Modell.* **243**: 81–88.

Wang, H., Tomasch, J., Jarek, M., and Wagner-Döbler, I. (2014) A dual-species co-cultivation system to study the interactions between Roseobacters and dinoflagellates.

Front. Microbiol. **5**: 311.

Waterhouse, A.M., Procter, J.B., Martin, D.M.A., Clamp, M., and Barton, G.J. (2009) Jalview Version 2--a multiple sequence alignment editor and analysis workbench. *Bioinformatics* **25**: 1189–91.

Webb, L. and Johannes, R.E. (1967) Studies of the release of dissolved free amino acids by marine zooplankton. *Limnol. Ocean.* **12**: 376–382.

Wee, S. and Wilkinson, B.J. (1988) Increased outer membrane ornithine-containing lipid and lysozyme penetrability of *Paracoccus denitrificans* grown in a complex medium deficient in divalent cations. *J. Bacteriol.* **170**: 3283–3286.

Weissenmayer, B., Gao, J.-L., López-Lara, I.M., and Geiger, O. (2002) Identification of a gene required for the biosynthesis of ornithine-derived lipids. *Mol. Microbiol.* **45**: 721–33.

Wheeler, T.J. and Eddy, S.R. (2013) Nhmmer: DNA homology search with profile HMMs. *Bioinformatics* **29**: 2487–2489.

Whitfield, C. and Trent, M.S. (2014) Biosynthesis and export of bacterial lipopolysaccharides. *Annu. Rev. Biochem.* **83**: 99–128.

Wietz, M., Gram, L., Jørgensen, B., and Schramm, A. (2010) Latitudinal patterns in the abundance of major marine bacterioplankton groups. *Aquat. Microb. Ecol.* **61**: 179–189.

Wikström, M., Xie, J., Bogdanov, M., Mileykovskaya, E., Heacock, P., Wieslander, Å., and Dowhan, W. (2004) Monoglucosyldiacylglycerol, a foreign lipid, can substitute for phosphatidylethanolamine in essential membrane-associated functions in *Escherichia coli*. *J. Biol. Chem.* **279**: 10484–10493.

Wilkinson, S.G. (1997) Bacterial lipopolysaccharides - Themes and variations. *Prog. Lipid Res.* **35**: 283–343.

Williams, T.J., Long, E., Evans, F., Demaere, M.Z., Lauro, F.M., Raftery, M.J., et al. (2012) A metaproteomic assessment of winter and summer bacterioplankton from Antarctic Peninsula coastal surface waters. *ISME J.* **6**: 1883–900.

Williams, T.J., Wilkins, D., Long, E., Evans, F., Demaere, M.Z., Raftery, M.J., and Cavicchioli, R. (2013) The role of planktonic *Flavobacteria* in processing algal organic matter in coastal East Antarctica revealed using metagenomics and metaproteomics.

Environ. Microbiol. **15**: 1302–1317.

Wu, J., Sunda, W., Boyle, E.A., and Karl, D.M. (2000) Phosphate depletion in the western north Atlantic Ocean. *Science* **289**: 759–762.

Wymann, M.P. and Schneiter, R. (2008) Lipid signalling in disease. *Nat. Rev. Mol. Cell Biol.* **9**: 162–76.

Xing, P., Hahnke, R.L., Unfried, F., Markert, S., Huang, S., Barbeyron, T., et al. (2014) Niches of two polysaccharide-degrading *Polaribacter* isolates from the North Sea during a spring diatom bloom. *ISME J.* **9**: 1410-1422.

Yamada, T., Uchikata, T., Sakamoto, S., Yokoi, Y., Fukusaki, E., and Bamba, T. (2013) Development of a lipid profiling system using reverse-phase liquid chromatography coupled to high-resolution mass spectrometry with rapid polarity switching and an automated lipid identification software. *J. Chromatogr. A* **1292**: 211–8.

Yan, S., Fuchs, B.M., Lenk, S., Harder, J., Wulf, J., Jiao, N.Z., and Amann, R. (2009) Biogeography and phylogeny of the NOR5/OM60 clade of *Gammaproteobacteria*. *Syst. Appl. Microbiol.* **32**: 124–139.

Yancey, P.H. (2005) Organic osmolytes as compatible, metabolic and counteracting cytoprotectants in high osmolarity and other stresses. *J. Exp. Biol.* **208**: 2819–30.

Yang, K., Zhao, Z., Gross, R.W., and Han, X. (2009) Systematic analysis of choline-containing phospholipids using multi-dimensional mass spectrometry-based shotgun lipidomics. *J. Chromatogr. B.* **877**: 2924–36.

Yang, Z. (1993) Maximum-likelihood estimation of phylogeny from DNA sequences when substitution rates differ over sites. *Mol. Biol. Evol.* **10**: 1396–401.

Yooseph, S., Li, W., and Sutton, G. (2008) Gene identification and protein classification in microbial metagenomic sequence data via incremental clustering. *BMC Bioinformatics* **9**: 182.

Zavaleta-Pastor, M., Sohlenkamp, C., Gao, J.-L., Guan, Z., Zaheer, R., Finan, T.M., et al. (2010) *Sinorhizobium meliloti* phospholipase C required for lipid remodeling during phosphorus limitation. *Proc. Natl. Acad. Sci.* **107**: 302–7.

Zehr, J.P. and Kudela, R.M. (2011) Nitrogen cycle of the open ocean: from genes to ecosystems. *Ann. Rev. Mar. Sci.* **3**: 197–225.

Zhang, G., Meredith, T.C., and Kahne, D. (2013) On the essentiality of lipopolysaccharide to Gram-negative bacteria. *Curr. Opin. Microbiol.* **16**: 779–85.

Zhang, J. (2003) Evolution by gene duplication: An update. *Trends Ecol. Evol.* **18**: 292–298.

Zhang, X., Ferguson-Miller, S.M., and Reid, G.E. (2009) Characterization of ornithine and glutamine lipids extracted from cell membranes of *Rhodobacter sphaeroides*. *J. Am. Soc. Mass Spectrom.* **20**: 198–212.

Zhang, X., Tamot, B., Hiser, C., Reid, G.E., Benning, C., and Ferguson-miller, S. (2011) Cardiolipin Deficiency in *Rhodobacter sphaeroides* Alters the Lipid Profile of Membranes and of Crystallized Cytochrome Oxidase, but Structure and Function Are Maintained. *Biochemistry* **50**: 3879–3890.

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

Zimmerman, A.E., Martiny, A.C., Lomas, M.W., and Allison, S.D. (2014) Phosphate supply explains variation in nucleic acid allocation but not C : P stoichiometry in the western North Atlantic. *Biogeosciences* **11**: 1599–1611.

Zubkov, M. V, Mary, I., Woodward, E.M.S., Warwick, P.E., Fuchs, B.M., Scanlan, D.J., and Burkill, P.H. (2007) Microbial control of phosphate in the nutrient-depleted North Atlantic subtropical gyre. *Environ. Microbiol.* **9**: 2079–89.

Zuur, A.F., Hilbe, J.M., and Ieno, E.N. (2013) A beginner's guide to GLM and GLMM with R. Highland Statistics Ltd., Newburgh, UK.

Zwirglmaier, K., Heywood, J.L., Chamberlain, K., Woodward, E.M.S., Zubkov, M. V., and Scanlan, D.J. (2007) Basin-scale distribution patterns of picocyanobacterial lineages in the Atlantic Ocean. *Environ. Microbiol.* **9**: 1278–1290.

Appendices

Appendix 1

Putative genes involved in phosphorus-free lipid synthesis in the genomes of PlcP-containing marine isolates. The genomes of all marine isolates in the IMG database were BLASTP searched using MED193_17359 as a query (e-value < $1e^{-20}$) in order to construct a database of PlcP-containing marine isolates. To find genes in the vicinity of PlcP, the nucleotide sequence 5000 base pairs up- and downstream of PlcP was extracted. tBlastn searches were performed against these sequences using a $1e^{-20}$ E-value cutoff. To detect putative non-P synthesis genes in the genomes of these isolates we used BlastP searches with a $1e^{-20}$ E-value cutoff. Query sequences used: Agt - Agau_C200037; BtaB - Q93TQ0_RHOSH; OlsF - Spro_2569; SqdB - Q9L8S7_RHIML; Pgt - mlr5650

	non-P lipid synthesis genes in the vicinity of PlcP	**Presence of non-P lipid synthesis genes in the genome, and putative non-P lipid synthesized**		
taxon_name	**Neighbour**	**MGDG/ GADG**	**DGTS**	**SQDG**
Hoeflea phototrophica DFL-43	BtaBA	Agt	BtaB	
Phaeobacter daeponensis TF-218, DSM 23529 (scaffold version)	BtaBA		BtaB	
Phaeobacter sp. MED193	BtaBA		BtaB	
Sagittula stellata E-37	BtaBA		BtaB	SqdB
Planctomyces maris DSM 8797	BtaBA			
Rhodopirellula baltica SH 1	BtaBA			
Rhodopirellula baltica SH28	BtaBA			
Rhodopirellula baltica SWK14	BtaBA			
Rhodopirellula baltica WH47	BtaBA			
Saprospira grandis HR1, DSM 2844	BtaBA			
Saprospira grandis Lewin	BtaBA			
alpha proteobacterium SCGC AAA536-G10	Agt			
alpha proteobacterium SCGC AAA536-K22	Agt			
alpha proteobacterium sp. HIMB59	Agt			
Amorphus coralli DSM 19760	Agt			SqdB
Aurantimonas coralicida DSM 14790	Agt			
Aurantimonas manganoxydans SI85-9A1	Agt			
beta proteobacterium NB0016	Agt			
Candidatus Pelagibacter sp. HTCC7211	Agt			
Caulobacter crescentus CB15	Agt			
Caulobacter crescentus NA1000	Agt			
Citromicrobium bathyomarinum JL354	Agt		BtaB	

Citromicrobium sp. JLT1363	Agt		BtaB	
Cucumibacter marinus DSM 18995	Agt			SqdB
Desulfobulbus mediterraneus DSM 13871	Agt			
Erythrobacter litoralis HTCC2594	Agt			
Erythrobacter sp. NAP1	Agt		BtaB	
Erythrobacter sp. SD-21	Agt			
Fulvimarina pelagi HTCC2506	Agt			
gamma proteobacterium sp. HTCC5015	Agt			
Kordiimonas gwangyangensis DSM 19435	Agt			
Labrenzia aggregata IAM 12614	Agt		BtaB	SqdB
Labrenzia alexandrii DFL-11	Agt		BtaB	SqdB
Labrenzia sp. DG1229	Agt		BtaB	SqdB
Leucothrix mucor DSM 2157	Agt		BtaB	
Limnobacter sp. MED105	Agt			
Loktanella vestfoldensis SKA53	Agt		BtaB	
marine bacterium Betaproteobacteria HIMB624	Agt			
Maritalea myrionectae DSM 19524	Agt			
Methylophaga aminisulfidivorans MP, KCTC 12909	Agt			
Methylophaga frappieri JAM7	Agt			
Methylophaga nitratireducenticrescens JAM1	Agt			
Nisaea denitrificans DSM 18348	Agt		BtaB	
Nisaea sp BAL199	Agt		BtaB	
Nitrobacter sp. Nb-311A	Agt			
Pelagibacterium halotolerans B2	Agt			SqdB
Polynucleobacter necessarius asymbioticus QLW-P1DMWA-1	Agt			
Pseudomonas aeruginosa WC55	Agt			
Pseudovibrio sp. JE062	Agt		BtaB	SqdB
Pusillimonas sp. T7-7	Agt			
Roseibium sp. TrichSKD4	Agt		BtaB	
Rubritalea marina DSM 17716	Agt			
Sphingomonas sp. KC8	Agt			
Sphingomonas sp. S17	Agt			
Sphingomonas sp. SKA58	Agt		BtaB	
Sphingopyxis alaskensis RB2256	Agt			
Sphingopyxis baekryungensis DSM 16222	Agt		BtaB	
Stappia stellulata DSM 5886	Agt		BtaB	
Terasakiella pusilla DSM 6293	Agt			
Thalassobaculum salexigens DSM 19539	Agt		BtaB	
Thalassospira lucentensis DSM 14000	Agt			SqdB
Thalassospira profundimaris WP0211	Agt			
Thalassospira xiamenensis M-5, DSM 17429	Agt			

