# The Interplay Between DMSP and Glycine Betaine Regulation in *Labrenzia aggregata*

**Ocean Ellis**

University of East Anglia

A thesis submitted February 2023 to the University of East Anglia for the degree of Doctor of Philosophy. 

**Abstract**

Dimethylsulfoniopropionate (DMSP) is a key marine nutrient and forms part of the global biogeochemical sulfur cycle. Glycine betaine (GB) is a nitrogen-containing structural homolog of DMSP produced by a wide range of organisms. Both DMSP and GB are compatible solutes produced in response to osmotic stress. Despite the reported importance and abundance of these compounds, their proposed role in bacterial osmoprotection has so far not been confirmed via mutagenesis of synthesis genes. In this work, the model DMSP-producing bacterium *Labrenzia aggregata* was investigated as it produces both DMSP and GB. Mutagenesis, transcriptomics and metabolomics were used to understand the interplay between DMSP and GB production and their roles as osmoprotectants in *L. aggregata*.

*L. aggregata* growth conditions were screened for those that affected the production of DMSP and GB. Subsequent transcriptomic analysis determined that the key DMSP and GB synthesis genes (*dsyB* and *gsdmt*, respectively) were highly regulated by salinity. In order to assess the osmoprotective roles of these two compatible solutes, each synthesis gene was knocked out in *L. aggregata*. The Δ*gsdmt* strain had abolished GB production and a reduced growth rate when under osmotic stress. The Δ*dsyB* strain was unable to produce DMSP, however did not exhibit a reduced growth phenotype under high salinity conditions, in concurrence with the existing literature. However, DMSP was indicated to potentially have an osmoprotective role, as the reduced growth phenotype of an *L. aggregata* Δ*dsyB*Δ*gsdmt* double mutant in high salinity conditions was partially rescued through genetic complementation with *dsyB.*

Finally, transcriptomic analysis indicated a putative compatible solute regulator gene and several putative DMSP biosynthesis genes in *L. aggregata* which were partially characterised using mutagenesis. This study provides tentative support for the role of DMSP as an osmoprotectant and enhances understanding of the relationship between DMSP and GB in *L. aggregata*.

**Acknowledgements**

Firstly, a big thank you to Professor Jonathan Todd. Thank you for all of your help and guidance over the past four years. You always had time for me, no matter how busy you were.

Thank you to all of the members of the Todd lab, past and present for being the very best people to work with. Many thanks in particular to Andy, Ana, Peter and Beth for helping me so much when I first started and answering all of my seemingly endless questions - you all taught me so much.

I would like to thank Dr. Simon Moxon for his help with the RNA-Seq analysis and to Dr. Rocky Payet for his help with my protein work. Many thanks to Jinyan Wang for being a great neighbour, a ray of sunshine and for her help with the RT-qPCR work.

Special thanks to Ornella, Beth and Kasha for their help, not only with microbiology, but for being the most wonderful friends. I really would not have got this far without you. Thank you. Really.

Thank you also to those who have made Norwich a great place to be for the past four years, even throughout several lockdowns (Lauren, Alan and Tom). Thank you also to my parents and family for always supporting me. Finally, thank you to Jonny for being so kind and supportive, especially during the past few months of writing.

This work was funded by the UKRI-BBSRC Norwich Research Park Biosciences Doctoral Training Partnership.

**Author's Declaration**

The research described in this thesis was conducted at the University of East Anglia between October 2018 and February 2023. All data described here are original and were obtained by the author unless otherwise attributed in the text. No part of this thesis has previously been submitted for a degree at this or any other academic institution.

## List of Figures

## List of Tables

Table of Contents

# Chapter 1: Introduction

## 1.1 Stresses Facing Marine Microorganisms

Microorganisms are ubiquitous in marine environments; a single drop of seawater contains approximately a million microbial cells (Sunagawa et al., 2015). Despite their diminutive size, these microbes are vital for catalysing the chemical reactions that form parts of the carbon, nitrogen, sulfur and iron biogeochemical cycles (Dang et al., 2019). As such, marine microbes are indispensable for the maintenance of the habitability on Earth (Zehr et al., 2017).

As marine ecosystems cover 70 % of the Earth's surface, their environmental factors can vary wildly. For example, the temperature can range from ~2 °C in ice-covered seas to over 100 °C in hydrothermal vents (Cavicchioli et al., 2019; Sunagawa et al., 2015). However, on a smaller scale, there can still be extreme fluctuations in salinity, pressure and temperature within marine environments, and microorganisms must sense and respond to these fluctuations in order to survive. One way in which microorganisms achieve this is through the production of low molecular weight molecules, termed compatible solutes.

## 1.2 Compatible Solutes

When microorganisms are exposed to high osmolarity environments, water moves from the cell into the environment, causing a reduction in turgor and dehydration of the cytoplasm (Kempf & Bremer, 1998). To counteract this, microorganisms can utilise two strategies; the first is through the accumulation of inorganic ions, such as $K^+$ and $Na^+$, which can cause disruption to cell metabolism. Alternatively, the second strategy is to increase their intracellular solute pool by accumulating organic osmolytes, called compatible solutes. Fundamentally, compatible solutes can be described as organic osmolytes that are responsible for osmotic balance but are also compatible with the metabolism of cells (Galinski, 1993).

Compatible solutes serve a dual role in terms of osmoprotection in microorganisms. This is due to their ability to stabilise proteins and other cell components, acting to protect them from the denaturing effects of high salt concentrations. There are several theories that go some way to explain how these osmolytes stabilise proteins, including the 'preferential exclusion model', which states that compatible solutes are excluded from the hydration shell of proteins (the hydration shell is the interaction of the protein surface with the surrounding water, and is fundamental to protein activity) (Arakawa & Timasheff, 1985; Bohnert & Shen, 1998). According to this model, the compatible solutes interact with the bulk water in the cytoplasm, and not with the hydration shell of the proteins. As such, there is preferential hydration of the protein, and resultingly, it is forced to occupy a smaller volume in order to minimize its exposed surface. This forces a more compact, native-like conformation of the protein, and resultingly, an increase in stability (Kolp et al., 2006; Timasheff, 2002).

Due to their ability to stabilise cellular macromolecules, several compatible solutes can provide protection against other stresses, such as temperature and oxidative stress. The accumulation of compatible solutes by microorganisms is achieved either through synthesis or uptake from their environments. Usually, a range of compatible solutes is utilised for osmoregulation. This not only allows cells to withstand changes in osmolarity, but also expands the ecological niches in which these microorganisms can survive.

Different microorganisms accumulate different compatible solutes. For example, bacteria predominantly accumulate (either via synthesise or active uptake from the environment) compatible solutes including: sucrose, trehalose, glucosylglycerol, glutamate, proline, *N*-Acetylglutaminylglutamine amide, glycine betaine, ectoine/hydroxyectoine, and dimethylsulfoniopropionate (Kiene et al., 2000; Welsh, 2000). Whereas, commonly accumulated compatible solutes in micro-algae include: sucrose, glycerol, mannitol, proline, glycine betaine and dimethylsulfoniopropionate (Welsh, 2000).

Generally, compatible solutes have an overall neutral charge, being either zwitterionic or uncharged at physiological pH levels (Yancey, 2005). Further, compatible solutes can largely be categorised into four classes: sugars or sugar alcohols; quaternary ammonium

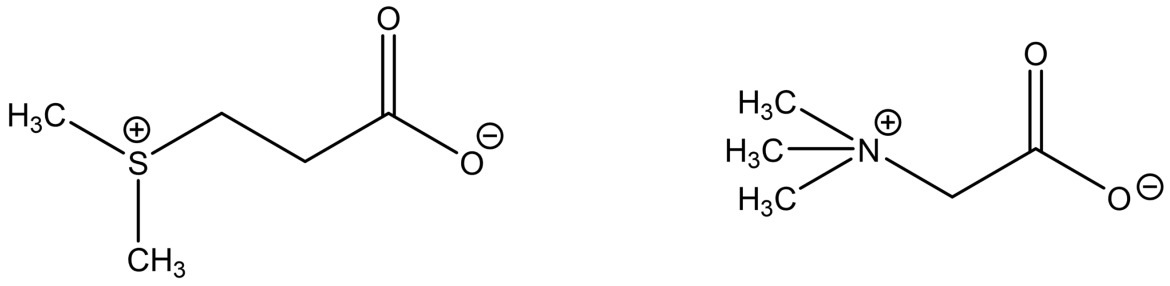
compounds; tertiary sulfonium compounds, and amino acids and their derivatives (Slama et al., 2015). The two major compounds discussed in this thesis are dimethylsulfoniopropionate (DMSP) and glycine betaine (GB), which fall into the classes of tertiary sulfonium compounds and quaternary ammonium compounds, respectively.

## 1.3 DMSP and GB – Overview and Structures

Within marine environments, the major osmolytes at play are the structurally homologous compatible solutes, DMSP and GB. DMSP is a hugely abundant organosulfur compound and has an important role in the biogeochemical sulfur cycle as it is also the major precursor of the climate-active gas, dimethylsulfide (DMS). It is produced by a wide range of organisms spanning different kingdoms of life, including certain bacteria, algae, phytoplankton, angiosperms and an animal (coral) (Curson et al., 2017; Raina et al., 2013; Stefels, 2000). Previous estimates of total annual DMSP production have reached figures as high as two billion tons (Ksionzek et al., 2016).

GB has been termed the only true universal compatible solute, due to its presence in all three domains of life; it is produced by species of bacteria, archaea, plants and algae (Empadinhas & da Costa, 2008).

Structurally, both DMSP and GB are betaines containing a carboxylic acid group. Both compatible solutes are five-carbon zwitterionic compounds, but DMSP contains a positively charged sulfur atom bound to two methyl groups, whereas GB contains a positively charged nitrogen atom bound to three methyl groups (Figure 1.1) (Yoo et al., 2011).




Dimethylsulfoniopropionate (DMSP) Glycine Betaine (GB)

Figure 1.1 The molecular structures of dimethylsulfoniopropionate and glycine betaine.

## 1.4 Roles of DMSP and GB

The similarities between DMSP and GB are not limited to structure, as they also carry out many of the same roles in the organisms that accumulate them. The roles of GB are perhaps better studied, and as such many of the roles of GB are conclusively experimentally confirmed. Conversely, many of the roles of DMSP remain proposed roles based on evidence related to the correlation of accumulation of the organosulfur compound under different environments. For the purpose of this introduction, the explanation of the following proposed roles of DMSP and GB focus mainly on prokaryotic studies.

### 1.4.1 Osmoprotection

DMSP's involvement in protecting against osmotic stress was first proposed in the macroalgae *Ulva lactuca*, which when subjected to hypo-osmotic shock, decreased its intracellular concentrations of DMSP (Dickson et al., 1982). Further studies in many phytoplankton species have shown accumulation of high intracellular DMSP concentrations in response to highly saline environments (Karsten et al., 1990). Other experiments have also been carried out in bacterial species, for example the growth of *Escherichia coli* under high osmotic conditions was shown to be enhanced by DMSP uptake (Cosquer et al., 1999; Wolfe, 1996).

GB is utilised as an osmoprotectant in many organisms, spanning all domains of life (Empadinhas & da Costa, 2008). There have been some reports that the widespread utilisation of GB is partly due to the superior osmoprotection offered by GB compared to other compatible solutes (Csonka, 1989).

### 1.4.2 Osmoprotection in nitrogen deplete conditions

In the marine environment, sulfur is highly abundant, in contrast to nitrogen which is thought to be a limiting factor for growth. Therefore, it has been predicted that in marine environments DMSP accumulation may be favoured over the accumulation of nitrogen containing osmolytes, such as GB (Andreae, 1986). Van Diggelen and colleagues observed a positive correlation between GB accumulation and nitrogen availability (Van Diggelen et

al., 1986). Contrastingly there was a negative correlation between DMSP accumulation and nitrogen availability in the grass *Spartina anglica* (Van Diggelen et al., 1986). Further studies have found increases in DMSP production under nitrogen deplete conditions in coccolithophore and diatom species (Kettles et al., 2014; Wördenweber et al., 2018).

### 1.4.3 Cryoprotection

The accumulation of highly soluble compatible solutes helps to protect prokaryotic cells from low temperatures by reducing the intracellular freezing point of the cytoplasm (Casanueva et al., 2010). Additionally, as aforementioned, these compounds also maintain the hydration sphere of proteins and therefore stabilise the structure of enzymes, adding to their protection properties in low temperatures (Lyon & Mock, 2014; Welsh, 2000). Evidence to suggest a role for DMSP in cryoprotection includes many studies that show the accumulation of the molecule is influenced by temperature. For example, temperature-dependent accumulation of DMSP has been observed in several types of algae, such as the marine unicellular alga, *Tetraselmis subcordiformis,* which achieved higher intracellular DMSP concentrations when they were cultured in lower temperatures (Karsten et al., 1996; Sheets & Rhodes, 1996). Further, DMSP levels were found to be much higher in macroalgae species from Antarctica than in similar species from temperate regions (Karsten et al., 1990). In addition to these accumulation-based studies, DMSP has been shown to stabilize enzymes at low temperatures in a polar alga (Karsten et al., 1996).

GB was first reported to be used as a bacterial cryoprotectant in the opportunistic pathogen *Listeria monocytogenes,* as it was rapidly imported from growth media and accumulated in response to cold stress (Ko et al., 1994; Mendum & Smith, 2002). In the time since this finding, GB has been proposed as an effective, universal cryopreservative that can be used for the preservation of prokaryotes in lab-based settings (Cleland et al., 2004).

### 1.4.4 Antioxidant

The first indication that compatible solutes may have antioxidant activity, was a study in 1989 which found that several compatible solutes had hydroxyl radical scavenging activity (Smirnoff & Cumbes, 1989). This is the case for DMSP which acts as a scavenger of several reactive oxygen species (ROS) (Sunda et al., 2002). Further, intracellular DMSP concentrations have been shown to increase under oxidative stress (Sunda et al., 2002). However, GB has reported to be an ineffective scavenger of reactive oxygen species, including hydroxyl radicals (Smirnoff & Cumbes, 1989). In plants, it is predicted that GB protects against oxidative stress by having an activation role in other pathways used for ROS defence (Fariduddin et al., 2013). However, exactly how DMSP and GB function to protect against oxidative stress is not understood. Further, the antioxidant functions of DMSP and GB are mainly proposed due to accumulation of these molecules under conditions known to cause oxidative stress.

### 1.4.5 Barotolerance

A possible role for compatible solutes, including GB, in bacterial barotolerance was first proposed in 2004 (Smiddy et al., 2004). Evidence for GB's role in barotolerance include the heterologous expression of a betaine uptake transport system (*betL* from *E.coli)* in *Lactobacillus salivarius* conferring a significant increase in resistance to high pressures (Sheehan et al., 2006).

DMSP was first implicated in protection against high pressure in 2020, when several DMSP producing bacterial species were isolated from the Mariana Trench. When DMSP production was knocked out in these species via mutagenesis, the mutants were much less tolerant of deep ocean pressures (60 MPa) than the wild type strains (Zheng et al., 2020). Further, the transcripts of genes involved in DMSP synthesis were found to increase with depth, as pressure increased (Zheng et al., 2020). This study laid out clear links between DMSP and protection against hydrostatic pressure, however the mechanism via which DMSP offers protection against high pressure was not discussed.

### 1.4.6 Nutrient Source

Both DMSP and GB are important sources of nutrients, and therefore energy. In fact, DMSP is thought to support up to 13 % of bacterial carbon demand in surface waters (Kiene et al., 1999). DMSP synthesis is estimated to account for 3 – 10 % of global marine primary production of organic carbon (Kiene et al., 2000). Additionally, in DMSP-producing species of algae such as the prymnesiophyte *Emiliania huxleyii,* DMSP can comprise 50 – 100 % of the total intracellular organic sulfur (Stefels, 2000). It is also reported that both DMSP and GB can be used by numerous bacterial species as a sole carbon source (Curson et al., 2011; Kortstee, 1970; Liu et al., 2022; Wargo et al., 2008). The various catabolism pathways of DMSP and GB, and details of how these contribute to various nutrient sources for prokaryotic cells are further discussed in sections 1.6 and 1.9 of this introduction.

## 1.5 DMSP Synthesis

Until relatively recently, it was thought that DMSP production was limited to marine eukaryotes such as phytoplankton (including coccolithophores, diatoms and dinoflagellates), and more complex organisms including some species of higher plants and corals (Barnard et al., 1984; Dickschat et al., 2015; Dickson et al., 1980; Gage et al., 1997; Hanson et al., 1994; Raina et al., 2013; Uchida et al., 1996). More recently, DMSP synthesis was found to occur in several bacterial species, the first discovered being the marine alphaproteobacterial *Labrenzia aggregata* (Curson et al., 2017). Following this finding, many more DMSP-producing bacterial species have been identified (Williams et al., 2019).

There are three known DMSP biosynthesis pathways, all of which begin with the amino acid methionine (Met) (Gage et al., 1997; Kocsis & Hanson, 2000; Rhodes et al., 1997; Uchida et al., 1996) (**Figure 1.2**). Different organisms utilise different DMSP biosynthesis pathways, each of which is named after the first step of the pathway (transamination, methylation, and decarboxylation). The existence of these three separate biosynthesis pathways, which all differ from each other at every step, suggests that DMSP biosynthesis must have evolved independently at least three times (Stefels, 2000), which is indicative that the compound confers considerable ecological benefits to the producer.

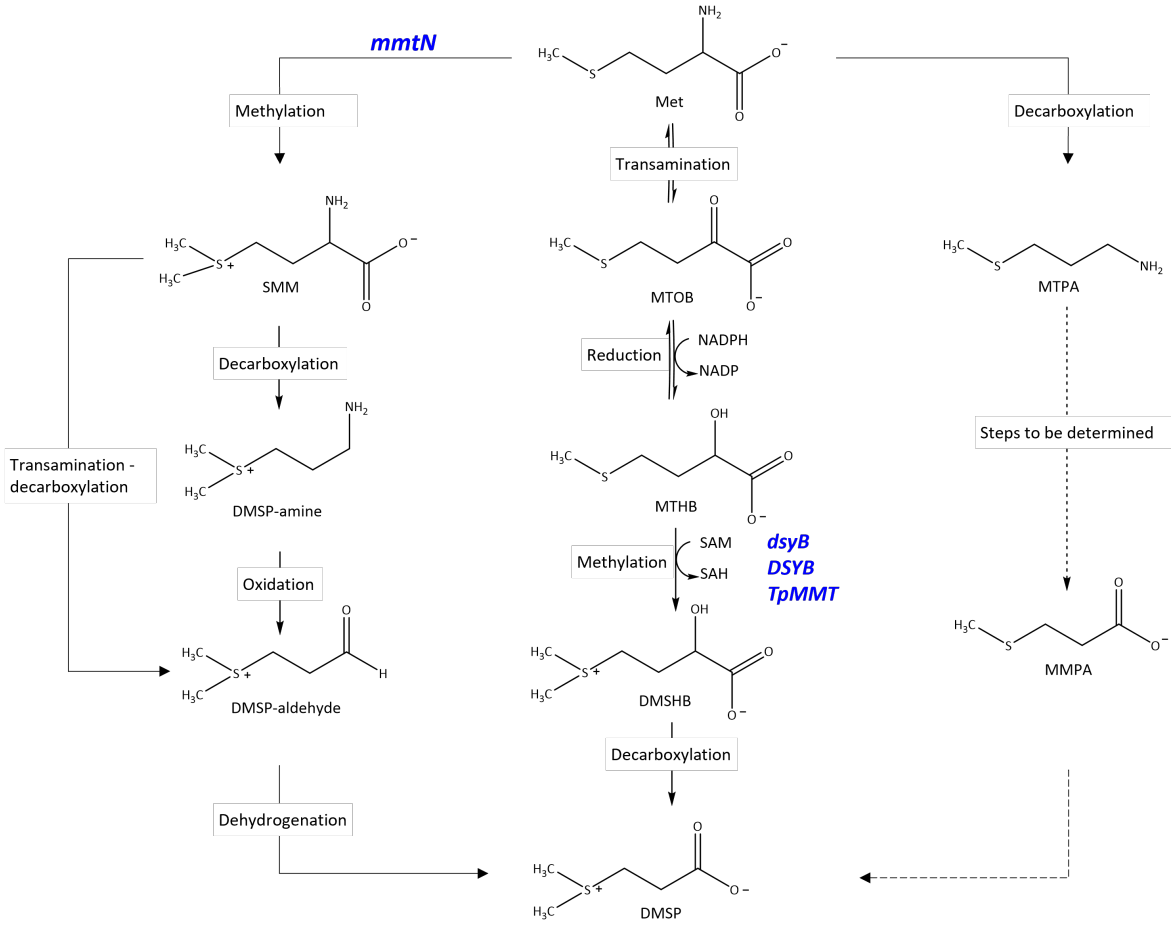



Figure 1.2 The three DMSP synthesis pathways.
The methylation pathway (left) is used by some species of bacteria and higher plants such as *Spartina*. In Wollastonia SMM is converted directly to DMSP-aldehyde. The transamination pathway (centre) is used by macroalgae, diatoms, prymnesiphytes, prasinophytes and bacteria (including *L. aggregata*). The methylation pathway (right) has so far been identified in one dinoflagellate. Figure adapted from (Curson et al., 2017). The genes known to be involved in the pathways are written in blue.

### 1.5.1 The Transamination Pathway

The transamination DMSP biosynthesis pathway is thought to be used by marine algae (both green and red), diatoms and some bacteria (Curson et al., 2017; Dickschat et al., 2015; Williams et al., 2019). This pathway was first elucidated using radiolabelled compounds and metabolite chasing experiments by Gage and colleagues (Gage et al., 1997). It begins with the transamination of methionine to form the unstable 2-oxo acid 4-methylthio-2-oxybutyrate (MTOB), the first intermediate of the pathway (Gage et al., 1997). This reaction is believed to be carried out by an NADPH-dependent aminotransferase, however the gene encoding this enzyme is yet to be identified in any organism. The second step in the pathway is the reduction of MTOB to 4-methylthio-2-

hydroxybutyrate (MTHB). Next, an *S*-methylation reaction converts MTHB to 4-dimethylsulfonio-2-hydroxybutyrate (DMSHB), which in turn undergoes an oxidative decarboxylation reaction to give rise to DMSP. The intermediates within this pathway were determined through radiolabelling experiments, in which $^{35}$S-labelled methionine was supplied to the green alga, *Enteromorpha intestinalis.* The intermediates MTHB and DMSHB were found to rapidly acquire $^{35}$S, confirming their role as DMSP synthesis intermediates. Gentle extraction methods confirmed that $^{35}$S was also incorporated into the first intermediate, MTOB. Whilst carrying out the aforementioned radiolabelling experiments, it was shown that MTHB could be converted back to methionine (Gag*e et al*., 1997).

Further work on the transamination pathway, carried out by Summers and colleagues, characterised the substrate specific activities of the enzymes responsible for catalysing the first three steps in the transamination synthesis pathway (Summers et al., 1998). However, this *in vitro* study on algal cell extracts did not identify the enzymes involved in this pathway. The amino acceptor of the methionine aminotransferase was found to be 2-oxoglutatarate, demonstrating that the enzyme responsible for the catalysis of the first reaction is a 2-oxoglutarate-dependent aminotransferase (Summers et al., 1998). Further, the MTOB reductase in the synthesis pathway is an NADPH-dependent reductase. The enzyme which catalyses the third step in the pathway; the methylation of MTHB to DMSHB was found to be an AdoMet dependent methyltransferase (Summers et al., 1998).

This work also highlighted that as the transamination of MTOB to Met is the last step in the ubiquitous Met salvage pathway, all algae have Met aminotransferase activity. The Met salvage pathway is found in many types of organisms and is responsible for the recycling of sulfur-containing metabolites into Met (Albers, 2009). However, the Met aminotransferase activity of DMSP-producing strains of algae was found to be up to 100-fold higher than the activity non-DMSP producing strains (Summers et al., 1998). Summers and colleagues also highlighted that the Met aminotransferase involved in DMSP synthesis is likely to be a novel enzyme rather than an overexpressed housekeeping gene as the estimated $K_m$ of the Met aminotransferase in the DMSP synthesis pathway is 30 μm, which is much lower than usual

for aminotransferases which usually have a $K_m$ of several mM (Jenkins, 1985; Summers et al., 1998).

Similarly, MTHB formation may be a side reaction of the Met salvage pathway as radiotracer findings in non-DMSP synthesizing algae have found MTHB to be a metabolite of Met (Pokorny et al., 1970; Summers et al., 1998). However, the methylation of MTHB to DMSHB is only known to be carried out in DMSP synthesis and is the committing step in this pathway (Gage et al., 1997).

Recently, Curson and colleagues identified that a number of species of heterotrophic bacteria, including the *Alphaproteobacterium Labrenzia aggregata,* produce DMSP likely through the same transamination synthesis pathway used by macroalgae and phytoplankton (Curson et al., 2017). This was an important finding as this indicates DMSP production may not be limited to the photic zone of marine environments and as such, estimates of DMSP production may have been underestimated.

In addition, the first ever DMSP synthesis gene was identified in this study. This gene, termed *dsyB*, encodes an AdoMet-dependent *S*-methyltransferase (DsyB). DsyB is responsible for the catalysis of the rate limiting and committing step of the pathway; the conversion of MTHB to DMSHB (Curson et al., 2017; Summers et al., 1998). Within this study, *Rhizobium leguminosarum* was used as a heterologous host and when *dsyB* was expressed in *R. leguminosarum*, it gained the ability to produce DMSP. This showed that surprisingly, *R. leguminosarum* had DMSHB decarboxylase activity and that unlike MTHB *S*-methylation, this step is not confined to DMSP producers. It also showed that the presence of *dsyB* can confer the ability to synthesise DMSP to certain organisms. Further, the *dsyB* gene was mutated in *L. aggregata* LZB033, resulting in a complete abolishment of DMSP synthesis. However, a growth phenotype was not found when the Δ*dsyB* mutant was grown under varying salinity, temperature, levels of oxidative stress or nitrogen availability (Curson et al., 2017).

Highly similar homologues (≥ 39 % identity) of the *L. aggregata* LZB033 *dsyB* gene were found to exist in more than fifty strains of *Alphaproteobacteria* and several of these were

experimentally proven to encode functional MTHB methyltransferases (Curson et al., 2017). Further, analysis of existing metagenomic datasets revealed that *dsyB* is present in 0.5% of the bacteria sampled. This, coupled with the knowledge that *dsyB* confers the ability to synthesise DMSP, means that DMSP is now thought to be produced by 0.5% of bacteria in marine environments (Curson et al., 2017).

Leading on from this study, functional eukaryotic *dsyB* homologues were identified in many phytoplankton and corals, termed *DSYB* (Curson et al., 2018). Several *DSYB* genes were cloned from corals, diatoms, dinoflagellates and prymensiophytes and found to have similar levels of MTHB methyltransferase activity to the bacterial *dsyB* gene from *L. aggregata*. The cloned eukaryotic *DSYB* genes were also able to fully complement bacterial *dsyB*$^-$ mutants, which otherwise had abolished DMSP production (Curson et al., 2018). To further study *DSYB,* several *DSYB* containing strains were grown under differing environmental conditions. For example, in *Prymnesium parvum* the expression of *DSYB* and the levels of DMSP production were increased by higher salinities, but were unaffected by other conditions tested, including temperature and nitrogen availability (Curson et al., 2018). Contrastingly, in the polar ice diatom *Fragilariopsis cylindrus, DSYB* expression and DMSP production were increased with nitrogen limitation and increased salinity. In all of the strains tested, DMSP concentration increased with *DSYB* transcription levels (Curson et al., 2018). Additionally, analysis of marine metatranscriptomes revealed that eukaryotic *DSYB* transcripts were approximately twice as abundant as those for bacterial *dsyB* (Curson et al., 2018).

In addition to this, the gene encoding the MTHB methyltransferase in the diatom *Thalassiosira pseudonana,* termed *TpMMT* has also been identified (Kageyama et al., 2018). This gene has low similarity to *dsyB* (12 %) and only has a close homologue in one other species (*Thalassiosira oceanica)* (Kageyama et al., 2018). The expression of *TpMMT* was found to increase when *T. pseudonana* was grown under salt stress and in nitrogen limited conditions (Kageyama et al., 2018). In all of the aforementioned organisms that utilise the transamination synthesis pathway, the genes encoding the remaining enzymes in the pathway are yet to be identified. In many cases, compatible solute synthesis genes are located in a gene cluster, however this does not appear to be the case for the genes

involved in the DMSP transamination synthesis pathway. The gene neighbourhood of *dsyB* has been described in many species of *Alphaproteobacteria,* and the surrounding genes do not appear to encode proteins with either Met aminotransferase, MTOB reductase or DMSHB decarboxylase activity (Curson et al., 2017). Thus, no candidate Met aminotransferase, MTOB reductase or DMSHB decarboxylase genes have been identified in any organism.

### 1.5.2 The Methylation Pathway

The methylation pathway was first elucidated in the DMSP-producing higher plant, *Wollostonia biflora* (commonly known as the 'sea daisy') (Hanson et al., 1994). Hanson and colleagues utilised radiolabelling of L-Met to define the first step of this synthesis pathway. It was determined that Met is methylated to *S*-methylmethionine (SMM) (Hanson et al., 1994). This is a S-adenosyl methionine (SAM) dependent reaction, catalysed by the enzyme, *S*-adenosylmethionine:methionine *S*-methyltransferase (MMT) (James et al., 1995). This methylation reaction takes place in the cytosol and the produced SMM is then transported to the chloroplast where it undergoes conversion to the next intermediate in the pathway, DMSP-aldehyde (DMSP-ald). This step in the pathway that is specific to DMSP synthesis. It is known that the conversion of SMM to DMSP-ald involves a transamination and decarboxylation reaction, but no more is known about these steps in the pathway (Hanson & Gage, 1996; Hanson et al., 1994; James et al., 1995). Finally, an oxidation reaction catalysed by a NAD-dependent dehydrogenase converts DMSP-ald to DMSP (Trossat et al., 1996).

In the saltmarsh grass *Spartina alterniflora*, a similar pathway has been elucidated although the genes involved in DMSP synthesis in plants are yet to be identified (Kocsis et al., 1998). In this pathway, DMSP-amine was identified as an intermediate between SMM and DMSP-aldehyde. It is predicted that a decarboxylase enzyme catalyses the conversion of SMM to DMSP-amine, and then an oxidase catalyses the turnover to DMSP-ald (Kocsis et al., 1998). The presence of DMSP-amine as an intermediate in *S. alterniflora,* and the lack of it in the *W. biflora* methylation pathway is indicative that the pathway from SMM to DMSP-ald has evolved independently in these organisms (Kocsis et al., 1998).

In 2018 Williams and colleagues also found that some bacterial species produce DMSP via the methylation pathway (Williams et al., 2019). The *Alphaproteobacterium*, *Novosphingobium* sp. BW1 was isolated from a salt marsh and found to produce DMSP independently of *dsyB*, and its DMSP production was significantly enhanced by the addition of the intermediates of the methylation pathway but not the transamination nor decarboxylation pathways. Screening of a BW1 genomic library led to the identification of the *mmtN* gene, which confers methionine *S*-methyltransferase (MMT) activity. The *mmtN* gene is the only identified gene in the methylation pathway and is thought to be a robust reporter gene for the ability to produce DMSP, with a large number of functionally ratified homologues in alphaproteobacteria, actinobacteria and one gammaproteobacterium. Further, a *mmtN* knockout mutant in the alphaproteobacterium, *Thalassospira profundimaris* caused the complete loss of DMSP production. However, in a similar manner to the LZB033 *dsyB*- mutant, no significant growth reduction or stress protection phenotype (including reduced salt tolerance) was found in this strain. This study predicted 1.1% of marine and saltmarsh bacteria are DMSP producers (containing *dsyB* or *mmtN*) (Williams et al., 2019).

### 1.5.3 The Decarboxylation Pathway

The least studied DMSP biosynthesis pathway is the decarboxylation pathway, which is thought to be carried out by one dinoflagellate, *Crypthecodinium cohnii* (Uchida et al., 1996). The first step in this reaction is catalyzed by a PLP-dependent methionine decarboxylase which converts methionine to methylthiopropylamine (MTPA) (Kitaguchi et al., 1999). The following steps in the pathway are yet to be determined, although likely involve 3-methylthiopropionate (MMPA) as a later intermediate (Uchida et al., 1996).

## 1.6 DMSP Catabolism

Much of the DMSP produced by microorganisms is eventually released into the marine environment as DMSP is released through cell lysis, for example the senescence of phytoplankton that occurs due to grazing. The released DMSP is then imported by a diverse range of algae and bacteria and subsequently utilized as a compatible solute, or a source of carbon, sulfur or energy (Curson et al., 2011; Kellogg et al., 1972; Simó & Pedrós-Alió,

1999; Sun et al., 2012; Tripp et al., 2008). Thus, that utilisation of DMSP's anti-stress properties is not limited to strains with the ability to biosynthesise the compound. For example, it is well known that many non-DMSP producing strains take up DMSP from their environment in response to salinity stress (Wolfe., 1996). Additionally, some non-DMSP producing strains (including members of the abundant SAR11 clade) are able to take up DMSP from the environment and metabolise it.

There are three pathways via which bacteria can catabolise DMSP: the demethylation pathway, the oxidation pathway and the cleavage pathway (Curson et al., 2011; Kiene, 1996; Thume et al., 2018) (**Figure 1.3**).

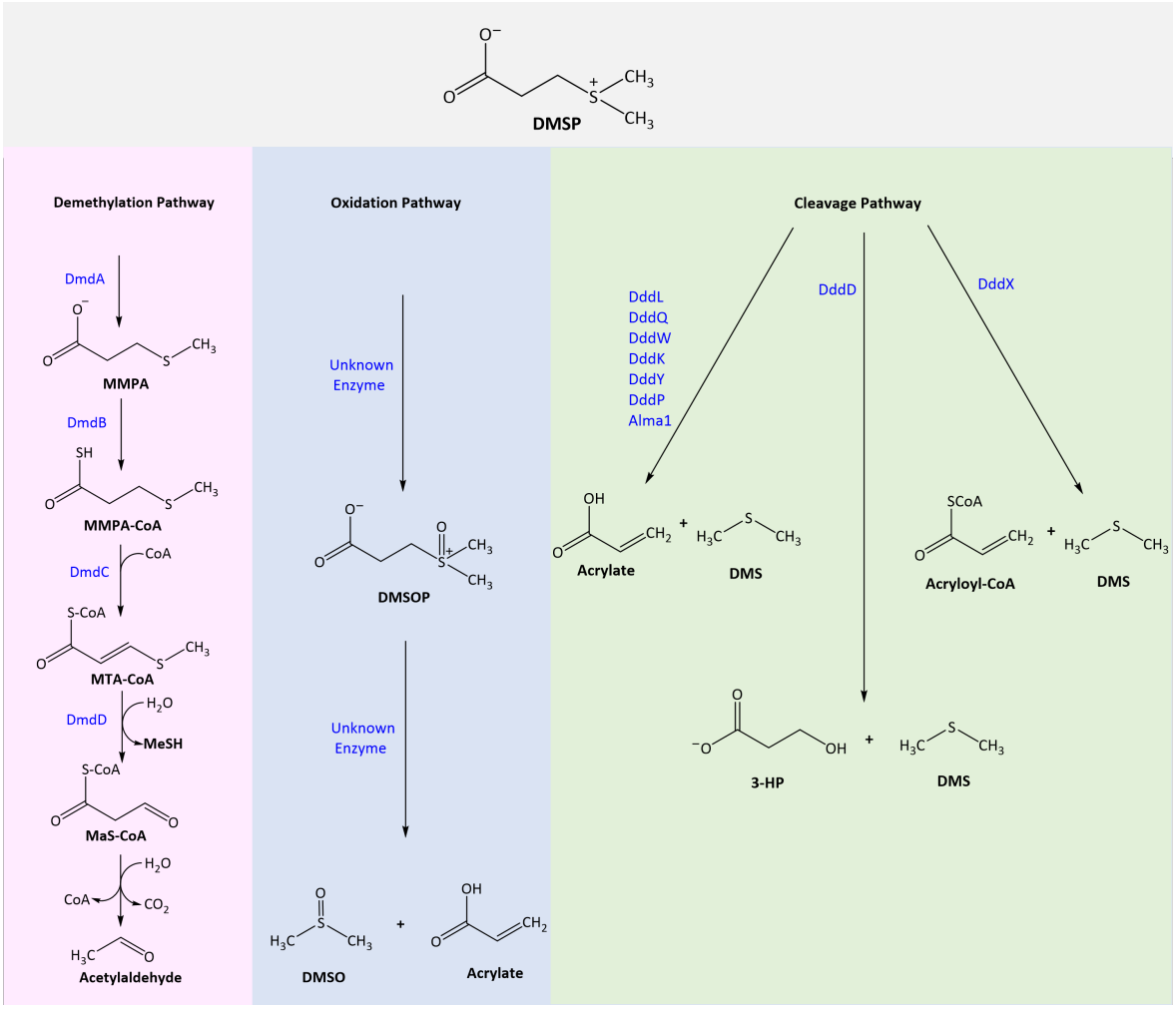



Figure 1.3 The three DMSP catabolism pathways.
The demethylation pathway (in pink) is catalyzed by the DmdABCD genes and produces MeSH and acetylaldehyde. The enzymes involved in the oxidation pathway (in blue) are currently unknown. The products of the cleavage pathway (in green) slightly differ depending on the enzyme catalysing the reaction, but DMS is always yielded from this pathway.

### 1.6.1 DMSP Demethylation

The majority of environmental DMSP (approximately 75%) is generally thought to be degraded via the demethylation pathway which is encoded for by the *dmdA*, *dmdB*, *dmdC* and *dmdD* genes. In this pathway (**Figure 1.3**), DMSP is first converted to 3-methiolproprionate (MMPA) via an initial demethylation reaction catalyzed by a DMSP demethylase encoded by the *dmdA* gene (Howard et al., 2006). This is the key step of the demethylation pathway and as such, *dmdA* is used as a reporter gene for DMSP demethylating bacteria. *dmdA* homologs are divided into five clades and 14 subclades (Varaljay et al., 2012). Analysis of metagenomic data has predicted *dmdA* to be widely distributed in marine environments and present in ~60% of marine bacteria (Howard et al., 2006; Howard et al., 2008)

In most cases, the produced MMPA then undergoes demethiolation, which involves a series of three reactions, all mediated by coenzyme A (CoA). This demethiolation pathway has been described as a fatty acid β-oxidation like process. The gene *dmdB* is responsible for the first reaction in which MMPA yields an MMPA-CoA thioester. Subsequently, *dmdC* is responsible for the second reaction in which a double bond is formed to create methylthioacrylyl (MTA)-CoA. Finally, *dmdD* is responsible for the final steps in this pathway. DmdD catalyses the hydration of MTA-CoA, which involves the incorporation of $H_2O$ and the liberation of methanethiol (MeSH), to form the intermediate, malonate semialdehyde (MAS)-CoA (Tan et al., 2013). Following this, a hydrolysis reaction liberates CoA from the rest of the molecule and it is predicted that MAS-CoA undergoes spontaneous decomposition to form an aldehyde, releasing $CO_2$ (Reisch et al., 2011; Tan et al., 2013).

Ultimately, MeSH is the sulfur-containing product of the DMSP demethylation pathway. The sulfur from MeSH is readily incorporated into proteins and as such this fate of DMSP is important in assimilatory metabolism (González et al., 1999; Moran et al., 2012).

There is an alternative fate for MMPA which has been found to occur in some cultures of marine bacteria in which MMPA undergoes a second demethylation to 3-mercaptoproprionate (MPA). However, these bacteria are far less abundant in seawater

compared to those that demethiolate MMPA to form MeSH (Moran et al., 2012; Visscher et al., 1992).

### 1.6.2 The DMSP Cleavage Pathway

The second pathway via which DMSP is catabolised is the cleavage pathway, of which DMS is always a product (**Figure 1.3).** The other product of this pathway is either 3-hydroxypropionate (3-HP) or acrylate, depending on the enzyme catalysing the reaction (Curson et al., 2011; Kirkwood et al., 2010; Todd et al., 2007).

In bacteria, the enzymes that catalyse the lysis of DMSP to DMS are all termed Ddd enzymes and seven *ddd* genes have been identified thus far. The first *ddd* gene to be identified was *dddD*, which differs from the other identified *ddd* genes as the products of the lysis reaction it catalyses are DMS and 3-HP, rather than acrylate (Todd et al., 2010; Todd et al., 2007). However, it is thought that 3-HP isn't the initial product in this reaction, as a CoA derivative of DMSP is likely formed before being converted to 3-HP. This prediction is based on the similarity of CoA-transferase domains in DddD to those in the *E.coli* protein CaiB (which converts CoA to cartinine in a similar way) (Todd et al., 2007). The *dddD* gene is present in several species of marine bacteria and in many cases, the gene is located nearby in the genome to predicted DMSP transporters (Sun et al., 2012).

The remaining DMSP lyases split DMSP into DMS and acrylate. DddK, DddL, DddQ, DddY and DddW are all small polypeptides which contain carboxy-terminal domains that likely form cupin pockets, and thus bind to transition metal ions (Curson et al., 2011; Li et al., 2017; Sun et al., 2016). DddL was the first cupin DMSP lyase to be identified, in the bacterium *Sulfitobacter* EE-36, and homologues have since been identified in many other marine α-proteobacteria species (including *Labrenzia aggregata,* the subject of this thesis) (Curson et al., 2008). DddY is another DMSP lyase of the cupin superfamily and was identified in the betaproteobacterium *Alcaligenes faecalis* (Li et al., 2017). In *A. faecalis,* the protein was shown to be associated with the cell surface, making it different from all of the other lyases, which are found in the cytoplasm. (Curson et al., 2011).

DddP is different from the other lyases as it belongs to a 'peptidase' super-family, although DddP itself cannot be classed as a peptidase, as it cleaves the C-S bond in DMSP as opposed to a peptide bond in a protein (Kirkwood et al., 2010; Todd et al., 2009). The *dddP* gene appears to be largely confined to the Roseobacters (the *Rhodobacteraceae* family of α-proteobacteria) and is very abundant in marine metagenomes (Curson et al., 2018; Todd et al., 2009).

The most recent DMSP lyase to be identified is DddX, which is an ATP-dependent DMSP lyase which was first isolated in *Psychrobacter sp.* D2 from Antarctic samples and has since been found in both Gram-negative and Gram-positive bacteria (Li et al., 2021). DddX acts differently to the aforementioned DMSP lyases as it catalyses a two-step reaction. The first is the ligation of DMSP and CoA, followed by the cleavage of the formed DMSP-CoA to produce DMS and acryl-CoA (Li et al., 2021).

The DMSP lyase present in eukaryotes has been termed Alma1, which was first identified from the algae *Emiliania huxleyi* and has been identified in several algae and corals (Alcolombri et al., 2015). Alma1 is a homotetramer and like most bacterial DMSP lyases, forms DMS and acrylate from DMSP (Alcolombri et al., 2015).

### 1.6.3 DMSP Oxidation

The final DMSP metabolic pathway is not catabolic as it proceeds via oxidation and was first described in 2018, when Thume and colleagues identified the structurally unusual metabolite, dimethylsulfoxonium propionate (DMSOP) (**Figure 1.3**) (Thume et al., 2018). The DMSP-producing bacteria *Pelagibaca bermudensis* was found to contain DMSOP and the biosynthesis and catabolism of the novel compound was studied using isotopically labelled DMSOP and DMSP. This revealed that DMSOP is likely formed via the direct oxidation of DMSP, and it is subsequently degraded to DMSO, although the enzymes involved in these processes remain unknown (Thume et al., 2018).

## 1.7 DMS and The Marine Organosulfur Cycle

DMSP has a vital role in global biogeochemical sulfur cycling, as it is the major precursor of the volatile, climate-active gas, DMS (**Figure 1.4**). Within the marine environment, DMS acts as a signalling molecule as it can be used by seabirds and marine mammals as a foraging cue (DeBose & Nevitt, 2008). There are then three potential fates of DMS: it can either be microbially degraded, undergo photolysis in the surface ocean or it can be transferred to the atmosphere (Andreae, 1990; Schäfer et al., 2010; Taalba et al., 2013).

The majority of DMS is catabolised via several microbial pathways. The most studied of these microbial degradation pathways are the two that lead to the utilization of DMS as a carbon and energy source: the DMS monooxygenase pathway and the DMS methyltransferase pathway (Schäfer et al., 2010). Both of these pathways involve the oxidation of DMS and production of MeSH (De Bont et al., 1981; Visscher & Taylor, 1993). The produced MeSH can then undergo further oxidation and the products of this can be assimilated directly into biomass (Schäfer et al., 2010). The genes in both aforementioned DMS degradation pathways have not yet been identified and so cannot be used as markers for the ability to degrade DMS. Nevertheless, large numbers of DMS degrading bacteria have been identified, from a wide range of environments, including multiple species of proteobacteria, firmicutes and actinobacteria (Raina et al., 2010; Schäfer et al., 2010).

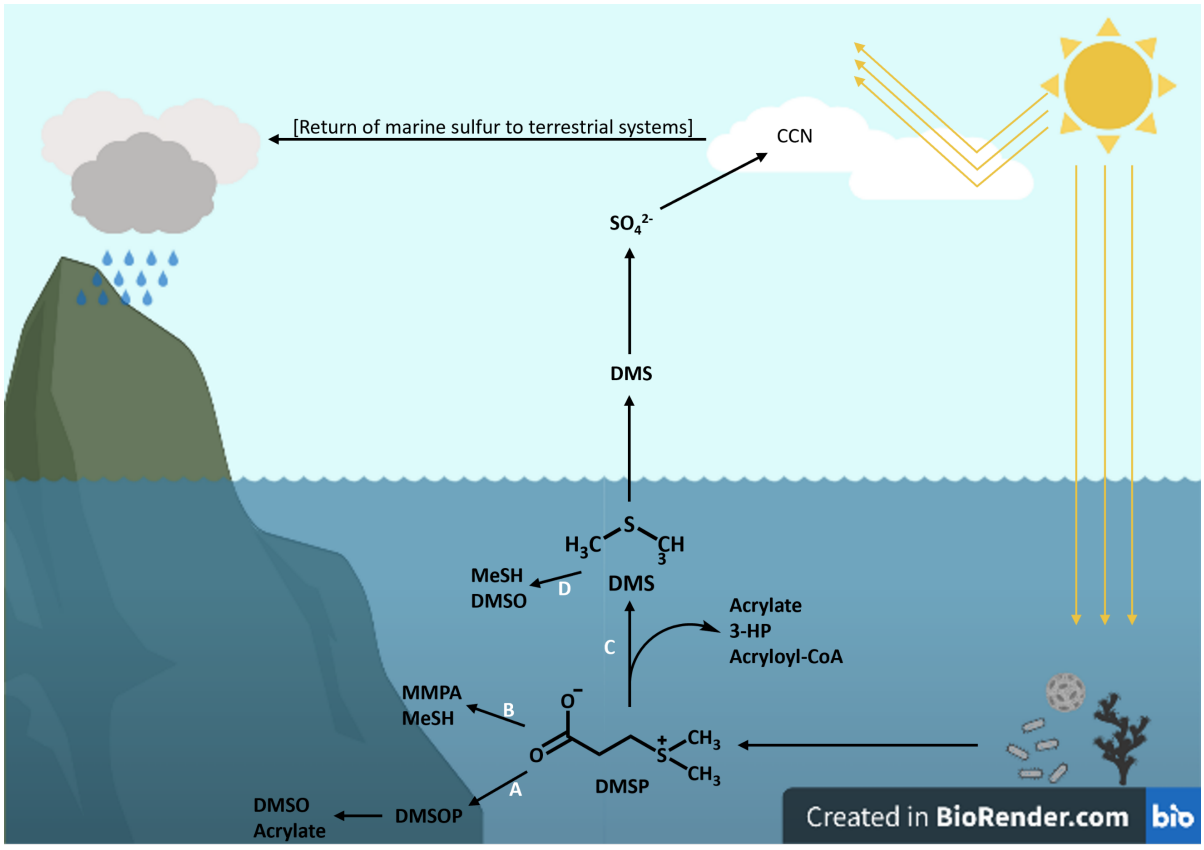



Figure 1.4 DMSP in the biogeochemical sulfur cycle.
DMSP is produced by organisms including phytoplankton, bacteria and corals in the marine environment. After its release into the ocean, it can undergo three pathways of microbial degradation. It can undergo oxidation (A) to form DMSOP, which can then be further degraded to DMSO and acrylate. Alternatively, DMSP can de demethylated (B) to form MMPA and MeSH or undergo lysis (C) to form DMS and either acrylate, 3-HP or acryloyl-CoA. 90% of the produced DMS undergoes further microbial degradation (D), but it can also be released into the atmosphere. The atmospheric DMS can then either be converted to DMSO and $SO_2$, which then leads to the formation of sulfate aerosol particles which can act as cloud condensation nuclei. This leads to increased cloud albedo, thus reflection of more solar radiation. Clouds also allow the return of marine sulfur to the land via wet deposition. Adapted from (Curson et al., 2011).