C

Thiomicrospira crunogena XCL-2	Agt			
Thiomicrospira kuenenii DSM 12350	Agt			
Verrucomicrobia bacterium SCGC AAA300-K03	Agt			
Citreicella sp. 357	Pgt			
Citreicella sp. SE45	Pgt			SqdB
Pelagibaca bermudensis HTCC2601	Pgt			
Rhodobacter sphaeroides KD131	Pgt		BtaB	SqdB
Algicola sagamiensis DSM 14643	OlsF			
Alteromonas macleodii AltDE1	OlsF			
Alteromonas macleodii ATCC 27126	OlsF			
Alteromonas macleodii Balearic Sea AD45	OlsF			
Alteromonas macleodii Black Sea 11	OlsF			
Alteromonas macleodii Deep ecotype, DSM 17117	OlsF			
Alteromonas sp. S89	OlsF			SqdB
Catenovulum agarivorans YM01	OlsF			
Enterovibrio calviensis DSM 14347	OlsF			
gamma proteobacterium BDW918	OlsF			SqdB
gamma proteobacterium IMCC1989	OlsF			
gamma proteobacterium IMCC3088	OlsF			
Gayadomonas joobiniege G7	OlsF			
Glaciecola agarilytica 4H-3-7+YE-5	OlsF			
Haliea rubra CM41_15a, DSM 19751	OlsF			
Melitea salexigens DSM 19753	OlsF			
Microbulbifer variabilis ATCC 700307	OlsF	Agt		SqdB
Pseudoalteromonas tunicata D2	OlsF			
Rheinheimera baltica DSM 14885	OlsF			
Shewanella algae ACDC	OlsF			
Shewanella baltica BA175	OlsF			
Shewanella baltica OS117	OlsF			
Shewanella baltica OS155	OlsF			
Shewanella baltica OS183	OlsF			
Shewanella baltica OS185	OlsF			
Shewanella baltica OS195	OlsF			
Shewanella baltica OS223	OlsF			
Shewanella baltica OS625	OlsF			
Shewanella frigidimarina NCIMB 400	OlsF			
Shewanella piezotolerans WP3	OlsF			
Shewanella sp. HN-41	OlsF			
Shewanella sp. MR-4	OlsF			
Shewanella sp. MR-7	OlsF			
Shewanella sp. W3-18-1	OlsF			

D

Shewanella violacea DSS12	OlsF			
Spongiibacter tropicus DSM 19543	OlsF			
Stenotrophomonas sp. SKA14	OlsF	Agt		
Cytophaga hutchinsonii ATCC 33406	Uncharacterised glycosyltransferase			
Eudoraea adriatica DSM 19308	Uncharacterised glycosyltransferase			
Flavobacteriaceae bacterium S85	Uncharacterised glycosyltransferase			
Flavobacterium sp. SCGC AAA536-P05	Uncharacterised glycosyltransferase			
Fulvivirga imtechensis AK7	Uncharacterised glycosyltransferase			
Gracilimonas tropica DSM 19535	Uncharacterised glycosyltransferase	Agt		
Microscilla marina ATCC 23134	Uncharacterised glycosyltransferase			
Owenweeksia hongkongensis DSM 17368	Uncharacterised glycosyltransferase			
Pedobacter sp. BAL39	Uncharacterised glycosyltransferase			
Polaribacter sp. MED152	Uncharacterised glycosyltransferase			
Prolixibacter bellariivorans ATCC BAA-1284	Uncharacterised glycosyltransferase			
Saccharicrinis fermentans DSM 9555	Uncharacterised glycosyltransferase			
Tenacibaculum ovolyticum DSM 18103	Uncharacterised glycosyltransferase			
Caminibacter mediatlanticus TB-2	Uncharacterised glycosyltransferase			
Blastopirellula marina SH 106T, DSM 3645	Uncharacterised glycosyltransferase			
Algoriphagus mannitolivorans DSM 15301	Uncharacterised glycosyltransferase			
Algoriphagus marincola DSM 16067	Uncharacterised glycosyltransferase			
Algoriphagus sp. PR1	Uncharacterised glycosyltransferase			
Algoriphagus vanfongensis DSM 17529	Uncharacterised glycosyltransferase			
Aquiflexum balticum BA160, DSM 16537	Uncharacterised glycosyltransferase			
Belliella baltica BA134, DSM 15883	Uncharacterised glycosyltransferase			
Cyclobacterium marinum Raj, DSM 745	Uncharacterised glycosyltransferase			
Echinicola pacifica DSM 19836	Uncharacterised glycosyltransferase			
Echinicola vietnamensis KMM 6221, DSM 17526	Uncharacterised glycosyltransferase			
Rhodonellum psychrophilum DSM 17998	Uncharacterised glycosyltransferase			

Rhodonellum psychrophilum GCM71, DSM 17998 (NZ_Draft)	Uncharacterised glycosyltransferase			
Rhodothermus marinus SG0.5JP17-171	Uncharacterised glycosyltransferase	Agt		
Rhodothermus marinus SG0.5JP17-172	Uncharacterised glycosyltransferase	Agt		
Croceibacter atlanticus HTCC2559	Uncharacterised glycosyltransferase			
Dokdonia sp. MED134	Uncharacterised glycosyltransferase			
Flavobacteria bacterium BBFL7	Uncharacterised glycosyltransferase			
Joostella marina En5, DSM 19592	Uncharacterised glycosyltransferase			
Kordia algicida OT-1	Uncharacterised glycosyltransferase			
Mesonia mobilis DSM 19841	Uncharacterised glycosyltransferase			
Robiginitalea biformata HTCC2501	Uncharacterised glycosyltransferase			
Mesoflavibacter zeaxanthinifaciens DSM 18436	Uncharacterised glycosyltransferase			
Aliagarivorans marinus DSM 23064	-			
Aliagarivorans taiwanensis DSM 22990	-			
alpha proteobacterium SCGC AAA536-B06	-	Agt		
Alteromonas sp. SN2	-			
Amphritea japonica ATCC BAA-1530	-			
Aquimarina latercula DSM 2041	-			
Aquimarina muelleri DSM 19832	-			
Arcobacter sp. CAB	-			
Arcobacter sp. L	-			
Cellulophaga algicola IC166, DSM 14237	-			
Coraliomargarita akajimensis DSM 45221	-	Agt		
Dokdonia sp. MED134	-			
Ensifer meliloti AK83, DSM 23913	-			SqdB
Ensifer meliloti CIAM1775	-		BtaB	SqdB
Flavobacterium frigidarium DSM 17623	-			
gamma proteobacterium IMCC2047	-			
gamma proteobacterium SCGC AAA076-P09	-			
gamma proteobacterium SCGC AAA076-P13	-			
Gilvimarinus chinensis DSM 19667	-			SqdB
Hirschia baltica ATCC 49814	-	Agt		
Hirschia maritima DSM 19733	-	Agt		
Idiomarina baltica OS145	-			
Idiomarina loihiensis L2TR	-			
Idiomarina sediminum DSM 21906	-			

Leeuwenhoekiella blandensis MED217	-			
Maribacter antarcticus DSM 21422	-			
Marinobacter adhaerens HP15, DSM 23420	-	Agt		
Marinobacter algicola DG893	-	Agt		
Marinobacter aquaeolei VT8	-	Agt		
Marinobacter daepoensis DSM 16072	-	Agt		
Marinobacter hydrocarbonoclasticus ATCC 49840	-	Agt		
Marinobacter manganoxydans MnI7-9	-	Agt		
Marinomonas mediterranea MMB-1, ATCC 700492	-			
Marinomonas posidonica IVIA-Po-181	-			
Marinomonas sp. MWYL1	-			
Marinomonas ushuaiensis DSM 15871	-			
Mesoflavibacter zeaxanthinifaciens S86	-			
Methylomonas methanica MC09	-			
Methylophaga thiooxydans DMS010	-			
Muricauda ruestringensis B1, DSM 13258	-			
Nitratireductor aquibiodomus RA22 (Draft1)	-		BtaB	SqdB
Nitratireductor indicus C115	-		BtaB	SqdB
Nitrosococcus halophilus Nc4	-	Agt		
Photobacterium angustum S14	-			
Photobacterium leiognathi mandapamensis svers.1.1	-			
Photobacterium sp. SKA34	-			
Pseudoalteromonas arctica A 37-1-2	-			
Pseudoalteromonas citrea NCIMB 1889	-			
Pseudoalteromonas flavipulchra 2ta6 (Draft assembly 1)	-			
Pseudoalteromonas haloplanktis ANT/505	-			
Pseudoalteromonas luteoviolacea 2ta16 (Draft assembly 1)	-			
Pseudoalteromonas marina mano4	-			
Pseudoalteromonas piscicida ATCC 15057	-			
Pseudoalteromonas piscicida JCM 20779	-			
Pseudoalteromonas rubra ATCC 29570	-			
Pseudoalteromonas sp. TW-7	-			
Pseudoalteromonas spongiae UST010723-006	-			
Pseudomonas putida CSV86	-	Agt		
Reinekea blandensis MED297	-			
Salinisphaera shabanensis E1L3A	-	Agt		SqdB
Salisaeta longa DSM 21114	-	Agt		
SAR86 cluster bacterium SAR86C	-			
Shewanella waksmanii ATCC BAA-643	-			SqdB
Simiduia agarivorans DSM 21679	-			

Simiduia agarivorans SA1	-			
Sulfurospirillum arcachonense DSM 9755	-			
Synechococcus sp. RCC 307	-	Agt		SqdB
Thiocapsa marina 5811, DSM 5653	-	Agt		
Thiorhodococcus drewsii AZ1	-	Agt		SqdB
Thiorhodovibrio sp. 970	-	Agt		
Verrucomicrobiales sp. DG1235	-	Agt		
Vibrio nigripulchritudo ATCC 27043	-			
Zunongwangia profunda SM-A87	-			

Appendix 2

Complete list of strains used to construct the 16S *Rhodobacteraceae* phylogeny in Chapter 4 (Figure 4.7) as well as the OlsB homologs used to construct the protein phylogeny in Figure 4.8. A locus tag is provide for each gene unless none was available, in which case the Integrated Microbial Genomes (IMG) ID was used instead. Also noted is whether or not the gene was adjacent to an OlsA (red shading) or BamE (blue shading) homolog. If neither were present the cells are unshaded.

Genome Name	Locus Tag	Synteny
Actibacterium mucosum KCTC 23349	ACMU_06785	BamE
Actibacterium mucosum KCTC 23349	ACMU_07225	OlsA
Ahrensia kielensis DSM 5890	C504DRAFT_01675	None assigned
Antarctobacter heliothermus DSM 11445	Ga0070504_105410	BamE
Antarctobacter heliothermus DSM 11445	Ga0070504_10836	OlsA
Celeribacter baekdonensis B30	B30_09173	BamE
Celeribacter baekdonensis B30	B30_09438	OlsA
Celeribacter indicus P73 (complete genome)	P73_2673	BamE
Celeribacter indicus P73 (complete genome)	P73_2314	OlsA
Citreicella sp. 357	C357_04337	BamE
Citreicella sp. 357	C357_05518	OlsA
Citreicella sp. SE45	CSE45_1672	BamE
Citreicella sp. SE45	CSE45_1564	OlsA
Defluviimonas sp. 20V17	U879_13075	BamE
Defluviimonas sp. 20V17	U879_12540	OlsA
Dinoroseobacter shibae DFL-12, DSM 16493	Dshi_1472	OlsA
Dinoroseobacter shibae DFL-12, DSM 16493	Dshi_1707	None assigned
Epibacterium ulvae U95	Ga0070586_103415	BamE
Epibacterium ulvae U95	Ga0070586_103167	OlsA
Falsirhodobacter sp. alg1	Ga0056889_1003306	BamE
Falsirhodobacter sp. alg1	Ga0056889_1003137	OlsA
Falsirhodobacter sp. alg1	Ga0056889_10411	None assigned
Gemmobacter nectariphilus DSM 15620	G481DRAFT_00424	OlsA
Gemmobacter nectariphilus DSM 15620	G481DRAFT_00941	None assigned