Due to the photolysis and microbial degradation of DMS, only ~10 % of total DMS is transferred from the ocean to the atmosphere as a gas. Despite this, DMS is the major natural source of reduced sulfur to the atmosphere and is thought to be responsible for ~28.1 Tg sulfur transfer from the ocean to the atmosphere annually (Andreae & Raemdonck, 1983; Lana et al., 2011). Once in the atmosphere, DMS undergoes oxidation to form products including dimethyl sulfoxide (DMSO) and sulfur dioxide ($SO_2$), which subsequently leads to the formation of sulfuric acid and methane sulfuric acid, which can then undergo condensation to form sulfate aerosol particles (Andreae et al., 1999; Barnes et al., 2006; Liss et al., 1997) (**Figure 1.4**). These aerosol particles can act as cloud

condensation nuclei (CCN), which provide surfaces upon which water droplets can form and lead to cloud formation. The importance of DMS in this process has been highlighted by one study that found DMS-derived sulfate can account for up to 32 % of CCN (Sanchez et al., 2018). The formation of clouds also facilitates the return of biogenic sulfur to terrestrial systems through wet deposition, and the sulfur can then be returned to the marine environment through run-off, thus completing the biogeochemical sulfur cycle (Sievert et al., 2007) (**Figure 1.4**).

Furthermore, as cloud albedo is linked to global mean temperature, due to its ability to deflect solar radiation, DMS has been implicated in climate change (Andreae, 1990). In fact, predictions have indicated that a change in DMS flux by an order of two can lead to a change of a few degrees in the global temperature (Legrand et al., 1988).

Although not discussed further in this thesis, it is pertinent to acknowledge the recent advances in understanding regarding the metabolism of organosulfur in marine bacteria. For instance, a limited but diverse number of bacterial taxa have recently been shown to degrade sulfoquinovose (SQ), and use the resulting reduced sulfur for assimilation, highlighting a previously undiscovered link in the marine organosulfur cycle (Liu et al., 2023; Tang and Liu., 2023).

### 1.7.1 The CLAW hypothesis

In 1987, the now controversial CLAW hypothesis was proposed, which suggests the existence of a feedback loop between global temperature and DMS production; increasing solar radiation, and thus temperature cause an increase in growth of DMSP-producing phytoplankton, and therefore total levels of DMSP (Charlson et al., 1987). Then, the increased release of DMS to the atmosphere causes a rise in CCN and thus, cloud albedo, leading to a decrease in solar radiation and temperature. However, this theory is now thought to be unlikely, due to the existence of major sources of CCN, which are not derived from DMS (Quinn & Bates, 2011). As such, the CLAW hypothesis is now contested.

## 1.8 Glycine Betaine Synthesis

There are two routes via which GB is synthesised, these vary depending on the organism carrying out the production and the enzymes involved (**Figure 1.5**).

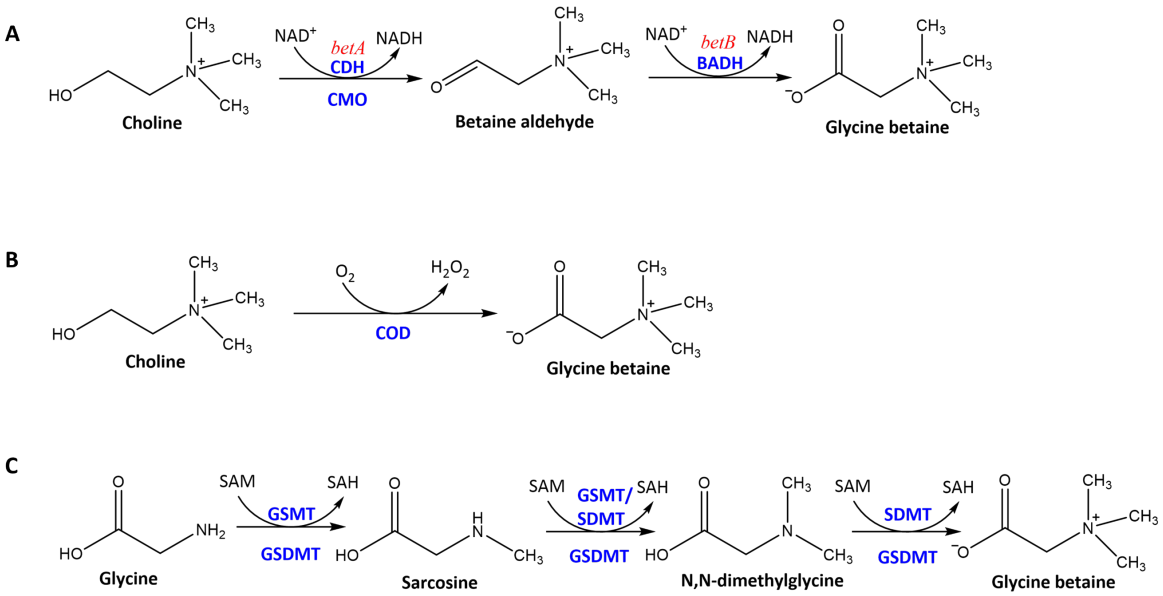



Figure 1.5 The synthesis pathways of glycine betaine
The choline synthesis pathway (A and B). In some bacteria, a CDH encoded for by *betA* and a BADH encoded for by *betB* catalyse the two reactions in this pathway (A). In plants the first step of the choline pathway is catalysed by CMO (A). In some species of bacteria and fungi, choline is oxidised to glycine betaine by a COD (B). The glycine synthesis pathway is either catalysed by a GSMT and SDMT or all three methylation reactions are catalysed by GSDMT (C).

### 1.8.1 The Choline-GB synthesis pathway

In the choline-GB synthesis pathway, choline is utilised as the initial substrate for GB synthesis (**Figure 1.5**). This choline-GB synthesis pathway is carried out by certain plants, animals and bacteria but has been most extensively studied in *E. coli* (Dragolovich, 1994; Hayashi & Murata, 1998; Landfald & Strøm, 1986). The choline undergoes an oxidation reaction catalysed by a choline dehydrogenase (CDH) enzyme, to form the intermediate betaine aldehyde (Landfald & Strøm, 1986). In *E. coli* this enzyme is encoded for by the *betA* gene. The next step in the synthesis pathway is the oxidation of betaine aldehyde to form glycine betaine. This second step can either be catalysed by the CDH or by a highly specific betaine aldehyde dehydrogenase (BADH), encoded by *betB* in *E. coli* (Falkenberg & Strøm, 1990; Landfald & Strøm, 1986).

In GB-synthesising plants, the choline-GB pathway differs, in that the first step is catalysed by choline monooxygenase (CMO) rather than by CDH (Burnet et al., 1995; Rathinasabapathi et al., 1997). Further, in some species of bacteria and fungi, such as *Athrobacter globiformas* and *Aspergillus fumigatus,* choline is oxidised to GB by just one enzyme, choline oxidase (COD) (**Figure 1.5B**). In these species, choline is oxidized in the presence of $O_2$, and hydrogen peroxide is produced in addition to GB (Ikuta et al., 1977; Lambou et al., 2013).

### 1.8.2 The Glycine-GB Synthesis Pathway

The second GB synthesis pathway was first reported to take place in two extremely halophilic bacteria, *Actinopolispora halophilia* and *Ectothiorhodospira halocloris* (Nyyssölä et al., 2000). This pathway begins with the amino acid glycine and is comprised of three successive methylation reactions, for which SAM is the methyl donor (**Figure 1.5C**). These methylation reactions are catalysed by two separate proteins with overlapping substrate specificity: a glycine sarcosine methyltransferase (GSMT), and a sarcosine dimethylglycine methyltransferase (SDMT) (Kimura et al., 2010; Nyyssölä et al., 2000). The first methylation reaction is the conversion of glycine to sarcosine, which is catalysed by GSMT. The second methylation yields N,N-dimethylglycine in a reaction that can be catalysed by either GSMT or SDMT. Subsequently, SDMT catalyses the final methylation reaction and GB is formed (Figueroa-Soto & Valenzuela-Soto, 2018; Kimura et al., 2010; Nyyssölä et al., 2000).

This pathway also occurs in the halophilic methanogen, *Methanohalophilus portucalansis,* but differs to the previously described pathway in the halophilic bacteria as there are three types of methyltransferases involved in catalysing the three methylation reactions. In addition to the GSMT and SDMT, *M. portucalansis* also contains a glycine sarcosine dimethylglycine methyltransferase (GSDMT) which is able to catalyse all three reactions in the synthesis pathway (Lai et al., 2006; Lai & Lai, 2011).

The *M. portucalensis* GSDMT is comprised of four subunits and can sequentially catalyse all three methyl transfers in this pathway. Therefore, GSDMT can directly synthesise glycine

betaine from glycine (Lai et al., 2006). More recently, the diatom *Thalassiosira pseudonana* was also found to contain GSDMT and produce GB via this pathway (Kageyama et al., 2018). Study of the *T. pseudonana* GSDMT revealed that it is comprised of two methyltransferase domains, which it uses independently to catalyse the three methylation reactions. The first methyltransferase domain of GSDMT catalyses the methylation of glycine to sarcosine, whilst the second domain catalyses the methylation of sarcosine to N-N-dimethylglycine and the subsequent methylation to glycine betaine (Kageyama et al., 2018).

## 1.9 Glycine Betaine Catabolism

As observed with DMSP degradation, many microorganisms have the ability to degrade GB and there are multiple pathways via which GB is catabolised (**Figure 1.6**). The catabolism of GB allows microorganisms to biosynthesise the essential primary metabolites glycine, choline and methionine. Further, one product of glycine betaine catabolism is methane, which is a major greenhouse gas.

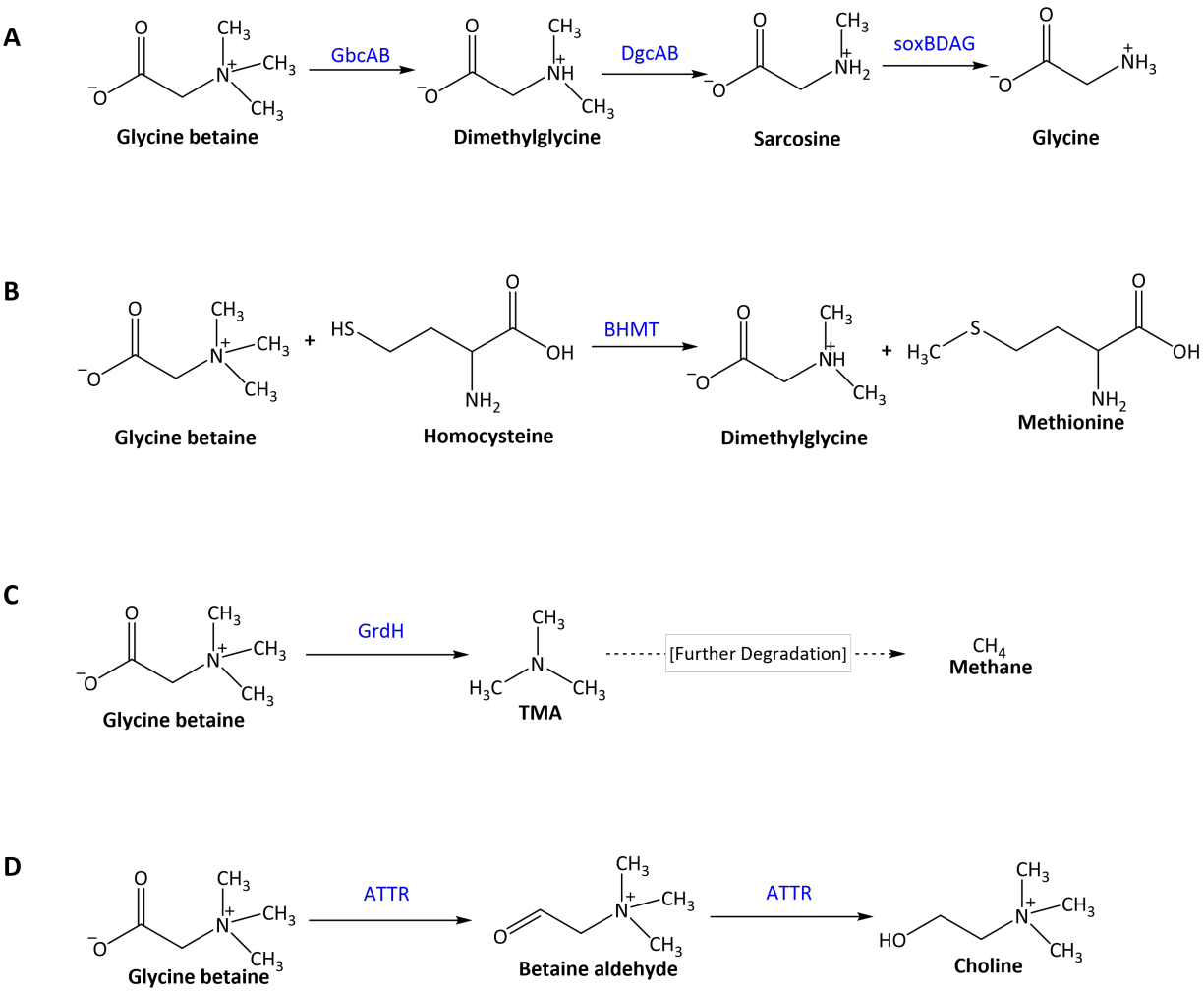



Figure 1.6 Glycine betaine catabolism pathways.
(A) The demethylation pathway as carried out by *Pseudomonas aerugionosa*. (B) The demethylation pathway used by other bacterial species such as *Sinorhizobium melioti*. (C) The pathway utilised by many marine bacteria, resulting in the formation of the climatically relevant gas, methane. (D) The pathway used by some fungal species to degrade GB to choline.

### 1.9.1 The Demethylation Pathway

The most well-known GB catabolism pathway is the conversion of GB to glycine via successive methyl group transfer. This pathway has been most extensively studied in *Sinorhizobium, Corynebacterium, Athrobacter* and *Pseudomonas* species (Meskys et al., 2001; Smith et al., 1988; Suzuki et al., 2005; Wargo et al., 2008). This pathway is the reverse of the Glycine-GB synthesis pathway, and thus involves serial demethylation reactions that form dimethylglycine, sarcosine and finally glycine (Barra et al., 2006; Wargo et al., 2008) (**Figure 1.6A and 1.6B**).

In some species, such as *Pseudomonas aerugionosa,* the GB demethylase (termed GbcAB) that converts GB to dimethylglycine also yields formaldehyde. Following this, dimethylglycine is further demethylated to sarcosine and subsequently glycine as the methyl groups are transferred to tetrahydrofolate and ultimately oxidized to carbon dioxide, in an energy yielding reaction (Boysen et al., 2022; Sun et al., 2011). In *P. aerugionosa,* the dimethylglycine demethylation is catalysed by DgcA and DgcB which form a heterodimeric flavin-linked oxidoreductase (Fitzsimmons et al., 2011; Wargo et al., 2008). The sarcosine demethylation is then catalyzed by the heterotetramic enzyme SoxBDAG, a sarcosine oxidase (Wargo, 2013). In other species, the demethylation of dimethylglycine is carried out by either a dimethylglycine dehydrogenase or dimethylglycine oxidase and the demethylation of sarcosine can be carried out by a sarcosine dehydrogenase or monomeric sarcosine oxidase (Boysen et al., 2022; Lahham et al., 2021; Wagner & Jorns, 2000).

In contrast, in some other species, such as *Sinorhizobium melioti,* the demethylation of GB to dimethylglycine is catalysed by a betaine homocysteine methyltransferase (BHMT) (Smith et al., 1988) (**Figure 1.6B**). The resulting methyl groups are then donated to the methionine cycle and are involved in the re-methylation of homocysteine (Zou et al., 2016). BHMT-dependant GB catabolism is widespread, occurring in both mammals and microorganisms, although there is rarely high similarity between the mammalian and microbial BHMTs (Serra et al., 2002; Smith et al., 1988; Wargo et al., 2008).

In organisms that possess the full GB demethylation pathway (**Figure 1.6A**), such as members of the *Alphaproteobacteria* SAR11 clade, the resulting glycine from this pathway can be used for respiration, protein synthesis and biosynthesis (Boysen et al., 2022; Noell & Giovannoni, 2019; Sun et al., 2011; Wargo, 2013). As such, GB can be utilized for the synthesis of many cellular components and some species have the ability to use GB as a sole carbon, nitrogen and energy source (I. D. Lidbury et al., 2015; Smith et al., 1988; Wargo, 2013; Welsh, 2000; Zou et al., 2016).

### 1.9.2 Alternative GB Catabolism Pathways

Many marine bacteria can also degrade GB to trimethylamine (TMA) via a glycine betaine reductase (GrdH) (**Figure 1.6C**), although abundance of the *grdH* gene in microbial metagenomic datasets is relatively low (Andreesen, 1994; Jameson et al., 2016). The generated TMA can then be used by anaerobic methanogens undergoing methanogenesis, implicating GB in the production of the climate affecting gas, methane (King, 1984). Recently, a novel family *Candidatus* 'Betaineceae' was discovered to be involved in methanogenesis from GBT, and these are widely distributed in salt marshes and coastal sediments (Jones et al., 2019). There are also some species of methanogens that produce methane via the direct demethylation of GB (Ticak et al., 2014; Watkins et al., 2014). Overall, it has been estimated that 90 % of methane emissions in marine coastal sediments result from the degradation of GB and TMA (Jones et al., 2019; Oremland et al., 1982).

Alternatively, in some fungal species, GB can be degraded to choline via the intermediate glycine betaine aldehyde in a reaction catalysed by a nonribosomal peptide synthetase (NRPS)-like GB reductase, termed ATTR (Hai et al., 2019) (**Figure 1.6D**). This pathway is essentially the reverse of the choline-GB synthesis pathway. As this degradation pathway was identified relatively recently, the presence of this pathway has not been explored in marine microorganisms (Boysen et al., 2022).

## 1.10 DMSP and GB Transport

As previously mentioned, not all species that degrade and utilise the protective properties of compatible solutes such as DMSP and GB, synthesise these compounds themselves. Instead, these species rely on uptake from the environment to transport the organic osmolytes into the cell. Further, microbial uptake of compatible solutes from their environment is often energetically favourable over synthesis, and often uptake of compatible solutes causes repression of genes involved in compatible solute synthesis (Oren, 1999; Ziegler et al., 2010).

Due to the structural and physiological similarities between DMSP and GB it is not surprising that there is substantial overlap in the transporters used. Both compatible

solutes are substrates of transporters belonging to the betaine-choline-carnitine-transporter (BCCT) and ATP binding cassette (ABC) transporter families (Dickschat et al., 2015; Kempf & Bremer, 1998; Ziegler et al., 2010).

Prokaryotic ABC transporters are comprised of a transmembrane protein (TMP), a nucleotide binding protein (NBP) and a substrate binding protein (SBP). The binding of ATP to the NBP causes a conformational switch in the NBP which results in the formation of a channel through the membrane for the substrate (Dickschat et al., 2015).

Two of the best studied ABC transporters with GB specificity are ProU in *E. coli,* which is comprised of subunits ProV, ProW and ProX; and OpuA in *Bacillus subtilis,* comprised of subunits termed OpuAA, OpuAB and OpuAC (Gowrishankar, 1989; Kempf & Bremer, 1995). These transporters were first identified as GB transporters, however both have since been shown to also channel DMSP through the membrane (Broy et al., 2015; Rudulier et al., 1996).

BCCTs are so named as they were first found to be transporters of GB, although they too have since been found to transport structurally similar molecules, including DMSP (Dickschat et al., 2015). The protein sequences of BCCTs are diverse, and the proteins have been given various names in different species, for example: BetU in *E. coli* and OpuD in *B. subtilis (Kappes et al., 1996; Wood, 2015).*

A DMSP transporter was identified in the marine bacterium *Marimonas* MWYL1, termed DddT. The *dddT* gene encoding the BCCT-like protein is located in close proximity to the *dddD* DMSP lysis gene in this organism (Todd et al., 2007). It has since come to light that in many species, *dddD* is near genes predicted to be DMSP transporters and some of these have been ratified experimentally (Sun et al., 2012). This suggests that the import and lysis of DMSP are transcriptionally linked. This study revealed a wide diversity of DMSP transporters and suggests that the genes encoding the transporters may have been acquired through horizontal gene transfer (Sun et al., 2012).

## 1.11 *Labrenzia aggregata* LZB033

The model organism that will be the focus of this thesis is the abundant marine *Alphaproteobacterium Labrenzia aggregata* LZB033. *L. aggregata* is a member of the *Rhodobacteraceae* family, which contains more than 100 genera. As a member of the *Labrenzia* genus, this strain is a Gram-negative, motile, aerobic rod-shaped bacterium. The genus is so far comprised of six species, all isolated from marine environments (Camacho et al., 2016)

The LZB033 strain studied here was first isolated from seawater of the East China Sea. This bacterium was screened for its ability to lyse DMSP to DMS, however was found to produce DMS when no DMSP was added to the culture. This led to the discovery that bacteria are in fact producers of DMSP (Curson et al., 2017). It was also the first species in which a DMSP synthesis gene was identified, *dsyB* (previously discussed in section 1.5.1). In addition to producing DMSP, *L. aggregata* LZB033 can also degrade it to DMS and acrylate, as it contains the DMSP lyase *dddL* (Curson et al., 2017). Genomic analysis of *L. aggregata* LZB033 has revealed that the strain has genes potentially encoding enzymes in two GB synthesis pathways: the choline oxidation pathway, and the glycine methylation pathway (**Figure 1.5**) (Cánovas et al., 1996; Nyyssölä et al., 2000; Zhong et al., 2021).

The site of *L. aggregata* LZB033 isolation is located close to the coast and two rivers (the Yangtze and Qiantang) and as such, this location likely receives large amounts of sediment from terrestrial origins (Xu et al., 2019). Further, this site is also influenced by the Taiwan Warm Current and Zhejiang Fujian Coastal Current and thus is subject to fluctuating environmental conditions. As such, *L. aggregata* LZB033 has to survive under environmentally complex and stressful conditions (Xu et al., 2019). There have since been multiple other *dsyB* containing, DMSP-producing *L. aggregata* strains isolated from other environments, for example the Mariana Trench (Zhong et al., 2021). Conditions in the Mariana Trench mean that these strains survive at depths up to 9,600 m and pressures up to 110 MPa (Jamieson et al., 2010).

*L. aggregata* LZB033 presented as an ideal model organism for the work that is detailed in this thesis due to its ability to produce the two compatible solutes of interest, its low doubling time and its genetic tractability.

## 1.12 Gaps in Knowledge and Aims of Thesis

Despite the abundance and importance of DMSP as a compatible solute, its assumed role in osmoprotection has not been ratified in any organism via mutagenesis of a DMSP-synthesis gene and subsequent phenotyping. In fact, the only role of DMSP that has been ratified in this way is its role in barotolerance. Further, most of the genes involved in the transamination synthesis pathway used by many organisms, including the model DMSP-producing bacterium *L. aggregata* LZB033, are yet to be identified.

Additionally, despite the hypothesised inverse relationship of GB and DMSP in response to fluctuations in nitrogen concentration, there has not yet been a study involving the relationship between the synthesis genes involved in production of both compatible solutes. It is important to rectify these gaps in research knowledge, to improve our understanding of the roles of DMSP and GB.

The following research aims will be discussed in this thesis:

1. Determine how DMSP and GB synthesis genes are regulated by differing environmental conditions in *L. aggregata* LZB033.
2. Establish whether mutagenesis can be used to determine whether DMSP and GB have a role in osmoprotection in *L. aggregata* LZB033.
3. Identify candidate DMSP synthesis genes responsible for catalysing the remaining uncharacterized steps in the transamination synthesis pathway in *L. aggregata* LZB033.

# Chapter 2: Materials and Methods

## 2.1 Chemical syntheses

Unless otherwise stated all chemicals were obtained from Sigma Aldrich. DMSP was synthesised as described in (Todd et al., 2010) from acrylic acid and DMS.

## 2.2 Growth conditions and media preparation

### 2.2.1 *L. aggregata* growth conditions

*Labrenzia aggregata* strains were grown in yeast-tryptone-sea-salt (YTSS) medium (González et al., 1996) or marine basal medium (MBM) 35 practical salinity units (PSU), unless otherwise stated. Where indicated, the salinity of MBM was adjusted to 5 PSU or 50 PSU by altering the amount of sea salts added. In MBM, 10 mM succinate was used as a carbon source and 0.5 mM or 10 mM $NH_4Cl$ was used as a nitrogen source. The N:P ratios were 2:1 and 40:1 respectively. Where indicated, methionine was added to the media at a final concentration of 0.5 mM. Where indicated, antibiotics were added to YTSS and MBM at the following concentrations: streptomycin (400 μg $ml^{-1}$), spectinomycin (200 μg $ml^{-1}$), rifampicin (20 μg $ml^{-1}$), gentamicin (80 μg $ml^{-1}$), kanamycin (40 μg $ml^{-1}$), tetracycline (5 μg $ml^{-1}$).

### 2.2.2 *E. coli* growth conditions

*Escherichia coli* strains were grown in Luria-Bertani (LB) complete medium ((Sambrook et al., 1989) at 37 °C. *E. coli* BL21 strains used for protein over-expression were grown in LB with the addition of 0.05 % glucose. Where indicated, antibiotics were added to LB media at the following concentrations: ampicillin (100 μg $ml^{-1}$), carbenicillin (100 μg $ml^{-1}$), gentamicin (10 μg $ml^{-1}$), kanamycin (20 μg $ml^{-1}$), streptomycin (400 μg $ml^{-1}$), spectinomycin (200 μg $ml^{-1}$), tetracycline (5 μg $ml^{-1}$). A full list of strains used in this study are denoted in **Table 1**.

Table 1: A list of strains used in this study

| Strain | Description | Reference |
|---|---|---|
| *Escherichia coli* 803 | Strain used for routine transformations | (Wood, 1966) |
| *E. coli* DH5-α | Strain used for transformations of pk18mobsacB clones | NEB |
| *E. coli* S17-1 | Strain used for conjugations of pk18mobsacB plasmids | (Simon et al., 1983) |
| *E. coli* BL21 | Strain used for over-expression of cloned genes in pET and pMAL-c2X vectors | (Studier & Moffatt, 1986) |
| *Labrenzia aggregata* LZB033 | Wildtype strain isolated from the East China Sea | (Curson et al., 2017) |
| *L. aggregata* Δ*dsyB* | Targeted deletion of *dsyB* | (Curson et al., 2017) |
| *L. aggregata* Δ*gsdmt* | Targeted deletion of putative *gsdmt* | This study |
| *L. aggregata* Δ*dsyB*Δ*gsdmt* | Targeted deletion of both *dsyB* and *gsdmt* | This study |
| *L. aggregata* Δ*aro8* | Targeted deletion of putative *aro8* | This study |
| *L. aggregata* Δ*at2* | Targeted deletion of putative *at2* | This study |
| *L. aggregata* Δ*aro8*Δ*at2* | Targeted deletion of both *aro8* and *at2* | This study |
| *L. aggregata* Δ*cosR* | Targeted deletion of putative *cosR* | This study |

## 2.3 Growth Curves

### 2.3.1 Growth curves of *L. aggregata* WT

Starter cultures of 5 ml YTSS were inoculated with a single colony of *L. aggregata* WT and incubated at 30 °C with shaking at 180 rpm for 14 hours. The $OD_{600}$ of the starter cultures were adjusted to 0.6. Triplicate 250 ml flasks containing 100 ml MBM (35 PSU, 5 PSU, Low N, + Met) were inoculated with 1 ml of the starter culture and incubated at 30 °C with shaking at 180 rpm. Every two hours, growth was assessed by removing a 1ml aliquot of each culture and recording the $OD_{600,}$ until the cultures reached stationary phase. In some cases, bi-hourly readings were not sufficient to determine the mid-exponential phase of the cultures, and in these cases hourly $OD_{600}$ readings were taken. Measurements were averaged and plotted on a line graph with the $OD_{600}$ readings plotted logarithmically. The growth rate of each culture was calculated from the gradient of line in exponential phase.

### 2.3.2 Growth curves for phenotyping mutant strains

Starter cultures of 5ml MBM 35 PSU were inoculated with each *L. aggregata* strain to be measured and incubated at 30°C with shaking at 180 rpm for 16 hours. The $OD_{600}$ of the cultures were measured and adjusted to 0.6. Each well in a colourless, lidded 24-well plate was filled with 1 ml MBM of varied salinities as stated. Where indicated, DMSP and GB were added to the media at a final concentration of 100 µM. A 20 µl aliquot of each $OD_{600}$ starter culture was used to inoculate triplicate wells of each salinity. Two initial $OD_{600}$ readings were taken of each well, one with path-check enabled (normalisation to pathlength), and one without. The plate was incubated at 30°C with 7 seconds of orbital shaking in between each reading and before every reading. The $OD_{600}$ of each well was measured every thirty minutes for 72 hours. The data was normalised to the pathcheck values and means were calculated and plotted on a line graph, with the $OD_{600}$ readings plotted logarithmically.

## 2.4 Transformations into *E. coli*

### 2.4.1 Making chemically competent cells

A starter culture of 5 ml LB was inoculated with a single colony of the *E. coli* strain to be used (803/DH5-α/S17) and incubated at 37 °C overnight. A 1 ml aliquot of this starter culture was used to inoculate 100 ml LB, which was then incubated at 37 °C at 200 rpm until the $OD_{600}$ reached 0.3 - 0.4. The *E. coli* culture was transferred to two sterile falcon tubes and centrifuged at 6000 rpm at 4 °C for 10 minutes. The supernatant was removed, and the cell pellets and re-suspended in 10 ml pre-cooled 0.1 M $CaCl_2$ and kept on ice for 30 minutes. The samples were centrifuged again at 6000 rpm at 4°C for 10 minutes and the supernatant was removed. Cell pellets were re-suspended in 2 ml cold 0.1 M $CaCl_2$ and stored at 4°C overnight.

### 2.4.2 Heat shock transformations

For transformations of whole vectors, 1 µl of DNA (~100 ng/µl) was added to 100 µl competent cells. For transformations of ligation reactions 16 µl of the ligation mixture was added to 100 µl competent cells. When carrying out heat shock transformations, the following controls were also used: a negative control of 100 µl competent cells, and a positive control of 100 µl competent cells and 1 µl of vector DNA. Samples were incubated on ice for an hour before being heat shocked in a 42 °C water bath for three minutes and then transferred to ice for 2 minutes. An addition of 500 µl pre-warmed LB was made to each of the samples which were then incubated at 37 °C for 60 – 90 minutes. The samples were centrifuged at 6000 rpm for three minutes and most of the supernatant was removed. The pellet was re-suspended in the remaining supernatant and plated on LB containing appropriate antibiotics. The plates were incubated at 37°C overnight, or for 48 hours if *E. coli* S17 was used.

## 2.5 Polymerase chain reaction (PCR)

Genes (and where stated, flanking regions of genes) were amplified using PCR in a Thermal Cycler using the cycling conditions detailed in **Table 2**. Typically, each 25 µl PCR reaction mix contained 12.5 µl MyFi™ DNA Polymerase master mix (Meridian Bioscience), 0.5 µl

forward and reverse primers (20 pmol), 0.5 µl template, and 11 µl sterile $H_2O$. A full list of the oligonucleotide primers used in this study are listed in **Table 3**. Primers were synthesised by Eurofins Genomics or Integrated DNA Technologies (IDT). In *L. aggregata* colony PCR reactions, a single colony was picked with a sterile toothpick and submerged into 100 µl sterile water, 0.5 µl of this mixture was used as template in colony PCR.

Table 2: PCR cycling conditions

| Temperature (°C) | Time | Number of cycles |
|---|---|---|
| 95 | 2 minutes | X1 |
| 95 | 45 seconds | X 35 |
| Annealing temp. (variable) | 45 seconds | |
| 72 | 20 seconds/kb | |
| 72 | 5 minutes | X1 |

Table 3: A list of primers used for PCR in this study

| Primer name | Sequence (5' to 3')* | Use |
|---|---|---|
| 27F | AGAGTTTGATCCTGGCTCAG | Forward primer used to amplify the 16S rRNA gene |
| 1492R | GGTTACCTTGTTACGACTT | Reverse primer used to amplify the 16S rRNA gene |
| M13 uni (-43) | AGGGTTTTCCCAGTCACGACGTT | Forward primer used to amplify inserts in pk18mobsacB |
| M13 rev (-29) | CAGGAAACAGCTATGACC | Reverse primer used to amplify inserts in pk18mobsacB |
| Gsdmt_flank1_FOR | GATCGAATTCCCCGATTTTTCCGAGACC | Forward primer used for amplifying flanking region 1 of *gsdmt* |
| Gsdmt_flank1_REV | GATCTCTAGACAACTCGTCCCACTTTTCCAC | Reverse primer used for amplifying flanking region 1 of *gsdmt* |
| Gsdmt_flank2_FOR | GATCTCTAGACTTCCGCCGACTATCTCG | Forward primer used for amplifying flanking region 2 of *gsdmt* |
| Gsdmt_flank_REV | GATCCTGCAGCTTCCTCGCCCTTGTCGC | Reverse primer used for amplifying flanking region 2 of *gsdmt* |
| Gsdmt_check_FOR | GTTTAATTACAAAGGCTTAACCAGC | Forward primer used to screen for *gsdmt* mutants in colony PCR |
| Gsdmt_check_REV | CTCGTTGCAGGCATAGCCGAACATG | Reverse primer used to screen for *gsdmt* mutants in colony PCR |
| Php45omega_FOR | GATCTCTAGAGGTGATTGATTGAGCAAG | Forward primer used for amplification of Ω region of php45Ω |
| Php45omega_REV | GATCTCTAGAGGTGATTGATTGAGCAAG | Reverse primer used for amplification of the Ω region of php45Ω |
| Aro8_check_FOR | GATCGGATCATACCCTGATGG | Forward primer used to screen for *aro8* mutants in colony PCR |
| Aro8_check_REV | GATCCAGCGTGATCTTGAAGTTGC | Reverse primer used to screen for *aro8* mutants in colony PCR |
| At2_check_FOR | GATCCGCGAGGAATTCGTGGTC | Forward primer used to screen for *at2* mutants in colony PCR |
| At2_check_REV | GATCGACCGGCATTTGATGCTTCAG | Reverse primer used to screen for *at2* mutants in colony PCR |
| CosR_check_FOR | GGCCTTGAGGATGACCTC | Forward primer used to screen for *cosR* mutants in colony PCR |
| CosR_check_REV | CCAGACCCTGAACGACTATGG | Reverse primer used to screen for *cosR* mutants in colony PCR |
| Aro8_pMAL_FOR | TGTGGTCTCAAGGTATGCGTGTCACTGGTAGCGGTCG | Forward primer to amplify *aro8* for golden gate cloning of into pMAL-c2X |
| Aro8_pMAL_REV | CGTGGTCTCAAAGCTTTCAGGAAGTCACGCAGGAAGGACAGG | Reverse primer to amplify *aro8* for golden gate cloning of into pMAL-c2X |
| BM0189 | GCGGTCGTCAGACTGTC | Forward primer to amplify pMAL-c2X inserts |
| BM0259 | ATTTGATGCCTGGCAGTTC | Reverse primer to amplify pMAL-c2X inserts |

*underlined sequences represent restriction sites.

## 2.6 DNA Visualisation and Purification

### 2.6.1 Gel electrophoresis

Gel electrophoresis was used to visualise PCR products and cloning inserts and vectors. Gels were comprised of 0.8 % (w/v) agarose in 1 x TAE Buffer (40 mM Tris base, 20 mM acetic acid, 1 mM EDTA in $H_2O$). Agarose was added to the buffer, melted and cooled to ~ 50 °C before 5 µl 10mg/ml ethidium bromide was added per 100 ml. Gels were poured into gel trays and left to set at room temperature. In each gel, 3 µl 1 KB Plus DNA ladder (Invitrogen) was loaded into a well. Gels were usually run for an hour at 90 V, and subsequently visualised using a UV gel imaging doc.

### 2.6.2 Gel extraction

Bands of interest were excised from the gel using a scalpel, weighed, and dissolved in the appropriate volume of Buffer QG (Qiagen). The gel extraction was then carried out using the QIAquick Gel Extraction Kit (Qiagen) following manufacturer's instructions. DNA was eluted from the QIAquick membrane in 30 µl sterile $H_2O$.

### 2.6.3 PCR purification

Purification of PCR products was carried out using the Roche High Pure PCR Product Purification Kit following manufacturer's instructions. All centrifugation steps were carried out at 13,000 rpm and the purified DNA was eluted in 50 µl sterile $H_2O$.

## 2.7 Plasmid extractions from bacterial cultures

### 2.7.1 Phenol chloroform minipreps

A 5 ml LB culture with the appropriate antibiotics was inoculated with a single colony of *E. coli* and incubated at 37 °C overnight. Cells were pelleted by the centrifugation of 3 ml of the culture. Centrifugation was carried out at 15,900 x *g* for two minutes. The supernatant was removed, and the cell pellet was resuspended in 250 µl P1 buffer (Qiagen) and 250 µl P2 buffer (Qiagen) was added. The sample was mixed by gentle inversion and incubated at room temperature for 5 minutes before 350 µl P3 buffer was added (Qiagen) and the

sample was incubated on ice for at least 5 minutes. The samples were centrifuged at 18,440 x *g* for 10 minutes and the supernatant was transferred to a clean 1.5 ml Eppendorf microcentrifuge tube. To each sample, 400 µl Phenol:Chloroform:Isoamyl Alcohol 25:24:1 (v/v) was added and the samples were vortexed for ~ 5 seconds. The samples were centrifuged at 18,440 x *g* for 2 minutes to cause layer separation and the top layer was transferred to a clean 1.5 ml microcentrifuge tube. To each sample, 700 µl ethanol was added and the samples were centrifuged at 18,440 x *g* for 15 minutes. The supernatant was removed and 500 µl ethanol (70 %) was added. The samples were centrifuged at 18,440 x *g* for two minutes and all supernatant was removed. The lid of the 1.5 ml microcentrifuge tube was left open at room temperature for 5 – 10 minutes until the pellet became transparent. The pellet was re-suspended in 50 µl $H_2O$.

### 2.7.2 Plasmid midipreps

To obtain high quantity and quality plasmid DNA, 100 ml LB culture of the *E. coli* strain was grown overnight at 37 °C. The Qiagen Plasmid Midiprep kit was used according to manufacturer's instructions. Following elution from the column, the DNA pellet was precipitated through the addition of 3.5 ml isopropanol and centrifuged at 15,900 x *g* for 30 minutes. The supernatant was removed, and the pellet was washed in 2 ml ethanol (70 %). Samples were centrifuged for 15,900 x *g* for 10 minutes and all ethanol supernatant was removed. The DNA pellet was dried in an open 1.5 ml microcentrifuge tube for ~ 10 minutes, and then re-dissolved in 50 µl $H_2O$. A full list of the plasmids used in this study is listed in **Table 4**

Table 4: A list of plasmids used in this study

| Plasmid | Description | Reference |
|---|---|---|
| pRK2013 | Helper plasmid used in triparental matings | (Figurski & Helinski, 1979) |
| pk18mobsacB | Plasmid used in making *L. aggregata* mutants | (Schäfer et al., 1994) |
| pRK415 | Wide host-range plasmid vector with IPTG-inducible *lac* promoter | (Keen et al., 1988) |
| pBIO2266 | *L. aggregata* IAM12614 *dsyB* cloned into pRK415 | (Curson et al., 2017) |
| p34S-Gm | Plasmid from which the gentamycin resistance cassette was cloned | (Dennis & Zylstra, 1998) |
| php45Ω | Plasmid from which the spectinomycin resistance cassette was cloned | (Prentki & Krisch, 1984) |
| pET16b | Plasmid vector for expression of cloned genes in *E.coli* | Merck Millipore |
| pET21a | Plasmid vector for expression of cloned gens in *E.coli* | Merck Millipore |
| pMAL-c2X | Plasmid for expression of cloned genes as fusions to maltose-binding protein | New England Biolabs |
| pBIO23(1) | *L. aggregata dsyB* cloned under the control of the *lacZ* promotor in PS4 | This study |
| pBIO23(2) | *gsdmt* flanking regions plus $Gent^R$ cassette in pk18mobsacB | This study |
| pBIO23(3) | *aro8* flanking regions plus the $Spec^R$ cassette in pk18mobsacB | This study |
| pBIO23(4) | *at2* flanking regions plus the $Spec^R$ cassette in pk18mobsacB | This study |
| pBIO23(5) | *aro8* flanking regions in pk18mobsacB | This study |
| pBIO23(6) | Codon-optimised* *L. aggregata aro8* cloned into pET16b | This study |
| pBIO23(7) | Codon-optimised* *L. aggregata at2* cloned into pET21a | This study |
| pBIO23(8) | Codon-optimised* *L. aggregata aro8* cloned into pMAL-c2X | This study |
| pBIO23(9) | Codon-optimised* *L. aggregata at2* cloned into pMAL-c2X | This study |
| pBIO23f(10) | *cosR* flanking regions in pk18mobsacB | This study |

*codon optimised for expression in *E. coli* using default settings in IDT Codon Optimisation Tool.

## 2.8 Cloning Techniques

### 2.8.1 Restriction Digestions

DNA digestion reactions were carried out using FastDigest restriction enzymes (Thermo Scientific). Depending on the concentration of DNA to be digested, 1 – 16 µl DNA was used in each reaction, with 1 µl of each FastDigest enzyme and 2 µl FastDigest Buffer. The total volume reaction was made up to 20 µl with distilled water and mixed via inversion and centrifugation at 15,900 x *g* for ten seconds. The digestion reaction mixture was incubated at 37 °C for one hour before inactivation of the reaction by incubation at 80 °C for ten minutes. The efficiency of digest reactions was assessed visually using gel electrophoresis. Where necessary, digested vector DNA was dephosphorylated using alkaline phosphatase (Promega), according to manufacturer's protocol.

### 2.8.2 Ligation Reactions

Ligation reactions were carried out using T4 DNA Ligase (Promega), according to manufacturer's protocols and incubated at room temperature overnight. Typically, 16 µl digested insert DNA and 2 µl dephosphorylated vector were used in each ligation reaction. When ligations were performed, a control was used in which $H_2O$ was used in place of insert DNA to determine if the digested vector was re-ligating.

### 2.8.3 Golden Gate Cloning into pMAL-c2X

The following reagents were used in golden gate cloning procedures: 1 µl pMAL-c2X (100 ng/ µl), 1 µl insert DNA (100 ng/ µl), 1.5 µl 10x T4 Buffer, 1.5 µl 10x BSA, 1 µl BsaI restriction enzyme, 1 µl T4 ligase and the mixture was made up to a final volume of 15 µl with $H_2O$. The Golden Gate reaction was carried out in a thermocycler, using the cycling conditions listed in **Table 5**.

Table 5: Golden Gate cycling conditions

| Temperature | Time | Number of Cycles |
|---|---|---|
| 37 °C | 3 minutes | X 25 |
| 16 °C | 4 minutes | |
| 50 °C | 5 minutes | X 1 |
| 80 °C | 5 minutes | X 1 |

## 2.9 Tri-parental crossings

Tri-parental matings were used to transfer plasmids from *E. coli* to *L. aggregata* strains. The crossings involved three strains: the *E. coli* donor strain containing the plasmid of interest; the *L. aggregata* host strain; and the helper strain of *E. coli* 803 (pRK2013), which is kanamycin resistant.

### 2.9.1 Patch crosses

Each of the three strains were streaked to fresh plates containing the appropriate antibiotics and incubated for 3 days at 28 °C (*L. aggregata*) or overnight at 37 °C (*E. coli*). A sterile loop was used to mix cells from each of these plates on a fresh YTSS plate. Controls were also used in which just the donor and helper strain, and just the host and helper strain were mixed on a YTSS plate. The plates were incubated at 28 °C for 3 days and then each patch was streaked on a YTSS plate containing appropriate antibiotics to select for *L. aggregata* transconjugants. The plates were incubated at 28 °C for 3 days until single colonies were present.

### 2.9.2 Filter crosses

A 5 ml YTSS culture of the *L. aggregata* host strain, a 5 ml LB culture of the *E. coli* donor strain and a 5 ml LB culture of the *E. coli* 803 (pRK2013) were incubated overnight with shaking at 180 rpm. Each culture was supplemented with the appropriate antibiotics. Aliquots of 500 µl of each of the *E. coli* strains and of 1 ml of the *L. aggregata* strains were centrifuged at 3,400 x *g* for three minutes. The supernatant was removed from all of the samples and each of the cell pellets were washed three times in 500 µl YTSS. After the

washes, each of the cell pellets were re-suspended in 100 µl YTSS. Ethanol-sterilised forceps were used to place a sterile filter on a YTSS agar plate containing no antibiotics. 100 µl aliquots of each of the three strains were added to the filter and mixed using a sterilised loop. Control crosses were set up using just the donor and helper strain, and just the host and helper strain. Plates were incubated at 28 °C for 2 days after which the filters were transferred to sterile universals. Cells were washed off the filter using 800 µl of sterile 50 % glycerol and serial dilutions were plated on YTSS plated containing the appropriate antibiotics to only allow for growth of *L. aggregata* transconjugants. Plates were incubated at 28 °C for approximately three days, until single colonies were present.

## 2.10 Mutagenesis of *L. aggregata* using pk18mobsacB

Traditional restriction enzyme cloning was used to create constructs of the pk18mobsacB backbone, with the flanking regions of the gene of interest cloned into the MCS. In most cases, an antibiotic resistance cassette (spectinomycin/gentamycin) was cloned into the middle of the flanking regions. This plasmid was transformed into *E. coli* DH5-α and a filter cross was used to transfer the plasmid to *L. aggregata.* The filter cross was plated on YTSS plates containing rifampicin and either spectinomycin or gentamycin (depending on the antibiotic resistance cassette present in the pk18mobsacB construct) to select for colonies that had undergone single homologous recombination events. A single colony was inoculated into 5 ml YTSS containing no antibiotics and incubated at 30 °C with shaking at 180 rpm for three days. To select for colonies that had undergone a double homologous recombination event, serial dilutions of this culture were plated on YTSS plates containing rifampicin, spectinomycin/gentamycin (depending on the antibiotic resistance cassette used) and 10 % sucrose (v/v). Plates were incubated at 28 °C for up to three days, until single colonies were present. The resulting colonies were screened via colony PCR using checking primers designed just outside of the flanking regions in the *L. aggregata* genome. A *L. aggregata* WT colony was used as a positive control in these reactions.

In cases where an antibiotic resistance cassette was not present in the middle of the flanking regions of the gene of interest within the pk18mobsacB construct, these plasmids were transformed into *E. coli* S17. A filter cross was carried out in the same way as

described for triparental mating, however a helper strain was not used. To select for *L. aggregata* colonies that had undergone a single cross-over, the filter crosses were plated on YTSS containing rifampicin and kanamycin. A single colony was used to inoculate 5 ml YTSS containing no antibiotics and incubated at 30 °C with shaking at 180 rpm for three days. A serial dilution of this culture was plated on YTSS plates containing rifampicin and 10 % sucrose (v/v). Plates were incubated at 28 °C for up to three days, until single colonies were present. The resulting colonies were screened via colony PCR in the same way as described above.