Haematobacter missouriensis CCUG 52307	Ga0056721_00398	OlsA
Halocynthiibacter namhaensis RA2-3	Ga0078356_16760	BamE
Halocynthiibacter namhaensis RA2-3	Ga0078356_14065	OlsA
Jannaschia sp. CCS1	Jann_1789	BamE
Jannaschia sp. CCS1	Jann_2436	OlsA
Ketogulonicigenium vulgare WSH-001	KVU_0880	BamE
Labrenzia aggregata IAM 12614	SIAM614_27747	None assigned
Labrenzia alexandrii DFL-11	ladfl_03551	None assigned
Leisingera methylohalidivorans MB2, DSM 14336	Leime_2599	BamE
Leisingera methylohalidivorans MB2, DSM 14336	Leime_2192	OlsA
Litoreibacter janthinus DSM 26921	Ga0069998_1858	BamE
Litoreibacter janthinus DSM 26921	Ga0069998_2422	OlsA
Litorimicrobium taeanense DSM 22007	Ga0070518_10192	BamE
Litorimicrobium taeanense DSM 22007	Ga0070518_101369	OlsA
Loktanella vestfoldensis SKA53	SKA53_05328	BamE
Loktanella vestfoldensis SKA53	SKA53_04393	OlsA
Mameliella alba UMTAT08	Ga0077690_10515	BamE
Mameliella alba UMTAT08	Ga0077690_10413	OlsA
Maribius pelagius DSM 26893	Ga0069994_101600	BamE
Maribius pelagius DSM 26893	Ga0069994_101814	OlsA
Marinosulfonomonas sp. PRT002 (unscreened)	Rhd02DRAFT_00243	None assigned
Marinovum algicola DG 898	MALG_01881	BamE
Marinovum algicola DG 898	MALG_01588	OlsA
Maritimibacter alkaliphilus HTCC2654	RB2654_04859	BamE
Maritimibacter alkaliphilus HTCC2654	RB2654_04516	OlsA
Nereida ignava DSM 16309	Ga0004566_00829	BamE
Nereida ignava DSM 16309	Ga0004566_01411	OlsA
Nesiotobacter exalbescens DSM 16456	G512DRAFT_00615	None assigned
Oceanibulbus indolifex HEL-45	OIHEL45_13790	BamE
Oceanibulbus indolifex HEL-45	OIHEL45_13125	OlsA
Oceanicola granulosus HTCC2516	OG2516_00210	BamE
Oceanicola granulosus HTCC2516	OG2516_03949	OlsA
Oceaniovalibus guishaninsula JLT2003	OCGS_0729	BamE
Oceaniovalibus guishaninsula JLT2003	OCGS_0141	OlsA
Octadecabacter antarcticus 307	OAN307_c26180	BamE
Octadecabacter antarcticus 307	OAN307_c27320	OlsA
Octadecabacter arcticus 238, DSM 13978	OA238_c22780	BamE
Octadecabacter arcticus 238, DSM 13978	OA238_c18660	OlsA

Octadecabacter temperatus SB1	osb_18580	BamE
Octadecabacter temperatus SB1	osb_15520	OlsA
Paenirhodobacter enshiensis DW2-9	Ga0056890_100150	BamE
Paenirhodobacter enshiensis DW2-9	Ga0056890_100921	OlsA
Palleronia marisminoris DSM 26347	Ga0070011_11178	BamE
Palleronia marisminoris DSM 26347	Ga0070011_10114	OlsA
Pannonibacter phragmitetus DSM 14782	H163DRAFT_01743	None assigned
Paracoccus aminophilus JCM 7686	JCM7686_0884	BamE
Paracoccus aminophilus JCM 7686	JCM7686_0670	OlsA
Paracoccus denitrificans PD1222	Pden_1892	BamE
Paracoccus denitrificans PD1222	Pden_4506	OlsA
Paracoccus pantotrophus J40	ParpaJ40DRAFT_00009110	BamE
Paracoccus pantotrophus J40	ParpaJ40DRAFT_00011040	OlsA
Paracoccus zeaxanthinifaciens ATCC 21588	F804DRAFT_01967	OlsA
Paracoccus zeaxanthinifaciens ATCC 21588	F804DRAFT_01595	None assigned
Pelagibaca bermudensis HTCC2601	R2601_25991	BamE
Pelagibaca bermudensis HTCC2601	R2601_22467	OlsA
Pelagicola litoralis DSM 18290	BP08DRAFT_02238	BamE
Pelagicola litoralis DSM 18290	BP08DRAFT_03012	OlsA
Phaeobacter gallaeciensis DSM 26640	Gal_02360	BamE
Phaeobacter gallaeciensis DSM 26640	Gal_01844	OlsA
Phaeobacter inhibens T5, DSM 16374	Phain_01121	BamE
Phaeobacter inhibens T5, DSM 16374	Phain_01555	OlsA
Planktomarina temperata RCA23, DSM 22400 (RCA23)	RCA23_c16700	BamE
Planktomarina temperata RCA23, DSM 22400 (RCA23)	RCA23_c12810	OlsA
Pleomorphobacterium xiamenense DSM 24423	Ga0070523_102391	OlsA
Poseidonocella sedimentorum DSM 29315	Ga0074845_103166	BamE
Poseidonocella sedimentorum DSM 29315	Ga0074845_11277	OlsA
Pseudodonghicola xiamenensis DSM 18339	G455DRAFT_00300	BamE
Pseudodonghicola xiamenensis DSM 18339	G455DRAFT_00486	OlsA
Pseudooceanicola batsensis HTCC2597	OB2597_02452	BamE
Pseudooceanicola batsensis HTCC2597	OB2597_02887	OlsA
Pseudorhodobacter ferrugineus DSM 5888	G560DRAFT_01277	BamE
Pseudorhodobacter ferrugineus DSM 5888	G560DRAFT_01817	OlsA
Pseudoruegeria lutimaris DSM 25294	Ga0069989_100454	BamE
Pseudoruegeria lutimaris DSM 25294	Ga0069989_102342	OlsA
Pseudovibrio sp. FO-BEG1	PSE_0378	None assigned

Rhodobacter capsulatus SB1003	RCAP_rcc01906	BamE
Rhodobacter capsulatus SB1003	RCAP_rcc02998	OlsA
Rhodobacter sphaeroides 2.4.1, ATCC BAA-808	RSP_2618	BamE
Rhodobacter sphaeroides 2.4.1, ATCC BAA-808	RSP_3826	OlsA
Rhodobacteraceae bacterium KLH11	RKLH11_1542	BamE
Rhodobacteraceae bacterium KLH11	RKLH11_693	OlsA
Rhodobacteraceae sp. HIMB11	HIMB11_02712	BamE
Rhodobacteraceae sp. HIMB11	HIMB11_01468	OlsA
Rhodobacterales sp. HTCC2083	RB2083_1057	OlsA
Rhodobacterales sp. HTCC2083	RB2083_1198	None assigned
Rhodobacterales sp. HTCC2150	RB2150_03499	BamE
Rhodobacterales sp. HTCC2150	RB2150_01909	OlsA
Rhodobacterales sp. HTCC2255 (original sequence, contaminants removed)	2517701447	OlsA
Rhodobacterales sp. HTCC2255 metagenome (RAST annotated metagenome)	2517860771	None assigned
Rhodobacterales sp. Y4I	RBY4I_3283	BamE
Rhodobacterales sp. Y4I	RBY4I_1748	OlsA
Roseibium sp. TrichSKD4	TRICHSKD4_2022	None assigned
Roseicitreum antarcticum CGMCC 1.8894	Ga0070192_11244	BamE
Roseicitreum antarcticum CGMCC 1.8894	Ga0070192_104250	OlsA
Roseivivax isoporae LMG 25204	Ga0055015_100265	BamE
Roseivivax isoporae LMG 25204	Ga0055015_103144	OlsA
Roseobacter denitrificans OCh 114	RD1_3150	BamE
Roseobacter denitrificans OCh 114	RD1_2656	OlsA
Roseobacter litoralis Och 149	RLO149_c023790	BamE
Roseobacter litoralis Och 149	RLO149_c019070	OlsA
Roseobacter sp. AzwK-3b	RAZWK3B_16280	BamE
Roseobacter sp. AzwK-3b	RAZWK3B_15203	OlsA
Roseobacter sp. CHAB-I-5 SCGC AAA076-I17 (contamination screened)	2627886070	OlsA
Roseobacter sp. MED193	MED193_22491	BamE
Roseobacter sp. MED193	MED193_03912	OlsA
Roseovarius mucosus DSM 17069	rosmuc_01464	BamE
Roseovarius mucosus DSM 17069	rosmuc_01394	OlsA
Roseovarius nubinhibens ISM	ISM_16360	BamE
Roseovarius nubinhibens ISM	ISM_16095	OlsA
Roseovarius tolerans DSM 11457	Ga0070503_10811	BamE
Roseovarius tolerans DSM 11457	Ga0070503_13013	OlsA
Rubellimicrobium thermophilum DSM 16684	ruthe_00991	OlsA

Ruegeria lacuscaerulensis ITI-1157	SL1157_2600	BamE
Ruegeria lacuscaerulensis ITI-1157	SL1157_2839	OlsA
Ruegeria pomeroyi DSS-3	SPO2489	BamE
Ruegeria pomeroyi DSS-3	SPO1980	OlsA
Ruegeria sp. R11	RR11_602	BamE
Ruegeria sp. R11	RR11_1273	OlsA
Ruegeria sp. TM1040	TM1040_0919	BamE
Ruegeria sp. TM1040	TM1040_1208	OlsA
Salinihabitans flavidus DSM 27842	Ga0074816_10592	BamE
Salinihabitans flavidus DSM 27842	Ga0074816_12711	OlsA
Salipiger mucosus DSM 16094 (scaffold version)	salmuc_01869	BamE
Salipiger mucosus DSM 16094 (scaffold version)	salmuc_01906	OlsA
Sediminimonas qiaohouensis DSM 21189	G568DRAFT_00486	BamE
Sediminimonas qiaohouensis DSM 21189	G568DRAFT_01508	OlsA
Shimia haliotis DSM 28453	Ga0070219_101473	BamE
Shimia haliotis DSM 28453	Ga0070219_101685	OlsA
Silicibacter sp. TrichCH4B	SCH4B_4304	BamE
Silicibacter sp. TrichCH4B	SCH4B_4677	OlsA
Stappia stellulata DSM 5886	G572DRAFT_3756	None assigned
Sulfitobacter donghicola DSW-25, KCTC 12864	Z948DRAFT_03006	BamE
Sulfitobacter donghicola DSW-25, KCTC 12864	Z948DRAFT_03159	OlsA
Sulfitobacter litoralis DSM 17584	Ga0070569_11563	BamE
Sulfitobacter litoralis DSM 17584	Ga0070569_104140	OlsA
Thalassobacter stenotrophicus	Ga0069543_102283	BamE
Thalassobacter stenotrophicus	Ga0069543_104104	OlsA
Thalassobium sp. R2A62	TR2A62_0498	BamE
Thalassobium sp. R2A62	TR2A62_0759	OlsA
Thalassobius aestuarii DSM 15283	Ga0070018_104121	BamE
Thalassobius aestuarii DSM 15283	Ga0070018_11349	OlsA
Thalassococcus halodurans DSM 26915	Ga0069995_1583	BamE
Thalassococcus halodurans DSM 26915	Ga0069995_1728	OlsA
Thioclava pacifica DSM 10166	TP2_13125	BamE
Thioclava pacifica DSM 10166	TP2_10750	OlsA
Tranquillimonas alkanivorans DSM 19547	Ga0070241_103267	BamE
Tranquillimonas alkanivorans DSM 19547	Ga0070241_109102	OlsA
Tropicibacter multivorans DSM 26470	Ga0070013_104333	BamE
Tropicibacter multivorans DSM 26470	Ga0070013_105255	OlsA
Tropicimonas isoalkanivorans DSM 19548	Ga0070515_10453	BamE
Tropicimonas isoalkanivorans DSM 19548	Ga0070515_10189	OlsA

Wenxinia marina DSM 24838	A37WDRAFT_02909	BamE
Wenxinia marina DSM 24838	A37WDRAFT_03264	OlsA
Yangia pacifica CGMCC 1.3455	Ga0070148_102472	BamE
Yangia pacifica CGMCC 1.3455	Ga0070148_10868	OlsA

N

The ISME Journal (2016) 10, 968–978

www.nature.com/ismej

OPEN

ORIGINAL ARTICLE

Lipid remodelling is a widespread strategy in marine heterotrophic bacteria upon phosphorus deficiency

Marta Sebastián[1,9], Alastair F Smith[2,9], José M González[3], Helen F Fredricks[4], Benjamin Van Mooy[4], Michal Koblížek[5], Joost Brandsma[6], Grielof Koster[6], Mireia Mestre[1], Behzad Mostajir[7], Paraskevi Pitta[8], Anthony D Postle[6], Pablo Sánchez[1], Josep M Gasol[1], David J Scanlan[2] and Yin Chen[2]

[1]*Departament de Biologia Marina i Oceanografia, Institut de Ciències del Mar, CSIC, Barcelona, Spain;* [2]*School of Life Sciences, University of Warwick, Coventry, UK;* [3]*Department of Microbiology, University of La Laguna, La Laguna, Spain;* [4]*Department of Marine Chemistry and Geochemistry, Woods Hole Oceanographic Institution, Woods Hole, MA, USA;* [5]*Institute of Microbiology, Center Algatech, Třeboň, Czech Republic;* [6]*Clinical and Experimental Sciences, Faculty of Medicine, University of Southampton, Southampton, UK;* [7]*Center of Marine Biodiversity, Exploitation and Conservation (MARBEC), UMR 9190, CNRS – Université de Montpellier – IRD – IFREMER, Place Eugène Bataillon, Université de Montpellier, Case 93, Montpellier, France and* [8]*Hellenic Centre for Marine Research, Oceanography Institute, Heraklion, Greece*