A schematic describing the mutagenesis technique carried out in this thesis (an example for mutagenesis of the *gsdmt* gene) is detailed in **Figure 2.1.**

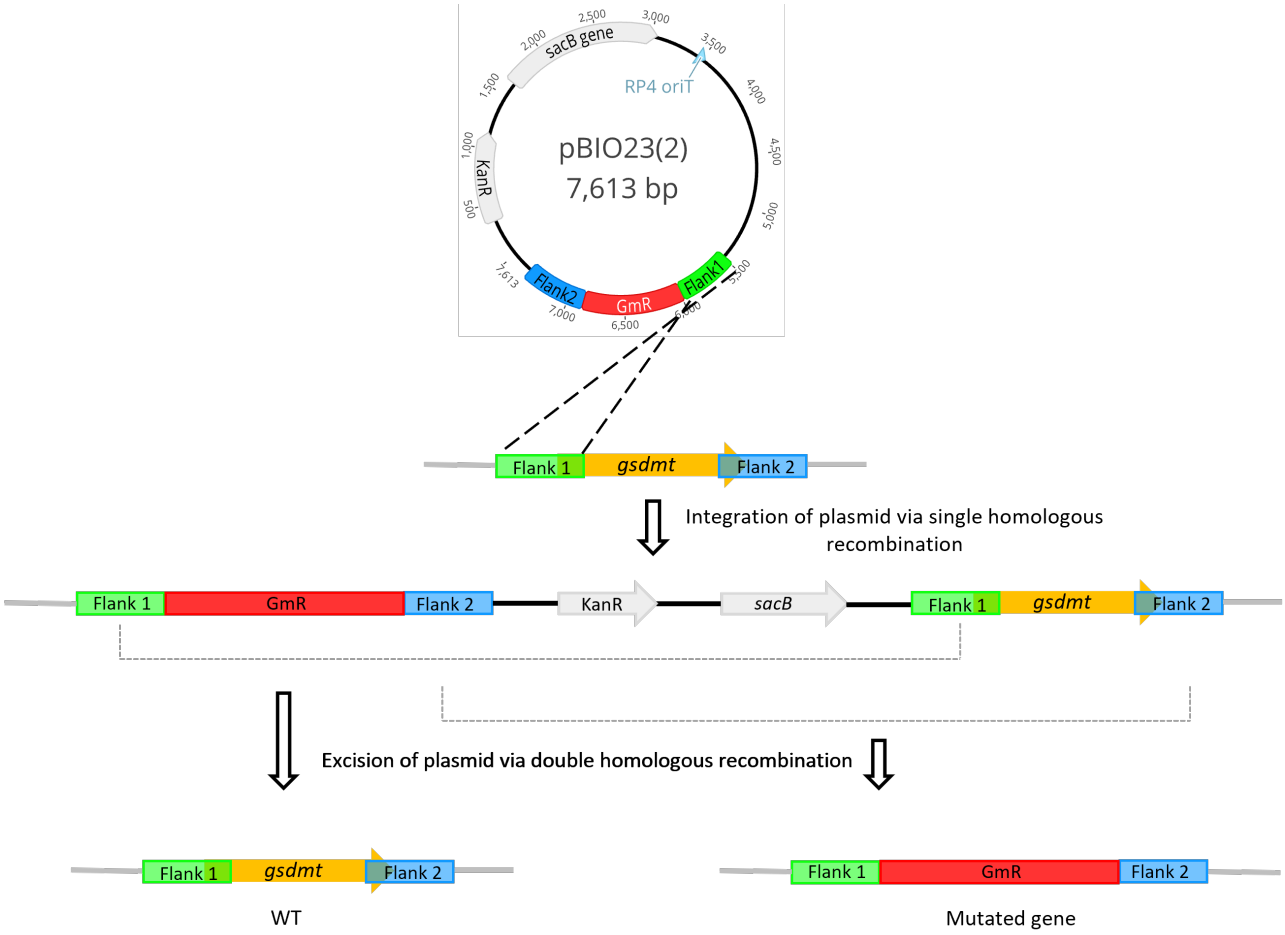



Figure 2.1 Schematic of targeted mutagenesis using pk18mobsacB.
The pBIO23(2) plasmid undergoes single homologous recombination with either Flank 1 or Flank 2 of the target gene causing integration of the plasmid into the *L. aggregata* genome. When double homologous recombination occurs, the plasmid is excised. Depending on the location of this second homologous recombination event, the strain is either restored to WT or the middle portion of the target gene is excised and replaced with an antibiotic resistance cassette. GmR = gentamycin resistance cassette. Adapted from (Biswas, 2015).

## 2.11 Quantification of compatible solutes

### 2.11.1 Quantification of DMSP by gas chromatography

The DMSP concentration of cultures was determined by gas chromatography (GC). The DMSP was lysed to DMS via alkaline lysis, with the addition of 100 µl of 10 M NaOH to 200 µl culture in 2 ml glass vials. The vials were immediately crimp sealed with PTFE/rubber crimp caps and incubated in the dark, at room temperature, for at least 16 hours. Measurements were carried out using a flame photometric detector (Agilent 7890A GC fitted with a 7693 autosampler) and a HP-INNOWax 30 m x 0.320 mm capillary column (Agilent Technologies J&W Scientific). To generate calibration curves, DMSP standards of different concentrations (0.015, 0.03, 0.06, 0.15, 0.6, 1.5, 3 and 6 nmol) were made up to a final volume of 200 µl in $H_2O$ and treated as described above (**Figure 2.2**).

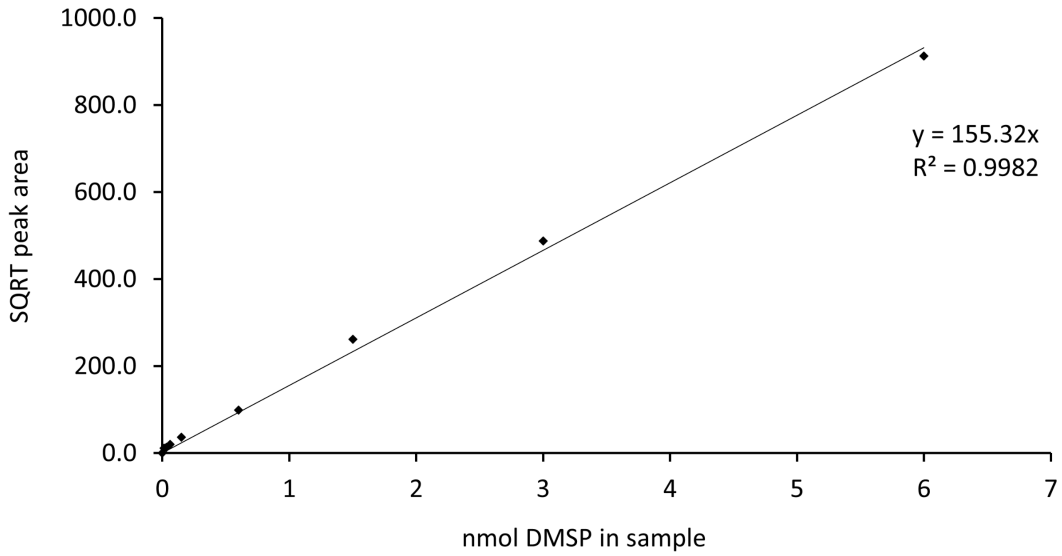



Figure 2.2 Calibration curved used to quantify DMSP by gas chromatography.
Eight 200 µl DMSP standards of varying concentrations ranging from 0.015 nmol to 6 nmol underwent alkaline lysis to DMS through the addition of 100 µl 10 M NaOH. Vials were immediately crimped and incubated at room temperature in the dark for 16 hours.

### 2.11.2 Quantification of GB via Nuclear Magnetic Resonance (NMR)

Cultures of *L. aggregata* strains were grown under the required conditions and 50 ml culture was centrifuged at 4,350 x *g* for 10 minutes. Almost all of the supernatant was discarded and the pellet was re-suspended in the remaining 1 ml supernatant and transferred to a 1.5 ml microcentrifuge tube. Samples were centrifuged again at 4,350 x *g*

for 5 minutes and the supernatant was completely removed. The cell pellet was washed three times in 1 ml sterile distilled $H_2O$ and finally re-suspended in 1 ml sterile distilled $H_2O$. The samples were subjected to five rounds of sonication for thirty seconds, with incubation on ice for 1 minute between each round. The sonicated samples were centrifuged at 18,440 x *g* for 10 minutes and the supernatant was added to 1 ml of a $H_2O$: $D_2O$ mix (90:10 v/v). Of this sample, 600 µl was aliquoted into an NMR tube and analysed via $^1$H-NMR (500 MHz). GB standards of varying concentrations were prepared with the same $H_2O$: $D_2O$ mix and analysed in the same way.

### 2.11.3 Quantification of zwitterionic metabolites by UHPLC

Quantification of zwitterionic metabolites within this study were carried out by Dr Muhaiminatul Azizah (Pohnert Lab, Friedrich Schiller University Jena, Germany). The methods described in this section were also written by Dr Azizah.

Cell pellets of *L. aggregata* (15 ml) that had been re-suspended in 750 µl methanol were disrupted by sonication using 6 cycles, 10 second pulses with 40% intensity in a Bandelin Sonoplus ultrasound homogenizer (Bandelin). The samples were centrifuged for 5 minutes at 16,100 x *g*. 30 µl of supernatants were diluted with 70 µl of a mixture of acetonitrile and water (9:1 v/v) containing an aqueous solution of an internal standard mixture (D6-DMSA, D6-DMSP, D3-ectoine, and D3-gonyol with final concentration of 500 nM) and vortexed for 30 seconds. 5 µl of aliquots were directly submitted to UHPLC/HRMS for analysis.

Analytical separation and quantification were performed using a Dionex Ultimate 3000 system (Thermo Scientific) connected to a Q-Exactive Plus Orbitrap mass spectrometer (Thermo Scientific). The UHPLC column was a Sequent ZIC-HILIC column (2.1 × 150 mm, 5 µm) coupled with SeQuent ZIC HILIC guard column (2.1 × 20 mm, 5 µm) (Merck, Darmstadt, Germany). MS data were processed using the Xcalibur software. Electrospray ionization was carried out in positive mode ionization with the following parameters: capillary temperature, 380 °C; spray voltage, 3,000 V; sheath gas flow, 60 arbitrary units; and aux gas flow, 20 arbitrary units.

For the analysis of zwitterionic metabolites via UHPLC, the method of (Thume et al., 2018) was used. The mobile phase consisted of high-purity water (Th Geyer GmbH) with 2 % acetonitrile LC-MS grade (Th Geyer GmbH) and 0.1 % formic acid LC-MS grade (Thermo Scientific) (solvent A) and 90 % acetonitrile with 10 % 1 mmol of aqueous ammonium acetate LC-MS grade (LGC Promochem) (solvent B). A gradient elution was performed using isocratic elution of 100% solvent B for 1 minute, followed by a linear gradient from 100 % solvent B to 20 % solvent B for 5.5 minutes and a linear gradient from 20 % solvent B to 100 % solvent B for 0.6 minutes, and isocratic equilibration at 100 % solvent B for 2.9 minutes. The total run time was 10 minutes. The column was kept at 25 °C. The flow rate was set at 0.6 ml minute$^{-1}$. The injection volume was 5 µLl. Full scan mode was set from 75 to 200 *m/z* at a resolution of 70,000. Before running the samples, the UHPLC was controlled by repeatedly running blanks.

Identification of zwitterionic metabolites was performed by comparing MS and MS/MS of zwitterionic metabolites in the samples with commercial and synthetic standards. Commercially available standards used were glycine betaine (Sigma-Aldrich), choline (Sigma-Aldrich), L-carnitine (Sigma-Aldrich), sarcosine (ABCR GmbH), L-ectoine (Sigma-Aldrich). Other standards were obtained by synthesis as described in previous studies [1,2]. MS/MS fragmentation was performed with a collision energy of 35 V. For quantification, a calibration curve of the standards was recorded followed by comparisons of the peak area of the analytes with the peak area of the internal standard. As internal standards for quantification of DMSA and DMSOP, we used D6-DMSA. D6-DMSP and D3-ectoine were used as internal standards for quantification of DMSP and ectoine, respectively. For all other zwitterionic metabolites, D3-gonyol was used as the internal standard. D6-DMSP, D6-DMSA, D3-gonyol, and D3-ectoine were obtained by synthesis based on published procedures (Fenizia et al., 2020; Gebser & Pohnert, 2013).

### 2.11.4 Quantification of protein concentration by Bradford assay

To allow the concentration of compatible solutes to be normalised to cell growth, the protein content of each culture was measured. This was estimated using the Bradford method (BioRad). From each culture, 1 ml was centrifuged at 18,400 x *g* for three minutes

and all supernatant was removed. Each pellet was re-suspended in 500 µl distilled $H_2O$ and lysed using two rounds of sonication for 20 seconds each, with incubation on ice for one minute in between each round. Samples were centrifuged at 18,400 x *g* for 5 minutes and a 20 µl aliquot of the supernatant was added to 980 µl Bradford Reagent. The samples were transferred to cuvettes and the absorbance of each sample was measured using a spectrophotometer set to $OD_{595}$. A calibration curve was produced using four standards of varying concentrations of BSA (0 – 400 ng/ µl) and measured as described above. This calibration curve was used to enable calculation of the protein concentration in each sample.

#### 2.11.5 Statistics

Unless otherwise stated, measurements of compatible solute production are represented as the mean of three biological replicates per strain/condition tested. Error bars either denote the standard error or standard deviation, as stated. To identify statistically significant differences, two-tailed paired Student's t-tests were carried out using Microsoft Excel. Differences were reported as significant if $p < 0.05$.

## 2.12 Transcriptional Analysis

### 2.12.1 RNA-Seq

#### 2.12.1.1 Preparation of cultures for RNA extractions

A single colony was used to inoculate 5 ml YTSS medium and was grown at 30 °C with shaking at 180 rpm for 14 hours. The $OD_{600}$ was adjusted to 0.6 and a 1 ml aliquot was centrifuged at 3,400 x *g* for three minutes and the supernatant was discarded. The pellet was washed three times in MBM and final resuspension was in 1 ml MBM. The entire 1 ml was used to inoculate 100 ml MBM (in 250 ml flasks) of varying salinity/nitrogen conditions, as described in Section 2.2 (in biological triplicate). The cultures were grown to mid-exponential phase using previously constructed growth curves. A 50 ml aliquot of the culture was centrifuged at 4,350 x *g* for 10 minutes at 4 °C. All but 10 ml of the supernatant was removed and the pellet was re-suspended in the remaining 10 ml. To each sample, 20

ml RNA*later* Stabilization Solution (Invitrogen) was added and samples were stored at -20 °C for up to a week. Immediately prior to carrying out RNA-extractions, the samples were thawed on ice and centrifuged at 4,350 x *g* for 10 minutes at 4 °C and the supernatant was removed. The pellet was resuspended in 2 ml RNA*later* and a 500 µl aliquot was removed to a fresh tube and centrifuged at 9,400 x *g* for 3 minutes at 4 °C and all of the supernatant was removed.

#### 2.12.1.2 RNA extractions

RNA-extractions were carried out using the RNeasy Mini Kit (Qiagen). Cells were re-suspended in RLT buffer, according to the manufacturer's instructions and added to 300 µl worth of glass beads (≤ 106 µm) and the samples were lysed in a Ribolyser (FastPrep System, MP Biomedicals) in two rounds of 40 seconds at a speed setting of 6 m/s, before incubation on ice for 2 minutes. The sample underwent centrifugation at 9,400 x *g* for three minutes at 4 °C and 600 µl of the supernatant was transferred to a new tube. The remaining protocol was carried out following the manufacturer's instructions. The RNA was eluted from the RNeasy column with 100 µl RNAse free water. Genomic DNA was removed by treating the samples with two applications of TURBO DNA-free DNAse (Ambion) according to manufacturer's instructions. The samples were immediately flash frozen in liquid nitrogen and stored at -80 °C until needed. The RNA sequencing was carried out by Novogene.

#### 2.12.1.3 Analysis of RNA-Sequencing data

The quality checking and differential expression analysis of the RNA-Seq data was carried out by Dr. Simon Moxon (UEA). The program Kallisto was used to quantify the abundance of transcripts and mapping to the reference genome (Bray et al., 2016). Annotation of the *L. aggregata* LZB033 genome was performed with RAST (Aziz et al., 2008). The subsequent differential analysis was then carried out using Sleuth (Pimentel et al., 2017). The levels of differential expression were calculated as β-values (adjusted log2 fold change) and only reported as significant if $p < 0.05$. Heatmaps of differentially expressed genes were generated using the heatmap2 tool within Galaxy (Community, 2022).

### 2.12.2 Quantitative reverse transcription PCR (RT-qPCR)

RT-qPCR was carried out with Jinyan Wang (UEA/Ocean University of China). Genomic DNA elimination and reverse transcription of the samples were carried out as described in (Curson et al., 2017). Primers for RT-qPCR of *L. aggregata* LZB033 *dsyB, dddL* and housekeeping genes *recA* and *rpoD* were used as reported (Curson et al., 2017). Primers for RT-qPCR of *gsdmt*, *aro8* and *at2* were designed using Premier 6.0 (Premier Software Inc., Cherry Hill, NJ) and synthesised by IDT. A full list of the primers used in the RT-qPCR reactions are listed in **Table 6**.

qPCR was performed used an AriaMix Real-Time PCR System instrument (Agilent). The total volume of each reaction was 20 µl, and contained 10 ng cDNA, primers at concentrations of 400 nM (*dsyB, recA, gsdmt, aro8, at2*) or 300 nM (*rpoD, dddL*), with an elongation temperature of 60 °C.

Table 6 A list of primers used in RT-qPCR

| Primer | Sequence | Reference |
|---|---|---|
| LZB033_dddL_FOR_qpcr | CGCTCCTGAAACGCAGATA | (Curson et al., 2017) |
| LZB033_dddL_REV_qpcr | GGCAGTATGGCTACGAGAAA | (Curson et al., 2017) |
| LZB033_dsyB_FOR_qpcr | CTTGACGCCACAGCATGTTG | (Curson et al., 2017) |
| LZB033_dsyB_REV_qpcr | TCCGTCCTTTCACCAGAAAC | (Curson et al., 2017) |
| LZB033_gsdmt_FOR_qpcr | GAGGAGCTTCTCGCCAACTACG | This study |
| LZB033_gsdmt_REV_qpcr | TCAGACAGGCTTGCGGAACAG | This study |
| LZB033_aro8_FOR_qpcr | TCATCCGCCAGCTTGCCGATA | This study |
| LZB033_aro8_REV_qpcr | TGGAGCACTCACGCACAGACA | This study |
| LZB033_at2_FOR_qpcr | CGACAAGTGTAATGCCGCTGGT | This study |
| LZB033_at2_REV_qpcr | GCACAACGCCAAGCCTGAATC | This study |
| LZB033_recA_FOR_qPCR | CACTGGAAATTGCCGATACG | (Curson et al., 2017) |
| LZB033_recA_REV_qPCR | CACCATGCACTTCGACTTG | (Curson et al., 2017) |
| LZB033_rpoD_FOR_qPCR | ACAAGTTCTCCACCTATGCG | (Curson et al., 2017) |
| LZB033_rpoD_REV_qPCR | CGATTTCATGCAGCATCTGG | (Curson et al., 2017) |

### 2.12.3 β-galactosidase assays

*L. aggregata* WT and *L. aggregata*: pBIO23(1) were grown in 10 ml MBM cultures of varying salinities, as described in Section 2.2. Aliquots of 0.5 ml of each culture were added to 0.5 ml Z-buffer (1 ml 3M $Na_2HPO_4.7H_2O$, 0.5 ml 4 mM $NaH_2PO_4.H_2O$, 0.5 ml 1M KCl, 0.5 ml 0.1 M $MgSO_4.7H_2O$, 175 µl β-mercaptoethanol, adjusted to a final volume of 50 ml with $H_2O$). To each sample, 2 drops of chloroform and 1 drop of 0.1 % SDS were added, and samples were vortexed for ten seconds to lyse the cells. The samples were incubated at 28 °C for five minutes and then 200 µl 4 mg/ml ONPG (ortho-Nitrophenyl-β-galactoside) was added to each sample. The samples were immediately incubated at 28 °C until a sufficient yellow colour had developed, and then 500 µl $Na_2CO_3$ was added to stop the reaction and the time taken for the colour change to occur was recorded. The samples were centrifuged at 18,440 x *g* for 5 minutes and 1 ml supernatant was transferred to a cuvette and the absorbance at $OD_{420}$ was measured using a spectrophotometer. The $OD_{600}$ of the initial culture was also measured. All β-galactosidase assays were carried out in biological triplicate. The following formula (in which t = time, v = volume of culture used) was used to calculate the β-galactosidase activity (Miller Units) of each sample:

$$\text{Units of β-galactosidase activity (Miller Units)} = \frac{1000 \times OD_{420}}{t \times v \times OD_{600}}$$

## 2.13 Protein purification

### 2.13.1 *E. coli* induction (pET vectors)

5 ml of LB containing ampicillin were inoculated with a single colony of *E. coli* BL21 transformed with a pET16b or pET21a vector and incubated overnight at 37 °C with shaking at 180 rpm. A 2 ml aliquot of this starter culture was used to inoculate 200 ml LB with ampicillin, which was incubated at 37 °C with shaking at 200 rpm until an $OD_{600}$ of 0.6 was reached. Isopropylthio-β-galactoside (IPTG) was added to the culture, at a final concentration of 20 µM and the cultures were incubated at 16 °C overnight with shaking at 180 rpm.

### 2.13.2 Buffers used in protein purifications

A list of the buffers used in protein purifications can be found in **Table 7**.

Table 7: A list of buffers used in protein purification work in this study

| Buffer | Components |
|---|---|
| Extraction Buffer | 25 mM Tris-HCl (pH 7.6), 150 mM NaCl, 1 mM DTT |
| Equilibration Buffer | 25 mM Tris-HCl (pH 7.6), 150 mM NaCl, 15 mM $MgCl_2$ |
| Final Buffer | 250 mM Tris-HCl (pH 7.6), 10 mM DTT |
| Lysis Buffer | 20 mM Tris-HCl (pH 7.4), 200 mM NaCl, 1 mM EDTA, 1 mM DTT |
| Washing Buffer | 20 mM Tris-HCl (pH 7.4), 200 mM NaCl, 1 mM EDTA |
| Buffer 1 | 20 mM Tris-HCl (pH 7.4), 200 mM NaCl |
| Elution Buffer | Buffer 1, 10mM Maltose monohydrate |

### 2.13.3 Protein extraction (pET vectors)

The 200 ml induced cultures were centrifuged at 1,900 x *g* at 4 °C for 20 minutes to pellet the cells. The supernatant was discarded and the pellet re-suspended in 15 ml Extraction Buffer. The samples were lysed using a cell disruptor. A 100 µl aliquot was removed and stored for later analysis via SDS-PAGE. The lysed samples were centrifuged at 9000 x *g* for 30 minutes to pellet the insoluble material. The soluble fraction was transferred into a fresh tube and small aliquots of both fractions were stored for later SDS-PAGE analysis. The soluble fraction was applied to a spin concentrator and centrifuged for 30 minutes at 4 °C until the volume was reduced to 3 ml. Five rounds of 5 ml Equilibration Buffer were run through a PD10 column (Cytiva) and then 2.5 ml of the concentrated protein sample was added to the column and allowed to run through. Protein samples were eluted from the column in 3.5 ml Equilibration Buffer and collected. To this, 400 µl Final Buffer and 100 µl $H_2O$ was added and the protein samples were separated into 100 µl aliquots, flash frozen in liquid nitrogen and stored at -80 °C.

### 2.13.4 *E. coli* induction (pMAL-c2X)

5 ml of LB containing carbenicillin were inoculated with a single colony of *E. coli* BL21 transformed with a pMAL-c2X vector and incubated overnight at 37 °C with shaking at 180 rpm. The starter culture was used to inoculated 500 ml LB + 0.05 % glucose + carbenicillin and was incubated at 37 °C with shaking at 180 rpm until an $OD_{600}$ of 0.6 was reached. IPTG

was added to the culture, at a final concentration of 500 µM. and the cultures were incubated at 28 °C for two hours with shaking at 180 rpm.

### 2.13.5 Protein extraction and purification (pMAL-c2X)

The induced cultures were centrifuged at 1,900 x *g* at 4 °C for 20 minutes to pellet the cells. The supernatant was discarded and the pellet re-suspended in 20 ml Lysis Buffer. The cells were lysed using a cell disruptor. The samples were centrifuged at 9000 x *g* for 30 minutes to pellet the insoluble material. For each sample to be processed, a 1 ml aliquot of amylose resin was added to a 2 ml tube. To remove the ethanol in which the amylose resin is stored, the 2 ml tube was centrifuged at 100 x *g* for 1 minute, and the supernatant was removed. 1.5 ml Washing Buffer was added to the resin and centrifuged at 100 x *g* for 1 minute and the supernatant was removed. This step was repeated four times to equilibrate the resin. The resuspended, equilibrated resin was added to the previously stored soluble protein sample, transferred to a 15 ml tube, and incubated at 4 °C with shaking. The sample was centrifuged at 100 x *g* for 1 minute and the supernatant was removed. The resin was re-suspended in 10 ml Washing Buffer and centrifuged at 100 x *g* for 1 minute and the supernatant was removed, the spins and re-suspension were repeated three times.

The resin was then transferred to a PD10 column (Cytiva) and 10 ml Buffer 1 was also added to the column. A stopper was placed on the PD10 column, and 1 ml Elution Buffer was added and incubated for 5 minutes before the stopper was removed the flowthrough was collected. A further 1 ml Elution Buffer was added to the PD10 column and incubated for 5 minutes before the stopper was removed the flowthrough was collected. The samples were split into 100 µl aliquots which were flash-frozen in liquid nitrogen and stored at -80 °C.

## 2.14 Protein analysis by SDS-Polyacrylamide Gel Electrophoresis (SDS-PAGE)

The whole cell lysate, insoluble and soluble fractions of the protein extractions and purifications were analysed via SDS-PAGE. The SDS-PAGE gels were composed of a resolving gel and stacking gel, the components of which can be found in **Table 8**. The SDS-PAGE running buffer used was diluted from a 10 X stock, which was comprised of 3 % TRIS, 14.4 % glycine and 1 % SDS in Milli-Q water. The protein samples were diluted as necessary

and mixed with SDS-PAGE loading buffer (Invitrogen). In each gel, a well was loaded with PageRuler™ ladder (Thermo Scientific). Gels were resolved at 200 V for 30 minutes and then 120 V for 90 minutes and stained with Rapid Coomassie Blue overnight. Following staining, gels were washed in water with shaking for 1 hour before imaging.

Table 8: Components of an SDS-PAGE gel

| Component | Volume for Resolving Gel | Volume for Stacking Gel |
|---|---|---|
| 1.5 M TRIS (pH 8.8) | 2.5ml | 0 |
| 0.5 M TRIS (pH 6.8) | 0 | 750 µl |
| $H_2O$ | 3.95 ml | 1.65 ml |
| 10 % SDS | 100 µl | 30 µl |
| 30% Acrylamide/Bis solution (Severn Biotech) | 3.35 ml | 502.5 µl |
| TEMED (Sigma) | 5 µl | 3 µl |
| 10 % APS (0.5 g APS in 5 ml $H_2O$) | 100 µl | 30 µl |
| Total | 10 ml | 3 ml |

## 2.15 Detection of amino acids via HPLC

To determine if amino acids could be detected in future enzyme assays, solutions were made up containing 30 µl soluble fraction from protein purification attempts, 2 mM amino acids (methionine or glutamate), 2 mM 2-oxoglutarate, 0.15 mM pyridoxal phosphate (PLP), 50 mM Tris Base (pH 9) and made up to 200 µl with $H_2O$. These solutions were then subjected to a 1:5 dilution and mixed 1:1 with OPA (Phthaldialdehyde Reagent, Merck) and incubated for 5 minutes to derivatise. Standards of 2 mM amino acids were also treated in the same way to act as positive controls.

For HPLC detection, 20 µl of the derivatised solution was injected onto a Synergi 4 µm Hydro-RP 80 Å, LC Column 50 x 1 mm. The solvent system was Buffer A ($NaH_2PO_4$ adjusted to pH2 with phosphoric acid) and Buffer B (50:40:10 methanol: acetonitrile: water) The HPLC method was as follows: A gradient of 10 – 80 % Buffer B is run against Buffer A over 19 minutes (0.25 ml/minute). 80 % Buffer B is held for 5 minutes (0.25 ml/minute). Buffer B is dropped to 10 % over 30 seconds (0.25 ml/minute). Hold 10 % Buffer B for 1 minute (0.25 ml/minute).

# Chapter 3: Using Transcriptomics to Understand the Regulation of DMSP and GB Synthesis in *Labrenzia aggregata*

## 3.1 Introduction

### 3.1.1 Relationship between DMSP and GB production

As discussed in Section 1.4, organisms can accumulate DMSP and GB in response to different environmental conditions, such as increased salinity or other stresses. Additionally, in phytoplankton species that produce both DMSP and GB, it has been shown that the relative proportions of these compounds fluctuate, in response to changes in both salinity and temperature (Dickson & Kirst, 1986; Sheets & Rhodes, 1996).

In some phytoplankton, such as the unicellular alga *Tetraselmis subcordiformis,* an inverse relationship between intracellular DMSP concentrations and nitrogen availability has often been reported (Gröne & Kirst, 1992; Keller & Korjeff-Bellows, 1996). Conversely, a positive correlation has been observed between nitrogen availability and intracellular GB concentration (Keller et al., 1999; Keller et al., 2004). This relationship is particularly prevalent in the short-term, as addition of nitrogen to existing nitrogen-limited cultures resulted in short-term increases in GB synthesis (Keller et al., 1999).

Due to the structural similarity between DMSP and GB (**Figure 1.1**), it has been hypothesised that a reciprocal relationship may exist between these compounds, with DMSP synthesis being favoured over GB synthesis in nitrogen limiting conditions (Andreae, 1986). Whereas, in nitrogen replete conditions, GB synthesis is favoured over DMSP synthesis. Further, this hypothesis was extended to suggest that DMSP may be favoured over GB in marine environments where sulfate is a lot more abundant (~28 mmol $L^{-1}$) than nitrogen (1-10 µmol $L^{-1}$) (Andreae, 1986). When nitrogen is not limiting, the cells may begin to produce nitrogen-containing molecules such as GB, to replace the DMSP (Andreae, 1986). These hypotheses predict that GB is perhaps a favourable osmoprotectant over DMSP in marine environments.

However, the proposed inverse relationship between intracellular concentrations of DMSP and GB has been contested. For instance, Keller et al., studied this relationship in three species of marine phytoplankton (*Thalassiosira pseudonana, Emiliania huxleyi* and *Amphidinium carterae*), and only found evidence of this inverse relationship in one strain (*T. pseudonana*), where the intracellular concentrations of DMSP and GB where within the same order of magnitude. In the other strains tested, the dominance of DMSP made a completely reciprocal relationship with GB unlikely, as although they accumulated GB in the short term, following nitrogen addition, they likely did not eliminate equivalent amounts of DMSP (Keller et al., 1999).

### 3.1.2 Overview of DMSP and GB synthesis genes in *L. aggregata* LZB033

As a producer of both DMSP and GB, *L. aggregata* LZB033 is an ideal species to study the regulation of the synthesis of both compatible solutes, and the relationship between them. As a utiliser of the DMSP transamination synthesis pathway, *L. aggregata* LZB033 contains the *dsyB* gene, encoding the *S*-methyltransferase responsible for catalysing the rate limiting and committing step of DMSP synthesis (as previously discussed in Section 1.5.1). However, the identity of the genes responsible for the other three steps of DMSP synthesis via this pathway (**Figure 1.2**) are unknown. Analysis of the genetic regions surrounding *dysB* in *L. aggregata* did not identify these missing genes, with their products not likely to possess the activities required for the remaining steps in the transamination synthesis pathway. In addition to producing DMSP, *L. aggregata* can also catabolise DMSP via the lysis pathway using the DMSP lyase protein encoded by the *dddL* gene, but how this important gene is expressed in response to environmental stress conditions has received little attention (Curson et al., 2017).

Previous work by Dr. Ana Bermejo Martinez (Todd lab, UEA) indicated that *L. aggregata* possesses homologues of the *betA* and *betB* genes (involved in the choline pathway in *E. coli*) and a homologue of the *gsdmt* gene (involved in the glycine pathway in *T. pseudonana*) (Bermejo Martinez, 2019) as described in Sections 1.8.1 and 1.8.2. This is consistent with data reported by Zhong and colleagues, which also identified the genes involved in the two

GB synthesis pathways in the *L. aggregata* genome (Zhong et al., 2021). The proposed DMSP and GB synthesis pathways in *L. aggregata* are detailed in **(Figure 3.1)**.

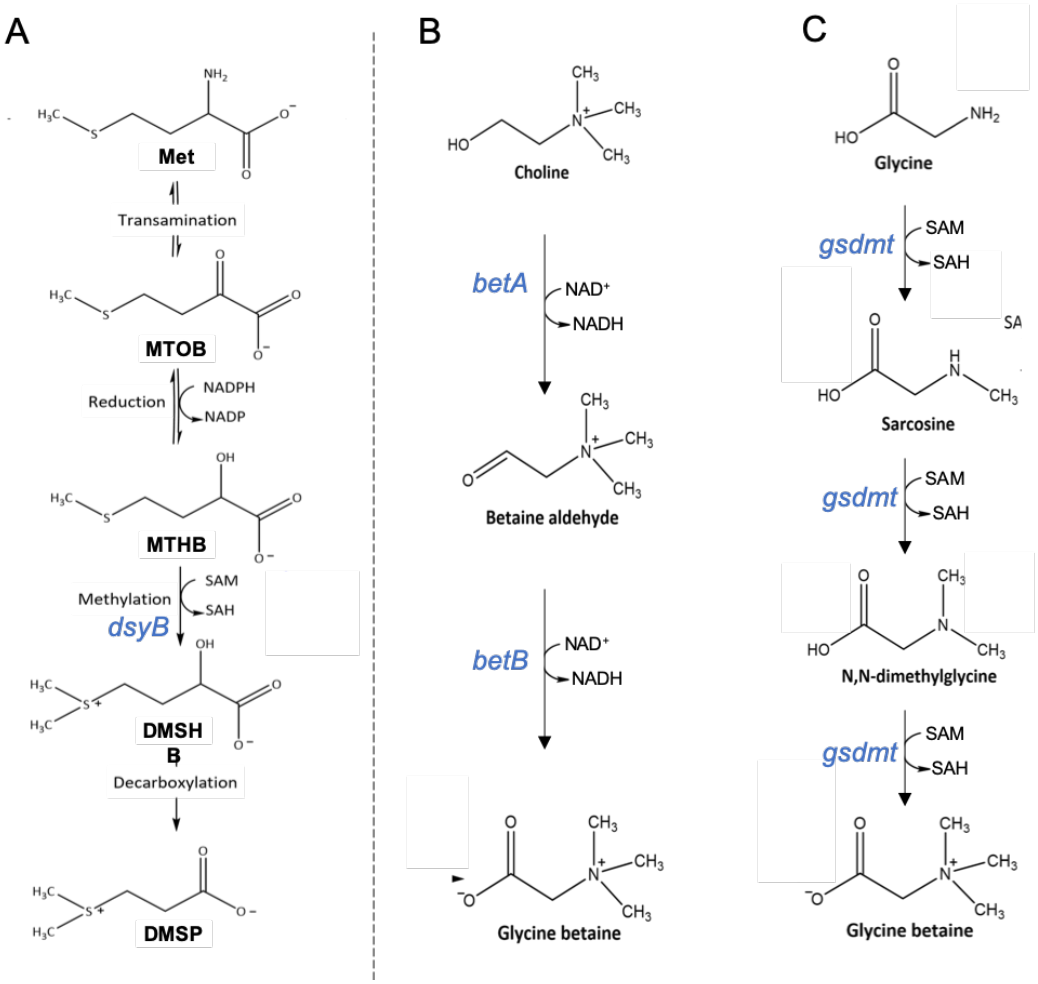



Figure 3.1 DMSP and GB synthesis pathways in *L. aggregata.*
A: The transamination DMSP synthesis pathway. B and C: The GB synthesis pathways.

In *L. aggregata*, the putative *bet* genes are located in close proximity to other GB metabolic gene homologues (**Figure 3.2**). The *betA* gene is directly upstream of the *betB* gene, and directly downstream of the known transcriptional regulator of the *betIBA* operon, *betI.* Previous work in Todd's lab using *betA-lacZ* fusions showed that *betA* transcription was upregulated by choline availability and not by salinity. Indeed, in other bacteria, such as the soil bacterium *Acinetobacter baylyi, betI* has been shown to be a choline sensing transcriptional repressor, which is released from the DNA in the presence of choline (Scholz et al., 2016). The three genes just upstream of this *betIBA* operon in *L. aggregata* are annotated as homologues of the components of an ABC transporter with high affinity to choline. The homologues of these genes have been described in species such as *Agrobacterium tumefaciens* and encode: the substrate binding protein ChoX, the permease

protein ChoW, and the ATP binding protein ChoV (Aktas et al., 2011). This ChoVWX ABC transporter has been shown to import choline into the cell (Aktas et al., 2011).

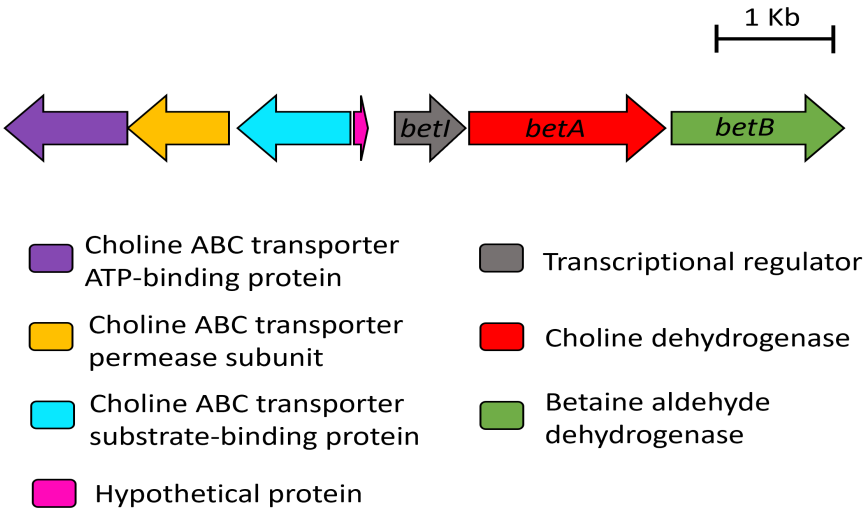



Figure 3.2 Gene neighbourhood of *betA* and *betB* homologues in *L. aggregata* LZB033.
The key denotes the protein encoded for by each of the genes shown, as predicted using blastX analysis and RASTtk annotation of the genome.

Conversely, the *L. aggregata gsdmt* homologue, proposed to be involved in the glycine GB synthesis pathway, does not appear to be located nearby to any other genes involved in GB metabolism. Therefore, this genetic analysis indicates that *L. aggregata* synthesises DMSP via the transamination pathway (*dsyB*) and synthesises GB using both the choline (*betAB*) and glycine (*gsdmt*) pathways.

### 3.1.3 Regulation of GB and DMSP Synthesis in *L. aggregata* LZB033

Taken individually, there has been limited research into the regulation of GB synthesis in *L. aggregata*. However, Curson and colleagues have reported several factors affecting *dsyB* transcription and DMSP production. DMSP production was increased in *L. aggregata* when the growth culture medium was supplemented with intermediates of the DMSP transamination synthesis pathway (Met, MTOB, MTHB and DMSHB) (Curson et al., 2017). Further, it was shown that DMSP production increased in high salinity, in low temperatures (16°C), under low nitrogen conditions and in stationary phase compared to exponential phase.

Within the Curson et al., study, the level of transcription of *dsyB* and *dddL* were also measured under the different conditions. As expected, the transcription of *dsyB* increased in all the aforementioned conditions (high salinity, low temperature, low nitrogen and in stationary phase) shown to increase DMSP synthesis (Curson et al., 2017). The transcription of *dddL* decreased in all of the conditions tested compared to the standard condition. This implies that the regulation of DMSP synthesis and catabolism in *L. aggregata* results in reduced catabolism of DMSP when the strain is under DMSP synthesis inducing conditions.

### 3.1.4 Interplay of DMSP and GB Synthesis Regulation in *L. aggregata*

In the existing literature, it is proposed that the production of the compatible solutes, DMSP and GB, is regulated by nitrogen availability. However, this proposed interplay has been contested (Section 3.1.1). Therefore, in this work, the production of compatible solutes by *L. aggregata* under different growth conditions will be assessed to indicate whether this interplay occurs.

Within this chapter, genome-wide transcriptional changes caused by DMSP synthesis inducing and repressing growth conditions were studied using RNA sequencing (RNA-Seq). The expression of DMSP (*dsyB*) and proposed GB (*betAB* and *gsdmt*) synthesis genes are further studied using reporter gene fusions and quantitative reverse transcription polymerase chain reaction (RT-qPCR). This analysis aimed to further improve our understanding of DMSP and GB synthesis regulation in *L. aggregata*.

### 3.1.5 RNA-sequencing

Since the development of RNA-Seq in 2007, use of the technique has become ubiquitous in molecular biology and is now the gold standard for measurement of global gene expression levels (Emrich et al., 2007; Lister et al., 2008; Mortazavi et al., 2008). Generally, the workflow begins with the extraction of mRNA from the biological sample of interest, the sequencing of the obtained mRNA and finally, computational analysis.

The sequencing of extracted RNA relies upon technology developed for high-throughput DNA sequencing. As such, the extracted mRNA is first converted to a library of cDNA fragments, with adaptor sequences attached to one or both ends (Wang et al., 2009). High throughput sequencing of these fragments then begins from either one end (single-end sequencing) or from both ends (pair-end sequencing). Depending on the sequencing technology being used, reads are ~30-400 bp in length (Wang et al., 2009). The resultant reads can then be aligned to a reference genome or assembled without the genomic sequence. Either way, the output of RNA-Seq is a transcription map that covers the entire genome and the read frequency is used to express gene expression levels (Wang et al., 2009).

RNA-Seq has advantages over other methods used to study bacterial transcription. These alternative methods, such as microarray studies, rely on the hybridization of targeted oligonucleotides to specific sequence regions. For example, primers are designed to bind to regions of interest in RT-qPCR and cDNA binds to labelled probes in microarray studies. RNA-Seq differs from the aforementioned transcriptional techniques as all cellular transcription is studied, without the use of specifically designed probes. Thus, RNA-Seq is an unbiased approach to transcriptional studies and is regarded as the gold standard methodology (Croucher & Thomson, 2010).

Another advantage of RNA-Seq's non-reliance upon targeted oligonucleotides, is that it does not suffer from background noise. Other techniques, such as microarrays, often suffer from non-specific binding (Kane et al., 2000). RNA-Seq can also register a much wider range of expression levels compared to other transcriptomic approaches such as microarrays. For example, in an RNA-Seq experiment carried out in the common yeast model organism, *Saccharomyces cerevisiae,* gene expression was measured over a range of approximately four orders of magnitude (Nagalakshmi et al., 2008). In contrast, microarrays have low sensitivity for genes expressed at both very low and very high levels, and thus obtain a much narrower range of expression levels (Wang et al., 2009).

RNA-Seq is useful for assessing genome-wide regulation and indicating a number of genes differentially expressed under experimental conditions (the regulon). RT-qPCR is

considered the best method for validation of gene expression data from high-throughput sequencing platforms. As such, many studies rely on the results of RT-qPCR of a set number of genes to confirm the accuracy of their RNA-Seq dataset (Everaert et al., 2017).

### 3.1.6 Chapter Aims

The overall aim of this chapter is to study the regulation of genes involved in DMSP and GB synthesis in *L. aggregata* LZB033 and determine whether there is interplay between the production of these compatible solutes. As such, the sub-aims are:

1) To identify growth conditions that regulate DMSP and GB synthesis in *L. aggregata* LZB033.
2) To perform osmolyte analysis of *L. aggregata* LZB033 under each of the conditions tested to confirm that DMSP, GB and/or other osmolyte production is regulated by these conditions.
3) To perform RNA extraction and subsequently analyse RNA-Seq results of *L. aggregata* LZB033 grown under each condition.
4) To undertake bioinformatic differential expression analysis of the RNA-Seq data in order to understand the regulation of DMSP and GB synthesis genes.
5) Validate the obtained RNA-Seq results using RT-qPCR.

## 3.2 Results

### 3.2.1 Identification of *L. aggregata* growth conditions affecting DMSP production

*L. aggregata* was screened under different growth conditions that affected DMSP production. It was assumed that as this would result in the regulated transcription of DMSP synthesis genes, these conditions may also regulate transcription of GB synthesis genes.

It has previously been shown that DMSP production in *L. aggregata* was affected by varying the levels of salinity and nitrogen availability (Curson et al., 2017). Growth conditions which were reported to enhance DMSP production in this publication were used. This included the growth of *L. aggregata* in MBM media under the following conditions: high salinity (50 Practical Salinity Units (PSU)), low nitrogen (0.5 mM) and in the presence of the starting substrate for DMSP synthesis, methionine (0.5 mM).

The DMSP production of *L. aggregata* LZB033 grown under these different conditions was measured by gas chromatography and normalised to protein content as described in materials and methods. *L. aggregata* produced approximately 3-fold more DMSP in high salinity compared to the control condition and 5-fold more DMSP under low nitrogen conditions compared to the control (**Figure 3.3**). However, the biggest increase in DMSP production (~ 40-fold) occurred when methionine was present in the growth medium (**Figure 3.3**). Therefore, this experiment identified three growth conditions (+Met, low N, 50 PSU) which may induce increased expression of DMSP synthesis genes in comparison to the control condition (35 PSU MBM with 10 mM $NH_4Cl$).

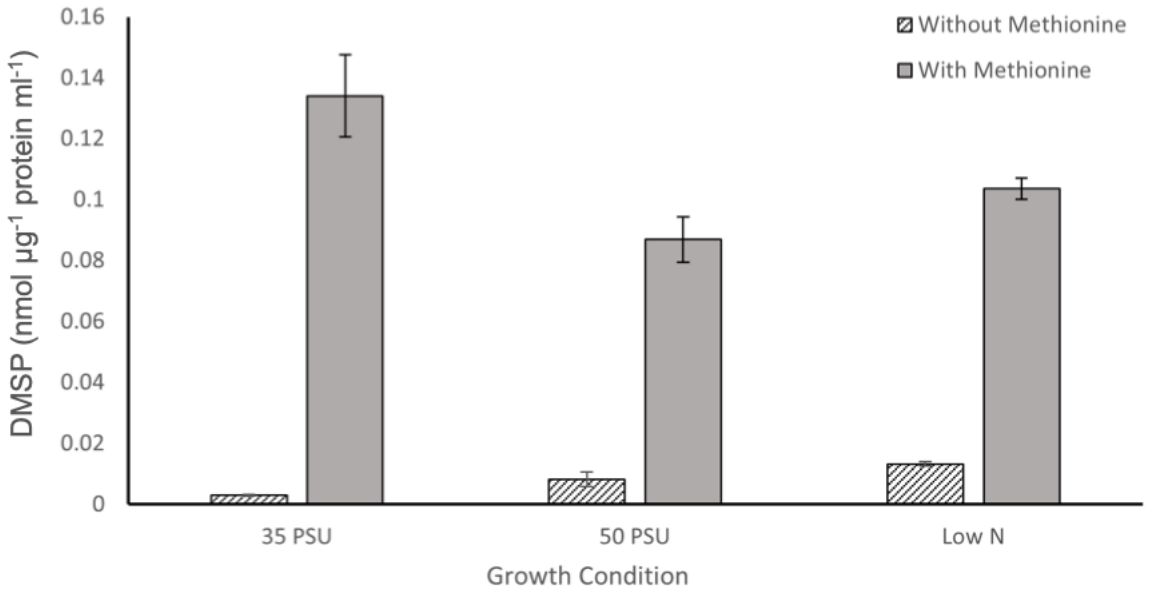



Figure 3.3 DMSP production of *L. aggregata* LZB033 under varying growth conditions.
Each condition was in MBM media and varied as follows: control condition (35 PSU, 10 mM N), high salinity (50 PSU, 10 mM N), Low Nitrogen (35 PSU, 0.5 mM N). Where indicated, methionine was added to each of the growth conditions, at a concentration of 0.5 mM. DMSP production was measuring using gas chromatography of biological triplicate cultures and normalised to the protein content (µg) of each of the cultures. Error bars represent the standard deviation of the mean.

### 3.2.2 Determination of *dsyB* regulation via β-galactosidase assays

Having identified conditions that induce DMSP production, it was necessary to determine whether these conditions enhanced the expression of the key DMSP synthesis gene *dsyB* prior to conducting the RNA-Seq. This was carried out by β-galactosidase assays, using the pBIO23(1) plasmid which features the promotor region of *dsyB* from *L. aggregata* LZB033 cloned upstream of, and thus in transcriptional control of, the *lacZ* (β-galactosidase) gene. This construct was previously produced by Dr. Andrew Curson (UEA). The pBIO23(1) plasmid was conjugated into *L. aggregata* LZB033 via tri-parental mating and the resultant strain was grown under varying salinity, low nitrogen and in the presence/absence of 0.5 mM methionine. A colorimetric assay was performed to quantitate the levels of expressed β-galactosidase under each of the conditions tested. Therefore, as the *dsyB* promotor was in control of *lacZ* expression, the levels of β-galactosidase activity can be directly related to the levels of *dsyB* transcription in each of the different conditions.

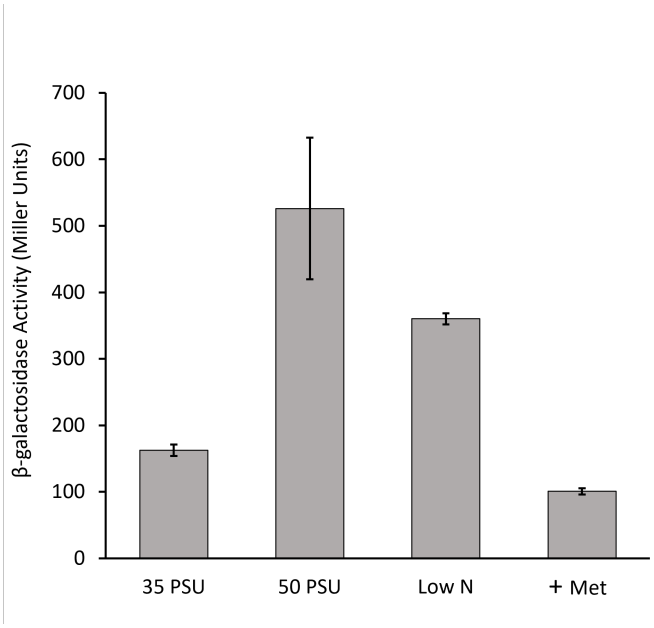



Figure 3.4 The regulation of *dsyB* as determined via β-galactosidase assays.
The assays were carried out on *L. aggregata* LZB033 transfected with pBIO(1). Each condition was in MBM media and varied as follows: control condition (35 PSU, 10 mM N), high salinity (50 PSU, 10 mM N), Low Nitrogen (35 PSU, 0.5 mM N). Where indicated, methionine was added at a concentration of 0.5 mM. N=3 biological replicates. Error bars represent the standard deviation of the mean.

Consideration of the β-galactosidase assay results (**Figure 3.4**) indicated that *dsyB* transcription is increased in high salinity and by low nitrogen availability. This was in concurrence with the regulation of *dsyB* as studied via RT-qPCR in Curson et al., 2017.

Conversely, the transcription of *dsyB* was slightly decreased when methionine was present in the growth media (**Figure 3.4**). This was unexpected, as methionine is the starting substrate for DMSP synthesis and previously caused large increases in DMSP production when present in the growth conditions (**Figure 3.3**). The effect of methionine addition to *dsyB* transcription was not studied by Curson and colleagues. Despite the unexpected effect of methionine condition on *dsyB* transcription, the β-galactosidase assays confirmed that *dsyB* transcription was regulated by each of the conditions to be used in the RNA-Seq experiment.

### 3.2.3 Identification of DMSP supressing growth conditions

For the RNA-Seq experiment, conditions were chosen that would cause both up- and down-regulation of DMSP synthesis genes. Having identified growth conditions inducing DMSP synthesis, there was a need to identify growth conditions that reduced DMSP synthesis. In the Curson et al., (2017) study, when *L. aggregata* LZB033 was grown under low salinity conditions (MBM 5 PSU), DMSP production was reduced so greatly it was below the detection limit of the gas chromatography used for DMSP detection. So, although not tested in this initial screening for conditions effecting the DMSP synthesis of *L. aggregata* LZB033, low salinity (5 PSU) was chosen as a condition to be used in the RNA-sequencing experiment.

Thus, the conditions chosen for the RNA-Seq experiment were as follows: control (MBM 35 PSU), low salinity (MBM 5 PSU), low nitrogen (35 PSU, 0.5 mM N source) and + Met (35 PSU + 0.5 mM methionine).

### 3.2.4 Growth of *L. aggregata* LZB033 under selected growth conditions

Previous studies involving other DMSP-producing organisms have determined that DMSP production is influenced by growth phase (Stefels, 2000). Further, in *L. aggregata* LZB033, *dsyB* transcription levels have been found to differ between stationary phase and exponential phase (Curson et al., 2017). Therefore, it was necessary to determine when *L. aggregata* reached mid-exponential phase under the selected growth conditions to ensure that gene expression was not affected by differing growth phases.

To determine the mid-exponential phase of the cultures, growth curves of *L. aggregata* LZB033 were constructed under each growth condition (**Figure 3.5**). To ensure consistency between the experiments, the growth curves were constructed using cultures of the same volume and exact same growth conditions to be used for the RNA-Seq experiment (see materials and methods for details).

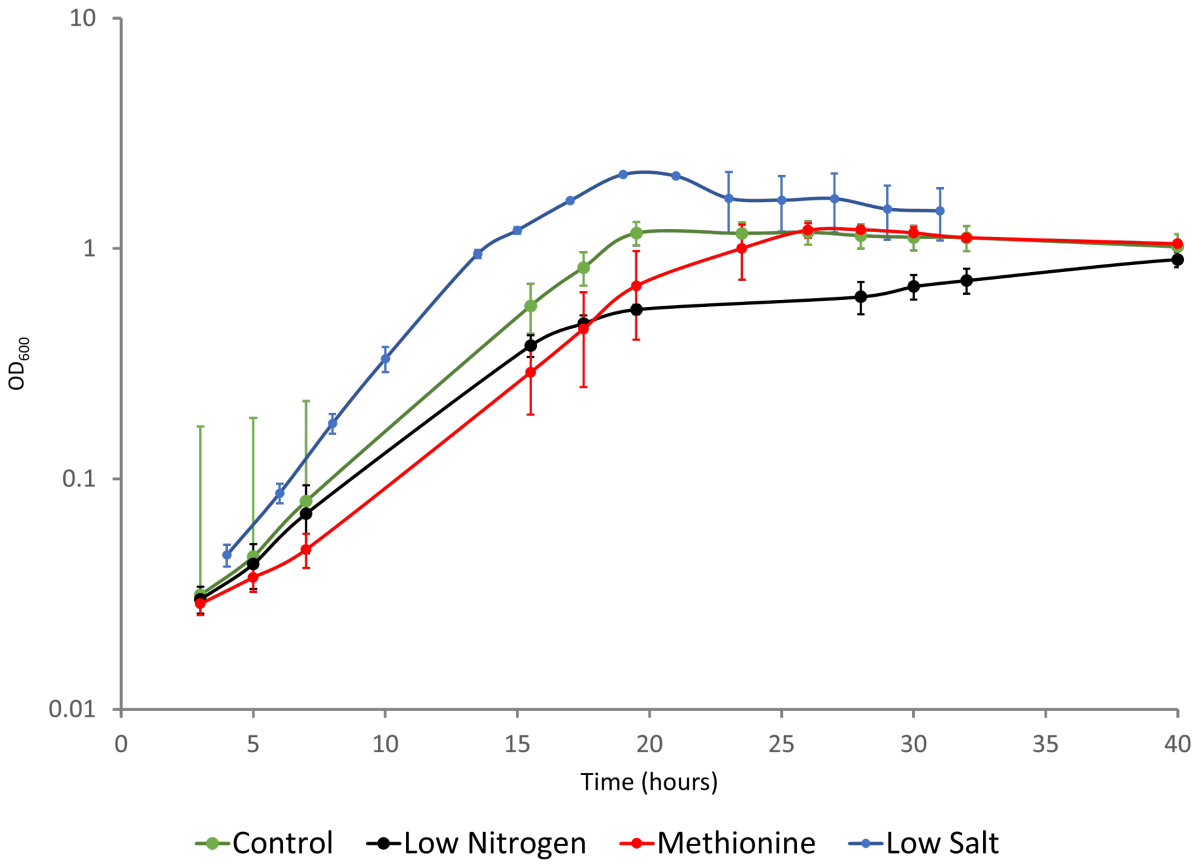



Figure 3.5 Growth curves of *L. aggregata* LZB033 under varying growth conditions.
Cultures of *L. aggregata* LZB033 were grown in 100 ml MBM media under the following conditions. Control: 35 PSU. Low Nitrogen: 35 PSU, 0.5 mM $NH_4Cl$ as the nitrogen source. Methionine: 35 PSU + 0.5 mM Met. Low Salt: 5 PSU. Unless otherwise stated, 10 mM succinate was used as the carbon source and 10 mM $NH_4Cl$ was used as the nitrogen source. The $OD_{600}$ of triplicate cultures was measured at regular intervals and error bars represent the standard deviation of the mean.

The growth curves are plotted in **Figure 3.5** and indicate that in the control, low nitrogen and +Met conditions, mid-exponential phase occurred around 13 hours after inoculation. This suggested that future cultures grown under these conditions should be sampled at 12-13 hours for RNA-Seq. Whereas, cultures grown under low salinity grew faster and reached mid-exponential phase at ~11.5 hours. Therefore, sampling cultures at these timepoints should ensure that they are all in mid-exponential phase and that the RNA-Seq preparations are directly comparable.

Analysis of the exponential phase growth rates calculated from **Figure 3.5** detailed how the growth rate of *L. aggregata* LZB033 is ~2 fold faster in low salinity compared to the control condition. Conversely, the low nitrogen and + Met condition decreased the growth rate

compared to the control condition by 36% and 50%, respectively. Additionally, the cultures grown under lower salinity reached a higher $OD_{600}$ than the other conditions. This suggested that *L. aggregata* LZB033 cultures grown in low salinity are under less stress than the other conditions.

### 3.2.5 Osmolyte analysis of *L. aggregata* LZB033 under varied growth conditions

Following the growth curve analysis, triplicate *L. aggregata* LZB033 cultures were grown to mid-exponential phase in each condition and RNA-stabilised cell pellets were stored for RNA extractions. To ensure that the samples to be used for the RNA-Seq analysis were producing differing amounts of DMSP, the content of the cultures at this mid-exponential phase was measured and compared again (**Figure 3.6**). This indicated that the low salinity samples were producing ~5-fold less DMSP than the control condition, the low nitrogen samples were producing ~7-fold more DMSP than the control condition and the +Met samples were producing ~70-fold more DMSP than the control (t-test, $p < 0.05$ in all cases). This provided confirmation that in mid-exponential phase, the DMSP production of the cultures used for RNA-Seq were regulated as expected in the different growth conditions. As such, genes involved in DMSP synthesis, including *dsyB,* should be up or down-regulated in each condition, accordingly.

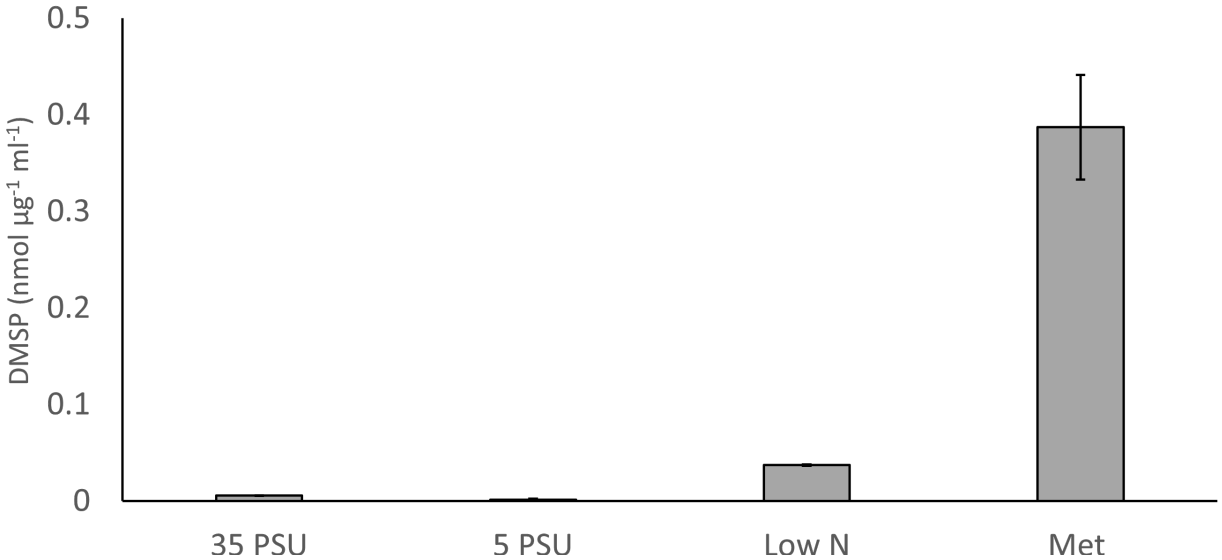



Figure 3.6 DMSP production of *L. aggregata* LZB033 cultures used for RNA-sequencing experiment. The DMSP production were measured in triplicate and normalised to the protein content of each of the cultures. Error bars denote the standard error of the mean.

In addition to using gas chromatography to assess the DMSP production of the cultures at mid-exponential phase in each condition, aliquots of these cultures were retained for assessment using liquid chromatography-mass spectrometry (LC-MS). This LC-MS analysis was conducted by Muhaiminatul Azizah (Friedrich Schiller University Jena, Germany). This LC-MS analysis provided the intracellular concentration of DMSP and GB in each of the conditions tested (**Figure 3.7**), in addition to the concentrations of other S- and N-containing zwitterionic metabolites present in the cultures (**Figure 3.8**).

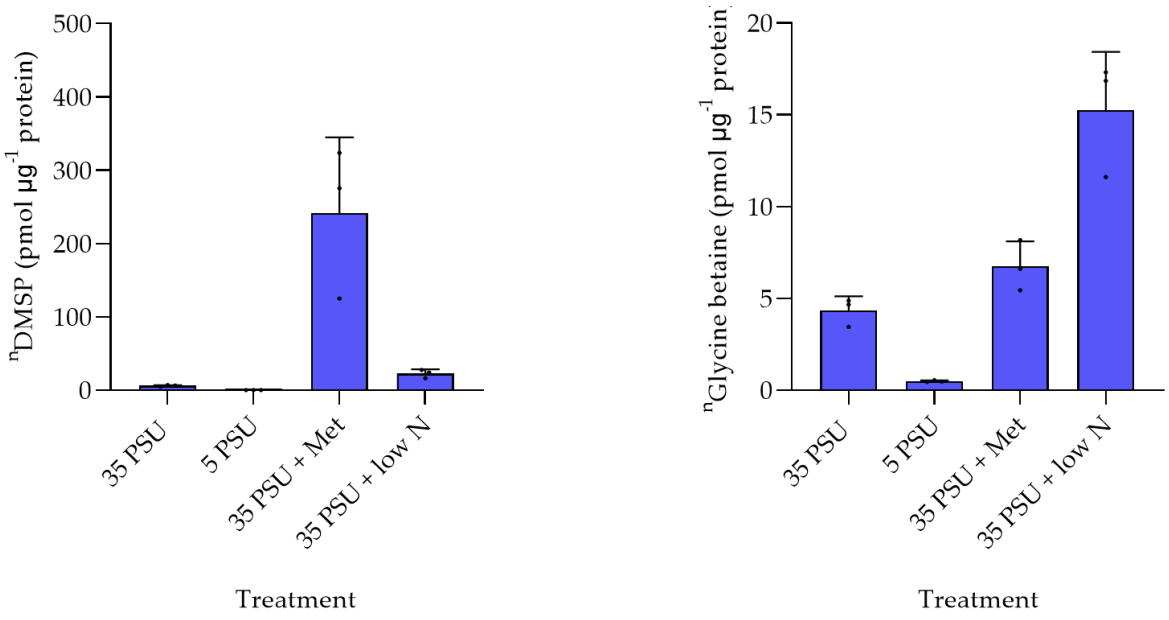



Figure 3.7 DMSP and GB concentration of *L. aggregata* LZB033 cultures sent for RNA-Seq, determined by LC-MS.
Concentration of each osmolyte was normalised to protein content and LC-MS was carried out on three biological replicates (n=3). Error bars represent the standard deviation of the mean. This LC-MS analysis was carried out by Muhaiminatul Azizah.

LC-MS analysis indicated that the amount of DMSP was greatest in the cultures grown with the addition of methionine. In comparison to the control culture, the production of DMSP was increased under low nitrogen and decreased under low salinity (**Figure 3.7**). Therefore, this LC-MS analysis is concurrent with the GC DMSP quantification (**Figure 3.6**) and provided further reassurance that DMSP production was regulated in each of the conditions used. Therefore, these exact samples could be prepared for RNA-Seq as the genes involved in DMSP synthesis production should be transcriptionally regulated.

The intracellular GB content of *L. aggregata* LZB033 under each condition was regulated as expected by salinity, as it was reduced at 5 PSU compared to the 35 PSU control (**Figure 3.7**). Surprisingly, the GB intracellular concentration was increased in the low nitrogen condition, which was unexpected due to GB being a nitrogen containing molecule, and the proposed inverse relationship between DMSP and GB, regulated by nitrogen availability. It is possible that the nitrogen concentration of the culture media would have to be nitrogen limiting in order to observe this proposed relationship. Further, GB concentration was also slightly increased by the addition of methionine to the media, which was unexpected as methionine is not involved in GB synthesis, as is the case with DMSP.

The LC-MS analysis also allows direct comparison between DMSP and GB production of *L. aggregata* LZB033 in different conditions. In the control condition, the mean intracellular concentration of DMSP was 6.0 pmol/µg protein compared to the mean GB concentration of 4.3 pmol/µg protein, and thus very similar. Similarly, in the low nitrogen condition, the DMSP concentration was 22.8 pmol/µg protein whereas the GB concentration was 15.3 pmol/µg protein.

However, the relative ratio of the compatible solutes varied under the other growth conditions. For example, the mean intracellular concentration of DMSP was significantly more increased in the + Met condition (241 pmol/µg protein) than the intracellular GB concentration (6.7 pmol/µg protein). Further, in the low salinity condition, the DMSP concentration was decreased to a lower level (0.001 pmol/µg protein) than the GB level (0.496 pmol/µg protein). Therefore, this LC-MS analysis may suggest that DMSP and GB confer similar advantages when *L. aggregata* is grown under low nitrogen conditions. However, when grown under low salinity conditions, GB may be a more important osmolyte.

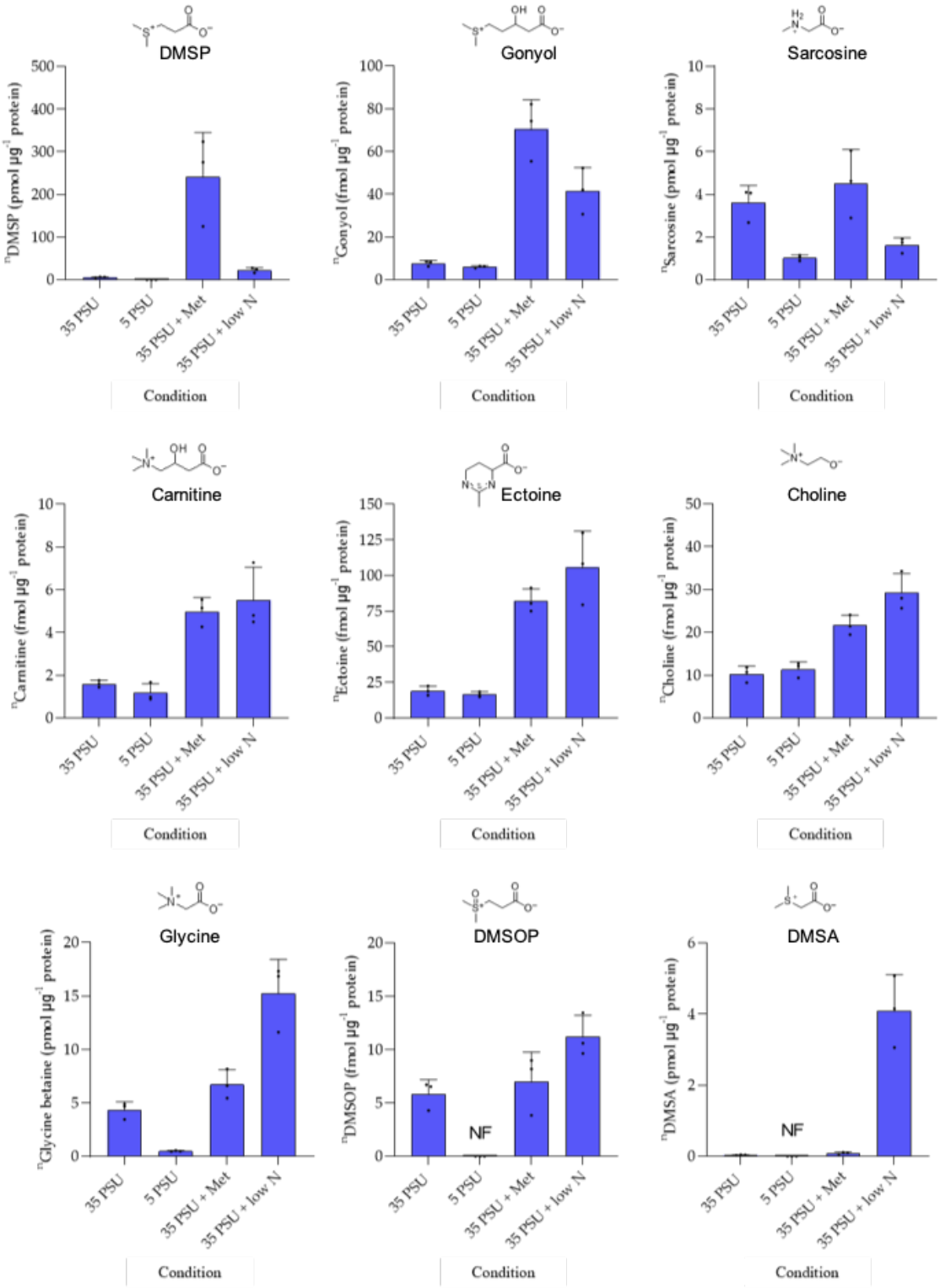



Figure 3.8 Intracellular concentration of zwitterionic metabolites of *L. aggregata* LZB033 as determined by LC-MS

Concentrations are normalised to protein content and error bars denote the standard deviation of the mean. Analysis carried out in biological triplicate (n=3). NF = not found (the compound concentration was below the limit of detection). LC-MS analysis carried out by Muhaiminatul Azizah.

This LC-MS analysis also allowed the detection and quantification of all sulfur and nitrogen containing zwitterionic osmolytes in the *L. aggregata* LZB033 cultures, under all of the tested conditions. All of these additional detected metabolites: gonyol, sarcosine, carnitine, ectoine, choline, dimethylsulfonioacetate (DMSA) and DMSOP, have been shown to be involved in stress responses in other organisms (Azizah & Pohnert, 2022; Cosquer et al., 1999). However, DMSP and GB, along with sarcosine were the only detected metabolites present in pmol concentrations, suggesting the importance of these metabolites in *L. aggregata* LZB033.

This work also represents the first report of, gonyol, carnitine, ectoine, DMSOP and DMSA produced in this bacterium. Furthermore, of these compounds, the synthesis genes are only known for gonyol, ectoine and carnitine (Bernal et al., 2007; Little et al., 2022; Louis & Galinski, 1997). Thus, the below RNA-Seq experiments also have the potential to identify the genes involved in the production of DMSOP and DMSA.

### 3.2.6 Extraction of RNA from *L. aggregata* LZB033 cultures

RNA was isolated from three biological replicates of *L. aggregata* LZB033 grown to mid-exponential phase in each of the four conditions used. The protocol for the isolation of RNA from *L. aggregata* LZB033 cultures required optimisation as standard methodology resulted in degraded RNA. The Qiagen RNeasy Mini Kit was used to extract RNA, however the volume of culture, volume of RNA stabilization solution and the method of lysing the cells required troubleshooting. A full description of the optimised protocol is detailed in Section 2.12.1.2.

Various forms of quality checks were used in order to confirm that the obtained extracted RNA was of suitably high quality for RNA-Seq. Firstly, two applications of TURBO$^{TM}$ DNAse treatment were required to fully remove DNA from the samples. The efficacy of the DNAse treatment was determined by performing a 16S PCR reaction on each of the samples (**Figure 3.9**). The lack of a PCR product in any of the samples indicated that all DNA was removed successfully. In addition to a negative (water only) and positive (*L. aggregata* gDNA) control, a third control was used. This control, a DNAse-treated RNA sample plus *L.*

*aggregata* gDNA was used to ensure that high concentrations of RNA were not preventing a successful PCR reaction when DNA was present. The lack of a PCR product in this control indicated that this was not the case and that the DNAse treatment was successful.

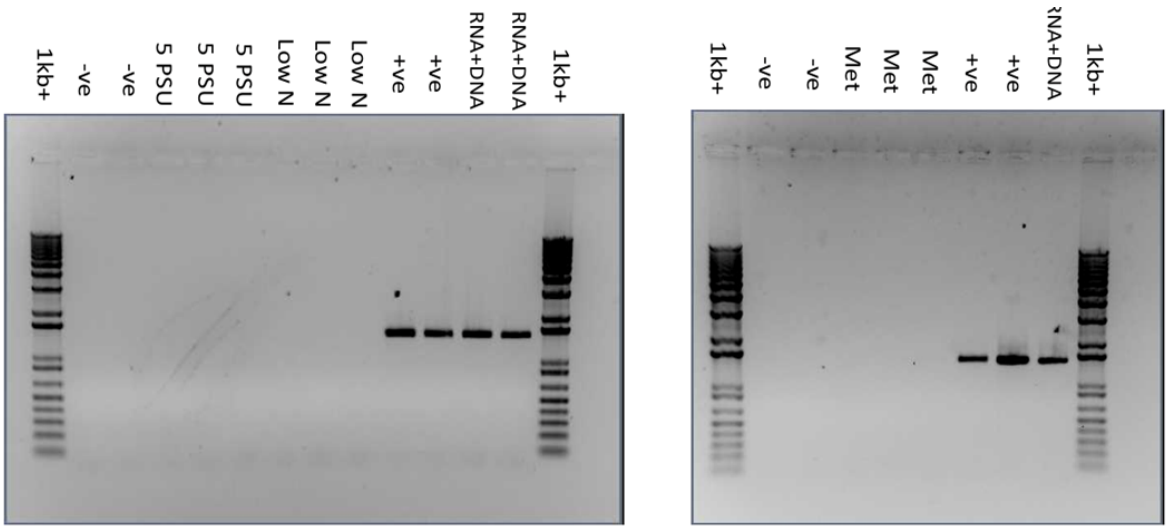



Figure 3.9 Gel image displaying successful DNAse treatment of isolated RNA.
Following two rounds of DNAse treatment, the isolated RNA from *L. aggregata* was used as template in a 16S PCR reaction. The negative (-ve) control was a control in which water was used as a template. The positive (+ve) control denotes where *L. aggregata* DNA was used as a template. The RNA + DNA control denote where both the extracted RNA and *L. aggregata* DNA were both used as templates in the same PCR reaction, to confirm that high concentrations of RNA were not interfering with DNA being amplified in the non-control reactions.

Following the DNAse treatment, the concentrations of the RNA samples were quantified using the Nanodrop and Qubit to ensure that there was sufficient quantities of material to submit for RNA-Seq. The final amount of RNA sent for RNA-Seq was >1400 ng in all cases.

Integrity of the RNA was checked using the Experion™ Automated Electrophoresis Station. Integrity of the RNA was assessed using the generated electropherogram and virtual gel image. These generated figures are based on the separation and subsequent fluorescence-based detection of the RNA sample.

In the electropherogram, non-degraded RNA presented as two clear peaks, each corresponding to the 16S and 23S prokaryotic ribosomal subunits (**Figure 3.10B**). In the virtual gel image, the samples were not smeared and there were distinct bands present at

~1500bp and ~2900 bp corresponding to the 16S and 23S prokaryotic ribosomal subunits, respectively (**Figure 3.10A**), indicating that the RNA samples were non-degraded. Therefore, these RNA samples were of high integrity.

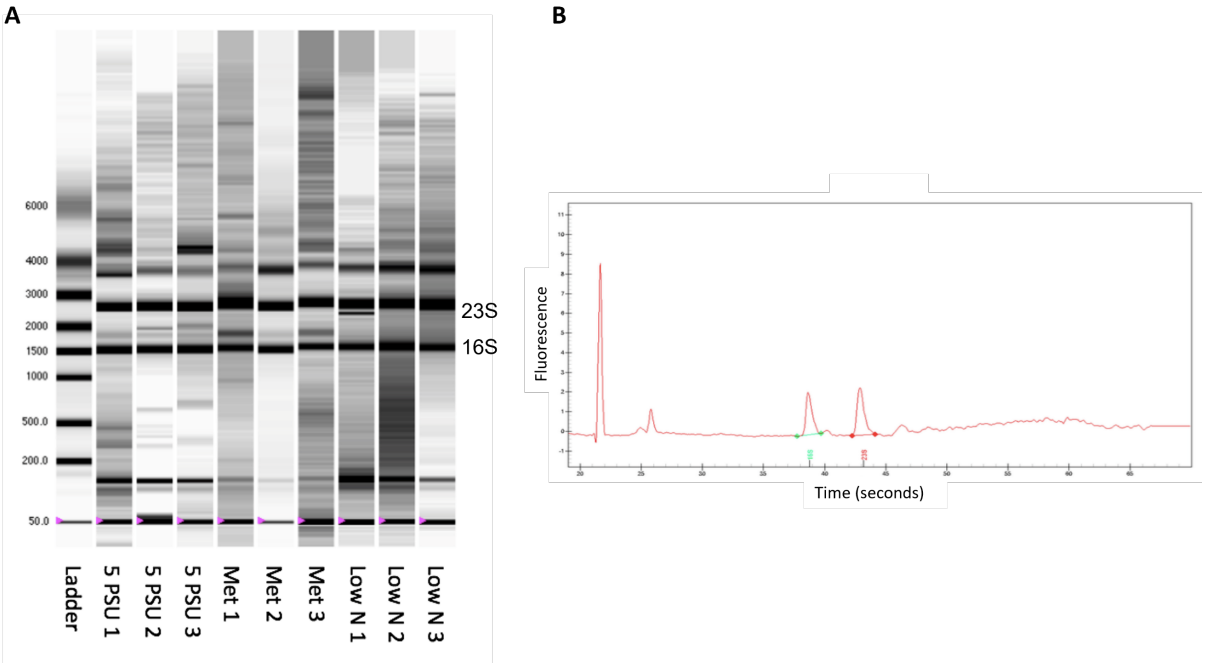



Figure 3.10 Virtual gel image of RNA integrity.
A: Virtual gel image rendered from analysis using the Experion™ Automated Electrophoresis Station. Lane 1 contains the RNA ladder and detection is fluorescence based. The bands corresponding to the size of the 16S and 23S prokaryotic ribosomal subunits are indicated. B: Electropherogram graph of one RNA sample analysed on the Experion™ Automated Electrophoresis Station. Peaks corresponding to the RNA subunits, used as a marker for RNA integrity are indicated.

Once the high concentration and integrity of the isolated RNA samples were confirmed, the samples were sent to Novogene for RNA-Seq. The sequencing was completed to a high standard, with the error rate ≤ 0.03 % in all samples (**Table 9**). Additionally, >97% of bases in all of the samples were given a Q20 score; indicating that the probability of an incorrect base being called was 1 in 100, thus the accuracy in these cases can be described as 99 %. Further, > 93 % of all bases were given a Q30 score; the probability of an incorrect base being called is 1 in 1000, as such, the accuracy is described as 99.9 %.

Table 9 Quality of RNA-sequencing data as performed by Novogene.
The number of clean reads in each sample are shown (reads which did not contain adaptor sequences as part of the Illumina sequencing process). Also shown is the sequencing error rate and the quality scores Q20 and Q30. The GC content is also displayed.

| RNA Sample | Clean Reads | Error Rate (%) | Q20 (%) | Q30 (%) | GC content (%) |
|---|---|---|---|---|---|
| 35 PSU | 8712471 | 0.03 | 97.78 | 93.62 | 58.59 |
| 35 PSU | 9735138 | 0.02 | 98.11 | 94.41 | 58.62 |
| 35 PSU | 10351547 | 0.02 | 98.12 | 94.41 | 58.55 |
| 5 PSU | 10157705 | 0.03 | 98.04 | 94.23 | 58.69 |
| 5 PSU | 8531270 | 0.03 | 98.07 | 94.30 | 58.57 |
| 5 PSU | 9408970 | 0.03 | 97.72 | 93.53 | 58.94 |
| Met | 9777657 | 0.03 | 97.92 | 94.03 | 58.64 |
| Met | 8689896 | 0.03 | 97.88 | 93.91 | 58.13 |
| Met | 10765616 | 0.03 | 97.77 | 93.67 | 58.67 |
| Low N | 7912777 | 0.03 | 97.78 | 93.72 | 58.30 |
| Low N | 10263966 | 0.03 | 97.71 | 93.59 | 58.57 |
| Low N | 8319642 | 0.03 | 97.80 | 93.76 | 58.48 |

### 3.2.7 Analysis of RNA-sequencing data

Before analysis of the RNA-Seq data could commence, the *L. aggregata* LZB033 genome was annotated using RASTtk (Brettin et al., 2015). The RNA sequence trimming, removal of adaptor sequences, mapping to the *L. aggregata* LZB033 genome, quantification of gene expression and subsequent differential analysis of the sequencing data was carried out by Dr. Simon Moxon (UEA).

The outcome of the bioinformatical analysis carried out by Dr. Simon Moxon was the normalised count data of each gene under each condition and the differential expression of each gene, as determined via a pairwise comparison between the expression levels of each gene in the control condition (35 PSU) and each of the other conditions. The degree of differential expression of each gene is displayed as a β-value (an adjusted log2 fold

change of each gene compared to the control condition). As such, a negative β-value represents a down-regulation of the gene compared to the control condition and a positive β-value represents an up-regulation of the gene compared to the control condition. A change in expression was considered significant if the adjusted p-value is ≤ 0.05.

In order to determine the similarity between the biological replicates and the similarity between the samples from different conditions, a principle component analysis (PCA) plot was generated by Dr. Simon Moxon (**Figure 3.11**). The PCA plot depicts the high degree of similarity between biological replicates in respect to the first two principal components and good separation between the different conditions (**Figure 3.11**). Further, it can be seen that the RNA-Seq data from the control condition (35 PSU) and the methionine condition are more similar to each other than the other conditions in respect to the first two principal components. This is also the case when looking at the number of statistically significant differentially expressed genes between these two conditions (2681) which is lower than the number of differentially expressed genes between the control and low nitrogen condition (4525) and the low salinity condition (3468). In all conditions, a high proportion of the total genes in the genome (5947) were differentially expressed compared to the control condition.

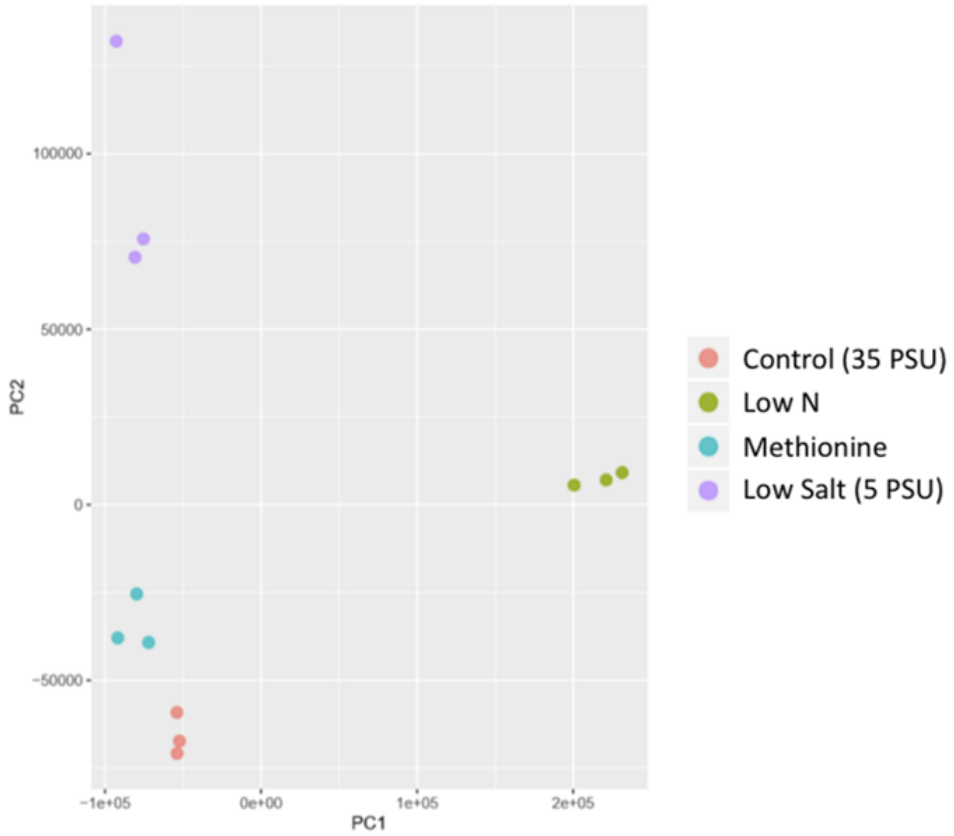



Figure 3.11 Principal component analysis plot of *L. aggregata* LZB033 RNA-sequencing data. The PCA plot was generated from regularized-logarithm transformed RNA-Seq data and the first two principal components are shown. PCA plot was generated by Dr Simon Moxon.

### 3.2.8 Analysis of RNA-sequencing differential gene expression analysis

A heatmap was generated of the ten most up- and down-regulated genes across each of the different conditions (**Figure 3.12**). As to be expected, many of the most differentially expressed genes between the control and low nitrogen condition are those involved in nitrogen metabolism. Five of the most up-regulated genes in the low nitrogen condition are annotated as nitrite or nitrate transporters, as is consistent with the bacterium sensing the low nitrogen concentration and upregulating systems to scavenge nitrogen from its environment.

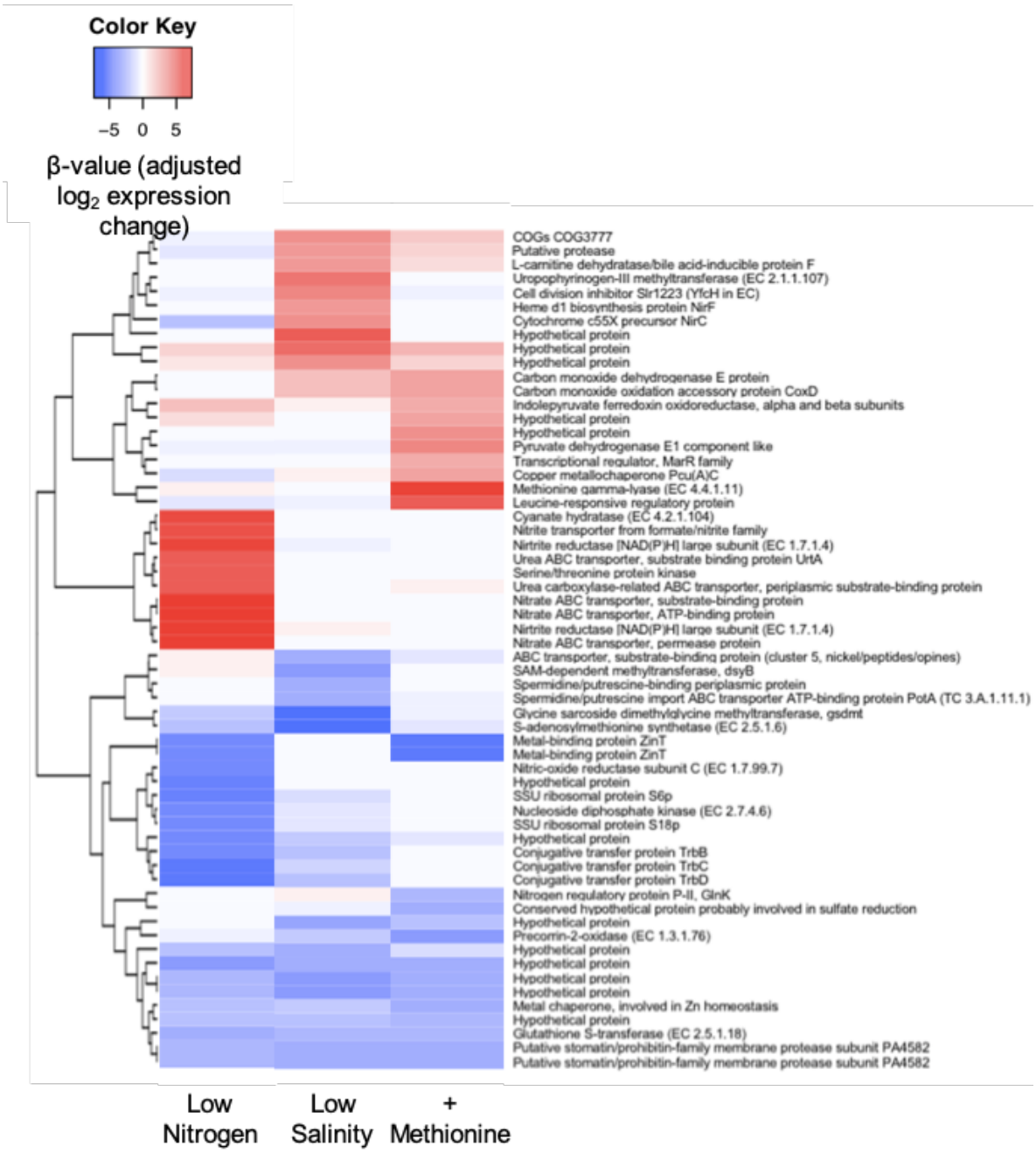



Figure 3.12 Heatmap of the most differentially expressed genes in the *L. aggregata* LZB033 RNA-Seq.
The ten most up- and down-regulated genes in each condition are shown. The degree of differential expression is displayed as β-values (adjusted $\log_2$ fold change).

When looking at the most differentially expressed genes under the added methionine condition, the processes in which the genes are involved in appear to be divergent from each other. For example, the most up-regulated gene in the methionine condition is *megL*, which encodes for the methionine-gamma lyase enzyme, responsible for the catalysis of the cleavage of methionine to form MeSH, ammonia and $\alpha$-ketobutyrate (Tanaka et al., 1985). The ability to produce methanethiol from methionine is widespread in marine heterotrophic bacteria (Drotar et al., 1987). It is possible that *megL* may have been so upregulated in the added methionine condition to catabolise methionine as a sulfur source. Additionally, methionine is potentially toxic to bacteria at high concentrations, so upregulation of *megL* may be a mechanism to remove excess methionine. Among the most downregulated genes in the + methionine condition was the gene encoding the enzyme precorrin-2-oxidase, which is involved in the production of siroheme, which in turn is involved in sulfur assimilation (Claud et al., 2013) . Thus, it is possible that this gene is regulated by the concentration of the sulfur-containing amino acid methionine.

The three most down-regulated genes in the low salinity condition include both *dsyB* and *gsdmt*, indicating that DMSP and glycine betaine are highly important osmolytes in *L. aggregata* LZB033. Interestingly, the most down-regulated gene under the low salinity condition encodes for a S-adenosylmethionine (SAM) synthetase. As both *dsyB* and *gsdmt* are SAM-dependent methyltransferases, this may explain why the SAM-synthetase was so down-regulated under low salinity.

### 3.2.9 Regulation of known DMSP and GB synthesis genes in *L. aggregata*

As some genes of interest to this study did not appear in the 10 most up and down regulated genes in any of the three conditions, a second heatmap was created to compare the differential expression of the DMSP and GB synthesis genes within the RNA-Seq data (**Figure 3.13**).

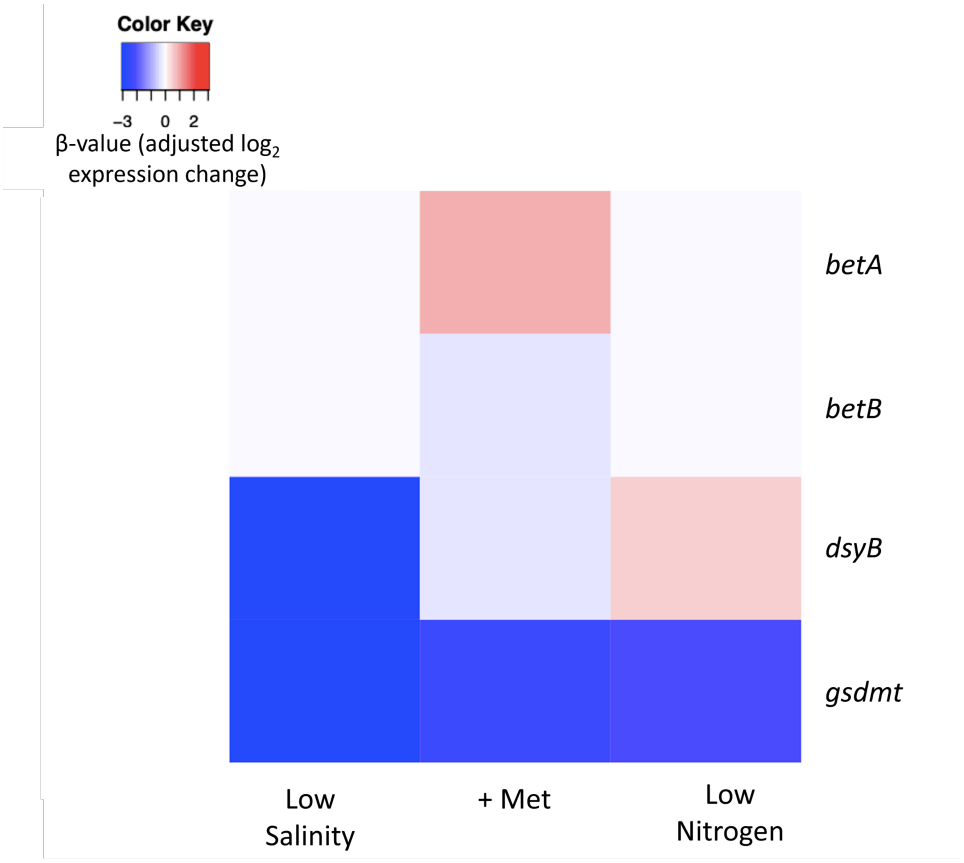



Figure 3.13 Heatmap of differential expression of the *L. aggregata* LZB033 DMSP and GB synthesis genes.
The β-value (adjusted $log_2$ fold change) is shown, this depicts the differential expression of each gene compared to expression in the control condition.

As expected, and in concurrence with the regulation studied in Curson et al. (2017), *dsyB* was significantly down-regulated by low salinity and up-regulated in low nitrogen conditions. The +Met condition caused down-regulation of *dsyB,* which was unexpected as Met is the starting substrate for DMSP synthesis. This down-regulation of *dsyB* in the presence of Met is supported by the earlier conducted β-galactosidase assays (**Figure 3.4**). As such, the Met-mediated down-regulation of *dsyB* has been confirmed via two methods of transcriptional analysis. It is likely that the enhanced DMSP production seen in *L. aggregata* when it was grown in the presence of Met, is likely just a result of enhanced substrate availability. It would be interesting in the future to test if any other DMSP intermediates enhance the transcription of *dsyB* or whether the expression of this gene is only responsive to environmental conditions.