Upon phosphorus (P) deficiency, marine phytoplankton reduce their requirements for P by replacing membrane phospholipids with alternative non-phosphorus lipids. It was very recently demonstrated that a SAR11 isolate also shares this capability when phosphate starved in culture. Yet, the extent to which this process occurs in other marine heterotrophic bacteria and in the natural environment is unknown. Here, we demonstrate that the substitution of membrane phospholipids for a variety of non-phosphorus lipids is a conserved response to P deficiency among phylogenetically diverse marine heterotrophic bacteria, including members of the *Alphaproteobacteria* and *Flavobacteria*. By deletion mutagenesis and complementation in the model marine bacterium *Phaeobacter* sp. MED193 and heterologous expression in recombinant *Escherichia coli*, we confirm the roles of a phospholipase C (PlcP) and a glycosyltransferase in lipid remodelling. Analyses of the Global Ocean Sampling and *Tara* Oceans metagenome data sets demonstrate that PlcP is particularly abundant in areas characterized by low phosphate concentrations. Furthermore, we show that lipid remodelling occurs seasonally and responds to changing nutrient conditions in natural microbial communities from the Mediterranean Sea. Together, our results point to the key role of lipid substitution as an adaptive strategy enabling heterotrophic bacteria to thrive in the vast P-depleted areas of the ocean.
The ISME Journal (2016) **10,** 968–978; doi:10.1038/ismej.2015.172; published online 13 November 2015

Introduction

Phosphorus (P) is required by all living organisms (Karl, 2014). In the oceans, P availability appears to exert a strong selective pressure on organisms growing in low P regions, driving ecotypic divergence in bacterioplankton populations (Coleman and Chisholm, 2010; Grote *et al.*, 2012). The bioavailable P concentration is often low enough to limit phytoplankton (Moore *et al.*, 2013) and heterotrophic bacteria production (Sebastián and Gasol, 2013). When P concentrations are low, heterotrophic bacteria appear to compete with phytoplankton for both organic and inorganic P sources (Havskum *et al.*, 2003; Zubkov *et al.*, 2007). The outcome of this struggle has potential implications for ecosystem productivity, as well as the balance between organic carbon consumption and sequestration (Havskum *et al.*, 2003).

When P is the limiting nutrient, the competitive ability of plankton groups should be inversely proportional to their cellular P quotas (Thingstad and Rassoulzadegan, 1999). Indeed, the imprint of this pressure for cost minimization is shown by the prevalence of streamlined genomes in diverse free-living bacterioplankton in the oligotrophic surface ocean (Swan *et al.*, 2013), noticeably the SAR11 clade representative *Pelagibacter ubique* (Giovannoni *et al.*, 2014) and the cyanobacterium *Prochlorococcus* (Swan *et al.*, 2013). Other than in nucleic acids, the major reservoir for organic P is in membrane phospholipids, which typically account for ~20% of the P content in heterotrophic bacteria and around 10% in some phytoplankton (Van Mooy *et al.*, 2008). Marine phytoplankton have evolved an adaptive strategy to minimize the P cost of their

Correspondence: M Sebastián, Departament de Biologia Marina i Oceanografia, Institut de Ciències del Mar-CMIMA, CSIC, Pg Marıtim de la Barceloneta 37-49, Barcelona E08003, Spain.
E-mail: msebastian@icm.csic.es
[9]These authors contributed equally to this work.
Received 4 August 2015; accepted 10 August 2015; published online 13 November 2015

membranes under low phosphate availability by replacing phospholipids with membrane lipids devoid of P (Van Mooy *et al.*, 2006 and 2009). Until recently, abundant marine heterotrophic bacteria were thought to lack the capacity for lipid remodelling in response to P deficiency (Van Mooy *et al.*, 2009) and the view that non-P lipids are primarily derived from phytoplankton has influenced the interpretation of some studies (Close *et al.*, 2014; Gašparović *et al.*, 2014). This view has since been challenged by the observation of glycolipids in the membranes of heterotrophic bacteria from the P-depleted Sargasso Sea (Popendorf *et al.*, 2011) as well as from an isolate of the SAR11 clade, which was originally isolated from P-depleted seawater (Carini *et al.*, 2015). However, it remains unclear whether lipid remodelling is a peculiarity of certain strains of the oligotrophic SAR11 clade or a more general adaptation among marine heterotrophic bacteria to growth in a variety of low P environments.

In this study, we demonstrate that the capacity to remodel lipid content under low P availability is widespread among phylogenetically diverse marine heterotrophic bacteria. We confirm the identity of key genes involved in this process. By deletion mutagenesis and complementation in the model marine heterotrophic bacterium *Phaeobacter* sp. MED193, we conclusively demonstrate that a recently identified phospholipase, PlcP (Zavaleta-Pastor *et al.*, 2010) is central for lipid remodelling. Moreover, analyses of the Global Ocean Sampling (GOS) and *Tara* Oceans metagenome data sets reveal that PlcP is more abundant in surface waters of P-depleted areas of the ocean, such as the Mediterranean Sea and the North Atlantic Ocean. We find that, in addition to its presence in some members of the SAR11 clade bacteria, PlcP is widely distributed across a number of abundant clades of marine bacteria with divergent life strategies, including other *Alphaproteobacteria* (for example, the SAR116 clade, and the marine *Roseobacter* clade), *Gammaproteobacteria* (for example, the SAR86 clade), *Flavobacteria* and *Verrucomicrobia*. By heterologous expression, we validate the activity of a SAR11 glycosyltransferase in the synthesis of glycolipids. Finally, we demonstrate for the first time that lipid remodelling occurs in nature in P-starved heterotrophic bacterial communities, which synthesize a wide array of non-P lipids, some of which were previously ascribed to phytoplankton.

Materials and methods

Phosphorus-starvation experiments with marine heterotrophic bacterial isolates

Growth conditions for marine heterotrophic bacteria under P-replete and deplete media are detailed in the Supplementary Materials and Methods. Reverse transcription–PCR (RT-PCR) experiments were performed as described elsewhere (Sebastián and Ammerman, 2009). RT-PCR primers are listed in Supplementary Table S3.

Searches for PlcP homologues

PlcP homologue sequences were retrieved from marine bacterial isolates using the Integrated Microbial Genomes database and marine metagenomes of the GOS (Rusch *et al.*, 2007) and the *Tara* Oceans (Sunagawa *et al.*, 2015) data sets, using PlcP (MED193_17359) of strain MED193 as a query (e-value $< 10^{-40}$). Details of the BLAST searches and the phylogenetic analysis can be found in Supplementary Materials and Methods. To search for genes involved in non-P lipid synthesis in the vicinity of *plcP*, the nucleotide sequences 5 kb upstream and downstream of PlcP in the genomes of the marine isolates were extracted. tBLASTn searches on these sequences were performed using characterized non-P lipid synthesis genes as queries (e-value $< 10^{-20}$; see Supplementary Table S2).

Construction and complementation of the plcP mutant in Phaeobacter sp. MED193 and heterologous expression of a SAR11 glycosyltransferase

Construction of a *plc*P deletion mutant of *Phaeobacter* sp. MED193 (MED193_17359) and complementation of the mutant with the native PlcP and SAR11 HTCC7211 PlcP (PB7211_983) were performed following Lidbury *et al.* (2014) and are described in Supplementary Materials and Methods. To characterize the SAR11 HTCC7211 glycosyltransferase, the *agt* homologue (PB7211_960) was chemically synthesized and codon optimized for overexpression in *Escherichia coli*. Glycolipids in recombinant *E. coli* were characterized by mass spectrometry as described in Supplementary Materials and Methods. Bacterial strains, plasmids and the primers used can be found in Supplementary Tables S4 and S5.

Collection of samples for intact polar diacylglycerolipids (IP-DAGs) in the environment

Samples from the LAMP2011 mesocosm experiment in the Eastern Mediterranean Sea (see Supplementary Material and Methods for details) were size-fractionated through 0.8-μm pore size filters to remove *Synechococcus* and picoeukaryote cells. Microbial communities were then collected onto 0.2-μm pore size filters. Flow cytometry analyses showed that the abundance of *Prochlorococcus* in the samples was negligible. Membrane lipids were extracted and quantified as detailed previously (Popendorf *et al.*, 2013), with the exception of the identification and characterization of glucuronic acid diacylglycerol (GADG) (see below). The bacterial community composition was evaluated by catalyzed reporter deposition fluorescence *in situ*

hybridization as described elsewhere (Sebastián *et al.*, 2012).

Natural seawater samples were also collected at the Blanes Bay Microbial Observatory (Western Mediterranean Sea) in August and September 2012 and processed as described above. Microbial community structure was investigated by pyrosequencing of bacterial 16S rRNA genes. Further details on this section can be found in Supplementary Materials and Methods.

Intact polar lipids measurement

Membrane lipids were analyzed using a high performance liquid chromatography–electrospray ionization–triplequadrupole mass spectrometry (HPLC-ESI-TQMS) method following the protocol detailed in Popendorf *et al.* (2013). Concentrations of GADG were obtained by HPLC/ESI–ion-trap MS (HPLC/ESI-IT-MS) as described in Sturt *et al.* (2004) and modified from Van Mooy and Fredricks (2010). Briefly, samples were injected onto a Princeton-SPHER (Princeton Chromatography Inc., Cranbury, NJ, USA) diol column ($2.1 \times 150\,mm^2$) operating in normal phase mode with a gradient from 100% A to 50% B at 20 min, then to 75% B at 25 min, hold at 75% B for 10 min and then back to 100% A. Flow rate for the gradient was $0.4\,ml\,min^{-1}$ and then the column was equilibrated with 100% A for 15 min at $1\,ml\,min^{-1}$. Eluent A was 80:20:0.1:0.04 of n-hexane: 2-propanol: formic acid: 14.8 N ammonium (aq) and eluent B was 90:10:0.1:0.04 of 2-propanol: water: formic acid: 14.8 N ammonium (aq). The electrospray source and the mass spectrometer (Thermo Scientific, Waltham, MA, USA) were configured as previously described (Sturt *et al.*, 2004) and programmed such that the base peak from alternating positive and negative ion full scans (250–2000 Da) was fragmented up to MS^2.

Molecular ions of *m*/*z* 760, 786, 788 and 814 that eluted at 14.7 min were identified as GADG (Supplementary Figure S4). The major peaks of the MS^2 spectrum of this group are typical of glycolipids such as monoglycosyl diacylglycerol (MGDG) but showed distinct headgroup neutral losses of *m*/*z* 193 and 211, which are *m*/*z* 14 greater than MGDG. The retention time of GADG was considerably longer than that of MGDG (8.3 min), indicating that it was more polar; this rules out the *m*/*z* 14 difference being due to an additional CH_2 group. An addition of an N atom is also unlikely because the positive and negative ion molecular ions showed clear evidence of ammonium (+) and formate (−) adducts, which would not occur if there was an N atom in the headgroup (Sturt *et al.*, 2004). This leaves the addition of a carboxylic acid oxygen atom as the remaining possibility, which is indeed consistent with a glucuronic acid headgroup moiety and an assignment of the lipid as GADG. As GADG lipids are not readily commercially available, their molecular ions were quantified against an MGDG calibration curve. Quantities were then normalized to internal standard DNP-PE (1,2-dipalmitoyl-sn-glycero-3-phosphoethanolamine-*N*-(2,4-dinitrophenyl)), which was added before extraction.

Results

plcP is expressed under P-deficiency conditions in marine heterotrophic bacteria and is essential for lipid remodelling

To test the hypothesis that phylogenetically diverse marine heterotrophic bacteria synthesize non-P lipids in response to P deficiency, we performed experiments with marine bacterial isolates whose genomes contain homologues to the newly discovered *plcP* from a terrestrial rhizobium (Zavaleta-Pastor *et al.*, 2010). We used the photoheterotrophic alphaproteobacterium *Erythrobacter* sp. NAP1 and two isolates from the Mediterranean Sea, the alphaproteobacterium *Phaeobacter* sp. MED193, which belongs to the *Roseobacter* clade, and the flavobacterium *Dokdonia* sp. MED134. These isolates harbour *plcP* genes sharing 50, 57 and 35% similarity with the rhizobial *plcP*, respectively. We subjected each strain to abrupt P starvation and studied the P starvation response. Alkaline phosphatase activity was used as a diagnostic for P deficiency (Thingstad and Mantoura, 2005) for the two Mediterranean strains, and activity was indeed induced in the P-starved cultures (Figures 1a and b). To investigate whether *plcP* was expressed following P limitation, we used RT-PCR to detect *plcP* transcripts. Expression of *plcP* was induced following P starvation and repressed when P was added back to the culture (Figures 1a–c), showing *plcP* is regulated by P availability.

Each of the strains synthesized a mixture of phospholipids and non-P lipids under both P-replete and P-deplete conditions (Figures 1d–f). Phospholipids detected were phosphatidylethanolamine (PE) and phosphatidylglycerol (PG), although the latter was absent in *Dokdonia* sp. MED134. Non-P lipids were diverse and differed between the three strains. In each case, the proportion of phospholipids decreased following P starvation, whereas non-P lipids became more abundant. Absolute amounts of phospholipids per cell were also determined for the Mediterranean isolates. In *Phaeobacter* sp. MED193, the phospholipid content decreased by half in response to P starvation, contributing $0.27 \pm 0.15 \times 10^6$ P atoms $cell^{-1}$ compared with $0.54 \pm 0.19 \times 10^6$ P atoms $cell^{-1}$ when P was not limiting (Supplementary Table S1). A similar response was observed in *Dokdonia* sp. MED134, with phospholipids reduced by around a third following P starvation, from $2.63 \pm 1.0 \times 10^6$ P atoms $cell^{-1}$ to $1.72 \pm 0.28 \times 10^6$ P atoms $cell^{-1}$ (Supplementary Table S1).

Different sets of non-P lipids became enriched following P starvation in each of the three isolates. The betaine lipid diacylglyceryl trimethylhomoserine (DGTS) accumulated as the major lipid in

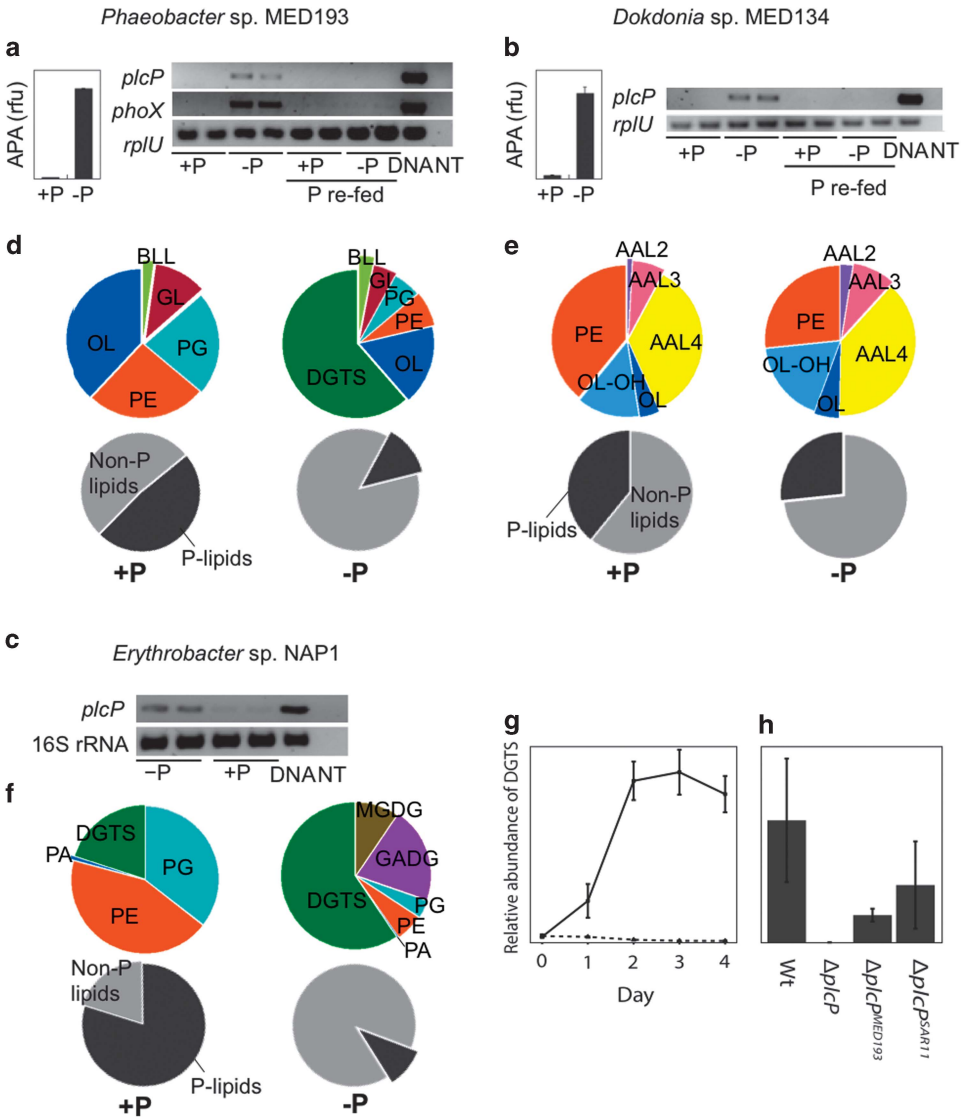


Figure 1 Phosphorus starvation induces alkaline phosphatase activity, *plcP* expression (**a–c**) and membrane lipid remodelling (**d–f**) in representative marine bacterial heterotrophs. RT-PCR results show *plcP* is transcribed under P-deplete conditions and repressed shortly (2 h) after P is added back to the media (P re-fed). *rplU* or 16S rRNA transcript levels were used as control for cDNA synthesis. Line graph (**g**) shows the accumulation of DGTS in *Phaeobacter* sp. MED193 over time in P-deplete (solid line) but not in P-replete (dashed line) media. DGTS synthesis is observed in wild type (Wt) but not in a *plcP* deletion mutant (Δ*plcP*) (**h**). Complementation with *plcP* from MED193 (Δ$plcP^{MED193}$) or from SAR11 (Δ$plcP^{SAR11}$) is sufficient to restore DGTS synthesis. Error bars represent the s.d. of three independent replicates. *phoX*: Alkaline phosphatase PhoX gene, DNA: PCR positive control, NT: no template control, APA: alkaline phosphatase activity, PA: phosphatidic acid, OL: ornithine lipid, GL: glutamine lipid, BLL: an uncharacterized betaine lipid, AAL2, AAL3, AAL4: uncharacterized aminolipids, OL-OH: a hydroxylated ornithine lipid.

Phaeobacter sp. MED193 (Figure 1d and g). DGTS also increased in abundance in the membrane of *Erythrobacter* sp. NAP1 when P was limiting, along with two glycolipids, MGDG and GADG (Figure 1f). Compared with the two alphaproteobacterial strains, the flavobacterium *Dokdonia* sp. MED134 exhibited a distinctive membrane lipid composition comprised of a number of different aminolipids (Figure 1e).

To determine whether PlcP is directly involved in lipid remodelling in *Phaeobacter* sp. MED193, we constructed a deletion mutant, Δ*plcP*. When grown under P-limiting conditions, the betaine lipid DGTS was not detectable in the mutant (Figure 1h). To eliminate the possibility of polar effects during mutant construction, we also verified that DGTS synthesis was restored by complementation of the Δ*plcP* mutant with the native *plcP* of *Phaeobacter* sp. MED193 (Δ$plcP^{MED193}$). Furthermore, complementation of the Δ*plcP Phaeobacter* sp. MED193 mutant with the *plcP* of the SAR11 HTCC7211 strain (Δ$plcP^{SAR11}$, Figure 1h) also restored the synthesis of DGTS, indicating that both enzymes have a similar functional behaviour.

PlcP is abundant and widely distributed among diverse phyla and also expressed in the marine environment

Having established that PlcP is central to the lipid remodelling response to P deficiency, we proceeded to investigate the abundance and distribution of PlcP in the marine environment. We mined the genomes of marine isolates in the Integrated Microbial Genomes database as well as the GOS (Rusch *et al.*, 2007) and *Tara* Oceans metagenomes (Sunagawa *et al.*, 2015) for PlcP homologues. BLAST searches retrieved 216 PlcP homologues in marine isolates (Supplementary Table S2) and >1100 unique environmental PlcP homologues in the GOS metagenomes. PlcP was also found to be ubiquitous in the *Tara* Oceans metagenomes. Phylogenetic analysis of these sequences provided evidence that PlcP is widespread among diverse bacterial phyla, including *Alphaproteobacteria*, *Gammaproteobacteria* and *Flavobacteria* (Figure 2a). Notably, PlcP is present in the two tropical ocean representatives of the ubiquitous SAR11 clade ('*Candidatus* Pelagibacter ubique' sp. HTCC7211 and HTCC7217), as has been recently shown by Carini *et al.* (2015). *Alphaproteobacteria* appeared to dominate the GOS and *Tara* Oceans data sets, accounting for around 75% of the hits (Figures 2b and c). Close relatives of the SAR11 clade PlcP, represented by strain HTCC7211, accounted for 54% and 30% of the hits in the GOS and *Tara* Oceans data sets, respectively. We also found PlcP homologues in single amplified genomes from a number of other abundant lineages of marine bacteria. These included the SAR116 clade of *Alphaproteobacteria*, the gammaproteobacterial SAR86, a flavobacterium and a verrucomicrobium (Dupont *et al.*, 2012; Swan *et al.*, 2013). As an adaptation to P deficiency, PlcP is expected to be more abundant in areas of the oceans where low P availability exerts a strong selective pressure. Analyses of the relative abundance of PlcP in the metagenomes from the GOS and the *Tara* Oceans expeditions showed that PlcP is indeed more abundant in the Mediterranean Sea and the North Atlantic Ocean (Supplementary Figure S1), which are characterized by low phosphate concentrations and high N/P ratios compared with other areas of the oceans (Fanning, 1992; Krom *et al.*, 2010; Wu, 2000). We also analyzed marine metatranscriptomic data sets and found that *plcP* is being transcribed in the marine environment by heterotrophic bacteria belonging to different phylogenetic groups (Supplementary Figure S2).

Genomic context of plcP

In order to shed further light on the role of PlcP, we analyzed its genomic context in marine isolates in the Integrated Microbial Genomes data set and environmental metagenomic scaffolds of the GOS data set (Figure 3). Genes involved in the phosphorus starvation response (Pho regulon) are regulated by the two-component system PhoR/PhoB and share a consensus sequence named Pho box in their promoter regions. Using bioinformatics approaches, we identified putative Pho-box binding sites in the PlcP gene clusters. Based on the genomic context of *plcP*, there are broadly six types of gene arrangements, and we detected a gene putatively involved in non-P lipid synthesis in the neighbourhood of *plcP* in around two-thirds of the strains investigated (Supplementary Table S2).

Type 1 is present in several strains, including *Phaeobacter* sp. MED193, in which the two genes necessary for the synthesis of the betaine lipid DGTS, *btaA* and *btaB* (Riekhof *et al.*, 2005), are found

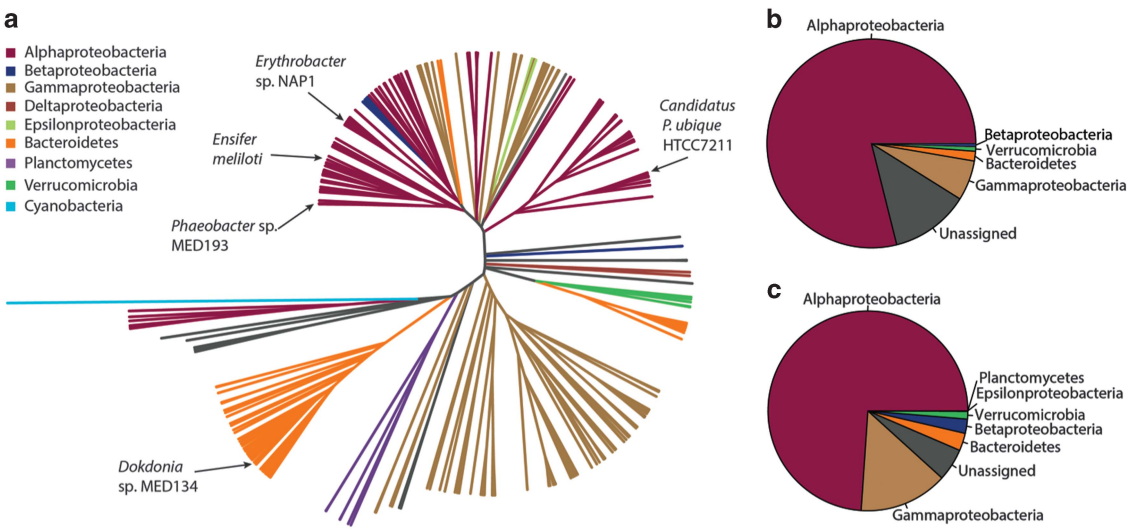


Figure 2 Distribution of PlcP in marine bacteria. (**a**) Phylogenetic relationships of PlcP homologues in marine bacterial isolates and all non-redundant PlcP sequences in the Global Ocean Sampling data set. Sequences were retrieved by BLASTP using *Phaeobacter* sp. MED193 PlcP (MED193_17359) as query (*e*-value $<10^{-40}$). Uncolored sequences are metagenome sequences that did not cluster alongside any isolate sequences. Right panel: Percentage of distribution of metagenome hits within each phylogenetic group in the Global Ocean Sampling (**b**) and *Tara* Oceans (**c**) data sets.