The regulation of the GB synthesis genes was also unexpected. Both genes involved in the choline-GB synthesis pathway, *betA* and *betB* were not significantly differentially expressed compared to the control condition, in either the low salinity or low nitrogen treatments. In contrast, GB production has been shown to be affected by these conditions (**Figure 3.7**). However, *gsdmt* was down-regulated in both the low salinity and low nitrogen condition. This correlated with the decreased production of GB in the low salinity condition but not with the increased GB synthesis in the low N condition (**Figure 3.7**). The reason for this is unknown and discussed in Section 3.3.

Further, the RNA-Seq normalised count data for the *betA* and *betB* genes were much lower across all of the conditions tested, than for the *gsdmt* gene. For example, in the control condition, the mean normalised counts for the *betA* and *betB* genes were 256 and 778, respectively. Whereas, the mean normalised count for the *gsdmt* gene in the control condition was 14,603. Thus, *gsdmt* appears to be transcribed at approximately 30-fold higher levels than the genes involved in the choline-GB pathway. The significant down-regulation of *gsdmt* in the low salinity condition and the higher normalised count data, suggests that GB is primarily produced via the glycine-GB pathway in *L. aggregata* LZB033 and as such, *gsdmt* is the key gene in GB synthesis.

To verify the accuracy of the obtained RNA-Seq dataset, RT-qPCR was also performed on a number of genes of interest. This RT-qPCR was performed by Jinyan Wang (UEA/Ocean University of China). This confirmed the regulation pattern of *dsyB, gsdmt* and the SAM-synthetase gene (*SAM-syn*), as the RT-qPCR data revealed that the direction of regulation (i.e. whether each gene was up- or down-regulated) under each condition was the same as that found by RNA-Seq (**Figure 3.14**).

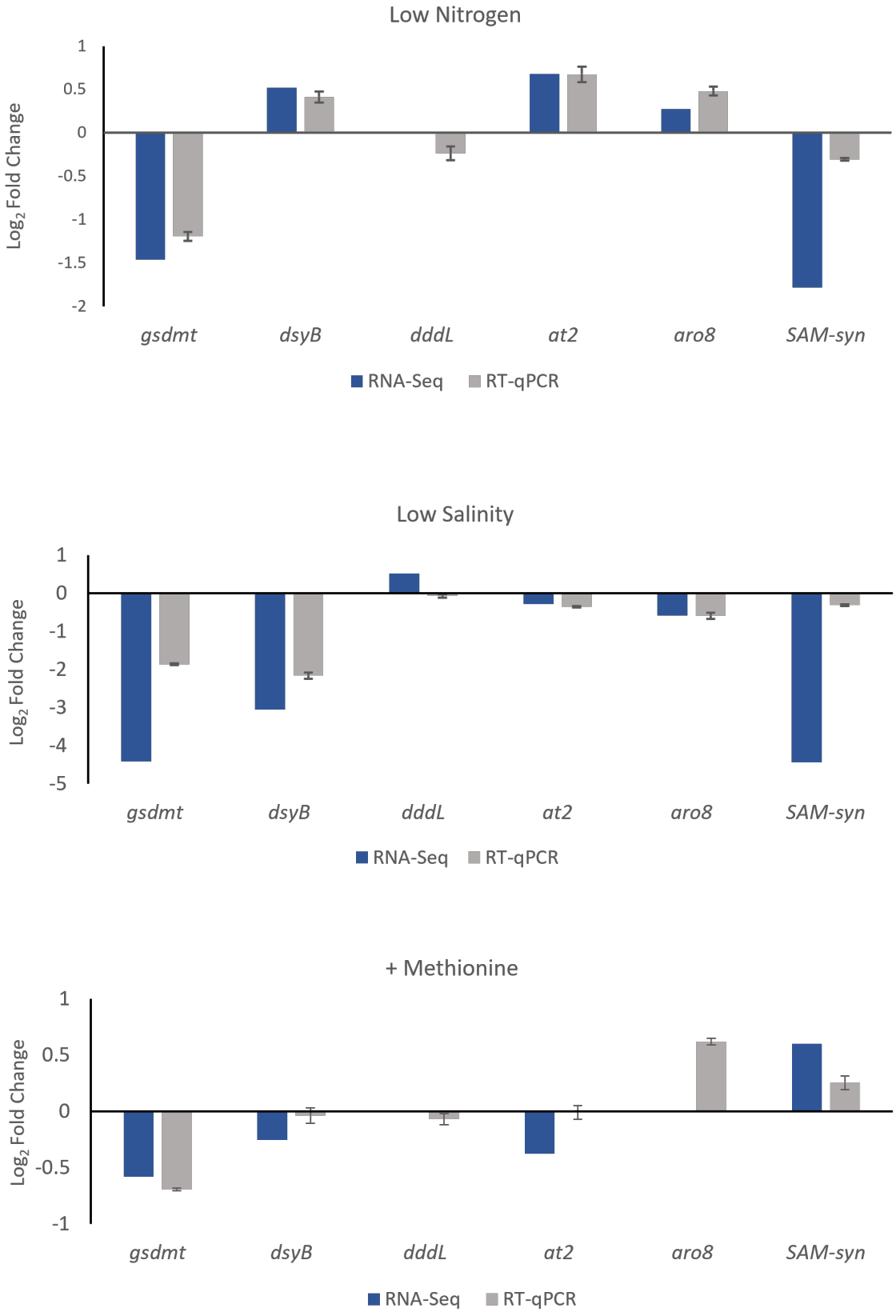



Figure 3.14 Expression changes of key *L. aggregata* LZB033 genes as determined by RT-qPCR and RNA-Seq.

Another of the genes that was studied by RT-qPCR in order to verify its regulation as determined by RNA-Seq, was the DMSP lyase *dddL*. The regulation of this gene in particular was studied to attempt to define the relationship between DMSP production and catabolism under varied conditions. However, the regulation of *dddL* as determined by RT-qPCR differed from the RNA-Seq results. For instance, in the RNA-Seq experiment, the differential expression change of *dddL* was found to not be significantly changed in the low nitrogen and +Met conditions. However, the RT-qPCR determined that *dddL* was down-regulated by these conditions (**Figure 3.14**). Further, the low salinity condition was found to cause significant up-regulation of *dddL* in the RNA-Seq dataset, however the RT-qPCR determined that the gene was slightly down-regulated in this condition (**Figure 3.14**). The regulation of *dddL* in *L. aggregata* LZB033 had previously been studied using RT-qPCR by Curson et al., where they determined that the gene was down-regulated in low salinity and in low nitrogen conditions (Curson et al., 2017). This provides support for the expression changes of *dddL* determined by RT-qPCR in **Figure 3.14.**

Also shown in **Figure 3.14** is the transcriptional regulation of two aminotransferases (denoted as *at2* and *aro8*), which will be discussed further in Chapter 6 of this thesis. However, they were included here to determine the accuracy of the RNA-Seq dataset. As such, the RT-qPCR determined that the direction of expression change was the same as the RNA-seq for ~ 78% of the genes/conditions tested. However, the regulation of the major DMSP and GB synthesis genes in *L. aggregata* (*dsyB* and *gsdmt*, respectively) was consistent when investigated using RNA-Seq and RT-qPCR across every condition.

## 3.3 Summary and Discussion

Within this chapter, it has been confirmed that salinity, nitrogen availability and the presence/absence of Met all have an effect on intracellular DMSP and GB concentrations in *L. aggregata.* Further, an optimised protocol was developed for the extraction of high-quality RNA from cultures of *L. aggregata* grown under conditions known to affect the concentration of both compatible solutes. This allowed analysis of the differential expression of the known DMSP and GB synthesis genes in *L. aggregata* under each of the studied conditions.

The importance of both DMSP and GB as compatible solutes in *L. aggregata* can be inferred from the far greater intracellular concentration of these compounds compared to the majority of the other S- and N- containing zwitterionic metabolites detected by LC-MS. In fact, out of the nine detected metabolites, only DMSP, GB and sarcosine were detected at nmol levels, compared to the fmol concentrations of the remaining compounds. This is suggestive of the importance of sarcosine as a compatible solute in *L. aggregata*, although further study of sarcosine is not included in this thesis. Sarcosine is known to be a compatible solute, however it is also a known intermediate in GB synthesis and degradation (Bashir et al., 2014; Nyyssölä et al., 2000; Wargo et al., 2008). Thus, the involvement of sarcosine in GB metabolism may go some way to explain the high levels of the zwitterionic metabolite found in the *L. aggregata* cultures.

The importance of DMSP and GB in *L. aggregata* is further supported by the significant degree of regulation of both *dsyB* and *gsdmt* in response to salinity, as these genes were the second and third most down-regulated genes in the *L. aggregata* transcriptome in the low salinity condition. The down-regulation of these genes also correlated with decreased intracellular concentrations of both compatible solutes. Therefore, this provides strong evidence to suggest that DMSP and GB are the major compatible solutes in *L. aggregata*, especially in response to salinity stress.

The most down-regulated gene within the *L. aggregata* transcriptome under low salinity conditions was annotated as encoding a SAM-synthetase. As both DsyB and Gsdmt are

SAM-dependent methyltransferases, it is possible that this SAM-synthetase encoding gene is providing the SAM for these enzymes, and thus for the production of the two major compatible solutes in *L. aggregata.* Further work would be required to ratify this theory.

An unexpected finding was that *L. aggregata* may favour the glycine-GB synthesis pathway over the choline-GB pathway, as the *gsdmt* gene was transcribed at ~30 fold higher levels than the *betAB* genes involved in the choline pathway. Further, *gsdmt* expression was regulated by conditions shown to affect intracellular GB concentration - salinity and nitrogen concentration, however the expression of the *betAB* genes were not regulated by either of these conditions. Further work would be required to determine if the glycine pathway is the favoured GB-synthesis pathway in *L. aggregata*, and thus if *gsdmt* is the key GB synthesis gene – this will be addressed in Chapter 4.

One aim of this work was to determine whether an inverse relationship exists in *L. aggregata* between DMSP and GB, in response to nitrogen availability. As expected, and previously reported, DMSP production increased in response to low nitrogen conditions. In other organisms this regulation of DMSP production has been attributed to the preferential production of the sulfur-containing molecule over its structural nitrogen-containing homologue (GB) when nitrogen levels are low (Andreae, 1986). However, in this study, GB production was not regulated as expected by nitrogen concentration and when nitrogen levels were low the intracellular concentration of GB was found to increase. Thus, this does not support the hypothesis that a reciprocal relationship exists between the two compounds.

In many of the studies which report a reciprocal relationship between DMSP and GB, the focus has been on control vs. high nitrogen conditions (Keller et al., 1999). In these studies, the addition of nitrogen to a nitrogen-deplete culture resulted in a short-term increase in GB production (Keller et al., 1999). In addition, the low nitrogen condition used in this thesis was not a nitrogen limiting condition. Therefore, it is possible that as only control vs. low nitrogen conditions have been studied in this thesis, the reciprocal relationship may exist outside of the conditions studied. However, future work would be needed to confirm this.

The regulation of what is so far presumed to be the key GB synthesis gene, *gsdmt* is down-regulated in the low nitrogen condition, which has been experimentally confirmed via both RNA-Seq and RT-qPCR. Thus, this inconsistency between the intracellular GB concentration and expression levels of the synthesis genes suggests that there is possibly a need to repeat the quantification of GB in varied nitrogen conditions. The activity of GSDMT could also be investigated in *L. aggregata* cultures grown under the differing conditions, to further investigate the regulation of this enzyme.

In summary, although future work is needed to confirm the reported nitrogen-regulated relationship between DMSP and GB, both compatible solutes are present at nmol concentrations within *L. aggregata*. In addition, the key genes involved in the synthesis of both molecules are highly regulated by stress factors, especially salinity.

The level of information provided by the RNA-Seq experiment covers much more than what can be described in this thesis. However, another aim of the RNA-Seq experiment was to identify the remaining genes in the DMSP synthesis pathway, and a potential osmolyte regulator. Thus, the RNA-Seq data will be revisited in later chapters.

# Chapter 4: Mutagenesis of the key DMSP and GB synthesis genes in *Labrenzia aggregata*

## 4.1 Introduction

As discussed in Section 1.4, many of the roles of DMSP and GB in organisms remain as 'proposed' roles as they have not been confirmed via mutagenesis of synthesis genes. Instead, the evidence for the roles of DMSP and GB is based on accumulation of the compatible solutes under different environmental conditions.

One of the major aims of this study was to mutate the key DMSP and GB synthesis genes in *L. aggregata* and discover a phenotype for the loss of production of these compatible solutes. As such, the 'proposed' role of DMSP and GB would become a confirmed role.

### 4.1.1 Mutagenesis of DMSP-synthesis genes

Several mutants have been made in DMSP synthesis genes, which result in the complete abolishment of DMSP synthesis (Curson et al., 2017; Williams et al., 2019; Zheng et al., 2020). The first mutagenesis of a DMSP synthesis gene, was the generation of an *L. aggregata* Δ*dsyB* mutant, via suicide plasmid mutagenesis. No impairment in growth phenotype of this mutant was identified when this strain was grown with differing amounts of nitrogen and salt, and under different temperatures. Further, the tolerance to freezing and to oxidative stress of the mutant was also compared to the *L. aggregata* WT, and again a phenotype was not found under these conditions (Curson et al., 2017).

Williams and colleagues studied an isolate that synthesised DMSP through the methylation synthesis pathway, and thus contained the key gene in this pathway, *mmtN.* When *mmtN* was knocked out via suicide plasmid mutagenesis in the DMSP-producing species *Thalassospira profundimis,* DMSP production was abolished. Once again, a phenotypic growth difference was not observed when the *T. profundimis ΔmmtN* was challenged with the aforementioned conditions described for DsyB, above. In both cases *L. aggregata* and *T. profundimis* produce other osmolytes in addition to DMSP, including GB (Williams et al., 2019). In fact, in the *T. profundimis* Δ*mmtN* mutant, more GB was produced than the WT

strain, perhaps suggesting that other osmolytes, in particular GB may compensate for the loss of DMSP (Williams et al., 2019).

The only reported phenotype of a mutant deficient in DMSP production was in relation to barotolerance. Zheng and colleagues found that when *dsyB* was mutated in two DMSP-producing strains, *Pelagibaca bermudensis* and *Marinibacterium sp.* La6, the mutants were less tolerant of deep ocean pressures compared to the WT strains (Zheng et al., 2020). Further, the tolerance to high pressure was restored by either chemical addition of DMSP or through complementation of the mutants with a cloned *dsyB* gene (Zheng et al., 2020). Despite the many proposed roles of DMSP, abolishment of production has so far only been proven to affect barotolerance.

### 4.1.2 Mutagenesis of GB-synthesis genes

The proposed roles of GB in bacteria have also been studied via mutagenesis of synthesis genes and subsequent abolishment of GB production. However, similarly to the DMSP mutagenesis studies described above, this has not always led to discovery of a phenotype (Munro et al., 1989; Ongagna-Yhombi & Boyd, 2013). As such, several studies are discussed below in which GB production was abolished through knock out mutagenesis of GB synthesis genes, but a growth phenotype was either not explored or did not exist.

For example, previous work on the choline-GB pathway in *Sinorhizobium meliloti* found that a knockout mutation in the *betA* gene resulted in the loss of conversion of choline to GB in high salinity. Further, genetic complementation of this mutant restored the betaine aldehyde dehydrogenase activity, and choline was metabolised to GB, as in the WT strain (Østeras et al., 1998; Pocard et al., 1997). However, further phenotyping of the mutant was not discussed.

Further, the GB-producing strain *R. pomeroyi* has been shown to use the choline pathway to produce GB and the *bet* genes involved in this pathway were deleted. The mutagenesis of *betA* resulted in the complete abolishment of *R. pomeroyi* growth on choline as the sole carbon source. The mutagenesis of *betB* resulted in greatly reduced growth on choline

compared to the WT. However, in this study, the effect on GB production and any other growth phenotypes of these mutants was not explored further (I. Lidbury et al., 2015).

In an *E. coli ΔbetA* strain there was an observed decrease in survival of this strain compared to the WT *E. coli* strain. In these survivability tests, both strains had previously been grown in low and high osmolarities, before being inoculated into artificial seawater for seven days, whilst the colony forming units/ml were measured. This decrease of the survivability of the *ΔbetA* was more pronounced when the strains had previously been grown in high osmolarity conditions. As such, this provided a case in which GB synthesis provided an advantage in terms of survival following exposure to high salinity conditions, but a growth phenotype for the ΔbetA strain was not explored in this study (Munro et al., 1989).

Interestingly, in the halophilic strain, *Vibrio parahaemolyticus,* when a GB-synthesis defective *ΔbetA* strain was grown in high salinity conditions (6 % NaCl w/v), there was no significant difference in growth compared to the WT strain. However, *V. parahaemolyticus* is also a producer of the compatible solute, ectoine and when the key ectoine synthesis gene, *ectA* was knocked out, the strain could not grow in the high salinity condition. When the production of both compatible solutes were knocked out in a double mutant strain (*ΔbetA*Δ*ectA*), the addition of exogenous glycine betaine restored the growth to WT levels, indicating that although knocking out GB production didn't have a phenotype, GB is an important osmolyte in this strain (Ongagna-Yhombi & Boyd, 2013).

Another case in which the *betA* gene was knocked out and resulted in the loss of GB production was in the opportunistic pathogen, *Acinetibacter baumannii.* The growth rates of the WT and *ΔbetA* mutant were similar in both standard and high salinity conditions. However, the growth rate of the WT strain in high salinity was increased by the addition of choline. The addition of choline to the *ΔbetA* mutant did not enhance growth under high salinity as it was unable to convert choline to GB. However, the exogenous addition of GB resulted in a similar growth rate increase for both the WT and *ΔbetA* strains. These data again confirm the role of GB as an important osmoprotectant (Breisch et al., 2022).

In all of these cases where GB production was abolished via mutagenesis, the mutants were not demonstrated to have a lower growth rate than the WT in response to high salinity. Further, in all of the aforementioned cases, the mutagenesis that has resulted in abolished GB production has been involved in the choline-GB synthesis pathway (*betA* or *betAB*).

Regarding the glycine-GB synthesis pathway, there have been studies which involve site-directed mutagenesis of genes involved in this pathway, to determine the residues involved in substrate binding (Waditee et al., 2003). However, there are no reported cases of any of the genes involved in the glycine-GB pathway being knocked out via mutagenesis.

However, there is a case of a GB mutant having an established growth phenotype other than salinity protection. The *betA* and *betB* genes were both knocked out in the fish pathogen strain, *Vibrio anguillarum*, resulting in reduced growth when the strain was grown under cold stress. As such, this provided proof that GB acts as an effective cold stress protectant (Ma et al., 2017).

In summary, the only role of GB in bacteria that has been confirmed via mutagenesis of GB-synthesis genes, is the protection against cold stress. The only role that has been confirmed for DMSP via mutagenesis is the protection against high pressures. Therefore, despite the abundance of DMSP and GB, and their roles as potentially important anti-stress compounds and nutrients in diverse environments, their widely-reported role as osmoprotectants is but yet to be confirmed via mutagenesis of synthesis genes, for either compatible solute.

### 4.1.3 Suicide plasmid mutagenesis

Curson and colleagues knocked out *dsyB* in *L. aggregata* via homologous recombination using a derivative of the suicide plasmid pk19mob (Curson et al., 2017; Schäfer et al., 1994). A similar mutagenesis technique was used in this thesis in order to knock out GB synthesis genes in *L. aggregata,* but the suicide plasmid used was pk18mob*sacB* (Schäfer et al., 1994).

The pk18mob*sacB* plasmid is itself a derivative of a pK plasmid, and as such contains a multiple cloning site (MCS) and binding sites for the standard universal M13 primers. However, pk18mob*sacB* also contains *sacB* from *Bacillus subtilis* which encodes the enzyme lavansucrose, the expression of which confers sucrose sensitivity and is mostly lethal in the presence of sucrose (Jäger et al., 1992; Schäfer et al., 1994). pk18mobsacB also contains a gene encoding kanamycin resistance ($Kan^R$) and the broad host-range transfer machinery of another plasmid, RP4 (Schäfer et al., 1994).

These features allow pk18mob*sacB* to be used to force double recombination events, which differ from the single integration mutants previously constructed to study *mmtN* (Williams et al., 2019), for example. Mutagenesis proceeds via the cloning of flanking regions of the gene of interest into the MCS of pk18mob*sacB*. This construct is then conjugated into the target strain. Then, when a single cross-over event has occurred between the flanking regions of the plasmid and the homologous genomic regions, these transconjugants can be selected for by the gain of $Kan^R$. Following this, the transconjugant can then be grown in non-selective conditions to induce a double cross-over event. This double cross-over event results in excision of the plasmid, and thus can be selected for by the loss of $Kan^R$ and loss of sucrose sensitivity. The double-cross over either results in restoration of the WT strain, or with a mutant with the middle portion of the gene excised, thus the resulting colonies need to be screened to determine whether they are WT or mutant strains (Schäfer et al., 1994).

The forcing of a double homologous recombination event results in the formation of a stable mutant, which cannot be reverted by loss of the plasmid (as can be the case in methods which utilise unstable single cross-overs).

### 4.1.4 Does the production of one compatible solute compensate for the loss of the other?

In all cases of reported bacterial mutagenesis that have led to the abolishment of DMSP or GB production, a slower growth in higher salinity compared to the corresponding WT strain has never been observed. This has led to the hypothesis that the loss of one of the compatible solutes may be compensated for by the production of the other, in species that produce both. *L. aggregata* and *T. profmundis* are producers of both DMSP and GB, and showed no reduced growth phenotype in high salinity when DMSP synthesis was abolished (Curson et al., 2017; Williams et al., 2019).

The LC-MS work carried out on *L. aggregata* and discussed in Chapter 3 demonstrated that DMSP and GB were the major compatible solutes present in this strain. Therefore, *L. aggregata* was an ideal organism in which to knock-out both DMSP and GB production simultaneously, and to investigate whether the compatible solutes compensate for one another.

### 4.1.5 Chapter Aims

The overall aim of this chapter is to determine whether the loss of DMSP production in the *L. aggregata ΔdsyB* mutant was being compensated for by the production of GB. As such, the aims of this chapter are:

1. Create an *L. aggregata* mutant in the key GB synthesis gene (*gsdmt*) to abolish GB production.
2. Create an *L. aggregata* double mutant in which both DMSP and GB production are abolished.
3. Determine whether DMSP production is increased in the *L. aggregata* Δ*gsdmt* mutant, and whether GB production is increased in the *L. aggregata ΔdsyB* mutant.
4. Assess whether the generated mutants exhibit phenotypes, in particular for growth in salinity.

## 4.2 Results

### 4.2.1 Creating an *L. aggregata* Δ*gsdmt* and an *L. aggregata* Δ*dsyB*Δ*gsdmt* double mutant

The *gsdmt* gene was implied as the key GB synthesis gene in *L. aggregata* by the RNA-Seq data (**Section 3.2.9**). Therefore, in order to confirm this theory, an *L. aggregata* mutant was created in which *gsdmt* was knocked out. Additionally, in order to determine whether the loss of DMSP production in the *L. aggregata ΔdsyB*- mutant was being compensated for by the production of GB, a mutant was produced in which both DMSP and GB production was abolished in order to determine whether any phenotype was exacerbated.

For both mutant strains, mutagenesis was achieved utilising the suicide vector pk18mob*sacB*. Flanking regions of *gsdmt* were amplified via PCR and cloned into the pk18mob*sacB* plasmid. However, the conjugation of the pk18mob*sacB* plasmid into *L. aggregata* required the use of the kanamycin resistant strain *E. coli* 803, containing the helper plasmid pRK2013. As such, kanamycin could not be relied upon for the single crossover selection in this case. In an attempt to overcome this, nitrofurantoin was added to the crosses at a concentration of 5 μg/ml in order to eliminate the *E. coli* from the crosses, however this proved unsuccessful.

Therefore, there was a need to introduce another antibiotic resistance cassette into the middle of the pk18mob*sacB* flanking regions as this would aid with selection by successfully removing *E. coli* 803 (pRK2013) from the crosses. The gentamycin resistance cassette ($Gent^R$) was subcloned from the plasmid p34S-Gm (Dennis & Zylstra, 1998) and cloned between two *gsdmt* flanking regions within pk18mob*sacB*. This created the construct pBIO23(2) (**Figure 4.1**).


Figure 4.1 Plasmid map of pBIO2 used for *gsdmt* mutagenesis.
The pk18mobsacB backbone is shown in black and grey. The cloned in flanking regions are shown in blue and the cloned gentamycin resistance cassette (GmR) is shown in red.

The flanking regions (~500 bp each) were designed in such a way that they each spanned one of the termini of *gsdmt* and proximal sequence from the *L. aggregata* LZB033 genome. This is of consequence, as when double homologous recombination occurred between the flanking regions and the *L. aggregata* genome, the middle portion (1441 bp) of the *gsdmt* gene was excised and replaced with the gentamycin resistance cassette.

The pk18mob*sacB* backbone contains the gene encoding SacB, which confers sucrose sensitivity. Therefore, selection for colonies that had undergone the required double homologous recombination event was carried out by plating on sucrose and gentamycin. This is because *L. aggregata* Δ*gsdmt*- mutants would not be sensitive to sucrose, due to loss of *sacB* on the pk18mob*sacB* backbone, but retain the Gent$^{R}$ cassette in place of the middle region of the *gsdmt* gene. A schematic of this mutagenesis technique is detailed in **Figure 2.1** in Materials and Methods.

Following the sucrose and gentamycin selection, the resultant colonies were screened via colony PCR. For this checking PCR, primers were designed to the *L. aggregata* genome, just outside of the flanking regions. As can be seen in **Figure 4.2**, when the PCR was performed with WT *L. aggregata* as a template (+ve control), the PCR product was ~2.5 Kb. In PCR reactions in which the transconjugants had successfully undergone double homologous recombination, the expected size of the PCR product was slightly smaller at ~2.2 Kb. Thus, 9 of the 10 colonies (all apart from number 2) screened appeared to have underdone double homologous recombination with pBIO23(2) and thus were likely *L. aggregata* Δ*gsdmt* mutants. This process for *gsdmt* mutagenesis was repeated in the *L. aggregata* Δ*dsyB* mutant, creating a Δ*dsyBΔgsdmt* double mutant. This mutant was confirmed using the same colony PCR process and primers.

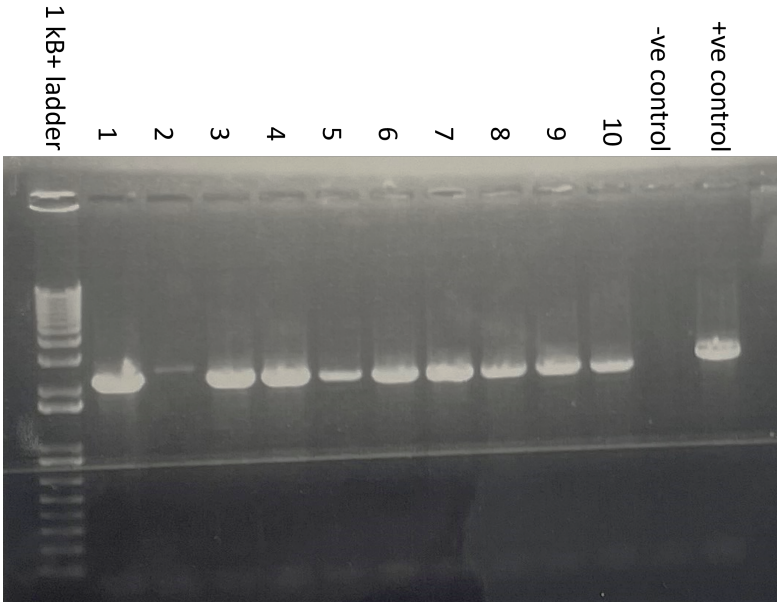



Figure 4.2 Agarose gel image of PCR Screening for *L. aggregata gsdmt*- mutants.
The first lane contains the 1 Kb+ DNA ladder. Lanes labelled 1-10 contain the PCR product of potential *L. aggregata* Δ*dsyB* mutants. The -ve (negative) control used water in place of template. The +ve (positive) control used *L. aggregata* WT as the template in the PCR reaction.

Following the production of an *L. aggregata Δgsdmt* mutant and a *L. aggregata ΔdsyBΔgsdmt* double mutant, the effect of knocking out this gene on GB production was determined via nuclear magnetic resonance (NMR) (**Figure 4.3**). The $^{1}$H-NMR spectrum obtained from a GB standard, indicates two diagnostic singlet peaks with shifts of 3.16 ppm and 3.81 ppm. Both peaks were clearly visible in the cell lysate samples of the *L. aggregata*

WT and the *L. aggregata ΔdsyB,* confirming that both of these strains produced GB, as to be expected (**Figure 4.3**).

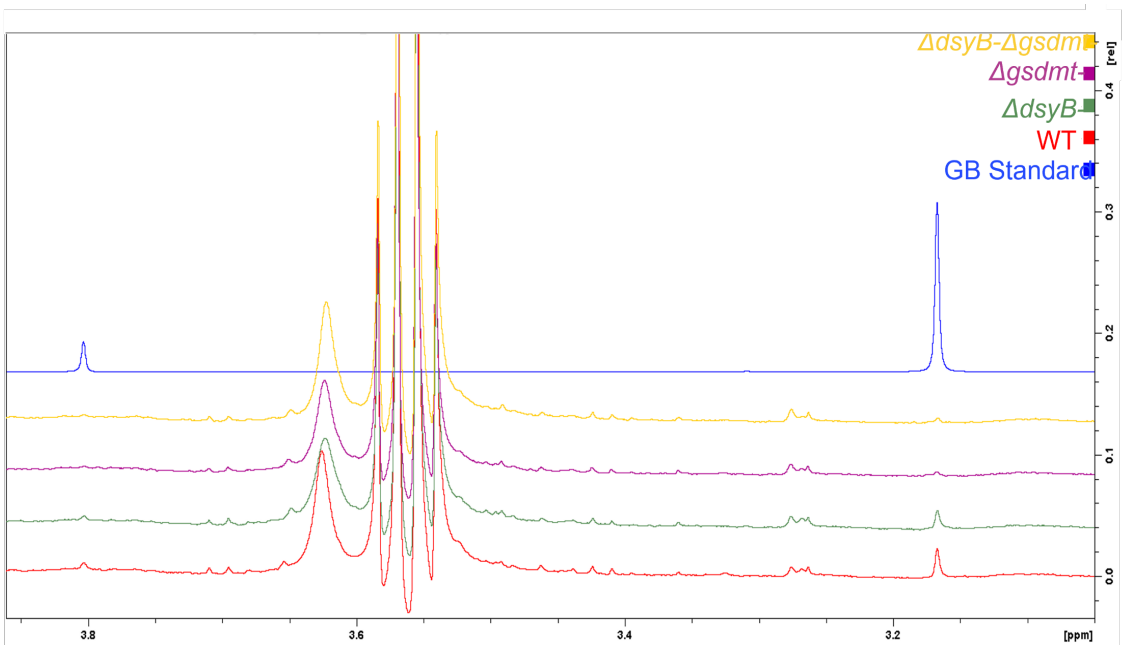



Figure 4.3 NMR Spectra of GB production of *L. aggregata* WT and *L. aggregata* mutant strains. The key denotes the cell lysate in each spectrum. In blue is the NMR spectrum of a 5 mM GB standard.

However, the 3.81 ppm peak was not visible above the background in the spectra of the *L. aggregata* Δ*gsdmt* and the *L. aggregata ΔdsyB*Δ*gsdmt* cell lysate samples, indicating that these samples did not contain GB. Therefore, it can be determined that knocking out *gsdmt* in *L. aggregata* results in abolishment of GB production. As such, this provides further evidence that *gsdmt* is the major GB synthesis gene and that the glycine-GB synthesis pathway is favoured over the choline-GB in *L. aggregata* under the conditions tested.

As mutagenesis of the *gsdmt* gene provided a mutant in which GB synthesis was abolished (or at least greatly reduced to below the $^{1}$H-NMR limit of detection), it was decided that the *betA* and *betB* genes would not be mutated. Production of the *L. aggregata ΔdsyB*, *L. aggregata* Δ*gsdmt* and *L. aggregata* Δ*dsyB*Δ*gsdmt* mutants would be sufficient to determine whether the loss of either of the compatible solutes was being compensated for by the production of the other.

### 4.2.2 Does the loss of one compatible solute result in increased production of the other?

To begin to investigate whether the loss of one compatible solute is compensated for by the increased production of the other, the intracellular concentration of the remaining compatible solute was measured in each of the single mutants.

The GB levels of cell lysate were determined by $^{1}$H-NMR using a calibration curve constructed by Vanja Kortaras (UEA), using the peak area of the 3.16 ppm diagnostic peak of known concentrations of GB standards. As such, the GB production levels could be compared between the *L. aggregata* WT and *L. aggregata ΔdsyB* mutant and standardised to protein content. The GB levels between WT and *ΔdsyB* strains were found to not be statistically different from each other (t-test, p=0.648) (**Figure 4.4**). As such, the *L. aggregata ΔdsyB-* mutant does not appear to compensate for the loss of DMSP production via increased GB production. This analysis could not be carried out for the other mutant stains used in this study, as GB was below the $^{1}$H-NMR limit of detection in the *L. aggregata Δgsdmt-* or *L. aggregata ΔdsyB- Δgsdmt-* strains (data not shown).

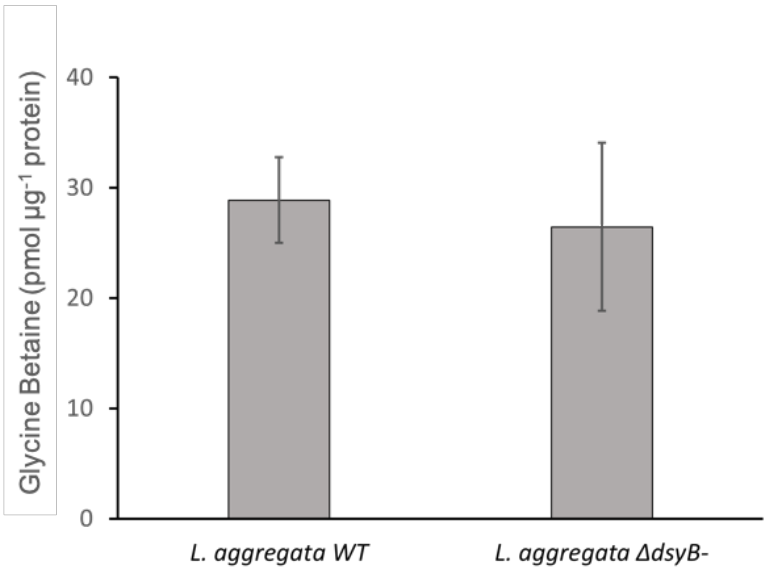



Figure 4.4 Glycine betaine production of *L. aggregata* WT and *L. aggregata ΔdsyB-* mutant, quantified via NMR.
The glycine betaine production of cultures grown at 35 PSU were measured and normalised to protein content. N=3 biological replicates and error bars represent the standard error of the mean.

The DMSP production of *L. aggregata* WT and *L. aggregata Δgsdmt* mutant were also compared under standard salinity conditions (35 PSU) and low salinity (5 PSU) using gas chromatography with normalisation to protein content (**Figure 4.5**). There was no statistical difference between the DMSP production of the strains at 35 PSU (t-test, $p > 0.05$). However, the *L. aggregata Δgsdmt* mutant did produce significantly more DMSP than the WT at low salinity (t-test, $p = 0.01$). This suggests, that at lower salinities, the loss of GB production may be compensated for by an increase in DMSP production.

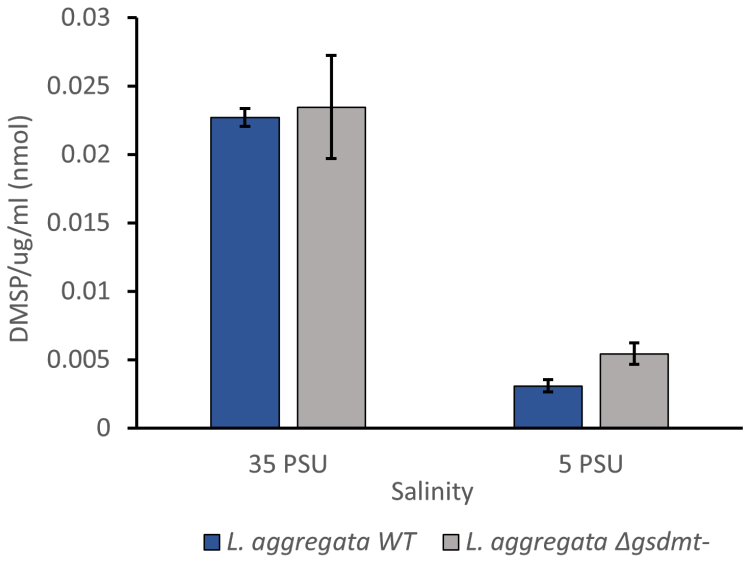



Figure 4.5 DMSP Production of *L. aggregata* WT and *L. aggregata Δgsdmt-*.
The DMSP concentration of the cultures was determined by gas chromatography and normalised to protein content. N=3 biological replicates, and error bars represent the standard error of the mean.

### 4.2.3 Genetic complementation of the Δ*dsyB* and Δ*dsyB*Δ*gsdmt* mutants

Following the generation of the mutants, genetic complementation with *dsyB* from *L. aggregata* was achieved through conjugation of the pBIO2266 plasmid (Curson et al., 2017) into the *L. aggregata* Δ*dsyB* and *L. aggregata ΔdsyBgsdmt* mutants. The pBIO2266 plasmid is comprised of the *dsyB* gene from *L. aggregata* IAM12614 in the wide host-range plasmid vector, pRK415 (Keen et al., 1988). As pRK415 encodes for tetracycline resistance ($Tet^R$), allowing the complemented mutants to be selected for via gain of $Tet^R$. The selected colonies were then confirmed as complemented mutants as ability to produce DMSP was restored (**Figure 4.6**).

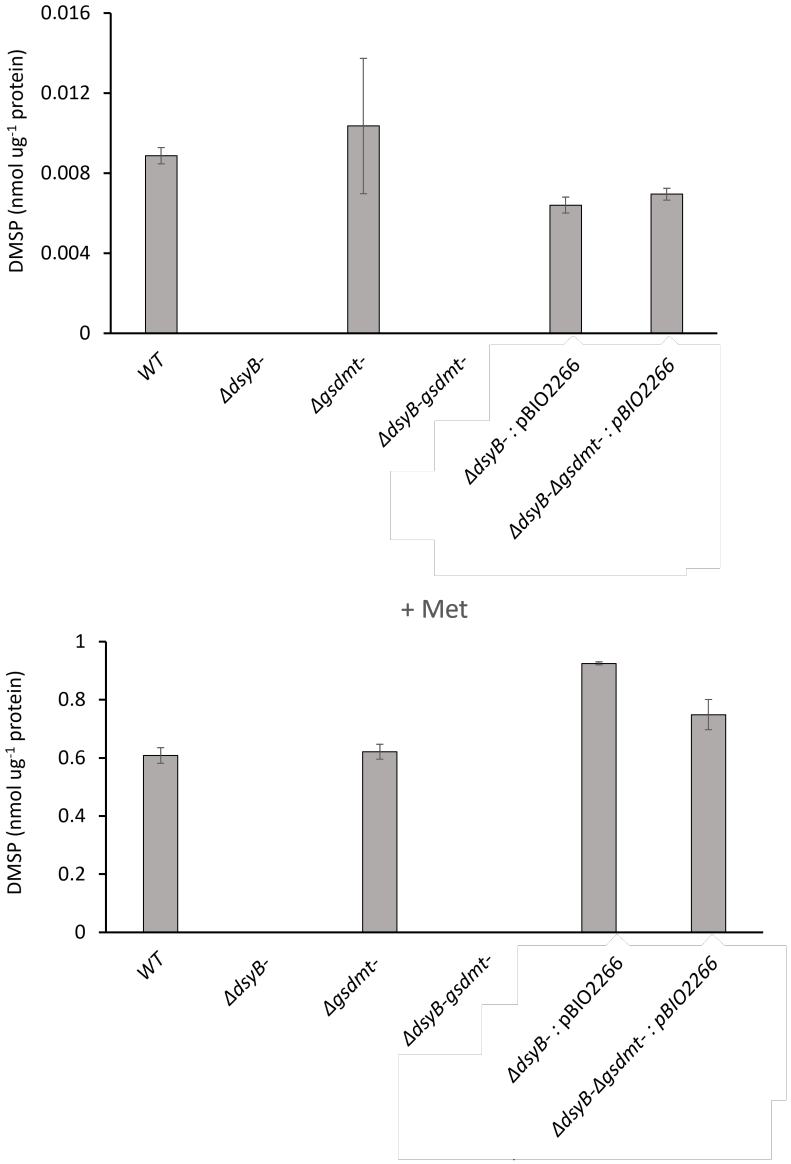



Figure 4.6 DMSP production of genetically complemented *L. aggregata ΔdsyB* and *L. aggregata ΔdsyBΔgsdmt* mutants.
DMSP was measured via gas chromatography and normalised to protein content. N=3 biological replicates and error bars represent the standard deviation of the mean.

The complemented Δ*dsyB* and Δ*dsyB*Δ*gsdmt* both gained the ability to produce DMSP, but at slightly lower levels than the WT when grown in MBM at 35 PSU (**Figure 4.6**). Further, when grown in the presence of 0.5 mM methionine, both complemented strains produced slightly more DMSP than the WT (**Figure 4.6**).

Several attempts were made to clone *gsdmt* into pRK415 to allow for genetic complementation of the Δ*gsdmt* and Δ*dsyB*Δ*gsdmt* mutants. However, this was not completed successfully so complementation of these mutants was achieved chemically in the later experiments detailed in this chapter.

### 4.2.4 Determining a growth phenotype for the *L. aggregata Δgsdmt* and *ΔdsyBΔgsdmt* mutants.

Curson and colleagues did not identify a growth phenotype for the *L. aggregata ΔdsyB* mutant (Curson et al., 2017). In order to determine a growth phenotype for the *L. aggregata Δgsdmt* and Δ*dsyB*Δ*gsdmt* strains and to further investigate whether the loss of either compatible solute was being compensated for by the production of the other compatible solute, the growth rate of each mutant was measured and compared to the WT. From herein, growth rate (μ) refers to the mean (n=3) increase in $OD_{600}$ per hour during logarithmic growth. As both *dsyB* and *gsdmt* were shown to be highly regulated by salinity (see section 3.2.9), growth curves of each of the mutant strains were performed across varying salinities (**Figure 4.7**).

In standard salinity conditions (35 PSU), the growth rates of *L. aggregata* WT (μ= 0.041) and *L. aggregata ΔdsyB* (μ= 0.046) were not significantly different to each other (t-test, p=0.16) (**Figure 4.7**). The growth rates of the Δ*gsdmt* (μ= 0.03) and Δ*dsyB*Δ*gsdmt* (μ=0.03) strains were highly similar to each other (t-test, p=0.15), but significantly slower than that of the WT (t-test, p=0.004). Thus, this suggests that loss of the *gsdmt* gene, and therefore GB sythesis, results in slower growth in standard salinity conditions.

In higher salinity conditions (50 PSU), the growth rates of the *L. aggregata* WT (μ = 0.22) and *L. aggregata ΔdsyB* (μ = 0.24) strains were not significantly different to each other (t-test, p=0.072). (**Figure 4.7**). This supports the finding that the *ΔdsyB* mutant does not exhibit a reduced growth phenotype in response to salinity (Curson et al., 2017). When grown at 50 PSU, the Δ*gsdmt* mutant (μ = 0.16) had a significantly reduced growth rate compared to the WT strain (t-test, p = 0.01). This disparity between the Δ*gsdmt* and WT growth rates was increased under this high salinity condition, as compared to the 35 PSU results, which further supports the role of GB in osmoprotection in *L. aggregata.*

Interestingly, the Δ*dsyBΔgsdmt* double mutant had a slightly reduced growth rate compared to the Δ*gsdmt* single mutant when grown in higher salinity, however this difference was not statistically significant (p=0.08) (**Figure 4.7**). This suggests that GB is more important than DMSP in terms of osmoprotection in *L. aggregata.*

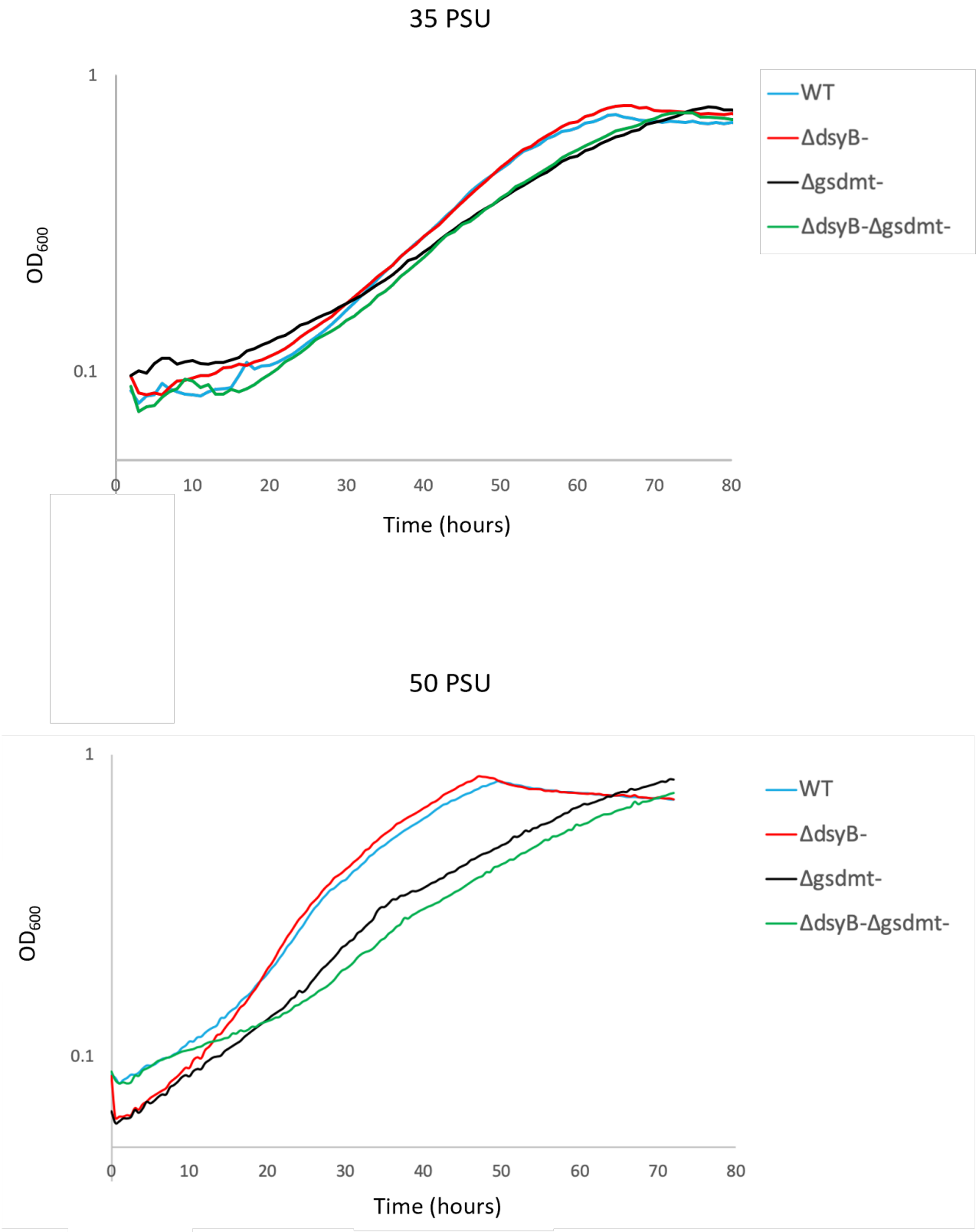



Figure 4.7 Growth curves of *L. aggregata* WT and mutant strains in standard and high salinity. The growth curves were constructed via plate reader automatic readings, with $OD_{600}$ measurements taken at 30-minute intervals. The Y-axis is presented as a logarithmic scale. N=3 biological replicates.