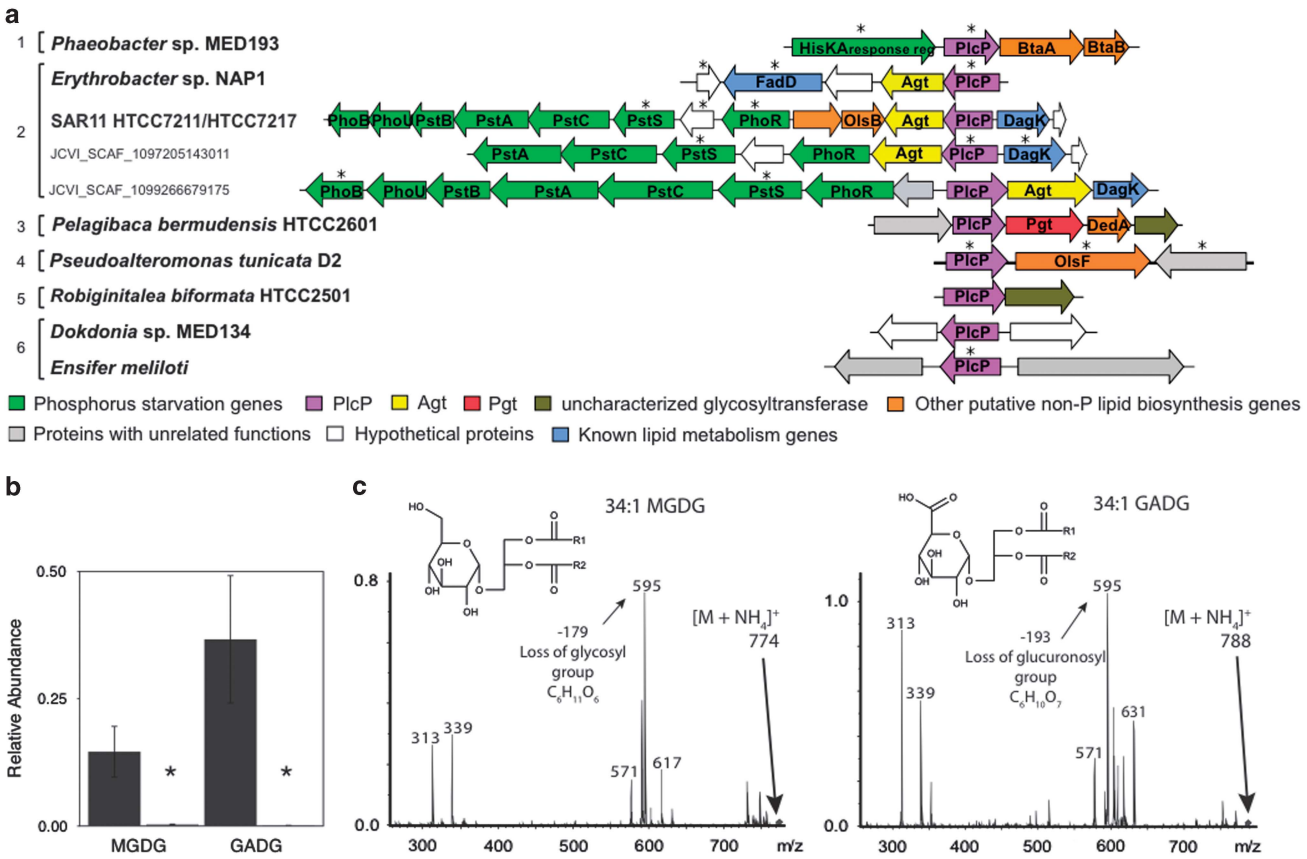


Figure 3 PlcP often appears in an operon with a glycosyltransferase. (**a**) Genomic context of *plcP*. Numbers denote the six types of *plcP* genomic neighbourhood in representative marine heterotrophic bacteria (see Supplemantary Table S2 for further information). Asterisks indicate those genes with an upstream PhoB-binding site (Pho box) and therefore putatively expressed upon phosphorus deficiency. (**b**) Heterologous expression of the putative glycosyltransferase *agt* from SAR11 HTCC7211 results in the accumulation of two glycolipids: MGDG and GADG (dark grey bars) but not in the control harbouring an empty expression plasmid (light grey bars). Abundance is expressed as the peak area of glycolipid relative to that of a phosphatidylcholine internal standard. (**c**) Fragmentation spectra for representative MGDG and GADG species obtained from recombinant *E. coli* harbouring the HTCC7211 *agt* homologue. The two species differ in the neutral loss corresponding to the polar head group (179 and 193 *m/z* for the loss of hexosyl and hexuronic acid groups, respectively). In each case, this loss yields diacylglycerol (595 *m/z*). Two peaks corresponding to monoacylglycerol with 18:1 and 16:0 fatty acids (339 and 313 *m/z*, respectively) are also present.

downstream of *plcP*. This is consistent with the lipid results obtained in strain MED193 cultures and its mutants (Figures 1d, g and h). The most commonly observed gene in the vicinity of *plcP* (type 2, present in >25% of the bacterial isolates analyzed, Supplementary Table S2) was a putative glycosyltransferase, homologous to the promiscuous glycosyltransferase *agt* just recently identified in *Agrobacterium tumefaciens* (Semeniuk *et al.*, 2014). Strains with *agt* homologues downstream of *plcP* include the environmentally relevant '*Candidatus* Pelagibacter ubique' sp. HTCC7211 and HTCC7217 and *Erythrobacter* sp. NAP1 (Figure 3a). In *A. tumefaciens*, this gene was found to be required for the synthesis of GADG and MGDG under P deficiency (Semeniuk *et al.*, 2014), which is consistent with our finding of those lipids in *Erythrobacter* sp. NAP1 when grown in P-limiting conditions (Figure 1f) and the recent findings of Carini *et al.* (2015). To confirm the role of the glycosyltransferase *agt* in the synthesis of these lipids, we chemically synthesized the gene (PB7211_960) from '*Candidatus* Pelagibacter ubique' sp. HTCC7211 and expressed it in *E. coli*. This resulted in the accumulation of MGDG and GADG (Figures 3b and c), which were not observed in lipid extracts from the same strain transformed with an empty vector (Figure 3b), confirming that this *agt* homologue in HTCC7211 is sufficient and responsible for the synthesis of the two glycolipids MGDG and GADG.

Types 3 and 5 include a number of marine heterotrophic bacteria where *plcP* is located immediately upstream of some putative glycosyltransferases, one of which has been confirmed to be responsible for the synthesis of di- and triglycosyldiacylglycerols (Devers *et al.*, 2011). However, in types 4 and 6, *plcP* does not seem to form an operon with its neighbouring genes although genes predicted to be involved in the synthesis of non-P lipids (for example, *btaBA*) are, in many cases, found elsewhere in their genomes (Supplementary Table S2). Together, our analyses

of the genomic context of *plcP* support its role in the synthesis of non-P glycolipids and betaine lipids in marine heterotrophic bacteria (Supplementary Figure S3).

Membrane lipid remodelling in environmental samples

To assess whether natural marine heterotrophic bacterial communities indeed replace phospholipids with a variety of non-P lipids upon P deficiency, we analyzed their membrane lipid composition using a size-fractionation approach. We first performed a mesocosm experiment with waters from the Eastern Mediterranean Sea, which is considered one of the most P-starved systems on Earth (Tanaka *et al.*, 2007). Non-P lipids represented >80% of the membrane lipids of heterotrophic bacteria in this system (Figures 4a and b). Four mesocosms were set up, two of which were enriched with phosphate (100 nM), while the other two were left un-amended, as controls. Six days after enrichment, phospholipids had greatly increased in proportion, to represent almost half of the total lipids, a 2–5-fold increase in the membrane lipid P content relative to P-starved controls (Figure 4b, Supplementary Table S1). Despite this shift in the lipid profile, there was little variation in bacterial community composition between treatments at the broad phylogenetic scale (Figure 4c). Interestingly, the most abundant non-P lipid in the P-starved mesocosm was GADG (fourfold higher compared with the P-enriched treatments) (Figure 4a), which has not been reported in marine surface waters previously (Popendorf *et al.*, 2011). GADG is likely synthesized by the promiscuous glycosyltransferase, *agt*, from SAR11 (Figures 3b and c), which dominated the heterotrophic bacterial communities during the mesocosm experiments (Figure 4c).

We also analyzed the membrane lipids of heterotrophic bacterial communities in the Western Mediterranean Sea. Samples were collected in August, when P limitation was expected to be severe (Pinhassi *et al.*, 2006), and in September, when P limitation commonly starts to be alleviated. In agreement with our results from the mesocosm experiments, there was a substantial change in membrane lipid composition, despite little change in bacterial community composition between samplings (Figure 5). Non-P lipids represented >70% of the polar lipids in August but only around 30% in September. Non-P lipids in the bacterial fraction were composed of glycolipids and betaine lipids. Among the glycolipids, the sulpholipid, sulphoquinovosyl diacylglycerol (SQDG), diglycosyl diacylglycerol (DGDG) and GADG were the most abundant, although the amount of MGDG was also significantly higher in August. The betaine lipid, diacylglyceryl hydroxymethyl trimethyl-β-alanine (DGTA), was also relatively more abundant in August. These two environmental studies provide direct support for our predictions, based on work with cultured isolates and comparative genomics, that lipid remodelling is a general response by marine heterotrophic bacteria to P deficiency.

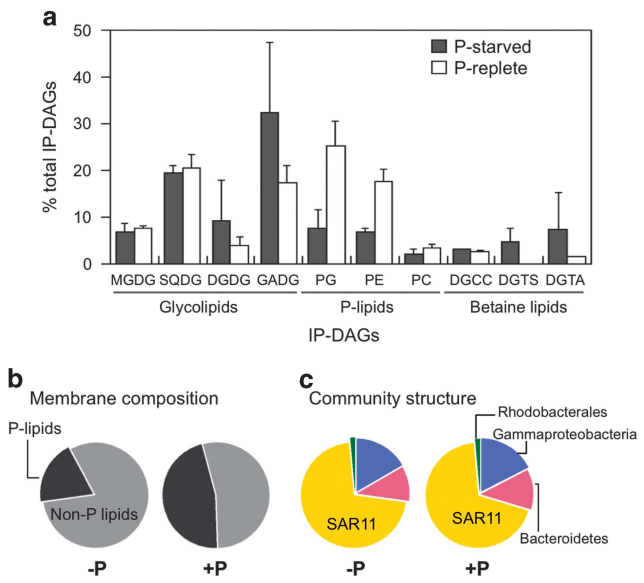


Figure 4 Polar lipids from the heterotrophic bacterial community (0.2–0.8-μm size fraction) in the Eastern Mediterranean during the mesocosm experiment. (**a**) Contribution of IP-DAGs to the membrane composition in the P-starved (control) and P-replete treatments. (**b**) Percentage of contribution of non-P versus P-lipids. (**c**) Bacterioplankton community structure at the time of sampling. PC: phosphatidylcholine, DGCC: diacylglycerylcarboxy-*N*-hydroxymethyl-choline.