Further evidence for the reduced osmoprotection phenotypes of the *Δgsdmt* and the *ΔdsyBΔgsdmt* strains is provided by the rescue of the observed decreased growth rates through chemical complementation (**Figure 4.8**). These growth curves were measured under high salinity (50 PSU) conditions, and neither were significantly different to the growth rate of the WT strain grown in the presence of DMSP (t-tests p=0.37 and p=0.09, respectively). Thus, confirming that the loss of GB production in these strains was the cause of decreased growth rate in higher salinities.

Although the *ΔdsyBΔgsdmt* strain had a statistically similar growth rate to the *Δgsdmt* strain in the aforementioned experiments, some evidence was provided in support of DMSP's role as an osmoprotectant in *L. aggregata.* This is because, complementation of the *ΔdsyBΔgsdmt* double mutant with *dsyB* resulted in partial restoration of the growth rate (**Figure 4.8**). In this experiment, the growth rate of *ΔdsyBΔgsdmt:*pBIO2266 was significantly higher than the growth rate of the uncomplemented *ΔdsyBΔgsdmt* strain (t-test, p=0.004). Thus, this provides some preliminary evidence for an osmoprotective role of *dsyB* (and therefore DMSP) in *L. aggregata* under high salinity, when GB is not being produced.

Future experiments would be required to determine whether the reduced growth phenotype observed in the *Δgsdmt* strain could be rescued via genetic complementation. Additionally, this study did not determine whether chemical complementation with either just DMSP or GB would restore the decreased growth phenotype exhibited by the *ΔdsyBΔgsdmt* strain.

Preliminary experiments were carried out in order to determine whether the *Δgsdmt* and *ΔdsyBΔgsdmt* mutant strains had phenotypes in other conditions such as recovery after freezing. In preliminary freezing assays (data not shown) the CFU/mL of the strains were calculated before and after being frozen for 5 days. The CFU/mL of the WT and single mutant strains increased after freezing (perhaps due to some time at room temperature post-thawing). However, CFU/ml of the *L. aggregata ΔdsyBΔgsdmt* strain decreased by ~ 50 % after freezing. This suggests a potential freezing phenotype for the *L. aggregata ΔdsyBΔgsdmt* strain. However, the results of these preliminary studies were not conclusive

and as such, these experiments should be optimised and repeated before conclusions can be drawn regarding additional phenotypes of these mutants.

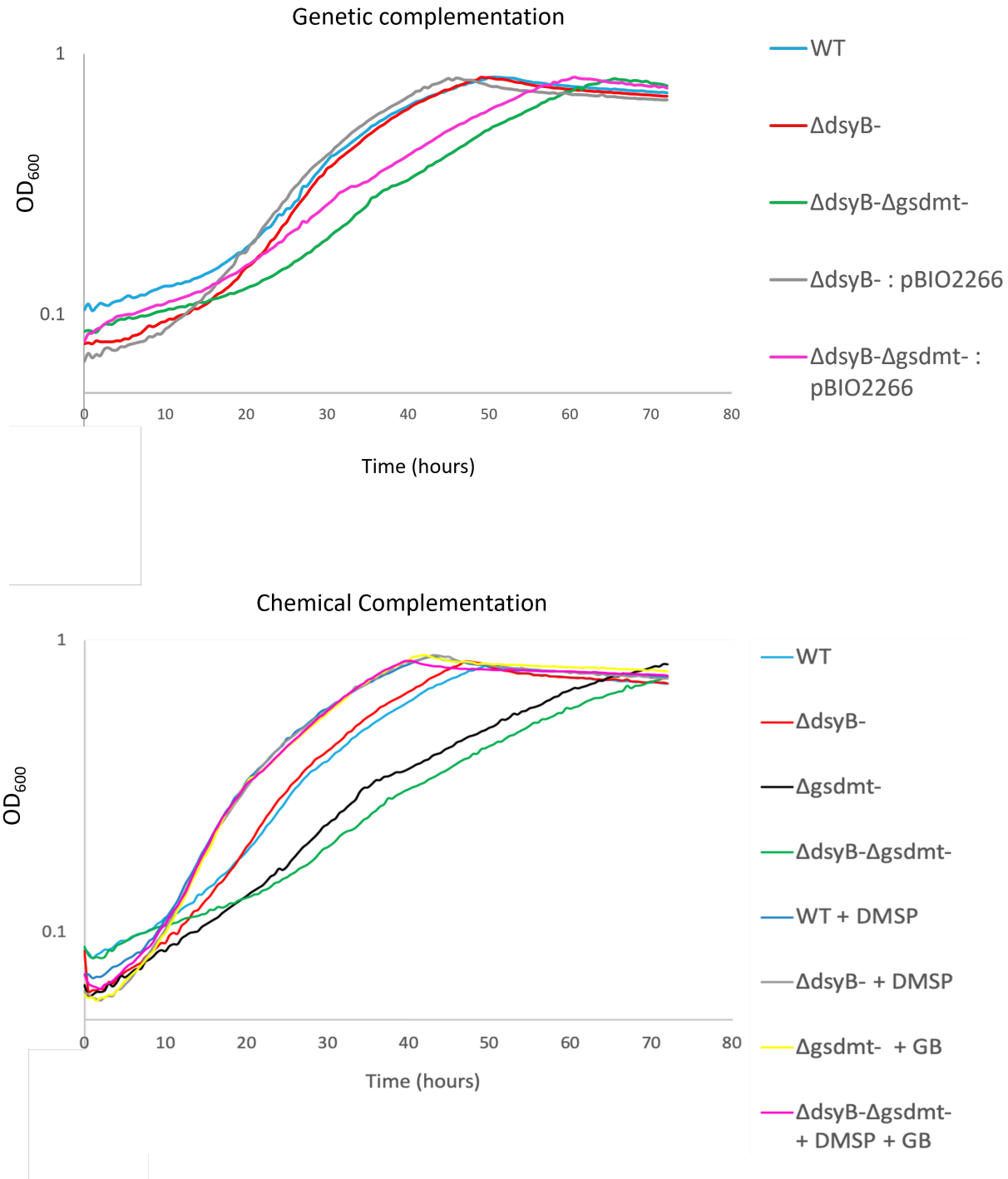



Figure 4.8 *L. aggregata* WT and mutant growth curves with genetic and chemical complementation. Growth curves were constructed at 50 PSU and $OD_{600}$ readings were taken at 30-minute intervals using a plate reader. Where indicated, DMSP and GB were added at 100 µM. Y-axis is presented as a logarithmic scale. N=3 biological replicates.

### 4.2.5 Prevalence of *gsdmt* homologues in marine bacteria.

The previously discussed growth experiments highlighted the importance of *gsdmt* in GB production (and the importance of GB in osmoprotection in *L. aggregata*). As *gsdmt* was found to be responsible for GB production in *L. aggregata,* it is possible that it could be used as a marker gene for GB producing organisms. Homologues of the *L. aggregata* Gsdmt sequence are present in 99 bacterial strains (percentage identity > 65%, BlastP analysis), suggesting that this method of GB production is relatively widespread. However, this cannot be confirmed until the homologues are experimentally ratified as functional. Further, preliminary searching against the *Tara* Oceans Metagenome and Metatranscriptome datasets (Vernette et al., 2022; Villar et al., 2018) revealed more hits to *L. aggregata gsdmt* homologues than to *L. aggregata dsyB* homologues. As *dsyB* is thought to be present in 0.5% of bacteria in marine metagenomes (Curson et al., 2017), this potentially suggests a high abundance of *gsdmt* in the environment.

## 4.3 Summary and Discussion

Within this chapter, *gsdmt* was confirmed as the major GB synthesis gene in *L. aggregata* as the mutation of this gene resulted in abolishment of GB production. Mutagenesis of *gsdmt* also resulted in a slower rate of growth compared to the WT strain in higher salinities, which could be restored by the exogenous addition of GB. Thus, it can be concluded that GB is an osmolyte in *L. aggregata* and facilitates improved growth in high salinity. According to a literature search, this is the first time that bacterial mutagenesis of a GB synthesis gene has directly led to a decreased growth rate in high salinity compared to wildtype strains. Therefore, this is the first direct evidence that GB functions as an osmoprotectant in bacteria.

Due to the abolishment of GB production in the *L. aggregata Δgsdmt* strain, the *betA* and *betB* genes of the choline-GB synthesis pathway were not targeted for mutagenesis. In *L. aggregata* LZB033, the *betA* and *betB* genes are located in the same operon as the known regulator of these genes, *betI.* BetI is a transcriptional repressor which is released from the DNA in the presence of choline (Scholz et al., 2016). As such, it is possible that the *betAB* genes would have had a more evident role in GB synthesis if choline had been added to the growth medium.

When the key DMSP and GB synthesis genes were knocked out simultaneously, a mutant was created in which both DMSP and GB production were abolished. However, there was not a signigifant difference between the growth rates of the *L. aggregata ΔdsyBΔgsdmt* and *L. aggregata Δgsdmt* strains in high salinity conditions. Complementation of the double mutant with *dsyB*, resulted in partial restoration (perhaps due to the different levels of *dsyB* expression in pBIO2266 compared to the genomic *dsyB*) of the growth rate to levels similar to that of the *Δgsdmt* mutant. As such, this provides some partial evidence of the role of *dsyB* (and DMSP production) in osmoprotection in *L. aggregata.* This may also suggests that the *L. aggregata* Δ*dsyB* mutant does not have a reduced growth rate in high salinity conditions because it is compensated for by the presence of GB.

As mutagenesis of Δ*gsdmt* resulted in a reduced growth rate in higher salinities, whereas mutagenesis of Δ*dsyB* did not, this perhaps suggests that GB is a more important osmolyte than DMSP in *L. aggrgegata.* It is possible that DMSP acts as a minor osmoprotectant in *L. aggregata* however other potential roles of the compatible solute cannot be disregarded. The production of GB may be the reason why other DMSP mutants do not display a reduced growth rate in high salinity. For example, the *T. profundimis* Δ*mmtN* mutant has completely abolished DMSP synthesis but no reduced growth rate in high salinity, which may be due to its increased production of GB (Williams et al., 2019).

Despite indications of GB compensating for the loss of DMSP production in the *L. aggregata ΔdsyB* strain, GB production was not increased in this strain compared to the WT. This suggests that the compensation provided by GB is not due to an increase in production of this compatible solute. The superior osmoprotection provided by GB production is therefore perhaps due to the differing structural characteristics of the osmolytes, rather than being concentration dependent. However, the GB production of the *L. aggregata ΔdsyB* strain was not measured and compared to the WT under high salinity conditions, which would be needed to further investigate this hypothesis.

More research is required to determine whether the *L. aggregata* Δ*gsdmt* and the *L. aggregata ΔdsyBΔgsdmt* strains have reduced growth or survivability compared to the WT in other conditions. Reduced survivability after freezing were indicated for the *L. aggregata ΔdsyBΔgsdmt* double mutant, however these assays would need to be repeated to confirm this phenotype. However, this is an initial indication of the role of DMSP and GB in protection against freezing why may be physiologically relevant for bacteria in polar regions.

Additional experiments could be carried out to determine other phenotypes for the generated mutants. For example, the growth rates of the mutants could be compared against the WT in response to different temperatures, nitrogen levels and under oxidative stress to determine whether phenotypes existed in any of the tested conditions. This would indicate whether DMSP or GB conferred any other protective qualities. Competition assays could also be carried out in natural seawater to determine the survivability of each of the

mutant strains compared to the WT. This would indicate whether DMSP and GB confer a fitness advantage that enhances *L. aggregata* survival in seawater.

The remaining osmolytes detected by LC-MS (**Figure 3.8**) have not been investigated further. In the cases where the synthesis genes are known (e.g. the *ect* genes involved in ectoine biosynthesis (Louis & Galinski, 1997)), homologues for these genes could be searched for in the *L. aggregata* genome. This would open up the possibility of mutagenesis of the identified compatible solute genes in *L. aggregata,* which could be knocked out singly, and in combination with *dsyB* and *gsdmt*. This would allow investigation of the roles of these compatible solutes in *L. aggregata.*

In summary, this is the first time that the osmoprotective role of GB has been confirmed in marine bacteria. This was achieved via mutagenesis of the key compatible solute synthesis genes. The results discussed within this chapter also provide some initial evidence for the role of DMSP as an osmoprotectant in *L. aggregata.*

# Chapter 5: Identification of candidate DMSP synthesis genes and a potential global osmolyte regulator

## 5.1 Introduction

### 5.1.1 Missing genes in the DMSP transamination synthesis pathway

The transamination synthesis pathway has previously been described in Section 1.5.1 (**Figure 5.1**). The intermediates in the pathway were identified by Gage and colleagues (Gage et al., 1997). Following this, the enzyme activities responsible for catalysing the first three steps of the pathway were identified in *E. intestinalis* cell-free extracts (**Figure 5.1**) (Summers et al., 1998).

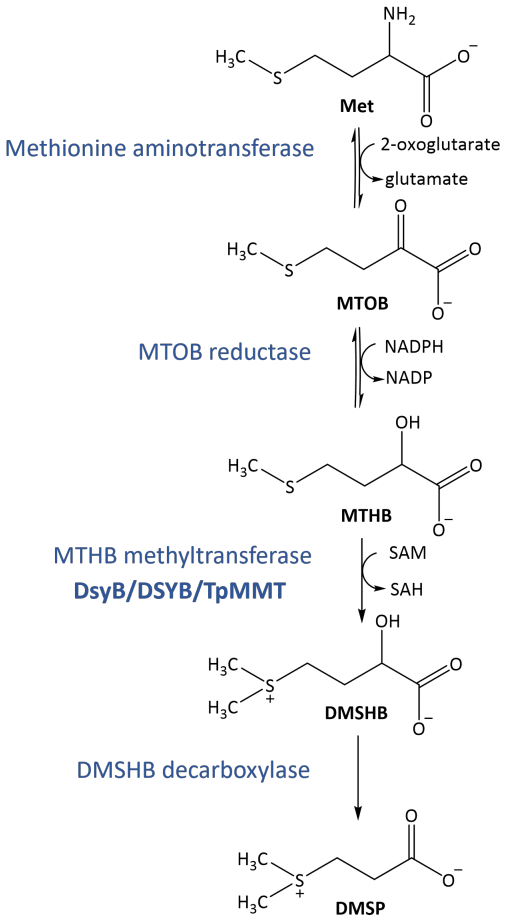



Figure 5.1 The activities of the enzymes involved in the DMSP transamination synthesis pathway. The enzyme activities involved in each step are shown in blue, with the characterised enzymes (DsyB, DSYB and TpMMT) displayed in bold. Met – methionine, MTOB – 4-methylthio-2-oxobutyrate, MTHB – 4-methylthio-2-hydroxybutyrate, DMSHB – dimethylsulfonio-2-hydroxy-butyrate, DMSP – dimethylsulfoniopropionate.

The only genes in the bacterial transamination synthesis pathway to be identified so far are *dsyB, DSYB* and *TpMMT* (Curson et al., 2017; Curson et al., 2018; Kageyama et al., 2018). These genes encodes a SAM-dependent methyltransferase which catalyses the committing and rate-limiting step in the pathway, the methylation of MTHB to DMSHB (Curson et al., 2017). *dsyB* has since been used as marker gene for DMSP synthesis in bacteria and has led to the identification of novel bacterial DMSP producers. This is also the case for the eukaryotic homologue, *DSYB*, which is present in many phytoplankton and corals (Curson et al., 2018). TpMMT was identified in the diatom *Thalassiosira pseudonana* and a homologue has so far been identified in one other diatom species. Thus, currently the identity is known of just one of the four transamination pathway enzymes. Therefore, identification of the missing genes involved in the transamination synthesis pathway would allow biochemical characterisation of the pathway and may also lead to the identification of other DMSP-producing organisms.

Often bacterial genes involved in the production of secondary metabolites such as DMSP- are clustered in gene operons, as is the case for antibiotic production in actinobacteria, for example. However, analysis of the *L. aggregata* LZB033 genome revealed that there are no genes in close proximity to *dsyB* which are predicted to encode enzymes with either methionine aminotransferase, MTOB reductase, or DMSHB decarboxylase activity (**Figure 5.2**).Curson and colleagues also studied the gene synteny of *dsyB* in selected DMSP-producing bacteria, including several members of *Rhodobacterales* (Curson et al., 2017). However, searching of the predicted gene products nearby to *dsyB* in these species also did not reveal any candidate DMSP synthesis genes.

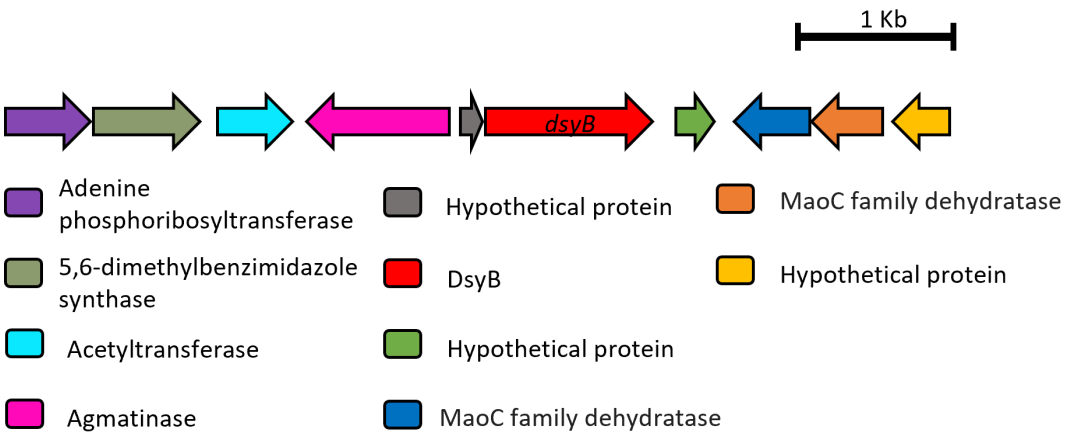



Figure 5.2 Genomic location of *dsyB* in *L. aggregata* LZB033.
The key denotes the predicted protein products of each of the genes, as annotated by RASTtk or BlastP.

In this chapter, one of the main aims was to identify candidate genes responsible for encoding an enzyme with Met aminotransferase activity, and thus a candidate gene for the first step of the transamination synthesis pathway.

### 5.1.2 The methionine aminotransferase reaction

Transamination reactions of amino acids involve the transfer of an amino group to an α-ketoacid, with the involvement of a glutamate/ α-ketoglutarate pair and are always pyridoxal-5'-phosphate (PLP) dependent (Bhagavan & Ha, 2015). In the case of the transamination pathway, the amino group is transferred from methionine to the α-ketoacid (Gage et al., 1997). However, the glutamate/α-ketoglutarate pair remained unknown until Summers and colleagues determined that the methionine aminotransferase reaction occurring in *Enteromorpha intestinalis* extracts had a strong preference for 2-oxoglutarate as the amino acceptor (Summers et al., 1998). Summer and colleagues also carried out enzyme assays in the reverse direction (MTOB to Met) and found that glutamate was the best amino donor in this case. This provided more evidence for 2-oxoglutarate dependency of the methionine aminotransferase involved in the transamination DMSP synthesis pathway. As such, the methionine aminotransferase catalyses the reaction depicted in **Figure 5.3.**

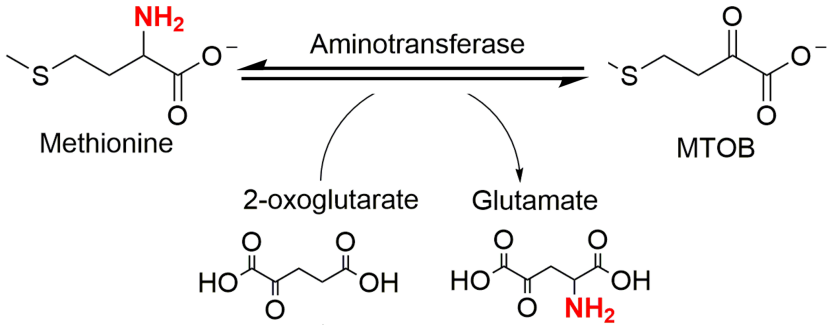



Figure 5.3 Mechanism of the methionine aminotransferase in the transamination DMSP synthesis pathway.
MTOB - 4-methylthio-2-oxobutyrate. The amino group that is transferred in this reaction is depicted in red.

The gene encoding the methionine aminotransferase in the DMSP transamination synthesis pathway has not yet been discovered in any organism. However, aminotransferase proteins have been biochemically characterised from the ubiquitous Met salvage pathway, which catalyses the conversion of MTOB to Met (i.e. running the proposed DMSP synthesis step 'in reverse' - **Figure 5.3**). The Met salvage pathway is responsible for the recycling of sulfur-containing metabolites into Met and is found in all types of organisms. Sulfur containing metabolites, such as 5'-methylthioadenosine (MTA) are produced in the cell as products of primary metabolism, including glycolysis and the tricarboxylic acid (TCA) cycle. These metabolites are then converted back to methionine via numerous reactions involving different intermediates and enzymes (Albers, 2009). The entire pathway has been characterized in many organisms, including *B. subtilis* and *Klebsiella pneumoniae.* In relation to this study, it is the final reaction of the Met salvage pathway that is of interest.

In *B. subtilis*, the products of several genes have been determined to be able to catalyse the final reaction in the Met salvage pathway. One of these genes is the putative aspartate aminotransferase *ykrV,* which was cloned and the enzyme it encoded was shown to be able to convert MTOB to Met (Berger et al., 2003). The other enzyme in *B. subtilis* that was found to have this activity was a putative branched chain aminotransferase, YbgE (Berger et al., 2003).

In *K. pneumoniae,* the MTOB transamination reaction was found to be carried out by the tyrosine aminotransferase (TyrAT) encoded by *tyrB* (Heilbronn et al., 1999). In enzyme assays utilising this enzyme, the highest rate of reaction occurred when aromatic amino acids (phenylalanine, tryptophan and tyrosine) and glutamate were used as the amino donors, and thus these amino acids were assumed to be the preferred amino donors for this reaction (Heilbronn et al., 1999).

As transamination reactions, by their nature, are reversible, it can be assumed that the aminotransferases with proven MTOB to Met activity, also have the ability to convert Met to MTOB. Therefore, it was logical to search the *L. aggregata* LZB033 genome for homologues of *ykrV,* YbgE and *tyrB.* However, highly similar homologues to these genes were not present in the *L. aggregata* genome (all hits were below 39 % identity) and as such, candidate methionine aminotransferases involved in DMSP synthesis were identified using a different approach.

### 5.1.3 Regulation of DMSP and GB Synthesis Genes

Another aim of this chapter was to identify potential transcriptional regulators of DMSP and GB synthesis genes. As discussed in Section 3.1.2, the *betA* and *betB* genes are in the same locus as the *betI* transcriptional regulator in *L. aggregata*. In *Acinetobacter baylyi,* it has been shown that BetI is a choline-dependent repressor of the *betIBA* operon, which is released from its regulatory position on DNA in response to choline (Scholz et al., 2016). This results in the choline-GB synthesis pathway being induced when choline is present in the environment (Scholz et al., 2016). However, in this thesis, the key GB synthesis gene in *L. aggregata* was found to be *gsdmt,* of the glycine-GB synthesis pathway. As shown in this study, *gsdmt* expression is regulated by salinity and nitrogen availability but a transcriptional regulator for the *gsdmt* gene has not yet been identified.

Further, a transcriptional regulator of the key DMSP synthesis gene, *dsyB* has not been identified in any of the organisms that have the gene, and none of the genes adjacent to *dsyB* in *L. aggregata* are proposed transcriptional regulators (**Figure 5.2**). As found in this study and previous work by Curson and colleagues, *dsyB* expression is regulated by salinity,

nitrogen availability, temperature, oxidative stress and growth phase (Curson et al., 2017). This regulation corresponds to the conditions known to regulate DMSP production. It should be noted that there may be a combination of many regulators which control *dsyB* expression in response to these diverse conditions.

### 5.1.4 A global regulator of bacterial osmotic stress response

A multiple antibiotic resistance (MarR)-type regulator, CosR has been identified as a global regulator of osmotic stress response and is widespread throughout *Alpha-, Beta-* and *Gamma-proteobacteria* (Gregory et al., 2020). The *cosR* gene was first identified in *Vibrio cholerae*, when a microarray approach was used to identify transcriptional regulators that were differentially expressed in response to salinity (Shikuma et al., 2013). The expression of *cosR* was found to correlate with salinity and following construction of a *V. cholerae* Δ*cosR* mutant, the expression of genes involved in compatible solute synthesis and transport were compared between the mutant and the wildtype strain. In higher salinities (0.2 M and 0.5 M NaCl), the expression of the compatible solute, ectoine biosynthesis genes (*ectA, ectB* and *ectC)* in the Δ*cosR* strain were upregulated compared to the wildtype (Shikuma et al., 2013). This suggested that CosR is a repressor of compatible solute synthesis and transport genes. As both *cosR* and the compatible solute synthesis genes in *V. cholerae* were found to be upregulated in high salinity, it was curious that CosR was a repressor of the synthesis genes. To explain this, a rapid negative feedback loop was suggested, in which CosR regulates its own expression (Shikuma et al., 2013).

Another study linking *cosR* to compatible solute regulation took place in the halophilic marine bacterium, *Vibrio parahaemolyticus* which is known to synthesise the compatible solutes ectoine, utilising the *ectABC* operon and GB, utilising the *betIBA* operon (Gregory et al., 2019; Naughton et al., 2009; Ongagna-Yhombi & Boyd, 2013). Mutagenesis of *cosR* in *V. parahaemolyticus* and subsequent transcriptional analysis revealed that CosR was also a repressor of the ectoine and GB biosynthesis genes in *V. parahaemolyticus* (Gregory et al., 2020; Gregory et al., 2019). Gregory and colleagues also identified that the ectoine and glycine betaine synthesis operons were both directly regulated by the quorum sensing regulators AphA and OpaR. Both AphaA and OpaR were also shown to activate

transcription of CosR, which in turn represses the ectoine and GB synthesis genes (Gregory et al., 2020; Gregory et al., 2019). In this instance at least, the transcriptional regulation of compatible solutes is complex, and this was suggested to be due to the energy cost associated with producing compatible solutes necessitating multiple levels of regulation in order to tightly control the production (Gregory et al., 2019).

CosR homologues (73% to 98% identity) are present in over 50 phylogenetically divergent *Vibrio* species. *cosR* is often located nearby to genes encoding ectoine and GB synthesis enzymes and transporters (Gregory et al., 2020). Therefore, it was suggested that the compatible solute regulatory function of CosR is conserved in these *Vibrio* species, despite the phylogenetical distance between the species (Gregory et al., 2020). CosR homologues have also been found in numerous species of *Alphaproteobacteria, Betaproteobacteria* and *Gammaproteobacteria* (Czech et al., 2018; Gregory et al., 2020).

Therefore, due to the wide-spread presence of this global osmolyte regulator in bacteria, it presented as a good candidate for a compatible solute regulator in *L. aggregata.* BlastP analysis revealed an *L. aggregata* LZB033 *cosR* homologue with 44 % identity to the amino acid sequence of the *V. parahaemolyticus* cosR. This putative *L. aggregata cosR* will be explored within this chapter as a potential compatible solute regulator.

Identification of a transcriptional regulator responsible for the regulation of DMSP synthesis genes would be a novel discovery. The identification of the regulon may also lead to the discovery of genes involved in DMSP synthesis or related processes.

### 5.1.5 Chapter Aims

The major aims of this chapter are to identify and test candidate DMSP synthesis genes and a potential osmolyte regulator in *L. aggregata.* As such, the aims of this chapter are:

1) Identify candidate methionine aminotransferases from the *L. aggregata* RNA-Seq data.
2) Mutate the candidate aminotransferases in *L. aggregata* and determine how this affects DMSP production.
3) Purify the candidate aminotransferases and determine whether they have methionine aminotransferase activity via *in vitro* enzyme assays.
4) Examine the regulation of the *cosR* homologue in *L. aggregata*.
5) Mutate *cosR* in *L. aggregata* and determine the effect on DMSP production

## 5.2 Results

### 5.2.1 Identification of candidate DMSP synthesis genes from the RNA-Seq data

The conditions used in the RNA-Seq experiment in Chapter 3 were chosen as they caused maximal differences in DMSP production, and thus in the differential expression of potential DMSP synthesis genes. Therefore, genes that were significantly up-regulated in low nitrogen and in the + Met condition, and down-regulated in the low salinity condition were searched for potential DMSP synthesis genes. However, the high number of differentially expressed genes across the different conditions made this a difficult task. To overcome this, a Venn diagram was constructed to identify genes that had a significant change in expression in more than one condition, specifically those that were up-regulated in low nitrogen and + Met conditions, and down-regulated under low salinity (**Figure 5.4**).

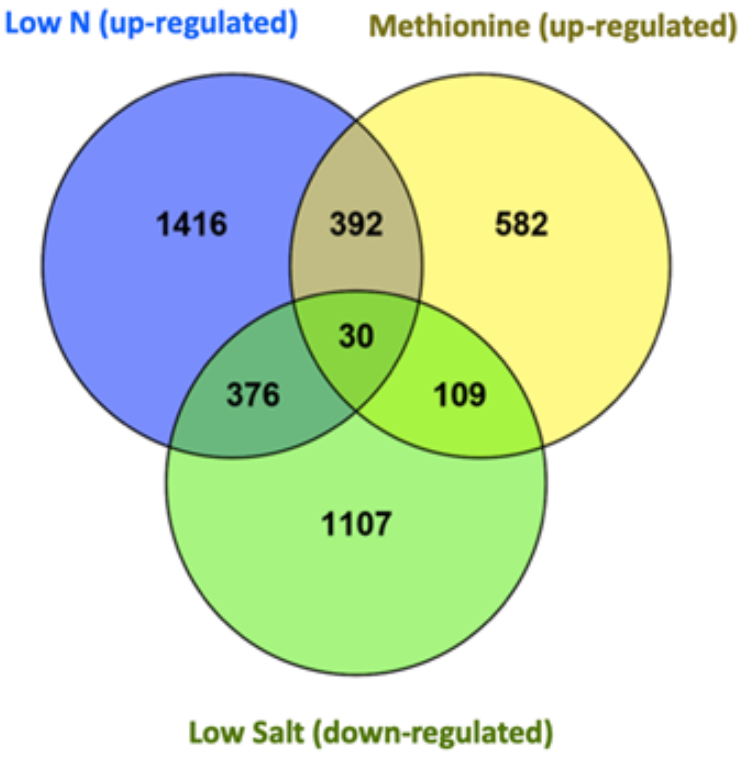



Figure 5.4 Venn diagram constructed from RNA-Seq differential expression data.
The number of genes which were significantly down-regulated in the low salinity condition and up-regulated in the low nitrogen condition and in the added methionine condition are shown.

The produced Venn diagram reduced the number of regulated genes from > 3500 to just 30 that were up-regulated in the Low N and methionine whilst also being down-regulated under low salinity. However, RASTtk annotation and BLAST analysis suggested that none of these 30 genes encoded putative Met aminotransferase, MTOB reductase or DMSHB decarboxylate proteins. This was not unexpected as *dsyB* itself was not up-regulated by the presence of methionine (**Figure 3.13**) so itself did not appear in the generated list of thirty genes.

Therefore, due to *dsyB* not being significantly up-regulated in the presence of methionine, it is possible that the remaining genes in the pathway will not be either. Therefore, the decision was made to look for candidate DMSP synthesis genes in the 376 genes that were significantly up-regulated in the low nitrogen and significantly down-regulated in the low salinity condition. The RASTtk annotation of these genes was searched for enzymes which could potentially have aminotransferase, reductase or oxidative decarboxylase activity (which may represent candidate Met aminotransferase, MTOB reductase and DMSHB decarboxylase proteins, respectively). Genes which were annotated as hypothetical proteins were subject to BlastP analysis to identify any domains which may indicate they have one of the three aforementioned activities. The identified candidate genes are presented in **Table 10**.

None of the 376 genes analysed were annotated to have decarboxylase activity, so a candidate DMSHB decarboxylase was not identified using this approach. Contrastingly, twenty of the 376 genes were predicted to encode enzymes with reductase activity. Only two of the 376 genes were annotated as encoding aminotransferase enzymes (**Table 10**). As such, these two genes were investigated as candidate methionine aminotransferase genes which may encode enzymes that catalyse the first step of the transamination synthesis pathway.

Table 10: Candidate DMSP synthesis genes identified from *L. aggregata* RNA-Seq data.
All genes presented below were significantly up-regulated under low nitrogen and significantly down-regulated in low salinity. Candidates were included in the table if the RASTtk/BlastP annotations indicated the gene encoded an enzyme with aminotransferase or reductase activity.

| Enzyme Activity | RASTtk/BlastP Annotation | Accession Number |
|---|---|---|
| Aminotransferase | PLP-dependent aminotransferase family protein | WP_023002694.1 |
| | aspartate aminotransferase family protein | WP_075282955.1 |
| Reductase | molybdopterin-dependent oxidoreductase | WP_075283544.1 |
| | SDR family oxidoreductase | WP_023002200.1 |
| | quinone oxidoreductase | WP_031270065.1 |
| | NAD(P)-binding domain-containing protein | WP_075282879.1 |
| | SDR family oxidoreductase | WP_152506038.1 |
| | Xanthine dehydrogenase | WP_075284654.1 |
| | Ribonucleotide reductase of class II | WP_075282093.1 |
| | Oxidoreductase | WP_075281664.1 |
| | SDR family oxidoreductase | WP_075282856.1 |
| | 2OG-Fe(II) oxygenase | WP_075284608.1 |
| | NAD(P)/FAD-dependent oxidoreductase | WP_075282856.1 |
| | Oxidoreductase | WP_023002201.1 |
| | SDR family oxidoreductase | WP_062488241.1 |
| | Aldehyde dehydrogenase | WP_075283090.1 |
| | GcvT family protein | WP_075282361.1 |
| | Gfo/Idh/MocA family oxidoreductase | WP_075283638.1 |
| | NAD(P)-dependent glycerol-3-phosphate dehydrogenase | WP_023002899.1 |
| | SDR family oxidoreductase | WP_075283479.1 |
| | FAD-binding oxidoreductase | WP_075283856.1 |
| | NAD(P)H-quinone oxidoreductase | WP_075282851.1 |

The two aminotransferases identified from the 376 genes are referred to as ARO8 (WP_023002694.1) and AT2 (WP_075282955.1) hereinafter (**Table 11**). The ARO8 aminotransferase was so-named because it contains an ARO8 (aromatic amino acid) domain. Both genes were positively regulated in low nitrogen conditions and downregulated in the low salinity condition. This expression pattern corresponds with the level of DMSP production and *dsyB* expression in these conditions.

The RNA-Seq data was analysed to determine the expression changes of these candidate methionine aminotransferase genes under conditions known to affect DMSP production. In the +Met condition, the *at2* gene was downregulated and there was no significant expression change of the *aro8* gene (**Table 11**). It may be expected that the methionine aminotransferase of the DMSP transamination synthesis pathway would have increased expression in the presence of its substrate, methionine. However, *dsyB* was down-regulated in the presence of methionine despite it causing a substantial increase in DMSP production (**Table 11**). Therefore, the methionine aminotransferase involved in DMSP synthesis may not be up-regulated by methionine.

Table 11: The differential expression of the two candidate DMSP-synthesis aminotransferases. The β-value is shown as a positive value to indicate up-regulation of the gene, and shown as a negative value to indicate down-regulation of the gene. In the case where the differential expression was not significant compared to the control condition, the β-value was recorded as 0.

| | | β-value (RNA-Seq) | | |
|---|---|---|---|---|
| | BlastP Annotation | Low N | 5 PSU | Met |
| AT2 | Aspartate aminotransferase family protein | +0.680 | -0.292 | -0.377 |
| ARO8 | PLP-dependent aminotransferase family protein | +0.274 | -0.600 | 0 |

### 5.2.2 Mutagenesis of the candidate aminotransferases

Following identification of the two candidate methionine aminotransferases, both were knocked out in *L. aggregata* to determine the effect on DMSP production. If one of the aminotransferases was involved in DMSP production, it is likely that the DMSP production would be reduced or abolished when the gene was mutated. This was achieved using suicide plasmid pk18mob*sacB*. In this process, flanking regions of the *at2* gene and the *aro8*

gene were synthesised with an *EcoRI* restriction site at the 5' end and a *PstI* restriction site at the 3' end. This allowed for sub-cloning from of the synthesised flanking regions into pk18mob*sacB*. An *XbaI* restriction site was also introduced in the middle of the flanking regions.

As described previously, the Kan$^R$ of the helper strain, *E. coli* 803 (pRK2013) used to facilitate the conjugation of the pk18mob*sacB* plasmid from the host *E. coli* DH5-α into *L. aggregata* presented a problem. This is because the Kan$^R$ encoded on the backbone on pk18mob*sacB* could not be used for the selection of *L. aggregata* colonies that had undergone single cross-over with the aminotransferase flanking regions in the pk18mob*sacB* plasmid. To overcome this, a spectinomycin resistance (Spec$^R$) cassette was cloned into the middle of the flanking regions. This Spec$^R$ cassette was amplified from the plasmid php45Ω, with primers that allowed for the addition of an *XbaI* restriction site on both the 5' and 3' end of the sequence.

As such, the following constructs were made. pBIO23(4) comprises of the pk18mobsacB backbone with flanking regions of the *at2,* with the Spec$^R$ cassette in the centre of the two flanking regions (**Figure 5.5**). pBIO23(3) comprises of pk18mob*sacB* with flanking regions of the *aro8* gene and the Spec $^R$ cassette in the centre (**Figure 5.5**).

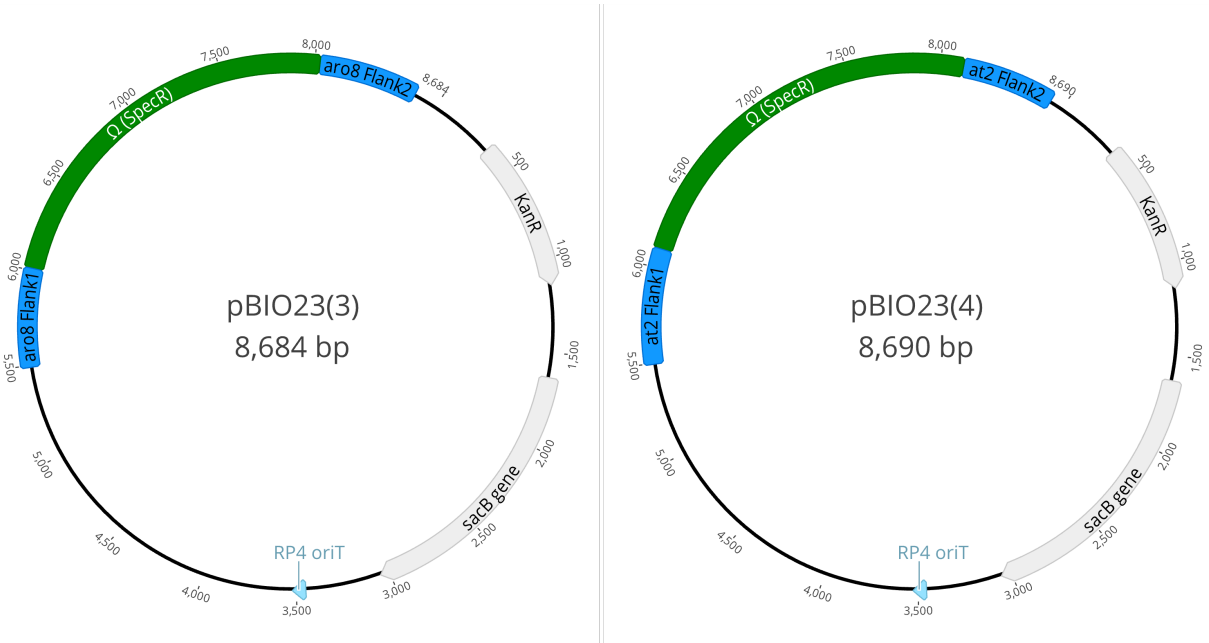



Figure 5.5 Plasmid maps of pBIO23(3) and pBIO23(4).

Following construction of the pBIO23(3) and pBIO23(4) plasmids, both constructs were crossed into the *L. aggregata* WT strain. Colonies that had undergone double homologous recombination were selected for by the loss of sucrose sensitivity and the gain of $Spec^R$ resistance. Mutants were confirmed through the PCR of each strain with checking primers designed outside of the flanking regions of both aminotransferases. The resulting PCR products were then analysed by gel electrophoresis (**Figure 5.6**).

In the case of the *aro8* mutagenesis, checking primers were designed to give a WT band of 1.6 Kb, and the mutant strain would give a larger band of 3 Kb . Similarly, in the case of the *at2* mutagenesis, the checking primers were designed to give a WT band of 1.6 Kb and the mutant strain would give a larger band of 3 Kb. As can be seen in **Figure 5.6**, these predicted band sizes correspond to the obtained PCR products, and thus it was confirmed that an *L. aggregata* Δ*aro8* and an *L. aggregata* Δ*at2* mutant had been constructed.

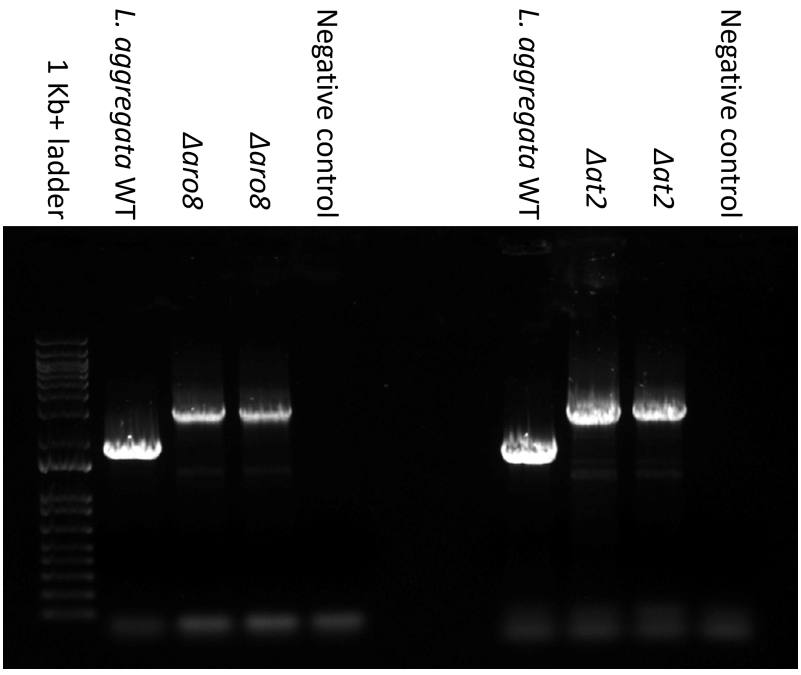



Figure 5.6 Gel image of checking PCR to determine success of aminotransferase mutagenesis. The first lane contains the 1 Kb+ DNA ladder. The next four lanes contain the PCR products of reactions carried out with aro8 checking primers. The remaining four lanes contain PCR products of reactions carried out with the at2 checking primers. *L. aggregata* WT was used as a positive control and the negative controls used water in place of template.

### 5.2.3 DMSP production of the *L. aggregata* aminotransferase mutants

To determine the effect of knocking out each of the candidate methionine aminotransferases, each of the mutants and the WT strain were grown in standard growth conditions (MBM 35 PSU) and the DMSP production of each was measured (**Figure 5.7**). There was a significant reduction in the DMSP production of both aminotransferase mutants compared to the WT (t-test, $p<0.05$). Therefore, this suggests both aminotransferases may have a role in the transamination of methionine which comprises the first reaction of the DMSP synthesis pathway. However, mutagenesis of the aminotransferases did not result in complete abolishment of DMSP synthesis.

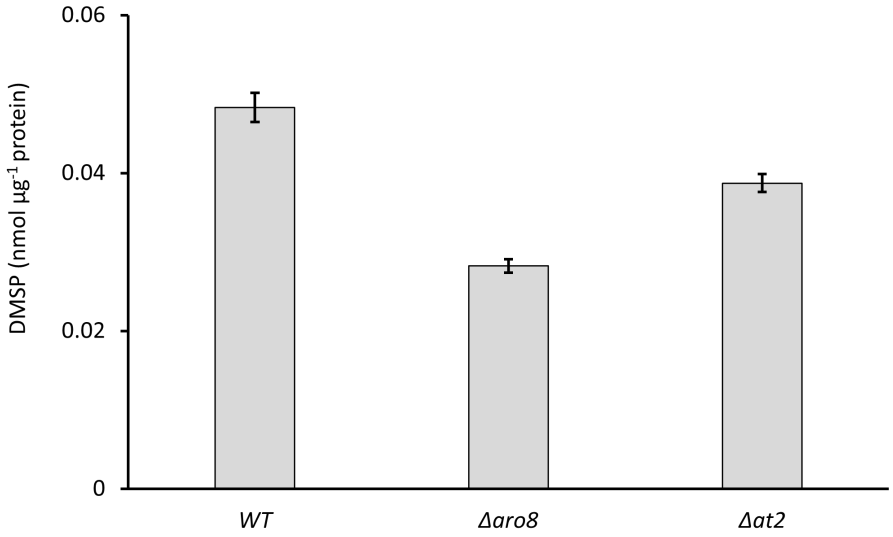



Figure 5.7 DMSP production of *L. aggregata* WT and single aminotransferase mutant strains. The DMSP production of each sample was measured via gas chromatography and normalised to protein content. N=3 biological replicates and error bars represent the standard error of the mean.

### 5.2.4 Creating an *L. aggregata* double aminotransferase mutant

As there was not a complete abolishment of DMSP synthesis in either the *L. aggregata* Δ*aro8* or the Δ*at2* mutant, but both appeared to reduce DMSP synthesis, it was hypothesised that the aminotransferases may be compensating for the loss of each other. To investigate this, a double mutant was created in which both aminotransferases were knocked out. This double mutant is denoted as *L. aggregata* Δ*aro8*Δ*at2* hereinafter.

To create the *L. aggregata* Δ*aro8*Δ*at2* strain, pk18mobsacB containing the flanking regions of *aro8* (pBIO23(5)) needed to be conjugated into the *L. aggregata* Δ*at2* strain. To aid with selection, many attempts were made to clone the $Gent^R$ cassette from p34S-Gm into the middle of the flanking regions in this construct, however this was not successfully achieved. Instead, to create the double mutant, pBIO23(5) was transformed into *E. coli* S17, which contains a chromosomally integrated RP4 plasmid to aid conjugation (Strand et al., 2014). This removed the need to use the $Kan^R$ *E. coli* 803 (pRK3013) in the cross to conjugate pBIO23(5) into *L. aggregata* Δ*at2*. Therefore, pBIO23(5) could be conjugated into *L. aggregata* Δ*at2*, and a single cross-over could be selected for by the gain of $Kan^R$. After selection for double cross-overs, using loss of sucrose-sensitivity for selection, the resulting colonies were screened for *L. aggregata* Δ*aro8*Δ*at2* mutants.

This screening was again carried out via PCR, using the same checking primers designed to bind just outside of the *aro8* flanking regions in **Figure 5.6**. Therefore, when the PCR products were analysed by gel electrophoresis, colonies were screened for a band either the size of the WT gene (~1.6 Kb) or a smaller band of ~ 1 Kb, corresponding to the size of the flanking regions (**Figure 5.8**). As can be seen in the gel image, 8 out of the 11 colonies screened were *L. aggregata* Δ*aro8Δat2* mutants.

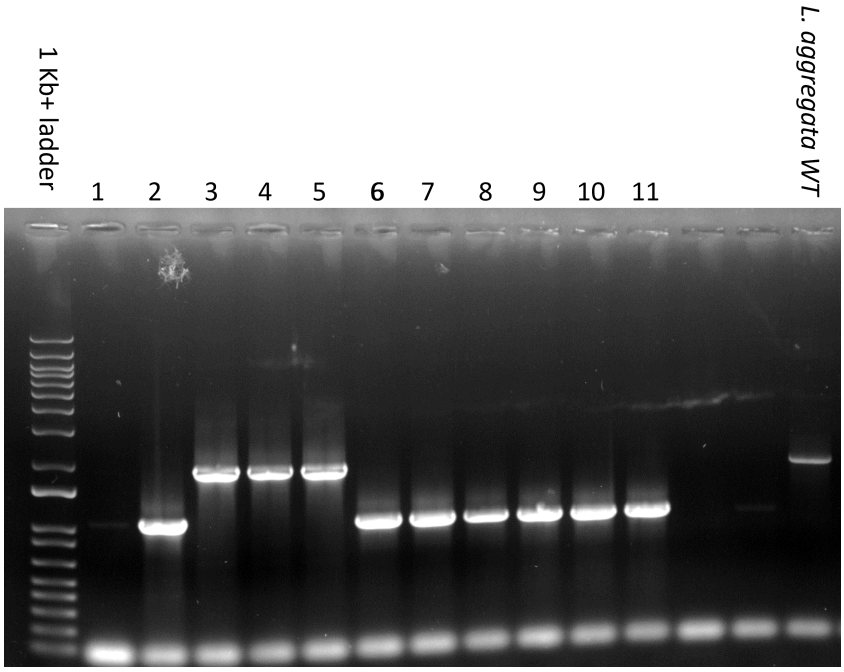



Figure 5.8 Gel image of colony PCR used for screening for double aminotransferase mutants. The first lane contains 1 Kb+ DNA ladder. The remaining lanes are PCR products from colony PCR screening used for detecting double aminotransferase mutants. The last lane contains the PCR product of *L. aggregata* WT.

### 5.2.5 DMSP production of the *L. aggregata* single and double aminotransferase mutants

Following the production of the double mutant, the DMSP production of all of the aminotransferase mutants was compared to the WT when grown in differing salinities, which had previously been shown to regulate DMSP production in the WT strain (**Figure 5.9**).

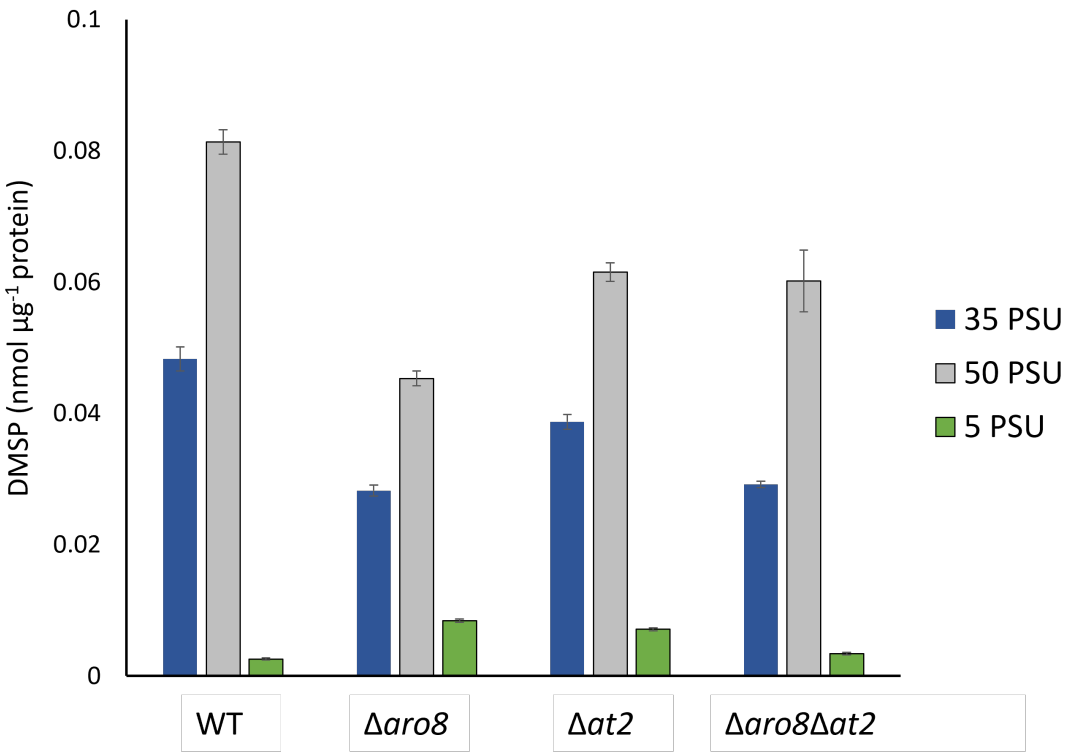



Figure 5.9 DMSP production of *L. aggregata* and single and double aminotransferase mutants in different salinities.
The DMSP production was measured by gas chromatography and normalised to the protein content of the cultures. N=3 biological replicates and error bars represent the standard error of the mean.

All of the mutant strains produced less DMSP than the WT in both the standard salinity condition (35 PSU) and the high salinity condition (50 PSU). However, the Δ*aro8*Δ*at2* mutant did not produce significantly less DMSP than the single mutants in either salinity condition. As such, this disproves the hypothesis that the At2 enzyme is compensating for loss of DMSP production in the Δ*aro8* mutant, and that the Aro8 enzyme is compensating for the loss of DMSP production in the Δ*at2* mutant.

To determine whether the reduction of DMSP production by the mutants would be more pronounced in the presence of the proposed starting substrate of the aminotransferases, the strains were grown in the presence of 0.5 mM methionine, and normalised to protein content (**Figure 5.10**). There was no significant difference between the DMSP production of the WT and any of the mutant strains in the presence of methionine. As will be discussed further in Section 5.3, this may be due to several more aminotransferases having the ability to convert methionine to MTOB.

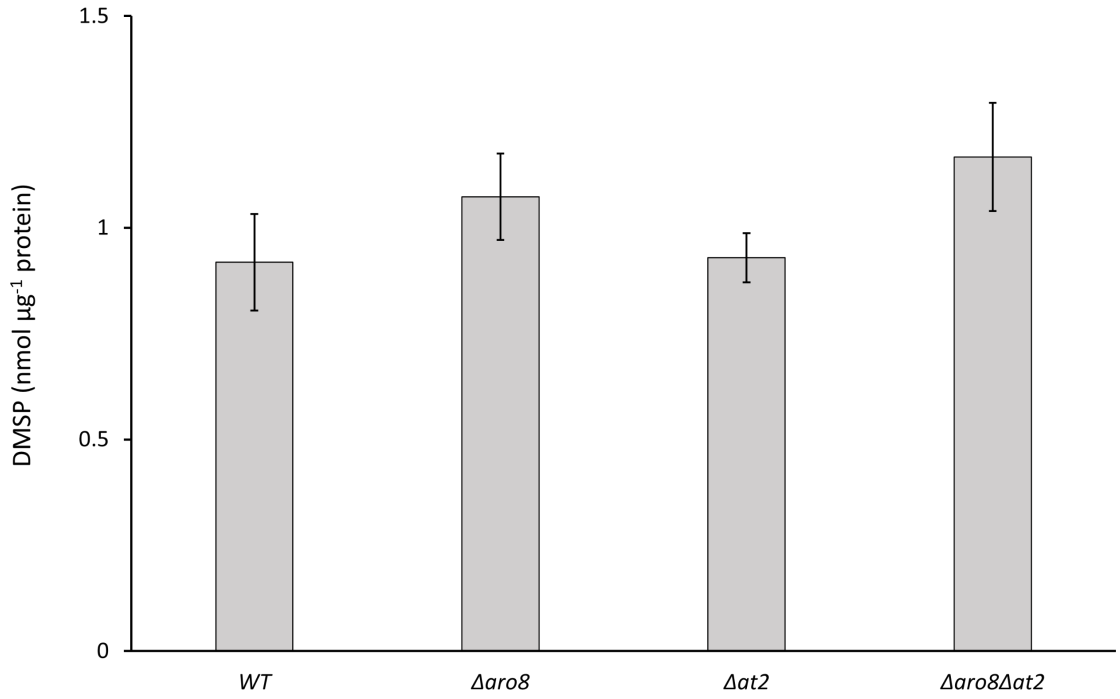



Figure 5.10 DMSP production of *L. aggregata* aminotransferase mutants in the presence of methionine.
DMSP production of each culture grown in the presence of 0.5 mM methionine was measured using gas chromatography and normalised to protein content. N=3 biological replicates, error bars represent the standard error of the mean.

### 5.2.6 Attempted protein purification of the candidate aminotransferases

As there were indications of the role of the aminotransferases in DMSP synthesis from the mutagenesis experiments, confirmation of these roles was attempted via the purification of these proteins and their *in vitro* characterisation. In an initial attempt to overexpress the Aro8 aminotransferase protein, the gene was codon-optimized for expression in *E. coli* and synthesised, with an *NdeI* restriction site at the 5' end an *EcoRI* site at the 3' end of the gene. The inclusion of these restriction sites allowed for sub-cloning of the gene from the supplied vector into the widely used expression plasmids pET21a and pET16b. Therefore, the following constructs were made pBIO23(6) (pET16b:*aro8*) and pBIO23(7) (pET21a:*aro8*).

Following transformation into the *E. coli* expression strain BL21(DE3), and growth and induction with IPTG, soluble and insoluble fractions were analysed via SDS-polyacrylamide

gel electrophoresis (SDS-PAGE) (**Figure 5.11**). Although this was a preliminary test, it was evident that there was not a band corresponding to overexpressed Aro8 protein in the soluble fraction of either pET16a or pET16b plasmid. However, there did appear to be two bands in the non-soluble fractions that appeared in the pBIO23(6) and pBIO23(7) constructs, but not the empty plasmid controls (denoted by red arrows in **Figure 5.11**). However, both bands were between 35 – 40 kDA, which were smaller than the expected size (~50 kDa) of the Aro8 protein. At this stage, it was unknown whether one of these bands was the target Aro8 protein and further optimisation was required to express these proteins in the soluble fraction.

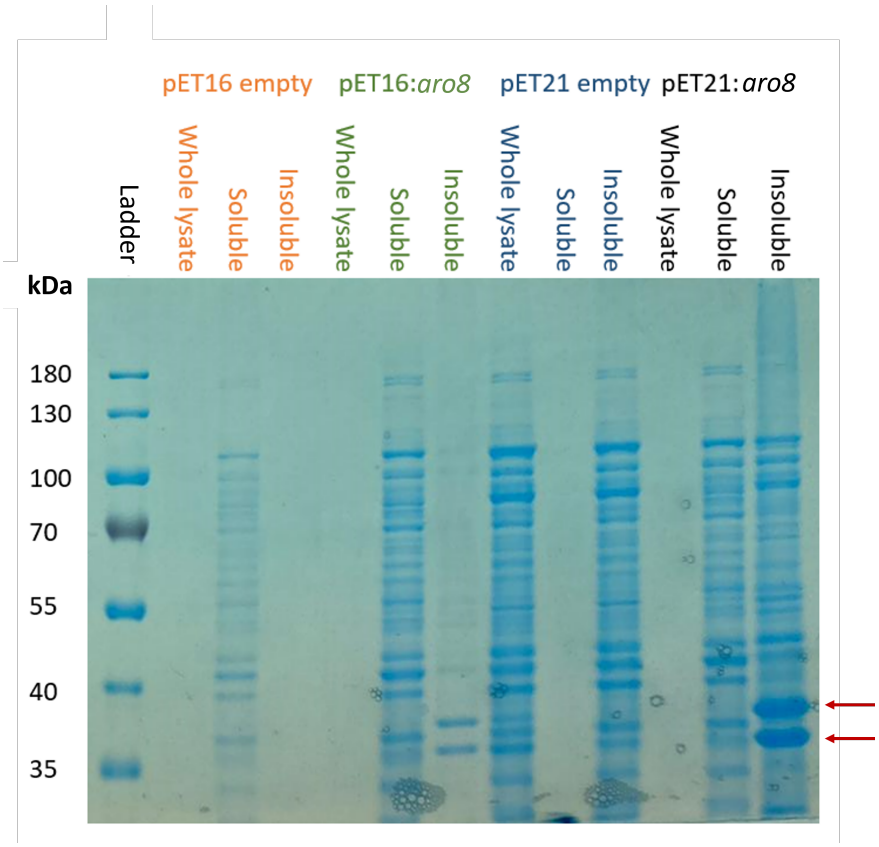



Figure 5.11 SDS-PAGE image of attempted Aro8 over-expression in pET16b and pET21a plasmids. The first lane contains the PageRuler ladder. pET16 refers to the expression plasmid pET16b and pET21 refers to the expression plasmid pET21b. The empty vector samples in lanes labelled in blue and orange were included as negative controls. The expected size of Aro8 protein is 52 KDa.

As the production of soluble Aro8 protein was not successfully carried out in either pET16b or pET21b, the decision was made to utilise a different vector which may support enhanced protein expression and/or solubility. As such, pMAL-c2X was used as it encodes a maltose

binding protein (MBP), which is fused to the N-terminus of the target protein. MBP is a useful tag for the purification of proteins, as it often enhances the solubility of otherwise insoluble proteins (Kapust & Waugh, 1999). Further, the MBP tag can function as an affinity tag and allow for the purification of the target protein via binding to an amylose column.

The nucleotide sequence of the *at2* gene was codon-optimised for expression in *E. coli,* in the same way as previously described for the *aro8* gene. In both cases, internal *BsaI* sites were removed, to allow for the Golden Gate cloning (using *BsaI*) of each gene into pMAL-c2X. Following the golden gate cloning and transformation into *E. coli* 803, colonies were screened for successful clones via PCR using primers designed to the backbone of the pMAL-c2X plasmid (**Figure 5.12**). Using these primers, empty pMAL-c2X vectors would produce a PCR product of ~500 bp. However, successful constructs would give a band corresponding to the size of the gene plus 500 bp. Successful pMAL-c2X:*aro8* (pBIO23(8)) and pMAL-c2X:*at2* (pBIO23(9)) constructs were detected via this PCR.

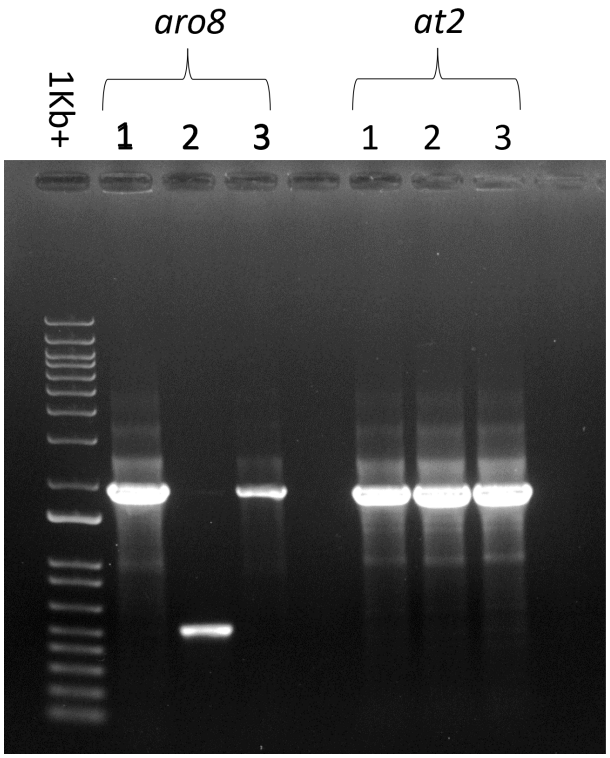



Figure 5.12 Agarose gel image confirming cloning of *aro8* and *at2* into pMAL-c2X 1039.
The first lane contains the 1Kb+ DNA ladder. The next three lanes contain the PCR products of colonies screened for successful pMAL-c2X 1039:*aro8* constructs. The last three lanes contain the PCR products of colonies screened for successful pMAL-c2X 1039:*at2* constructs.

To further confirm that pBIO6 and pBIO7 were correctly formed constructs, both were sequenced using the aforementioned primers. The obtained sequencing results confirmed that pBIO23(8) contained the codon-optimized sequence of *aro8* and pBIO23(9) contained the codon-optimised sequence of *at2.*

Following confirmation of the constructs, pBIO23(8) and pBIO23(9) were transformed into *E. coli* BL21 and cultures were grown and induced. Protein purification of each aminotransferase was attempted using affinity chromatography with amylose resin and PD-10 desalting columns were used for buffer exchange. The resulting protein preparations were then analysed using SDS-PAGE (**Figure 5.13**). As the MBP (40.3 KDa) would be fused to the aminotransferase proteins of interest, the expected size of the At2 (48.2 KDa) tagged with MBP would be 88.5 kDa and the expected size of the Aro8 protin (51.9 KDa) tagged with MBP would be 92.2 kDa.

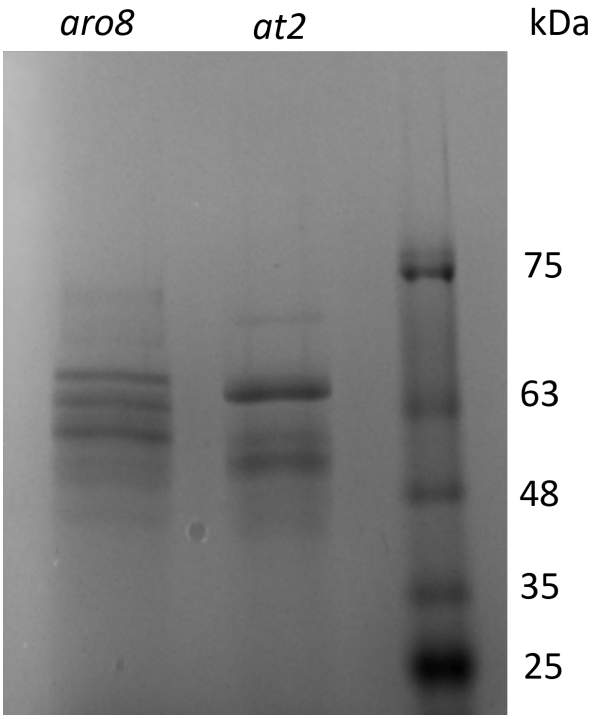



Figure 5.13 SDS-PAGE of attempted pBIO(6) and pBIO(7) protein purification.
The ladder on the right is the PageRuler™ ladder. The left lane was loaded with the sample obtained from the BL21:pBIO(6) strain and the middle lane was loaded with the sample obtained from the BL21:pBIO(7) strain. Expected protein sizes of the MBP-tagged proteins were 92.2 KDa (Aro8) and 88.5 KDa (At2).

The SDS-PAGE analysis of the purified protein samples determined that the aminotransferase proteins were not over-expressed, or not in the soluble fraction. In this case, the protein band would have been ~ 90 kDa, due to the size of the aminotransferase protein plus 40 kDa corresponding to the fused MBP. Although not shown, the non-soluble fraction was also run on a separate SDS-PAGE gel and again there was not a band present at 90 kDa. Therefore, it could be confirmed that it was not a case of the protein being insoluble but rather a result of the protein not being expressed.

Therefore, as the pBIO23(8) and pBIO23(9) constructs had been confirmed as correct via PCR and sequencing, there was an attempt to optimise the conditions to induce expression of the aminotransferase-MBP constructs. This optimisation involved different incubation times and varying the concentrations of IPTG used for induction, however when cell lysate was analysed via SDS-PAGE, there did not appear to be a band present at 90 kDa in either the soluble or insoluble fractions (data not shown). Due to timing constraints, this optimisation could not be continued, and purified aminotransferase proteins were not obtained. Therefore, purification and subsequent activity assays remain future work that need to be carried out in order to determine if Aro8 and At2 have the methionine aminotransferase activity of the DMSP transamination synthesis pathway.

### 5.2.7 Development of methionine aminotransferase activity assays

Whilst optimisation of the aminotransferase protein purification was being carried out, an assay was devised that would allow the detection of methionine aminotransferase activity. The development of this assay was aided by Dr Rocky Payet (UEA). These assays depended on the detection and measurement of the amino acids, methionine and glutamate. In principle, the incubation of a methionine aminotransferase enzyme with methionine, and the required cofactors (PLP and 2-oxoglutarate), would result in the amino group being transferred from methionine to 2-oxoglutarate and therefore the formation of glutamate (**Figure 5.3**). This experiment required the derivatisation of amino acids by *o*-phthaldialdehyde (OPA) to allow for their detection by high-performance liquid chromatography (HPLC). As such, a functional methionine aminotransferase could be

characterised by the HPLC detection of glutamate and decreased concentrations of methionine.

The preliminary OPA derivatization and HPLC was carried out by Dr. Rocky Payet, and successfully allowed detection of methionine and other amino acids needed for the assays (**Figure 5.14).** As such, aminotransferase assays have been developed but could not be used do to the lack of Aro8 and At2 protein. In the future, these aminotransferase assays could also be carried out on cell extracts of *L. aggregata* WT and aminotransferase mutant strains. This work could not be carried out due to time constraints but is discussed further in Section 5.3.

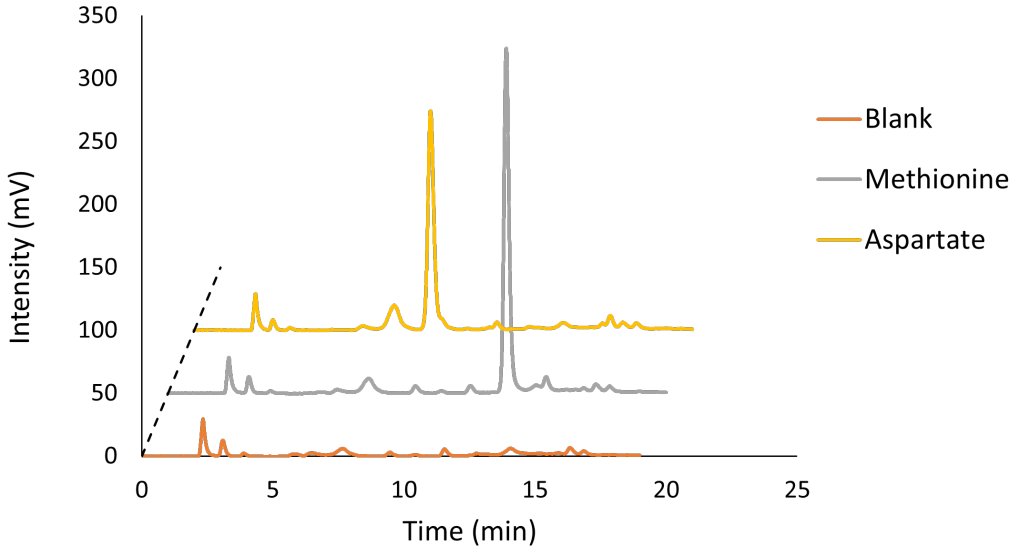



Figure 5.14 HPLC chromatogram of OPA-derivatised amino acids, as carried out by Dr. Rocky Payet.

### 5.2.8 Identification of a candidate global osmolyte regulator

Attempts have also been made to identify potential regulators of DMSP and GB production in *L. aggregata.* The RNA-Seq dataset contained a large quantity of transcriptional regulators that were significantly differentially expressed across the different conditions known to regulate DMSP and GB production. For example, there were 170 genes annotated as transcriptional regulators, that had significant differential expression changes between the control and low salinity (5 PSU condition). The top 5 genes encoding transcriptional

regulators that were upregulated and downregulated in low salinity are presented in **Table 12.**

Analysis of the highly regulated transcriptional regulators revealed that one of the genes (WP_023001062.1), annotated as a MarR type transcriptional regulator, had a 44 % amino acid identity to CosR, the global compatible solute regulator in *V. parahaemolyticus.* Therefore, this *L. aggregata* gene was identified as a potential *cosR* homologue and is referred to as *cosR* hereinafter.

Table 12 Highly differentially expressed genes annotated as transcriptional regulators in *L. aggregata* RNA-Seq
Pairwise comparison between the expression in the control and low salinity condition were carried out and the differential expression is represented as β-values (adjusted $Log_2$ fold change).

| Accession Number | Annotation | β-Value |
|---|---|---|
| WP_062486745.1 | Transcriptional regulator, LacI family | -1.67 |
| WP_023001062.1 | Transcriptional regulator, MarR family (putative CosR) | -1.46 |
| WP_083660770.1 | Transcriptional regulator, LysR family | -1.30 |
| WP_055655398.1 | Transcriptional regulator, AcrR family | -1.20 |
| WP_040440215.1 | Transcriptional regulator, Xre family | -1.19 |
| WP_022999787.1 | Transcriptional regulatory protein | +2.84 |
| WP_023003882.1 | Transcriptional regulator containing an amidase domain | +1.80 |
| WP_075281542.1 | Transcriptional regulator, PadR family | +1.41 |
| WP_075284623.1 | Transcriptional regulator, LysR family | +1.38 |
| WP_075283580.1 | Transcriptional regulator, AcrR family | +1.29 |

Upon further analysis of this potential *cosR* gene in the RNA-Seq dataset, it was found to be down-regulated in all of the conditions tested (low salinity, low nitrogen and + Met) compared to the control condition (**Table 13**). CosR has been shown to be a transcriptional repressor of compatible solute synthesis genes in studies carried out in *Vibrio* species (Gregory et al., 2020). Therefore, the *L. aggregata* CosR homologue is also likely to be a transcriptional repressor. It may be expected that a transcriptional repressor of DMSP synthesis genes would be downregulated in conditions shown to increase DMSP production, as is the case with downregulation of *cosR* in the low N and + Met conditions.

In contrast, *cosR* was downregulated under low salinity, a condition known to cause reduced DMSP and GB synthesis. This can perhaps be explained by the rapid negative self-feedback loop suggested for CosR in *V. cholerae,* where CosR was found to be a repressor of compatible solute synthesis but itself was down-regulated in low salinity conditions (Shikuma et al., 2013).

Interestingly, the putative *cosR*-like gene in *L. aggregata* was in the top 5 % of the genes that were significantly down-regulated under the low salinity conditions. This further suggests that the putative *cosR* is highly regulated by salinity, and therefore it is possible that CosR controls the transcription of genes involved in osmotic stress protection.

Table 13: Differential expression analysis of the *cosR* homologue in *L. aggregata*

| | | β-value (RNA-Seq) | | |
|---|---|---|---|---|
| | Blast Annotation | Low N | 5 PSU | Met |
| *cosR* | Transcriptional regulator, MarR family | -1.24 | -1.46 | -0.38 |

### 5.2.9 Mutagenesis of the *cosR* homologue in *L. aggregata*

To determine whether CosR is a transcriptional regulator of compatible solute synthesis genes in *L.* aggregata, the *cosR* gene was knocked out using pk18mobsacB. The flanking regions of *cosR* were synthesised directly into pk18mobsacB, to form construct pBIO23(10). pBIO23(10) was transformed into *E. coli* S17 before being conjugated into *L. aggregata* LZB033. Single crossovers were selected for by the gain of $Kan^R$. Double crossovers were then selected for via the loss of sucrose sensitivity and screened via colony PCR, using primers designed to bind to outside the flanking regions in the *L. aggregata* genome. The primers were designed to give a product of 1350 bp when amplifying from *L. aggregata* WT and a band of 1100 bp when amplifying from *L. aggregata* gDNA that had undergone a double crossover. Around 40 colonies were screened and analysis of the resulting PCR products identified one colony that was a successful *L. aggregata* Δ*cosR* mutant (**Figure 5.15**).

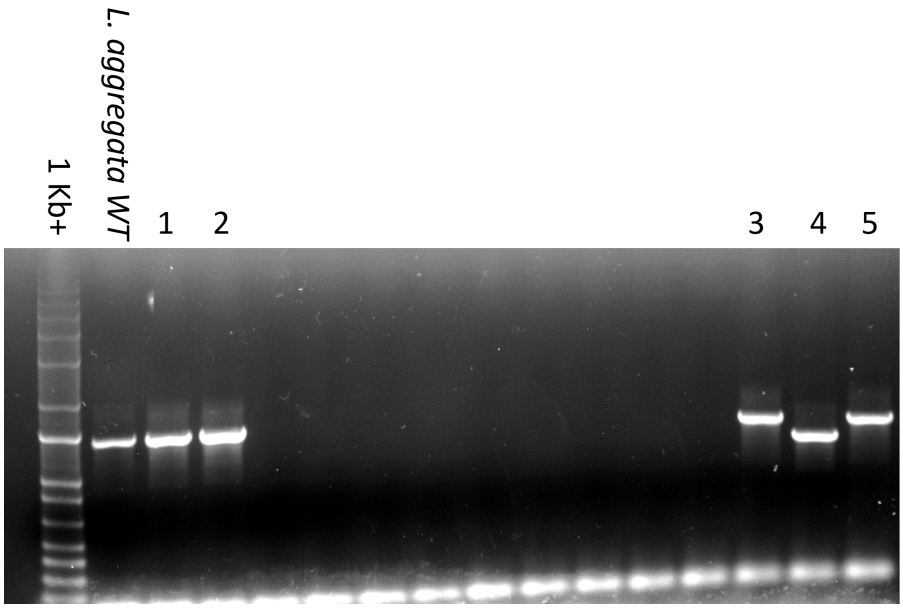



Figure 5.15 Agarose gel image of screening for *cosR* mutants using colony PCR.
The first lane contains the 1 Kb+ DNA ladder. The next lane contains the PCR product when the *cosR* checking primers were used against *L. aggregata* WT. Samples 1-5 are the PCR products obtained from the colony PCR screening. Sample 4 was identified as *L. aggregata* Δ*cosR.*

### 5.2.10 DMSP production of *L. aggregata* Δ*cosR*

Following the production of the *L. aggregata* Δ*cosR* mutant, the DMSP production of this strain was compared to the WT (**Figure 5.16**). The DMSP production in this strain was reduced compared to the WT in every condition tested, suggesting the involvement of CosR in the regulation of DMSP synthesis. However, the DMSP production of *L. aggregata* Δ*cosR,* was regulated in a similar fashion to the WT strain; production was decreased to below the detection limit in low salinity, and production was increased compared to the control condition in high salinity and low nitrogen conditions. Therefore, knocking out *cosR* did not result in the complete abolishment of DMSP regulation. This may suggest that CosR is not the only transcriptional regulator involved in DMSP synthesis. It is also important to note that the DMSP production was not normalised to the protein content of these samples, thus the amounts of DMSP produced by each strain cannot reliably be compared in this particular experiment. The effect of *cosR* mutagenesis on *dsyB* regulation could be further investigated via RT-qPCR to give a more definitive explanation of the role CosR has on the regulation of DMSP synthesis genes.

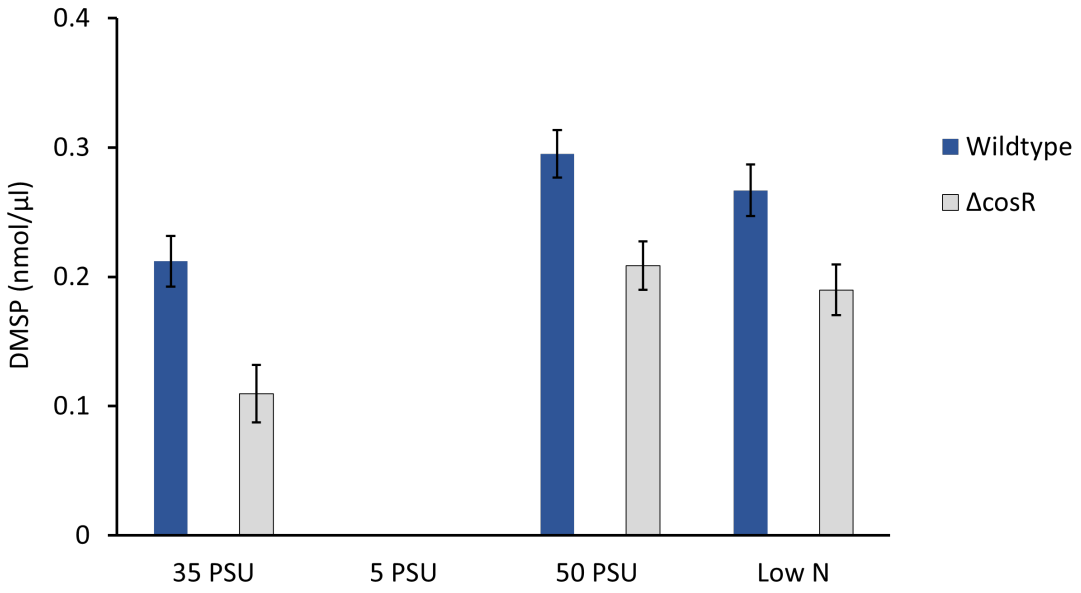



Figure 5.16 DMSP production of *L. aggregata ΔcosR* and *L. aggregata* WT.
DMSP was quantified via gas chromatography. N=3 biological replicates and error bars represent the standard deviation of the mean.

CosR is thought to be a global osmolyte regulator, and has been shown to regulate ectoine and GB synthesis genes in *Vibrio* species (Gregory et al., 2020). As such, cultures of the *L. aggregata* Δ*cosR* mutant (and WT strain as a control) were grown to mid-exponential phase in different conditions known to regulate compatible solute synthesis in *L. aggregata* (5 PSU, 50 PSU, Low N) and processed for future LC-MS analysis. This analysis has not yet been carried out but will provide the intracellular concentration of each N and S containing zwitterionic metabolite in the *L. aggregata* WT and Δ*cosR* cultures. Therefore, the role of CosR in *L. aggregata* compatible solute regulation will be able to be more thoroughly investigated.

## 5.3 Summary and Discussion

Within this chapter, RNA-Seq analysis identified two *L. aggregata* genes as candidate DMSP synthesis genes, as they potentially encoded enzymes with methionine aminotransferase activity. These two aminotransferase genes, *at2* and *aro8* were identified as candidate DMSP synthesis genes as the expression changes of these genes correlated with those expected of genes involved in DMSP synthesis.

As well as being one of only two genes annotated as encoding an aminotransferase that was upregulated under low nitrogen conditions and downregulated in low salinity, the ARO8 domain of the aminotransferase encoded for by the *aro8* gene was a positive indication in terms of its role as a candidate DMSP synthesis gene. This is due to the *ARO8* and *ARO9* genes of *Saccharomyces cerevisiae* being found to catalyse the reversible Met to MTOB reaction in the eukaryotic methionine salvage pathway (Deed et al., 2018; Pirkov et al., 2008). However, perhaps unsurprisingly there is low similarity between the amino acid sequences of the ARO8 proteins from *L. aggregata* LZB033 and *S. cerevisiae* (22 % identity). There are also multiple aminotransferases within the *L. aggregata* that are annotated as containing an ARO8 domain, although none of these have higher that 23 % identity to the characterised *S. cerevisiae* ARO8.

Following identification of the *aro8* and *at2* genes, both were mutated in *L. aggregata* and both the *Δaro8* and *Δat2* strains produced less DMSP in standard and high salinity conditions, indicating a role for these genes in DMSP synthesis in *L. aggregata.* To determine if the loss of one aminotransferase was being compensated for by the presence of the other in the mutant strains, a double mutant was created in which both aminotransferase genes were knocked out. However, the *L. aggregata Δaro8Δat2* strain did not produce significantly less DMSP than the single mutants and therefore it became clear that the loss of one aminotransferase was not being compensated for by the presence of the other.

It is perhaps possible that there are more than two aminotransferases capable of catalysing the methionine aminotransferase reaction at the start of the DMSP transamination

synthesis pathway. Microbial aminotransferases often have overlapping substrate specificity and can have involvement in more than one pathway (Jensen et al., 1981). Therefore, it is possible that aminotransferases involved in other biochemical pathways in the cell may also be able to carry out the methionine aminotransferase reaction and may therefore compensate for both knocked out genes in the Δ*aro8*Δ*at2* double mutant.

To further determine if the At2 and Aro8 proteins were involved in DMSP synthesis, attempts were made to purify the proteins and assay for methionine aminotransferase activity. Enzyme assays were developed which would allow the quantification of methionine and glutamate. The *at2* and *aro8* genes were cloned into protein expression plasmids and expressed using the *E. coli* BL21 strain. There were attempts to purify the proteins using affinity chromatography, however, SDS-PAGE analysis indicated that the target proteins were not being expressed. Terminal tagging of both proteins with the MBP solubility tag did not enhance expression nor solubility. Due to time limitations, the attempted protein purification was not continued. However, in the future it may be possible to overcome these problems by optimising growth and induction conditions and the use of alternative expression vectors and expression strains. The protein purification of these candidate methionine aminotransferases and subsequent enzyme activity assays would allow determination of whether these proteins are involved in DMSP production.

In summary, further investigation is required to determine whether *aro8* and *at2* are involved in DMSP synthesis. The activity of the enzymes would need to be determined in enzyme assays before they could be confirmed as DMSP synthesis genes.

The aminotransferase assays developed in this chapter could also be used on cell extracts of *L. aggregata* cultures grown under differing environmental conditions (for example, differing salinity and nitrogen concentrations). This would allow determination of whether the Met aminotransferase activity in *L. aggregata* is regulated as expected by salinity and nitrogen availability. Additionally, the enzyme assays could be carried out on cell extracts of *L. aggrgegata* Δ*aro8, L. aggrgegata* Δ*at2* and *L. aggrgegata* Δ*aro8*Δ*at2* to determine whether the knocking out of these aminotransferases does indeed reduce Met

aminotransferase activity. The aminotransferase assays were not carried out as part of this thesis due to time constrains, and therefore remain future work.

As discussed in Section 5.1.2, the Met aminotransferase reaction at the beginning of the transamination synthesis pathway, is the reverse of the final step in the ubiquitous Met salvage pathway. The *L. aggregata* genome was searched for homologues to the known bacterial MTOB aminotransferases of the Met salvage pathway and hits of above 30 % amino acid identity were found to the MTOB aminotransferase genes of *B. subtilis* (*ykrV* and *ybgE*) and *K. pneumoniae* (*tyrB*) (Berger et al., 2003; Heilbronn et al., 1999). These genes were not identified as significantly upregulated in low nitrogen conditions and significantly downregulated in low salinity conditions the RNA-Seq and therefore were not investigated further in this thesis. However, the identified homologues present as interesting candidate Met aminotransferase genes in *L. aggregata.* A particularly good candidate is a gene in *L. aggregata* (annotated as a LL-diaminopimelate aminotransferase) with 39 % identity to the amino acid sequence of YkrV (the *B. subtilis* MTOB transaminase).

If the aforementioned future work leads to the confirmation of the aminotransferase genes as DMSP synthesis genes, strains containing close homologues of the aminotransferases could be tested for the ability to produce DMSP. This could potentially lead to the identification of novel DMSP producing species, as was the case in other studies that have identified bacterial DMSP synthesis genes (Curson et al., 2017; Williams et al., 2019).

Work in this chapter also investigated the role of CosR as a potential compatible solute regulator in *L. aggregata.* CosR is a reported wide-spread repressor of bacterial compatible solute synthesis and transport genes (Gregory et al., 2020). A homologue of CosR was identified in *L. aggregata* and mutagenesis of *cosR* resulted in a slight decrease in DMSP production compared to the WT in all of the conditions tested. However, the DMSP production of the Δ*cosR* strain remained regulated by varying salinity and nitrogen concentrations. This may suggest that there may be more than one transcriptional regulator controlling the expression of DMSP synthesis genes in *L. aggregata*. This is the case in *Vibrio* species in which CosR has been found to be one of three transcriptional regulators involved in the regulation of compatible solute synthesis and transport genes

(Gregory et al., 2020; Gregory et al., 2019). The other two transcriptional regulators identified to be involved in compatible solute regulation were quorum sensing regulators, OpaR and AphaA (Gregory et al., 2019). OpaR and AphA homologues were searched for in the *L. aggregata* genome, but no hits of above 40 % identity were discovered.

This work potentially implicates CosR in the regulation of DMSP synthesis. However, future work would need to be carried out involving comparing the regulation of *dsyB* in the Δ*cosR* and WT strains to determine if CosR truly is a repressor of the key DMSP synthesis gene. The confirmation of CosR as a DMSP synthesis regulator would be the first identification of a transcriptional regulator of DMSP synthesis.

The role of CosR in relation to the regulation of the production of other compatible solutes has not been investigated in this thesis. However, cultures of the *L. aggregata* WT and *ΔcosR* strains were grown under different salinity conditions and have been saved for future LC-MS analysis, which will enable the quantification of each compatible solute in the strains under each condition. Additionally, further work to characterise the regulon of CosR in *L. aggregata* may be achieved by performing an RNA-Seq analysis of the Δ*cosR* strain under differing salinities and comparing expression data to the WT strain. This would allow full determination of the genes under the regulation of CosR.

# Chapter 6: Final Summary and Discussion

## 6.1 Review of aims and gaps in knowledge

DMSP and GB are major compatible solutes within marine environments, with reported protective roles against stresses such as salinity, temperature, pressure and oxidative stress (Csonka, 1989; Dickson et al., 1982; Karsten et al., 1996; Mendum & Smith, 2002; Smiddy et al., 2004; Sunda et al., 2002; Zheng et al., 2020). Both compounds can also be used as a source of nutrients, with DMSP reported to support up to 13 % of bacterial carbon demand in surface waters (Kiene et al., 1999). Both compatible solutes are also implicated in climate change, with DMSP being the major precursor of the climate active gas, DMS which has a role in cloud formation and a reported climate cooling effect (Andreae, 1990). Conversely, one of the breakdown products of GB is trimethylamine (TMA), which is involved in formation of the greenhouse gas methane (King, 1984). In fact, the degradation of GB and TMA are responsible for 90 % of methane emissions in marine coastal sediments (Jones et al., 2019; Oremland et al., 1982).

There are several environmental factors shown to regulate the intracellular accumulation of DMSP and GB. For example, high salinity has been shown to result in accumulation of both compatible solutes (Karsten et al., 1990) There is also a reported relationship between DMSP and GB production in some species of phytoplankton that produce both compatible solutes. In these cases, DMSP production is thought to be favoured over nitrogen-containing GB, in nitrogen deplete conditions (Andreae, 1986). However, there are some studies in which this relationship is disputed (Keller et al., 1999). Therefore, ratification of this proposed interplay between DMSP and GB synthesis remained a gap in the knowledge of the field.

Despite the widely reported importance and abundance of DMSP, its proposed role as an osmoprotectant has not been ratified in any organism via mutagenesis of a DMSP synthesis gene and subsequent phenotyping. GB is also a widely reported osmolyte and is added to cultures in lab settings to protect against freezing (Cleland et al., 2004). However, bacterial mutagenesis of a GB synthesis gene has not resulted in a reduced growth phenotype in high salinity.

Within this thesis, the abundant *Alphaproteobacterium L. aggregata* LZB033 was used as a model organism as it is known to produce both DMSP and GB. In the multi-step transamination DMSP synthesis pathway in *L. aggregata*, only one gene has been identified; *dsyB* encodes the methyltransferase that catalyses the third reaction of the pathway - the methylation of MTHB to DMSHB (Curson et al., 2017). However, the genes responsible for the catalysation of the other steps in the pathway have not yet been identified in any organism.

Therefore, in this thesis, there were three major aims which intended to better understand the aforementioned gaps in our knowledge of the relationship between DMSP and GB. These aims were designed to allow the study of the interplay between DMSP and GB regulation in *L. aggregata.* This will include determining whether the proposed nitrogen-dependent relationship between DMSP and GB exists in *L. aggregata.* The aims were also designed to allow the confirmation of the so far 'proposed' osmoprotection role of these compounds in *L. aggregata*.

As such, the main aims of this thesis were:

1. To determine how DMSP and GB synthesis genes are regulated by differing environmental conditions in *L. aggregata* LZB033.
2. To use mutagenesis to determine whether DMSP and GB have a role in osmoprotection in *L. aggregata* LZB033.
3. To identify candidate DMSP synthesis genes responsible for catalysing the remaining uncharacterized steps in the transamination synthesis pathway in *L. aggregata* LZB033.

## 6.2 Major findings described in this thesis

### 6.2.1 DMSP and GB are two of the major compatible solutes in *L. aggregata*

Quantification of the intracellular concentration of all nitrogen and sulfur containing zwitterionic metabolites in *L. aggregata* determined that only DMSP, GB and sarcosine were present in pmol concentrations, whereas the remaining detected compounds were present at fmol concentrations (**Figure 3.8**). This led to the conclusion that DMSP, GB and sarcosine are the major compatible solutes at play in *L. aggregata*. This is the first time that a screening of compatible solutes has been reported in the model DSMP producer, *L. aggregata*. However, it is important to note that known compatible solutes such as sucrose, spermidine and putrescine were not measured within this study, and therefore cannot be excluded as potentially important compatible solutes in *L. aggregata.*

The regulation of expression of *dsyB* and the known GB synthesis genes; *betA, betB* and *gsdmt* were studied under differing environmental conditions using RNA-Seq (**Figure 3.13**). Under low salinity, the three most downregulated genes included both *dsyB* and *gsdmt,* further indicating the importance of DMSP and GB in this organism, in particular in salinity stress response. Therefore, both transcriptomic and metabolomic approaches highlighted the importance of DMSP and GB production in *L. aggregata*, especially in response to salinity.

The RNA-Seq experiment also allowed investigation of the proposed nitrogen-regulated inverse relationship between DMSP and GB production. *dsyB* was significantly upregulated by low nitrogen conditions, as was DMSP production. This data correlates with a study previously carried out by Curson and colleagues (Curson et al., 2017). However, the GB production of *L. aggregata* was unexpectedly found to increase under low nitrogen conditions (**Figure 3.8**). Therefore, this finding does not support a relationship between DMSP and GB that is regulated by nitrogen availability. However, it is possible that this relationship would exist if cultures were grown under nitrogen deplete conditions, as the cells may be forced to decrease GB production and favour DMSP, in order to utilise the severely limited nitrogen for other essential cell processes.

### 6.2.2 *gsdmt* is the key GB synthesis gene in *L. aggregata*

Initial indications that *gsdmt* was the key GB synthesis gene were revealed in the RNA-Seq dataset. *gsdmt* was significantly differentially expressed in all of the conditions that caused significant changes in GB production in *L. aggregata.* This was in contrast to the genes involved in the choline-GB pathway, *betA* and *betB,* which were not significantly regulated by salinity or nitrogen concentration. Further, the normalised count data for *gsdmt* was approximately 30-fold higher than that of the *bet* genes. Mutagenesis of *gsdmt* resulted in abolished GB production in *L. aggregata,* and as such it was confirmed that *gsdmt* is the key GB synthesis gene (**Figure 4.3**). This represents the first report of *gsdmt* as the major GB synthesis gene in *L. aggregata* and therefore indicates that the glycine-GB pathway is favoured in this strain.

### 6.2.3 A confirmed osmoprotection role for GB in *L. aggregata*

In previous studies, it was found that an *L. aggregata* Δ*dsyB* mutant had completely abolished DMSP synthesis but did not have reduced growth in higher salinities compared to the WT strain (Curson et al., 2017). This experiment was repeated in this thesis and the same outcome was observed.

Within this thesis, the key GB synthesis gene in *L. aggregata* was knocked out to create the Δ*gsdmt* mutant, which had abolished GB synthesis. Growth curves revealed that the Δ*gsdmt* mutant had a far reduced growth rate compared to the WT strain when grown in 50 PSU conditions (**Figure 4.7**). Further, the growth rate was restored through the exogenous addition of GB to the medium. As such, this provided proof for the osmoprotective role of GB in *L. aggregata* for the first time.

When both DMSP and GB production were knocked out in the *L. aggregata* Δ*dsyB*Δ*gsdmt* double mutant strain, the growth rate of this strain was similar to that of the Δ*gsdmt* mutant (**Figure 4.7**). When the *L. aggregata* Δ*dsyB*Δ*gsdmt* double mutant was complemented with a cloned copy of the *dsyB* gene, the growth rate was restored to a similar level to the Δ*gsdmt* mutant (**Figure 4.8**). Therefore, this may indicate that that DMSP

does have a role in osmoprotection in *L. aggregata.* Additionally, it is possible that a reduced growth phenotype is not observed in the *L. aggregata* Δ*dsyB* strain because the loss of DMSP production is compensated for by the production of GB.

It is interesting to note that the GB production of the *L. aggregata* Δ*dsyB* mutant was not significantly increased compared to the WT (**Figure 4.4**). This suggests that the loss of DMSP is not being compensated for through increased GB synthesis. Therefore, it is possible that the apparent superior osmoprotection provided for by GB is perhaps due to the differing structural characteristics of the compatible solutes, rather than being concentration dependent.