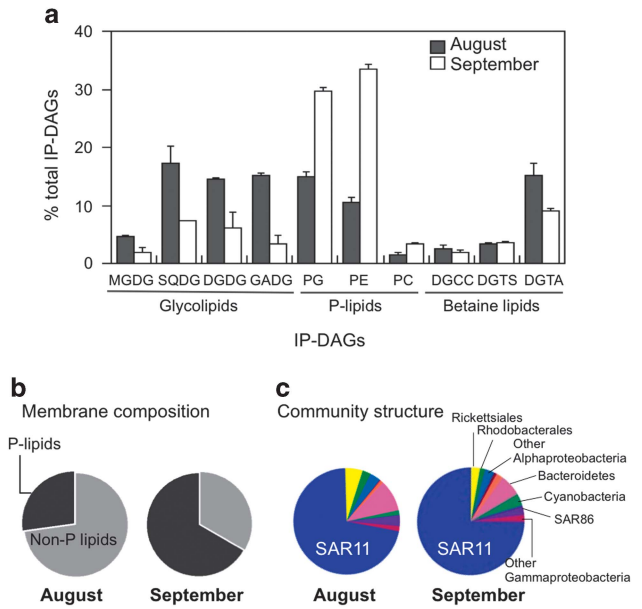


Figure 5 Polar lipids from the heterotrophic bacterial community (0.2–0.8-μm size fraction) at Blanes Bay (North-West Mediterranean) in August and September 2012. (**a**) Contribution of IP-DAGs to membrane composition. (**b**) Percentage of contribution of non-P versus P-lipids. (**c**) Bacterioplankton community structure at the time of sampling (right panel). Abbreviations are as described in Figure 4.

Discussion

Lipid remodelling is a widespread strategy in marine heterotrophic bacteria under low P conditions

The ability to synthesize non-P lipids seems to be widespread among marine heterotrophic bacteria adapted to low P environments. We found that in P-depleted regions, such as the North Atlantic Subtropical Gyre and the Mediterranean Sea, PlcP may be present in the majority of bacterial cells (Supplementary Figure S1). Although alphaproteobacterial PlcP sequences are the most abundant in marine metagenomes (Figure 2), PlcP appears to be broadly distributed among bacteria with distinct ecologies. Whereas the planktonic oligotrophic SAR11 clade are specialized for the high affinity uptake of small compounds (Giovannoni *et al.*, 2005), PlcP was also identified in the SAR86 clade, believed to dominate the uptake of high molecular weight organic matter (Dupont *et al.*, 2012). Members of groups with a particle-associated lifestyle, such as the *Roseobacter* clade and *Bacteroidetes* (Wagner-Döbler and Biebl, 2006; Fernández-Gómez *et al.*, 2013), also harboured PlcP homologues. Lipid remodelling mediated by PlcP therefore does not appear to be restricted to bacteria with particular lifestyles but rather constitutes part of a generic response to low P among marine bacteria from P-depleted environments.

One previous explanation for the accumulation of non-P lipids in marine heterotrophic bacteria is that they increase cell size without changing the cellular P quota (Thingstad *et al.*, 2005; Van Mooy *et al.*, 2009). However, we observed reductions in the amount of phospholipids per cell without obvious changes in cell size or in estimated cell volume (Supplementary Table S1). Together with the very low membrane P content found in environmental heterotrophic bacteria (Supplementary Table S1), our data suggest an alternative strategy whereby lipid remodelling is used to reduce the cellular P quota.

Enzymes involved in the synthesis of these non-P glycerolipids, including DGTS, SQDG, MGDG and GADG, require DAG as a substrate (Klug and Benning, 2001; Hölzl *et al.*, 2005; Semeniuk *et al.*, 2014). Bacterial lipid synthesis proceeds via a shared phosphatidate (phosphorylated DAG) intermediate (Parsons and Rock, 2013). The generation of DAG requires either the de-phosphorylation of phosphatidate or the removal of the polar head group from a phospholipid by the action of a phospholipase C (for example, PlcP). *Phaeobacter* sp. MED193 as well as the SAR11 representative, HTCC7211, appear to lack homologues of bacterial phosphatidate phosphatases (Icho and Raetz, 1983). Thus it is likely that the major means of generating DAG in marine heterotrophic bacteria is through the action of PlcP (Supplementary Figure S3). This is consistent with the abolition of DGTS synthesis in the *Phaeobacter* sp. MED193 Δ*plcP* deletion mutant (Figure 1h) and the complementation of the *Phaeobacter* sp. MED193 Δ*plcP* deletion mutant with the PlcP of SAR11 strain HTCC7211. On the other hand, in *E. coli*, DAG can be generated through the degradation of PG by MdoB (Parsons and Rock, 2013), which explains why glycolipids could accumulate in *E. coli* following overexpression of *agt* from '*Candidatus* Pelagibacter ubique' sp. HTCC7211 without an exogenous supply of DAG. In contrast to the widespread distribution of PlcP homologues across several phyla of marine heterotrophic bacteria, they are largely absent from the genomes of cyanobacteria (Figure 2), which are also capable of lipid remodelling (Van Mooy *et al.*, 2006, 2009). It is likely that cyanobacteria instead generate the required DAG through the action of phosphatidate phosphatases, which are found in their genomes (Nakamura *et al.*, 2007).

The ability to remodel lipids in response to P scarcity appears to be ubiquitous in phytoplankton (Van Mooy *et al.*, 2009), whereas in heterotrophic bacteria it seems to be more restricted to those adapted to low P conditions (Supplementary Figure S1). An explanation for this may be that the typical replacement lipids in cyanobacteria, SQDG and MGDG, are also required for the proper functioning of photosynthetic membranes, for example, in association with photosystem II (Umena *et al.*, 2011). Heterotrophic bacteria, by contrast, appear to require primarily PE and PG as bulk membrane lipids (Parsons and Rock, 2013). There is a selective pressure against the accumulation of genes that confer an insufficient advantage in bacterial genomes (Mira *et al.*, 2001), and this deletional bias has been invoked as an explanation for the patchy distribution of genes conferring an adaptation to P deficiency in the marine environment (Coleman and Chisholm, 2010). An example of this is found among the SAR11 bacteria. '*Candidatus* Pelagibacter ubique' sp. HTCC7211, isolated from the P-deplete Sargasso Sea, contains a higher number and diversity of P acquisition genes than other SAR11 strains (for example, *P. ubiqu*e HTCC1062), thereby increasing its ability to use alternative sources of P (Carini *et al.*, 2014), and is capable of lipid remodelling (Carini *et al.* 2015). In fact, HTCC7211 becomes dominant during the summer months in the Sargasso Sea (Brown *et al.*, 2012), when this region becomes P depleted. In contrast, *P. ubiqu*e HTCC1062, isolated from P-rich waters of the coastal North East Pacific, can only grow on phosphate (Carini *et al.*, 2014), and the lack of *plcP* in its genome suggests that it is not capable of P-lipid remodelling. Thus, it is likely that genes involved in the synthesis of non-P lipids are specifically acquired by marine heterotrophic bacteria as an adaptation to low P conditions. This may explain why non-P lipids in marine heterotrophic bacteria seem to be much more diverse than those in their photosynthetic counterparts.

Whereas sulpholipids are the preferred substitution lipid for cyanobacteria, and betaine lipids for eukaryotic phytoplankton (Van Mooy *et al.*, 2009),

heterotrophic bacteria synthesize a broad spectrum of lipids upon P deficiency, including betaine lipids (DGTS and DGTA), and a variety of glycolipids and putative aminolipids (Figures 2–5). Sulpholipids (SQDG), the glycolipid DGDG and betaine lipids (DGTS, DGTA) have been traditionally ascribed to phytoplankton (Van Mooy and Fredricks, 2010), although production of DGTA in dark seawater incubations using glucose as a precursor has been previously reported (Popendorf *et al*., 2011), supporting our suggestion of a heterotrophic origin. Furthermore, the gene *sqdB*, involved in SQDG biosynthesis, has recently been detected in environmental heterotrophic bacteria (Villanueva *et al*., 2014). However, sequences affiliated to heterotrophic bacteria appear to be rare in pelagic marine environments (Van Mooy *et al*., 2006). Similarly, homologues of glycosyltransferases known to be involved in DGDG synthesis such as *pgt* (Devers *et al*., 2011; Hölzl *et al*., 2005) are rare in the marine environment although found in some marine heterotrophic bacteria (for example, *Pelagibaca bermudensis*; Supplementary Table S2, Figure 3a). In our study, only small numbers of cyanobacteria were present in the fractions used for lipid analysis (<3% of total cells), suggesting that the detected lipids are of predominantly heterotrophic origin. How marine heterotrophic bacteria synthesize these lipids, that is, DGTA and DGDG, awaits further characterization.

The glycosyltransferase Agt is a key enzyme in the synthesis of glycolipids in marine heterotrophic bacteria

We also confirmed that the glycosyltransferase Agt is the enzyme mediating the synthesis of MGDG and GADG in SAR11 bacteria (Figure 3), and report that these lipids are abundant in the membrane lipids of natural bacterial communities in P-deplete waters of the Mediterranean Sea (Figures 4 and 5). To the best of our knowledge, our study constitutes the first report of glucuronic acid lipids in marine environments. Substitution of the acidic phospholipid PG by glucuronic acid-containing glycolipids has also been documented in P-limited cultures of *Brevundimonas diminuta* (formerly *Pseudomonas diminuta*) (Minnikin *et al*., 1974) and recently in the Sargasso sea SAR11 isolate HTCC7211 (Carini *et al.* 2015). Therefore, it seems that despite the metabolic diversity of the SAR11 clade (Schwalbach *et al*., 2010, Carini *et al.* 2014), SAR11 isolates from P-depleted systems behave similarly. Both the phospholipid PG and GADG are anionic under physiological conditions, so it may be that they could be interchanged while maintaining the biophysical properties of the membrane. A similar substitution of PG for the anionic SQDG has already been proposed for phytoplankton (Van Mooy *et al*., 2006). PlcP in *Sinorhizobium meliloti* only displayed activity towards the zwitterionic phospholipids PE and phosphatidylcholine (Zavaleta-Pastor *et al*., 2010) but the substrate specificity of SAR11 PlcP, as well as environmental PlcP homologues, warrants further investigation.

Physiological consequences of lipid remodelling

Despite the variety of surrogate non-P lipids, marine heterotrophic bacteria nonetheless seem to revert to phospholipid-dominated membranes under P sufficiency, similar to what has been observed in phytoplankton (Martin *et al*., 2011). Yet a study in yeast has shown that DGTS can functionally substitute for the phospholipid phosphatidylcholine without an apparent phenotype in culture (Riekhof *et al*., 2014). It is therefore puzzling why heterotrophic bacteria do not simply maintain a constant low P membrane composition. In plants, phospholipids have been proposed as P storage molecules, which can be mobilized upon P deficiency (Tjellström *et al*., 2008). Indeed, we observed that heterotrophic bacterial communities in the Mediterranean Sea accumulated phospholipids upon relief of P limitation (Figures 4 and 5), while there was no major change in community composition, suggesting that excess P is stored in phospholipids. This is consistent with the use of phospholipids as a P reservoir: on return to P limitation, PlcP-mediated degradation of membrane phospholipids might result in a net release of P for diversion to other cellular uses. An additional explanation for the prevalence of phospholipids under P sufficiency is that membrane proteins appear to have evolved in a phospholipid-dominated environment. For example, activity of certain membrane transporters can be enhanced by specific interactions with phospholipids (Laganowsky *et al*., 2014).

In summary, our results demonstrate that the ability to substitute phospholipids for non-P lipids under P deprivation is a common strategy in marine heterotrophic bacterial communities adapted to low P environments. Central to the remodelling process is a phospholipase, PlcP, which is widespread in marine surface waters where phosphorus is scarce. Our data point to lipid remodelling as an important ecological adaptation enabling natural heterotrophic bacteria to thrive in low P marine environments.

Conflict of Interest

The authors declare no conflict of interest.

Acknowledgements

This work was partially supported by grants STORM (CTM2009-09352/MAR), MALASPINA (CSD2008-00077), HOTMIX (CTM2011-30010/MAR), DOREMI (CTM2012-34294) and EcoBGM (CTM2013-48292-C3-3-R) funded by the Spanish Government, GAČR project GA13-11281S and MESOAQUA (228224) funded by the European Union Seventh Framework Program (FP7/2007-2013) and by

the Natural Environment Research Council (NERC), UK (NE/M002233/1). NERC is also acknowledged for a PhD scholarship to AFS and a Strategic Environmental Science Capital fund award for a liquid chromatography–mass spectrometry instrument to the University of Warwick. We thank J Pinhassi for providing the bacterial strains used in this study, K Kopejtka for his help with *Erythrobacter* sp. NAP1 cultures, A Jones and S Slade at the University of Warwick for helping with Q-TOF MS analyses, A Guerrero for his help with the laboratory experiments, I Forn for the image analyses and EL Sà for her participation in the field sampling.

References

Brown M V, Lauro FM, DeMaere MZ, Muir L, Wilkins D, Thomas T *et al.* (2012). Global biogeography of SAR11 marine bacteria. *Mol Syst Biol* **8**: 595.

Carini P, Van Mooy BAS, Thrash JC, White AE, Zhao Y, Campbell EO *et al.* (2015). SAR11 lipid renovation in response to phosphate starvation. *Proc Natl Acad Sci USA* **112**: 7767–7772.

Carini P, White AE, Campbell EO, Giovannoni SJ. (2014). Methane production by phosphate-starved SAR11 chemoheterotrophic marine bacteria. *Nat Commun* **5**: 4346.

Close HG, Wakeham SG, Pearson A. (2014). Lipid and 13C signatures of submicron and suspended particulate organic matter in the Eastern Tropical North Pacific: implications for the contribution of Bacteria. *Deep Sea Res Part I Oceanogr Res Pap* **85**: 15–34.

Coleman ML, Chisholm SW. (2010). Ecosystem-specific selection pressures revealed through comparative population genomics. *Proc Natl Acad Sci USA* **107**: 18634–18639.

Devers EA, Wewer V, Dombrink I, Dörmann P, Hölzl G. (2011). A processive glycosyltransferase involved in glycolipid synthesis during phosphate deprivation in *Mesorhizobium loti*. *J Bacteriol* **193**: 1377–1384.

Dupont CL, Rusch DB, Yooseph S, Lombardo M-J, Richter RA, Valas R *et al.* (2012). Genomic insights to SAR86, an abundant and uncultivated marine bacterial lineage. *ISME J* **6**: 1186–1199.

Fanning KA. (1992). Nutrient provinces in the sea: concentration ratios, reaction rate ratios, and ideal covariation. *J Geophys Res* **97**: 5693.

Fernández-Gómez B, Richter M, Schüler M, Pinhassi J, Acinas SG, González JM *et al.* (2013). Ecology of marine Bacteroidetes: a comparative genomics approach. *ISME J* **7**: 1026–1037.

Gašparović B, Frka S, Koch BP, Zhu ZY, Bracher A, Lechtenfeld OJ *et al.* (2014). Factors influencing particulate lipid production in the East Atlantic Ocean. *Deep Sea Res Part I Oceanogr Res Pap* **89**: 56–67.

Giovannoni SJ, Cameron Thrash J, Temperton B. (2014). Implications of streamlining theory for microbial ecology. *ISME J* **8**: 1553–1565.

Giovannoni SJ, Tripp HJ, Givan S, Podar M, Vergin KL, Baptista D *et al.* (2005). Genome streamlining in a cosmopolitan oceanic bacterium. *Science* **309**: 1242–1245.

Grote J, Thrash JC, Huggett MJ, Landry ZC, Carini P, Giovannoni SJ *et al.* (2012). Streamlining and core genome conservation among highly divergent members of the SAR11 clade. *MBio* **3**: e00252–12.

Havskum H, Thingstad TF, Scharek R, Peters F, Berdalet E, Sala MM *et al.* (2003). Silicate and labile DOC interfere in structuring the microbial food web via algal–bacterial competition for mineral nutrients: results of a mesocosm experiment. *Limnol Oceanogr* **48**: 129–140.

Hölzl G, Leipelt M, Ott C, Zähringer U, Lindner B, Warnecke D *et al.* (2005). Processive lipid galactosyl/glucosyltransferases from *Agrobacterium tumefaciens* and *Mesorhizobium loti* display multiple specificities. *Glycobiology* **15**: 874–886.

Icho T, Raetz CRH. (1983). Multiple genes for membrane-bound phosphatases in *Escherichia coli* and their action on phospholipid precursors. *J Bacteriol* **153**: 722–730.

Karl DM. (2014). Microbially mediated transformations of phosphorus in the sea: new views of an old cycle. *Ann Rev Mar Sci* **6**: 279–337.

Klug RM, Benning C. (2001). Two enzymes of diacylglyceryl-O-4'-(N,N,N,-trimethyl)homoserine biosynthesis are encoded by *btaA* and *btaB* in the purple bacterium *Rhodobacter sphaeroides*. *Proc Natl Acad Sci USA* **98**: 5910–5915.

Krom MD, Emeis KC, Van Cappellen P. (2010). Why is the Eastern Mediterranean phosphorus limited? *Prog Oceanogr* **85**: 236–244.

Laganowsky A, Reading E, Allison TM, Ulmschneider MB, Degiacomi MT, Baldwin AJ *et al.* (2014). Membrane proteins bind lipids selectively to modulate their structure and function. *Nature* **510**: 172–175.

Lidbury I, Murrell JC, Chen Y. (2014). Trimethylamine N-oxide metabolism by abundant marine heterotrophic bacteria. *Proc Natl Acad Sci USA* **111**: 2710–2715.

Martin P, Van Mooy BAS, Heithoff A, Dyhrman ST. (2011). Phosphorus supply drives rapid turnover of membrane phospholipids in the diatom *Thalassiosira pseudonana*. *ISME J* **5**: 1057–1060.

Minnikin DE, Abdolrahimzadeh H, Baddiley J. (1974). Replacement of acidic phosphates by acidic glycolipids in *Pseudomonas diminuta*. *Nature* **249**: 268–269.

Mira A, Ochman H, Moran NA. (2001). Deletional bias and the evolution of bacterial genomes. *Trends Genet* **17**: 589–596.

Moore CM, Mills MM, Arrigo KR, Berman-Frank I, Bopp L, Boyd PW *et al.* (2013). Processes and patterns of oceanic nutrient limitation. *Nat Geosci* **6**: 701–710.

Nakamura Y, Tsuchiya M, Ohta H. (2007). Plastidic phosphatidic acid phosphatases identified in a distinct subfamily of lipid phosphate phosphatases with prokaryotic origin. *J Biol Chem* **282**: 29013–29021.

Parsons JB, Rock CO. (2013). Bacterial lipids: metabolism and membrane homeostasis. *Prog Lipid Res* **52**: 249–276.

Pinhassi J, Gómez-Consarnau L, Alonso-Sáez L, Sala MM, Vidal M, Pedrós-Alio C *et al.* (2006). Seasonal changes in bacterioplankton nutrient limitation and their effects on bacterial community composition in the NW Mediterranean Sea. *Aquat Microb Ecol* **44**: 241–252.

Popendorf KJ, Fredricks HF, Van Mooy BAS. (2013). Molecular ion-independent quantification of polar glycerolipid classes in marine plankton using triple quadrupole MS. *Lipids* **48**: 185–195.

Popendorf KJ, Lomas MW, Van Mooy BAS. (2011). Microbial sources of intact polar diacylglycerolipids

in the Western North Atlantic Ocean. *Org Geochem* **42**: 803–811.

Riekhof WR, Andre C, Benning C. (2005). Two enzymes, BtaA and BtaB, are sufficient for betaine lipid biosynthesis in bacteria. *Arch Biochem Biophys* **441**: 96–105.

Riekhof WR, Naik S, Bertrand H, Benning C, Voelker DR. (2014). Phosphate starvation in fungi induces the replacement of phosphatidylcholine with the phosphorus-free betaine lipid diacylglyceryl-N,N,N-trimethylhomoserine. *Eukaryot Cell* **13**: 749–757.

Rusch DB, Halpern AL, Sutton G, Heidelberg KB, Williamson S, Yooseph S *et al.* (2007). The Sorcerer II Global Ocean Sampling expedition: northwest Atlantic through eastern tropical Pacific. *PLoS Biol* **5**: e77.

Schwalbach MS, Tripp HJ, Steindler L, Smith DP, Giovannoni SJ. (2010). The presence of the glycolysis operon in SAR11 genomes is positively correlated with ocean productivity. *Environ Microbiol* **12**: 490–500.

Sebastián M, Ammerman JW. (2009). The alkaline phosphatase PhoX is more widely distributed in marine bacteria than the classical PhoA. *ISME J* **3**: 563–572.

Sebastián M, Gasol JM. (2013). Heterogeneity in the nutrient limitation of different bacterioplankton groups in the Eastern Mediterranean Sea. *ISME J* **7**: 1665–1668.

Sebastián M, Pitta P, González JM, Thingstad TF, Gasol JM. (2012). Bacterioplankton groups involved in the uptake of phosphate and dissolved organic phosphorus in a mesocosm experiment with P-starved Mediterranean waters. *Environ Microbiol* **14**: 2334–2347.

Semeniuk A, Sohlenkamp C, Duda K, Hölzl G. (2014). A bifunctional glycosyltransferase from *Agrobacterium tumefaciens* synthesizes monoglucosyl and glucuronosyl diacylglycerol under phosphate deprivation. *J Biol Chem* **289**: 10104–10114.

Sturt HF, Summons RE, Smith K, Elvert M, Hinrichs K-U. (2004). Intact polar membrane lipids in prokaryotes and sediments deciphered by high-performance liquid chromatography/electrospray ionization multistage mass spectrometry–new biomarkers for biogeochemistry and microbial ecology. *Rapid Commun Mass Spectrom* **18**: 617–628.

Sunagawa S, Coelho LP, Chaffron S, Kultima JR, Labadie K, Salazar G *et al.* (2015). Structure and function of the global ocean microbiome. *Science* **348**: 1–10.

Swan BK, Tupper B, Sczyrba A, Lauro FM, Martinez-Garcia M, González JM *et al.* (2013). Prevalent genome streamlining and latitudinal divergence of planktonic bacteria in the surface ocean. *Proc Natl Acad Sci USA* **110**: 11463–11468.

Tanaka T, Zohary T, Krom MD, Law CS, Pitta P, Psarra S *et al.* (2007). Microbial community structure and function in the Levantine Basin of the eastern Mediterranean. *Deep Sea Res Part I Oceanogr Res Pap* **54**: 1721–1743.

Thingstad TF, Rassoulzadegan F. (1999). Conceptual models for the biogeochemical role of the photic zone microbial food web, with particular reference to the Mediterranean Sea. *Prog Oceanogr* **44**: 271–286.

Thingstad TF, Øvreås L, Egge JK, Løvdal T, Heldal M. (2005). Use of non-limiting substrates to increase size; a generic strategy to simultaneously optimize uptake and minimize predation in pelagic osmotrophs? *Ecol Lett* **8**: 675–682.

Thingstad TFR, Mantoura RFC. (2005). Titrating excess nitrogen content of phosphorus-deficient eastern Mediterranean surface water using alkaline phosphatase activity as a bio-indicator. *Limnol Oceanogr Methods* **3**: 94–100.

Tjellström H, Andersson MX, Larsson KE, Sandelius AS. (2008). Membrane phospholipids as a phosphate reserve: the dynamic nature of phospholipid-to-digalactosyl diacylglycerol exchange in higher plants. *Plant Cell Environ* **31**: 1388–1398.

Umena Y, Kawakami K, Shen J-R, Kamiya N. (2011). Crystal structure of oxygen-evolving photosystem II at a resolution of 1.9 Å. *Nature* **473**: 55–60.

Van Mooy BAS, Fredricks HF, Pedler BE, Dyhrman ST, Karl DM, Koblízek M *et al.* (2009). Phytoplankton in the ocean use non-phosphorus lipids in response to phosphorus scarcity. *Nature* **458**: 69–72.

Van Mooy BAS, Fredricks HF. (2010). Bacterial and eukaryotic intact polar lipids in the eastern subtropical South Pacific: water-column distribution, planktonic sources, and fatty acid composition. *Geochim Cosmochim Acta* **74**: 6499–6516.

Van Mooy BAS, Moutin T, Duhamel S, Rimmelin P, Van Wambeke F. (2008). Phospholipid synthesis rates in the eastern subtropical South Pacific Ocean. *Biogeosciences* **5**: 133–139.

Van Mooy BAS, Rocap G, Fredricks HF, Evans CT, Devol AH. (2006). Sulfolipids dramatically decrease phosphorus demand by picocyanobacteria in oligotrophic marine environments. *Proc Natl Acad Sci USA* **103**: 8607–8612.

Villanueva L, Bale N, Hopmans EC, Schouten S, Damsté JSS. (2014). Diversity and distribution of a key sulpholipid biosynthetic gene in marine microbial assemblages. *Environ Microbiol* **16**: 774–787.

Wagner-Döbler I, Biebl H. (2006). Environmental biology of the marine *Roseobacter* lineage. *Annu Rev Microbiol* **60**: 255–280.

Wu J. (2000). Phosphate depletion in the Western North Atlantic Ocean. *Science* **289**: 759–762.

Zavaleta-Pastor M, Sohlenkamp C, Gao J-L, Guan Z, Zaheer R, Finan TM *et al.* (2010). *Sinorhizobium meliloti* phospholipase C required for lipid remodeling during phosphorus limitation. *Proc Natl Acad Sci USA* **107**: 302–307.

Zubkov M V, Mary I, Woodward EMS, Warwick PE, Fuchs BM, Scanlan DJ *et al.* (2007). Microbial control of phosphate in the nutrient-depleted North Atlantic subtropical gyre. *Environ Microbiol* **9**: 2079–89.

Supplementary Information accompanies this paper on The ISME Journal website (http://www.nature.com/ismej)