### 6.2.4 Identification of candidate DMSP synthesis genes and a potential compatible solute regulator.

#### 6.2.4.1 Candidate DMSP synthesis genes

DMSP production was decreased in low salinity and increased in low nitrogen conditions. From the RNA-Seq dataset, two aminotransferases were identified that were upregulated by low nitrogen conditions and downregulated by low salinity. This pattern of regulation is as expected for DMSP synthesis genes and was observed for *dsyB.* Therefore, these two aminotransferases, termed *aro8* and *at2* were investigated as potential methionine aminotransferases in the transamination synthesis pathway which may catalyse the conversion of Met to MTOB.

To further explore the role of these genes in DMSP synthesis, they were mutated in *L. aggregata*. The DMSP production of the generated *L. aggregata* Δ*aro8* and *L. aggregata* Δ*at2* mutant strains were found to be significantly lower than that of the WT. However, as these strains were still producing DMSP, it was hypothesised that the loss of one aminotransferase may have been compensated for by the presence of the other aminotransferase. As such, a double mutant was created in which both aminotransferase genes were knocked out. However, the DMSP production of this *L. aggregata* Δ*aro8*Δ*at2* strain was not significantly lower than the single mutant strains (**Figure 5.9**). It is possible that other aminotransferases in the *L. aggregata* cell are also able to catalyse the

methionine aminotransferase reaction in question, as microbial aminotransferases often have overlapping substrate specificity and can have involvement in more than one pathway (Jensen et al., 1981).

To confirm the role of the Aro8 and At2 aminotransferases in DMSP synthesis, there was an attempt to purify the proteins for use in enzyme assays. However, as the protein purification was not successful, this remains future work that would need to be completed to confirm if these aminotransferases are involved in DMSP synthesis.

#### 6.2.4.2 Candidate compatible solute regulator

Existing studies had identified the transcriptional regulator CosR as a repressor of compatible solute synthesis genes in *Vibrio cholerae*, and was reported to be widespread in *Alphaproteobacteria*, *Betaproteobacteria*, and *Gammaproteobacteria* (Gregory et al., 2020). A CosR homologue was identified in *L. aggregata* (with 44 % amino acid identity to the characterised CosR sequence). The putative *L. aggregata cosR* gene was downregulated in the low salinity condition, as was the case in *V. cholerae.* To determine whether CosR is a regulator of DMSP synthesis, the gene was mutated. The *L. aggregata* Δ*cosR* strain produced slightly less DMSP than the WT strain (**Figure 5.16**). This was unexpected as in other organisms, CosR was found to be a repressor of compatible solute synthesis genes. Therefore, further work is needed to investigate whether CosR is a regulator of DMSP synthesis genes in *L. aggregata.*

## 6.3 Recommendations for future research

The results presented in this thesis have provided confirmation of DMSP and GB as the major compatible solutes in *L. aggregata* and confirmed the osmoprotection role of GB. The increased understanding of the regulation and production of the two solutes has led to more research questions to be answered. Also identified in this thesis are several candidate genes possibly involved in DMSP synthesis and compatible solute regulation. However, further work is required to confirm the predicted roles of these genes.

### 6.3.1 Further research on the relationship between DMSP and GB production

One finding of this thesis was the indication of the role of DMSP in osmoprotection. It is possible that the lack of a reduced growth phenotype in *L. aggregata* Δ*dsyB* could be explained by the continued production of GB in this strain. This compensation may also explain the lack of an osmoprotection phenotype in other DMSP synthesis mutants. For example, *T. profundimis* is known to produce GB as well as using the methylation pathway to synthesise DMSP (Williams et al., 2019). However, a *T. profundimis* Δ*mmtN* mutant (the key gene in the methylation DMSP synthesis pathway), with completely abolished DMSP synthesis did not have a phenotype when grown in high salinity (Williams et al., 2019). Therefore, it would be intriguing to mutate the GB synthesis genes in *T. profundimis* and in the *T. profundimis* Δ*mmtN* mutant to determine if the loss of DMSP is compensated for by GB production in other species. This work could also be widened to other organisms which produce both DMSP and GB (or DMSP and other compatible solutes, such as ectoine) and in which the synthesis genes for these compatible solutes have been identified. This could therefore provide evidence for the osmoprotection role of DMSP in a wide range of organisms and would thus provide more information on how these organisms respond to salinity stress.

Additionally, further phenotyping of the *L. aggregata* Δ*dsyB*Δ*gsmdt* double mutant could also be performed to confirm other stress protection roles of DMSP that so far have only been proposed from accumulation data. For example, the growth rate of the *L. aggregata* Δ*dsyB*Δ*gsmdt* strain could be compared to that of the WT and the Δ*dsyB and* Δ*gsmdt* single mutants under varying environmental conditions such as high pressure, low temperature, and under oxidative stress to determine if DMSP and GB have a role in protection against these conditions. Again, this work could be repeated in other organisms in which both the key DMSP and GB compatible solute synthesis genes are known. This would allow the determination of whether any other confirmed stress protection roles of these compounds are widespread.

### 6.3.2 *gsdmt* as a marker gene for GB synthesis

The key GB synthesis gene in *L. aggregata* was found to be *gsdmt* and homologues of more than 70 % amino acid sequence identity were found in ~80 members of Alphaproteobacteria. The high number of *gsdmt* homologues could indicate that *gsdmt* is the key GB synthesis gene in many organisms. Mutagenesis of a selection of the *gsdmt* homologues could determine whether *gsdmt* is involved in GB synthesis in these organisms, and whether it is the key gene in this process. Confirmation of this could allow *gsdmt* to be used as a marker gene for the ability to produce GB. Therefore, *gsdmt* could be searched for in various metagenomes to determine the abundance of this gene in different environments. It is possible this could lead to a change in GB production models, and perhaps could have an influence on climate change models as one of the breakdown products of GB is TMA, which is involved in the production of the major greenhouse gas, methane.

### 6.3.3 Further research on candidate DMSP synthesis genes and CosR

Further work is needed to confirm whether *at2* and *aro8* are DMSP synthesis genes. This work would require enzyme assays to be performed and has been more extensively discussed in Section 5.3. However, if these genes are indeed validated as encoding the methionine aminotransferase responsible for catalysing the first step of the DMSP transamination synthesis pathway, homologues of these genes could be searched for (as was carried out for *gsdmt*). The ratification of novel transamination DMSP synthesis genes could serve as new marker genes for the ability to produce DMSP.

If *aro8* and *at2* are ratified as DMSP synthesis genes, it is possible that at least one more aminotransferase-encoding gene is also involved in DMSP synthesis in *L. aggregata,* as the Δ*aro8*Δ*at2* strain did not have completely abolished DMSP synthesis. In several organisms which use the methionine salvage pathway, the gene encoding the aminotransferase enzyme that catalyses the final MTOB to Met reaction has been identified (Berger et al., 2003; Heilbronn et al., 1999). As aminotransferase reactions are reversible, it is likely that these genes can also catalyse the transamination reaction of Met to MTOB (which is also

the first step of the DMSP transamination synthesis pathway). The *L. aggregata* genome was searched for homologues to the known MTOB aminotransferases and hits of above 30 % amino acid identity were found to the MTOB aminotransferase genes of *B. subtilis* (*ykrV* and *ybgE*) and *K. pneumoniae* (*tyrB*) (Berger et al., 2003; Heilbronn et al., 1999). It is possible that the identified homologues may be involved in DMSP synthesis in *L. aggregata.* As such, these genes could be tested as candidate DMSP synthesis genes through mutagenesis, protein purification and enzyme assays.

CosR was also investigated as a potential global compatible solute regulator in *L. aggregata*. However, the Δ*cosR* mutant strain did not have unregulated DMSP production. This data suggests that DMSP synthesis is tightly regulated by more than one transcriptional regulator, as might be expected for such an important compatible solute. Samples of the WT and Δ*cosR* mutants grown in differing salinities have been stored for future LC-MS analysis, which will determine the intracellular concentration of all of the nitrogen and sulfur containing osmolytes in each of the strains. This will indicate the extent of global regulation of compatible solutes by CosR in *L. aggregata*.

A high quantity of genes annotated as transcriptional regulators were found to be differentially expressed under the conditions shown to regulate the synthesis of DMSP and other compatible solutes. As such, these genes could be investigated as encoding transcriptional regulators of compatible solute synthesis genes. Once a candidate gene is identified, mutagenesis of this gene and subsequent RNA-Seq experiment could be carried out on the WT and mutant strain grown in different salinities.

## 6.4 Concluding Remarks

In summary, this work represents the first confirmation in any organism of the role of DMSP as an osmolyte via mutagenesis of a synthesis gene. This work has also determined that the key GB synthesis gene in *L. aggregata* is *gsdmt*. Further, several candidate DMSP synthesis genes and a potential compatible solute regulator have been identified. Together, these findings contribute to our knowledge of the interplay between the two abundant and climatically relevant compatible solutes, DMSP and GB.

# Chapter 7: References


Aktas, M., Jost, K. A., Fritz, C., & Narberhaus, F. (2011). Choline uptake in Agrobacterium tumefaciens by the high-affinity ChoXWV transporter. *Journal of bacteriology*, *193*(19), 5119-5129.

Albers, E. (2009). Metabolic characteristics and importance of the universal methionine salvage pathway recycling methionine from 5'-methylthioadenosine. *IUBMB life*, *61*(12), 1132-1142.

Alcolombri, U., Ben-Dor, S., Feldmesser, E., Levin, Y., Tawfik, D. S., & Vardi, A. (2015). Identification of the algal dimethyl sulfide–releasing enzyme: A missing link in the marine sulfur cycle. *Science*, *348*(6242), 1466-1469.

Andreae, M. O. (1986). The ocean as a source of atmospheric sulfur compounds. In *The role of air-sea exchange in geochemical cycling* (pp. 331-362). Springer.

Andreae, M. O. (1990). Ocean-atmosphere interactions in the global biogeochemical sulfur cycle. *Marine Chemistry*, *30*, 1-29.

Andreae, M. O., Elbert, W., Cai, Y., Andreae, T. W., & Gras, J. (1999). Non-sea-salt sulfate, methanesulfonate, and nitrate aerosol concentrations and size distributions at Cape Grim, Tasmania. *Journal of Geophysical Research: Atmospheres*, *104*(D17), 21695-21706.

Andreae, M. O., & Raemdonck, H. (1983). Dimethyl sulfide in the surface ocean and the marine atmosphere: a global view. *Science*, *221*(4612), 744-747.

Andreesen, J. R. (1994). Glycine metabolism in anaerobes. *Antonie Van Leeuwenhoek*, *66*(1), 223-237.

Arakawa, T., & Timasheff, S. N. (1985). The stabilization of proteins by osmolytes. *Biophys J*, *47*(3), 411-414. https://doi.org/10.1016/s0006-3495(85)83932-1

Aziz, R. K., Bartels, D., Best, A. A., DeJongh, M., Disz, T., Edwards, R. A., Formsma, K., Gerdes, S., Glass, E. M., Kubal, M., Meyer, F., Olsen, G. J., Olson, R., Osterman, A. L., Overbeek, R. A., McNeil, L. K., Paarmann, D., Paczian, T., Parrello, B., . . . Zagnitko, O. (2008). The RAST Server: rapid annotations using subsystems technology. *BMC Genomics*, *9*, 75. https://doi.org/10.1186/1471-2164-9-75

Azizah, M., & Pohnert, G. (2022). Orchestrated Response of Intracellular Zwitterionic Metabolites in Stress Adaptation of the Halophilic Heterotrophic Bacterium Pelagibaca bermudensis. *Marine Drugs*, *20*(11), 727.

Barnard, W. R., Andreae, M. O., & Iverson, R. L. (1984). Dimethylsulfide and Phaeocystis poucheti in the southeastern Bering Sea. *Continental Shelf Research*, *3*(2), 103-113.

Barnes, I., Hjorth, J., & Mihalopoulos, N. (2006). Dimethyl sulfide and dimethyl sulfoxide and their oxidation in the atmosphere. *Chemical reviews*, *106*(3), 940-975.

Barra, L., Fontenelle, C., Ermel, G., Trautwetter, A., Walker, G. C., & Blanco, C. (2006). Interrelations between glycine betaine catabolism and methionine biosynthesis in Sinorhizobium meliloti strain 102F34. *Journal of bacteriology*, *188*(20), 7195-7204.

Bashir, A., Hoffmann, T., Smits, S. H., & Bremer, E. (2014). Dimethylglycine provides salt and temperature stress protection to Bacillus subtilis. *Applied and environmental microbiology*, *80*(9), 2773-2785.

Berger, B. J., English, S., Chan, G., & Knodel, M. H. (2003). Methionine regeneration and aminotransferases in Bacillus subtilis, Bacillus cereus, and Bacillus anthracis. *Journal of bacteriology*, *185*(8), 2418-2431.

Bernal, V., Sevilla, Á., Cánovas, M., & Iborra, J. L. (2007). Production of L-carnitine by secondary metabolism of bacteria. *Microbial Cell Factories*, *6*(1), 1-17.

Bhagavan, N. V., & Ha, C.-E. (2015). Chapter 15 - Protein and Amino Acid Metabolism. In N. V. Bhagavan & C.-E. Ha (Eds.), *Essentials of Medical Biochemistry (Second Edition)* (pp. 227-268). Academic Press. https://doi.org/https://doi.org/10.1016/B978-0-12-416687-5.00015-4

Biswas, I. (2015). Genetic tools for manipulating Acinetobacter baumannii genome: an overview. *Journal of medical microbiology*, *64*(7), 657-669.

Bohnert, H. J., & Shen, B. (1998). Transformation and compatible solutes. *Scientia Horticulturae*, *78*(1), 237-260. https://doi.org/https://doi.org/10.1016/S0304-4238(98)00195-2

Boysen, A. K., Durham, B. P., Kumler, W., Key, R. S., Heal, K. R., Carlson, L. T., Groussman, R. D., Armbrust, E. V., & Ingalls, A. E. (2022). Glycine betaine uptake and metabolism in marine microbial communities. *Environmental microbiology*, *24*(5), 2380-2403.

Bray, N. L., Pimentel, H., Melsted, P., & Pachter, L. (2016). Near-optimal probabilistic RNA-seq quantification. *Nature Biotechnology*, *34*(5), 525-527. https://doi.org/10.1038/nbt.3519

Breisch, J., Bendel, M., & Averhoff, B. (2022). The choline dehydrogenase BetA of Acinetobacter baumannii: a flavoprotein responsible for osmotic stress protection. *Environmental microbiology*, *24*(3), 1052-1061.

Brettin, T., Davis, J. J., Disz, T., Edwards, R. A., Gerdes, S., Olsen, G. J., Olson, R., Overbeek, R., Parrello, B., Pusch, G. D., Shukla, M., Thomason, J. A., Stevens, R., Vonstein, V., Wattam, A. R., & Xia, F. (2015). RASTtk: A modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes. *Scientific reports*, *5*(1), 8365. https://doi.org/10.1038/srep08365

Broy, S., Chen, C., Hoffmann, T., Brock, N. L., Nau-Wagner, G., Jebbar, M., Smits, S. H. J., Dickschat, J. S., & Bremer, E. (2015). Abiotic stress protection by ecologically abundant dimethylsulfoniopropionate and its natural and synthetic derivatives: insights from Bacillus subtilis. *Environmental microbiology*, *17*(7), 2362-2378. https://doi.org/https://doi.org/10.1111/1462-2920.12698

Burnet, M., Lafontaine, P. J., & Hanson, A. D. (1995). Assay, purification, and partial characterization of choline monooxygenase from spinach. *Plant Physiology*, *108*(2), 581-588.

Camacho, M., Redondo-Gómez, S., Rodríguez-Llorente, I., Rohde, M., Spröer, C., Schumann, P., Klenk, H.-P., & Montero-Calasanz, M. d. C. (2016). Labrenzia salina sp. nov., isolated from the rhizosphere of the halophyte Arthrocnemum macrostachyum. *International Journal of Systematic and Evolutionary Microbiology*, *66*(12), 5173-5180.

Cánovas, D., Vargas, C., Csonka, L. N., Ventosa, A., & Nieto, J. J. (1996). Osmoprotectants in Halomonas elongata: high-affinity betaine transport system and choline-betaine pathway. *Journal of bacteriology*, *178*(24), 7221-7226.

Casanueva, A., Tuffin, M., Cary, C., & Cowan, D. A. (2010). Molecular adaptations to psychrophily: the impact of 'omic'technologies. *Trends in microbiology*, *18*(8), 374-381.

Cavicchioli, R., Ripple, W. J., Timmis, K. N., Azam, F., Bakken, L. R., Baylis, M., Behrenfeld, M. J., Boetius, A., Boyd, P. W., Classen, A. T., Crowther, T. W., Danovaro, R., Foreman, C. M., Huisman, J., Hutchins, D. A., Jansson, J. K., Karl, D. M., Koskella, B., Mark Welch, D. B., . . . Webster, N. S. (2019). Scientists' warning to humanity: microorganisms and climate change. *Nature Reviews Microbiology*, *17*(9), 569-586. https://doi.org/10.1038/s41579-019-0222-5

Charlson, R. J., Lovelock, J. E., Andreae, M. O., & Warren, S. G. (1987). Oceanic phytoplankton, atmospheric sulphur, cloud albedo and climate. *Nature*, *326*(6114), 655-661.

Claud, E. C., Keegan, K. P., Brulc, J. M., Lu, L., Bartels, D., Glass, E., Chang, E. B., Meyer, F., & Antonopoulos, D. A. (2013). Bacterial community structure and functional contributions to emergence of health or necrotizing enterocolitis in preterm infants. *Microbiome*, *1*(1), 20. https://doi.org/10.1186/2049-2618-1-20

Cleland, D., Krader, P., McCree, C., Tang, J., & Emerson, D. (2004). Glycine betaine as a cryoprotectant for prokaryotes. *J Microbiol Methods*, *58*(1), 31-38. https://doi.org/10.1016/j.mimet.2004.02.015

Community, T. G. (2022). The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2022 update. *Nucleic acids research*, *50*(W1), W345-W351. https://doi.org/10.1093/nar/gkac247

Cosquer, A., Pichereau, V., Pocard, J.-A., Minet, J., Cormier, M., & Bernard, T. (1999). Nanomolar levels of dimethylsulfoniopropionate, dimethylsulfonioacetate, and glycine betaine are sufficient to confer osmoprotection to Escherichia coli. *Applied and environmental microbiology*, *65*(8), 3304-3311.

Croucher, N. J., & Thomson, N. R. (2010). Studying bacterial transcriptomes using RNA-seq. *Current Opinion in Microbiology*, *13*(5), 619-624. https://doi.org/https://doi.org/10.1016/j.mib.2010.09.009

Csonka, L. N. (1989). Physiological and genetic responses of bacteria to osmotic stress. *Microbiological reviews*, *53*(1), 121-147.

Curson, A., Rogers, R., Todd, J., Brearley, C., & Johnston, A. (2008). Molecular genetic analysis of a dimethylsulfoniopropionate lyase that liberates the climate-changing gas dimethylsulfide in several marine α-proteobacteria and Rhodobacter sphaeroides. *Environmental microbiology*, *10*(3), 757-767.

Curson, A. R., Liu, J., Bermejo Martínez, A., Green, R. T., Chan, Y., Carrión, O., Williams, B. T., Zhang, S.-H., Yang, G.-P., & Bulman Page, P. C. (2017). Dimethylsulfoniopropionate biosynthesis in marine bacteria and identification of the key gene in this process. *Nature microbiology*, *2*(5), 1-9.

Curson, A. R., Todd, J. D., Sullivan, M. J., & Johnston, A. W. (2011). Catabolism of dimethylsulphoniopropionate: microorganisms, enzymes and genes. *Nature Reviews Microbiology*, *9*(12), 849-859.

Curson, A. R., Williams, B. T., Pinchbeck, B. J., Sims, L. P., Martínez, A. B., Rivera, P. P. L., Kumaresan, D., Mercadé, E., Spurgin, L. G., & Carrión, O. (2018). DSYB catalyses the key step of dimethylsulfoniopropionate biosynthesis in many phytoplankton. *Nature microbiology*, *3*(4), 430-439.

Czech, L., Hermann, L., Stöveken, N., Richter, A. A., Höppner, A., Smits, S. H., Heider, J., & Bremer, E. (2018). Role of the extremolytes ectoine and hydroxyectoine as stress protectants and nutrients: genetics, phylogenomics, biochemistry, and structural analysis. *Genes*, *9*(4), 177.

Dang, H., Klotz, M. G., Lovell, C. R., & Sievert, S. M. (2019). The responses of marine microorganisms, communities and ecofunctions to environmental gradients. In (Vol. 10, pp. 115): Frontiers Media SA.

De Bont, J., Van Dijken, J., & Harder, W. (1981). Dimethyl sulphoxide and dimethyl sulphide as a carbon, sulphur and energy source for growth of Hyphomicrobium S. *Microbiology*, *127*(2), 315-323.

DeBose, J. L., & Nevitt, G. A. (2008). The use of odors at different spatial scales: comparing birds with fish. *Journal of chemical ecology*, *34*(7), 867-881.

Deed, R. C., Hou, R., Kinzurik, M. I., Gardner, R. C., & Fedrizzi, B. (2018). The role of yeast ARO8, ARO9 and ARO10 genes in the biosynthesis of 3-(methylthio)-1-propanol from L-methionine during fermentation in synthetic grape medium. *FEMS Yeast Research*, *19*(2). https://doi.org/10.1093/femsyr/foy109

Dennis, J. J., & Zylstra, G. J. (1998). Plasposons: modular self-cloning minitransposon derivatives for rapid genetic analysis of gram-negative bacterial genomes. *Applied and environmental microbiology*, *64*(7), 2710-2715.

Dickschat, J. S., Rabe, P., & Citron, C. A. (2015). The chemical biology of dimethylsulfoniopropionate. *Organic & Biomolecular Chemistry*, *13*(7), 1954-1968.

Dickson, D., Jones, R., & Davenport, J. (1980). Steady state osmotic adaptation inUlva lactuca. *Planta*, *150*(2), 158-165.

Dickson, D., & Kirst, G. (1986). The role of β-dimethylsulphoniopropionate, glycine betaine and homarine in the osmoacclimation of Platymonas subcordiformis. *Planta*, *167*(4), 536-543.

Dickson, D., Wyn Jones, R., & Davenport, J. (1982). Osmotic adaptation in Ulva lactuca under fluctuating salinity regimes. *Planta*, *155*(5), 409-415.

Dragolovich, J. (1994). Dealing with salt stress in animal cells: The role and regulation of glycine betaine concentrations. *Journal of Experimental Zoology*, *268*(2), 139-144. https://doi.org/https://doi.org/10.1002/jez.1402680211

Empadinhas, N., & da Costa, M. S. (2008). Osmoadaptation mechanisms in prokaryotes: distribution of compatible solutes. *Int Microbiol*, *11*(3), 151-161.

Emrich, S. J., Barbazuk, W. B., Li, L., & Schnable, P. S. (2007). Gene discovery and annotation using LCM-454 transcriptome sequencing. *Genome research*, *17*(1), 69-73.

Everaert, C., Luypaert, M., Maag, J. L. V., Cheng, Q. X., Dinger, M. E., Hellemans, J., & Mestdagh, P. (2017). Benchmarking of RNA-sequencing analysis workflows using whole-transcriptome RT-qPCR expression data. *Scientific reports*, *7*(1), 1559. https://doi.org/10.1038/s41598-017-01617-3

Falkenberg, P., & Strøm, A. R. (1990). Purification and characterization of osmoregulatory betaine aldehyde dehydrogenase of Escherichiacoli. *Biochimica et Biophysica Acta (BBA)-General Subjects*, *1034*(3), 253-259.

Fariduddin, Q., Varshney, P., Yusuf, M., Ali, A., & Ahmad, A. (2013). Dissecting the Role of Glycine Betaine in Plants under Abiotic Stress.

Fenizia, S., Thume, K., Wirgenings, M., & Pohnert, G. (2020). Ectoine from bacterial and algal origin is a compatible solute in microalgae. *Marine Drugs*, *18*(1), 42.

Figueroa-Soto, C. G., & Valenzuela-Soto, E. M. (2018). Glycine betaine rather than acting only as an osmolyte also plays a role as regulator in cellular metabolism. *Biochimie*, *147*, 89-97.

Figurski, D. H., & Helinski, D. R. (1979). Replication of an origin-containing derivative of plasmid RK2 dependent on a plasmid function provided in trans. *Proceedings of the National Academy of Sciences*, *76*(4), 1648-1652.

Fitzsimmons, L. F., Flemer Jr, S., Wurthmann, A. S., Deker, P. B., Sarkar, I. N., & Wargo, M. J. (2011). Small-molecule inhibition of choline catabolism in Pseudomonas aeruginosa and other aerobic choline-catabolizing bacteria. *Applied and environmental microbiology*, *77*(13), 4383-4389.

Gage, D. A., Rhodes, D., Nolte, K. D., Hicks, W. A., Leustek, T., Cooper, A. J., & Hanson, A. D. (1997). A new route for synthesis of dimethylsulphoniopropionate in marine algae. *Nature*, *387*(6636), 891-894.

Galinski, E. A. (1993). Compatible solutes of halophilic eubacteria: molecular principles, water-solute interaction, stress protection. *Experientia*, *49*(6), 487-496. https://doi.org/10.1007/BF01955150

Gebser, B., & Pohnert, G. (2013). Synchronized regulation of different zwitterionic metabolites in the osmoadaption of phytoplankton. *Marine Drugs*, *11*(6), 2168-2182.

González, J. M., Kiene, R. P., & Moran, M. A. (1999). Transformation of sulfur compounds by an abundant lineage of marine bacteria in the α-subclass of the class Proteobacteria. *Applied and environmental microbiology*, *65*(9), 3810-3819.

González, J. M., Whitman, W. B., Hodson, R. E., & Moran, M. A. (1996). Identifying numerically abundant culturable bacteria from complex communities: an example from a lignin enrichment culture. *Applied and environmental microbiology*, *62*(12), 4433-4440.

Gowrishankar, J. (1989). Nucleotide sequence of the osmoregulatory proU operon of Escherichia coli. *Journal of bacteriology*, *171*(4), 1923-1931.

Gregory, G. J., Morreale, D. P., & Boyd, E. F. (2020). CosR is a global regulator of the osmotic stress response with widespread distribution among bacteria. *Applied and environmental microbiology*, *86*(10), e00120-00120.

Gregory, G. J., Morreale, D. P., Carpenter, M. R., Kalburge, S. S., & Boyd, E. F. (2019). Quorum sensing regulators AphA and OpaR control expression of the operon responsible for biosynthesis of the compatible solute ectoine. *Applied and environmental microbiology*, *85*(22), e01543-01519.

Gröne, T., & Kirst, G. (1992). The effect of nitrogen deficiency, methionine and inhibitors of methionine metabolism on the DMSP contents of Tetraselmis subcordiformis (Stein). *Marine Biology*, *112*(3), 497-503.

Hai, Y., Huang, A. M., & Tang, Y. (2019). Structure-guided function discovery of an NRPS-like glycine betaine reductase for choline biosynthesis in fungi. *Proceedings of the National Academy of Sciences*, *116*(21), 10348-10353.

Hanson, A., & Gage, D. (1996). 3-Dimethylsulfoniopionate biosynthesis and use by flowering plants. In *Biological and environmental chemistry of DMSP and related sulfonium compounds* (pp. 75-86). Springer.

Hanson, A. D., Rivoal, J., Paquet, L., & Gage, D. A. (1994). Biosynthesis of 3-dimethylsulfoniopropionate in Wollastonia biflora (L.) DC.(evidence that S-methylmethionine is an intermediate). *Plant Physiology*, *105*(1), 103-110.

Hayashi, H., & Murata, N. (1998). Genetically engineered enhancement of salt tolerance in higher plants. *Stress response of photosynthetic organisms: Molecular mechanisms and molecular regulation*, 133-148.

Heilbronn, J., Wilson, J., & Berger, B. J. (1999). Tyrosine aminotransferase catalyzes the final step of methionine recycling in Klebsiella pneumoniae. *Journal of bacteriology*, *181*(6), 1739-1747.

Howard, E. C., Henriksen, J. R., Buchan, A., Reisch, C. R., Bürgmann, H., Welsh, R., Ye, W., González, J. M., Mace, K., & Joye, S. B. (2006). Bacterial taxa that limit sulfur flux from the ocean. *Science*, *314*(5799), 649-652.

Howard, E. C., Sun, S., Biers, E. J., & Moran, M. A. (2008). Abundant and diverse bacteria involved in DMSP degradation in marine surface waters. *Environmental microbiology*, *10*(9), 2397-2410.

Ikuta, S., IMAMURA, S., MISAKI, H., & HORIUTI, Y. (1977). Purification and characterization of choline oxidase from Arthrobacter globiformis. *The Journal of Biochemistry*, *82*(6), 1741-1749.

Jäger, W., Schäfer, A., Pühler, A., Labes, G., & Wohlleben, W. (1992). Expression of the Bacillus subtilis sacB gene leads to sucrose sensitivity in the gram-positive bacterium Corynebacterium glutamicum but not in Streptomyces lividans. *Journal of bacteriology*, *174*(16), 5462-5465.

James, F., Nolte, K. D., & Hanson, A. D. (1995). Purification and Properties of S-Adenosyl-L-methionine: L-Methionine S-Methyltransferase from Wollastonia biflora Leaves (∗). *Journal of Biological Chemistry*, *270*(38), 22344-22350.

Jameson, E., Doxey, A. C., Airs, R., Purdy, K. J., Murrell, J. C., & Chen, Y. (2016). Metagenomic data-mining reveals contrasting microbial populations responsible for trimethylamine formation in human gut and marine ecosystems. *Microbial genomics*, *2*(9).

Jamieson, A. J., Fujii, T., Mayor, D. J., Solan, M., & Priede, I. G. (2010). Hadal trenches: the ecology of the deepest places on Earth. *Trends in Ecology & Evolution*, *25*(3), 190-197.

Jenkins, W. (1985). Kinetics, equilibria, and affinity for coenzymes and substrates. *Transaminases*, 216-234.

Jensen, R. A., Calhoun, D. H., & Twarog, R. (1981). Intracellular roles of microbial aminotransferases: overlap enzymes across different biochemical pathways. *CRC Critical Reviews in Microbiology*, *8*(3), 229-266.

Jones, H. J., Kröber, E., Stephenson, J., Mausz, M. A., Jameson, E., Millard, A., Purdy, K. J., & Chen, Y. (2019). A new family of uncultivated bacteria involved in methanogenesis from the ubiquitous osmolyte glycine betaine in coastal saltmarsh sediments. *Microbiome*, *7*(1), 1-11.

Kageyama, H., Tanaka, Y., Shibata, A., Waditee-Sirisattha, R., & Takabe, T. (2018). Dimethylsulfoniopropionate biosynthesis in a diatom Thalassiosira pseudonana: Identification of a gene encoding MTHB-methyltransferase. *Archives of biochemistry and biophysics*, *645*, 100-106.

Kane, M. D., Jatkoe, T. A., Stumpf, C. R., Lu, J., Thomas, J. D., & Madore, S. J. (2000). Assessment of the sensitivity and specificity of oligonucleotide (50mer) microarrays. *Nucleic acids research*, *28*(22), 4552-4557.

Kappes, R. M., Kempf, B., & Bremer, E. (1996). Three transport systems for the osmoprotectant glycine betaine operate in Bacillus subtilis: characterization of OpuD. *Journal of bacteriology*, *178*(17), 5071-5079.

Kapust, R. B., & Waugh, D. S. (1999). Escherichia coli maltose-binding protein is uncommonly effective at promoting the solubility of polypeptides to which it is fused. *Protein science*, *8*(8), 1668-1674.

Karsten, U., Kück, K., Vogt, C., & Kirst, G. (1996). Dimethylsulfoniopropionate production in phototrophic organisms and its physiological functions as a cryoprotectant. In *Biological and environmental chemistry of DMSP and related sulfonium compounds* (pp. 143-153). Springer.

Karsten, U., Wiencke, C., & Kirst, G. (1990). The ß-dimethylsulphoniopropionate (DMSP) content of macroalgae from Antarctica and Southern Chile.

Keen, N., Tamaki, S., Kobayashi, D., & Trollinger, D. (1988). Improved broad-host-range plasmids for DNA cloning in gram-negative bacteria. *Gene*, *70*(1), 191-197.

Keller, M., Kiene, R., Matrai, P., & Bellows, W. (1999). Production of glycine betaine and dimethylsulfoniopropionate in marine phytoplankton. II. N-limited chemostat cultures. *Marine Biology*, *135*(2), 249-257.

Keller, M. D., & Korjeff-Bellows, W. (1996). Physiological aspects of the production of dimeyhtlsulfoniopropionate (DMSP) by marine phytoplankton. In *Biological and environmental chemistry of DMSP and related sulfonium compounds* (pp. 131-142). Springer.

Keller, M. D., Matrai, P. A., Kiene, R. P., & Bellows, W. K. (2004). Responses of coastal phytoplankton populations to nitrogen additions: dynamics of cell-associated dimethylsulfoniopropionate (DMSP), glycine betaine (GBT), and homarine. *Canadian Journal of Fisheries and Aquatic Sciences*, *61*(5), 685-699.

Kellogg, W. W., Cadle, R., Allen, E., Lazrus, A., & Martell, E. (1972). The Sulfur Cycle: Man's contributions are compared to natural sources of sulfur compounds in the atmosphere and oceans. *Science*, *175*(4022), 587-596.

Kempf, B., & Bremer, E. (1995). OpuA, an Osmotically Regulated Binding Protein-dependent Transport System for the Osmoprotectant Glycine Betaine in Bacillus subtilis (∗). *Journal of Biological Chemistry*, *270*(28), 16701-16713.

Kempf, B., & Bremer, E. (1998). Uptake and synthesis of compatible solutes as microbial stress responses to high-osmolality environments. *Arch Microbiol*, *170*(5), 319-330. https://doi.org/10.1007/s002030050649

Kettles, N. L., Kopriva, S., & Malin, G. (2014). Insights into the regulation of DMSP synthesis in the diatom Thalassiosira pseudonana through APR activity, proteomics and gene expression analyses on cells acclimating to changes in salinity, light and nitrogen. *PloS one*, *9*(4), e94795.

Kiene, R. P. (1996). Production of methanethiol from dimethylsulfoniopropionate in marine surface waters. *Marine Chemistry*, *54*(1-2), 69-83.

Kiene, R. P., Linn, L. J., & Bruton, J. A. (2000). New and important roles for DMSP in marine microbial communities. *Journal of Sea Research*, *43*(3-4), 209-224.

Kiene, R. P., Linn, L. J., González, J., Moran, M. A., & Bruton, J. A. (1999). Dimethylsulfoniopropionate and methanethiol are important precursors of methionine and protein-sulfur in marine bacterioplankton. *Applied and environmental microbiology*, *65*(10), 4549-4558.

Kimura, Y., Kawasaki, S., Yoshimoto, H., & Takegawa, K. (2010). Glycine betaine biosynthesized from glycine provides an osmolyte for cell growth and spore germination during osmotic stress in Myxococcus xanthus. *Journal of bacteriology*, *192*(5), 1467-1470.

King, G. M. (1984). Metabolism of trimethylamine, choline, and glycine betaine by sulfate-reducing and methanogenic bacteria in marine sediments. *Applied and environmental microbiology*, *48*(4), 719-725.

Kirkwood, M., Le Brun, N. E., Todd, J. D., & Johnston, A. W. (2010). The dddP gene of Roseovarius nubinhibens encodes a novel lyase that cleaves dimethylsulfoniopropionate into acrylate plus dimethyl sulfide. *Microbiology*, *156*(6), 1900-1906.

Kitaguchi, H., Uchida, A., & Ishida, Y. (1999). Purification and characterization of L-methionine decarboxylase from Crypthecodinium cohnii. *Fisheries science*, *65*(4), 613-617.

Ko, R., Smith, L. T., & Smith, G. M. (1994). Glycine betaine confers enhanced osmotolerance and cryotolerance on Listeria monocytogenes. *Journal of bacteriology*, *176*(2), 426-431.

Kocsis, M. G., & Hanson, A. D. (2000). Biochemical evidence for two novel enzymes in the biosynthesis of 3-dimethylsulfoniopropionate in Spartina alterniflora. *Plant Physiology*, *123*(3), 1153-1162.

Kocsis, M. G., Nolte, K. D., Rhodes, D., Shen, T.-L., Gage, D. A., & Hanson, A. D. (1998). Dimethylsulfoniopropionate Biosynthesis in Spartina alterniflora 1: Evidence That S-Methylmethionine and Dimethylsulfoniopropylamine Are Intermediates. *Plant Physiology*, *117*(1), 273-281.

Kolp, S., Pietsch, M., Galinski, E. A., & Gütschow, M. (2006). Compatible solutes as protectants for zymogens against proteolysis. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics*, *1764*(7), 1234-1242. https://doi.org/https://doi.org/10.1016/j.bbapap.2006.04.015

Kortstee, G. J. J. (1970). The aerobic decomposition of choline by microorganisms. *Archiv für Mikrobiologie*, *71*(3), 235-244. https://doi.org/10.1007/BF00410157

Ksionzek, K. B., Lechtenfeld, O. J., McCallister, S. L., Schmitt-Kopplin, P., Geuer, J. K., Geibert, W., & Koch, B. P. (2016). Dissolved organic sulfur in the ocean: Biogeochemistry of a petagram inventory. *Science*, *354*(6311), 456-459.

Lahham, M., Jha, S., Goj, D., Macheroux, P., & Wallner, S. (2021). The family of sarcosine oxidases: Same reaction, different products. *Archives of biochemistry and biophysics*, *704*, 108868.

Lai, M.-C., Wang, C.-C., Chuang, M.-J., Wu, Y.-C., & Lee, Y.-C. (2006). Effects of substrate and potassium on the betaine-synthesizing enzyme glycine sarcosine dimethylglycine N-methyltransferase from a halophilic methanoarchaeon Methanohalophilus portucalensis. *Research in microbiology*, *157*(10), 948-955.

Lai, S.-J., & Lai, M.-C. (2011). Characterization and regulation of the osmolyte betaine synthesizing enzymes GSMT and SDMT from halophilic methanogen Methanohalophilus portucalensis. *PloS one*, *6*(9), e25090.

Lambou, K., Pennati, A., Valsecchi, I., Tada, R., Sherman, S., Sato, H., Beau, R., Gadda, G., & Latgé, J.-P. (2013). Pathway of glycine betaine biosynthesis in Aspergillus fumigatus. *Eukaryotic cell*, *12*(6), 853-863.

Lana, A., Bell, T., Simó, R., Vallina, S., Ballabrera-Poy, J., Kettle, A., Dachs, J., Bopp, L., Saltzman, E., & Stefels, J. (2011). An updated climatology of surface dimethlysulfide concentrations and emission fluxes in the global ocean. *Global Biogeochemical Cycles*, *25*(1).

Landfald, B., & Strøm, A. R. (1986). Choline-glycine betaine pathway confers a high level of osmotic tolerance in Escherichia coli. *Journal of bacteriology*, *165*(3), 849-855.
Legrand, M., Delmas, R., & Charlson, R. (1988). Climate forcing implications from Vostok ice-core sulphate data. *Nature*, *334*(6181), 418-420.
Li, C.-Y., Wang, X.-J., Chen, X.-L., Sheng, Q., Zhang, S., Wang, P., Quareshy, M., Rihtman, B., Shao, X., & Gao, C. (2021). A novel ATP dependent dimethylsulfoniopropionate lyase in bacteria that releases dimethyl sulfide and acryloyl-CoA. *Elife*, *10*, e64045.
Li, C.-Y., Zhang, D., Chen, X.-L., Wang, P., Shi, W.-L., Li, P.-Y., Zhang, X.-Y., Qin, Q.-L., Todd, J. D., & Zhang, Y.-Z. (2017). Mechanistic insights into dimethylsulfoniopropionate lyase DddY, a new member of the cupin superfamily. *Journal of molecular biology*, *429*(24), 3850-3862.
Lidbury, I., Kimberley, G., Scanlan, D. J., Murrell, J. C., & Chen, Y. (2015). Comparative genomics and mutagenesis analyses of choline metabolism in the marine R oseobacter clade. *Environmental microbiology*, *17*(12), 5048-5062.
Lidbury, I. D., Murrell, J. C., & Chen, Y. (2015). Trimethylamine and trimethylamine N-oxide are supplementary energy sources for a marine heterotrophic bacterium: implications for marine carbon and nitrogen cycling. *The ISME journal*, *9*(3), 760-769.
Liss, P. S., Hatton, A. D., Malin, G., Nightingale, P. D., & Turner, S. M. (1997). Marine sulphur emissions. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *352*(1350), 159-169.
Lister, R., O'Malley, R. C., Tonti-Filippini, J., Gregory, B. D., Berry, C. C., Millar, A. H., & Ecker, J. R. (2008). Highly integrated single-base resolution maps of the epigenome in Arabidopsis. *Cell*, *133*(3), 523-536.
Little, R., Trottmann, F., Preissler, M., & Hertweck, C. (2022). An intramodular thioesterase domain catalyses chain release in the biosynthesis of a cytotoxic virulence factor. *RSC Chemical Biology*, *3*(9), 1121-1128.
Liu, L., Chen, X., Ye, J., Ma, X., Han, Y., He, Y. & Tang, K. (2023). Sulfoquinovose is a widespread organosulfur substrate for Roseobacter clade bacteria in the ocean. The ISME Journal, 17(3), 393-405.
Liu, J., Xue, C.-X., Wang, J., Crombie, A. T., Carrión, O., Johnston, A. W., Murrell, J. C., Liu, J., Zheng, Y., & Zhang, X.-H. (2022). Oceanospirillales containing the DMSP lyase DddD are key utilisers of carbon from DMSP in coastal seawater. *Microbiome*, *10*(1), 1-21.
Louis, P., & Galinski, E. A. (1997). Characterization of genes for the biosynthesis of the compatible solute ectoine from Marinococcus halophilus and osmoregulated expression in Escherichia coli. *Microbiology*, *143*(4), 1141-1149.
Lyon, B. R., & Mock, T. (2014). Polar microalgae: new approaches towards understanding adaptations to an extreme and changing environment. *Biology*, *3*(1), 56-80.
Ma, Y., Wang, Q., Gao, X., & Zhang, Y. (2017). Biosynthesis and uptake of glycine betaine as cold-stress response to low temperature in fish pathogen Vibrio anguillarum. *Journal of Microbiology*, *55*(1), 44-55.
Mendum, M. L., & Smith, L. T. (2002). Characterization of glycine betaine porter I from Listeria monocytogenes and its roles in salt and chill tolerance. *Appl Environ Microbiol*, *68*(2), 813-819. https://doi.org/10.1128/aem.68.2.813-819.2002
Meskys, R., Harris, R. J., Casaite, V., Basran, J., & Scrutton, N. S. (2001). Organization of the genes involved in dimethylglycine and sarcosine degradation in Arthrobacter

spp. Implications for glycine betaine catabolism. *European Journal of Biochemistry*, *268*(12), 3390-3398.

Moran, M. A., Reisch, C. R., Kiene, R. P., & Whitman, W. B. (2012). Genomic insights into bacterial DMSP transformations. *Annual review of marine science*, *4*, 523-542.

Mortazavi, A., Williams, B. A., McCue, K., Schaeffer, L., & Wold, B. (2008). Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nature methods*, *5*(7), 621-628.

Munro, P. M., Gauthier, M. J., Breittmayer, V., & Bongiovanni, J. (1989). Influence of osmoregulation processes on starvation survival of Escherichia coli in seawater. *Applied and environmental microbiology*, *55*(8), 2017-2024.

Nagalakshmi, U., Wang, Z., Waern, K., Shou, C., Raha, D., Gerstein, M., & Snyder, M. (2008). The transcriptional landscape of the yeast genome defined by RNA sequencing. *Science*, *320*(5881), 1344-1349.

Naughton, L. M., Blumerman, S. L., Carlberg, M., & Boyd, E. F. (2009). Osmoadaptation among Vibrio species and unique genomic features and physiological responses of Vibrio parahaemolyticus. *Applied and environmental microbiology*, *75*(9), 2802-2810.

Noell, S. E., & Giovannoni, S. J. (2019). SAR11 bacteria have a high affinity and multifunctional glycine betaine transporter. *Environmental microbiology*, *21*(7), 2559-2575.

Nyyssölä, A., Kerovuo, J., Kaukinen, P., von Weymarn, N., & Reinikainen, T. (2000). Extreme halophiles synthesize betaine from glycine by methylation. *Journal of Biological Chemistry*, *275*(29), 22196-22201.

Ongagna-Yhombi, S. Y., & Boyd, E. F. (2013). Biosynthesis of the osmoprotectant ectoine, but not glycine betaine, is critical for survival of osmotically stressed Vibrio parahaemolyticus cells. *Applied and environmental microbiology*, *79*(16), 5038-5049.

Oremland, R. S., Marsh, L. M., & Polcin, S. (1982). Methane production and simultaneous sulphate reduction in anoxic, salt marsh sediments. *Nature*, *296*(5853), 143-145.

Oren, A. (1999). Bioenergetic aspects of halophilism. *Microbiology and molecular biology reviews*, *63*(2), 334-348.

Østeras, M., Boncompagni, E., Lambert, A., Dupont, L., Poggi, M., & Le Rudulier, D. (1998). Isolation and molecular characterization of theSinorhizobium meliloti bet locus encoding glycine betaine biosynthesis. *Journal of biosciences*, *23*(4), 457-462.

Pimentel, H., Bray, N. L., Puente, S., Melsted, P., & Pachter, L. (2017). Differential analysis of RNA-seq incorporating quantification uncertainty. *Nature methods*, *14*(7), 687-690.

Pirkov, I., Norbeck, J., Gustafsson, L., & Albers, E. (2008). A complete inventory of all enzymes in the eukaryotic methionine salvage pathway. *The FEBS journal*, *275*(16), 4111-4120.

Pocard, J.-A., Vincent, N., Boncompagni, E., Smith, L. T., Poggi, M.-C., & Le Rudulier, D. (1997). Molecular characterization of the bet genes encoding glycine betaine synthesis in Sinorhizobium meliloti 102F34. *Microbiology*, *143*(4), 1369-1379.

Pokorny, M., Marčenko, E., & Keglević, D. (1970). Comparative studies of L-and D-methionine metabolism in lower and higher plants. *Phytochemistry*, *9*(10), 2175-2188.

Prentki, P., & Krisch, H. M. (1984). In vitro insertional mutagenesis with a selectable DNA fragment. *Gene*, *29*(3), 303-313.

Quinn, P. K., & Bates, T. S. (2011). The case against climate regulation via oceanic phytoplankton sulphur emissions. *Nature*, *480*(7375), 51-56.

Raina, J.-B., Dinsdale, E. A., Willis, B. L., & Bourne, D. G. (2010). Do the organic sulfur compounds DMSP and DMS drive coral microbial associations? *Trends in microbiology*, *18*(3), 101-108.

Raina, J.-B., Tapiolas, D. M., Forêt, S., Lutz, A., Abrego, D., Ceh, J., Seneca, F. O., Clode, P. L., Bourne, D. G., & Willis, B. L. (2013). DMSP biosynthesis by an animal and its role in coral thermal stress response. *Nature*, *502*(7473), 677-680.

Rathinasabapathi, B., Burnet, M., Russell, B. L., Gage, D. A., Liao, P.-C., Nye, G. J., Scott, P., Golbeck, J. H., & Hanson, A. D. (1997). Choline monooxygenase, an unusual iron-sulfur enzyme catalyzing the first step of glycine betaine synthesis in plants: prosthetic group characterization and cDNA cloning. *Proceedings of the National Academy of Sciences*, *94*(7), 3454-3458.

Reisch, C. R., Stoudemayer, M. J., Varaljay, V. A., Amster, I. J., Moran, M. A., & Whitman, W. B. (2011). Novel pathway for assimilation of dimethylsulphoniopropionate widespread in marine bacteria. *Nature*, *473*(7346), 208-211.

Rhodes, D., Gage, D. A., Cooper, A. J., & Hanson, A. D. (1997). S-Methylmethionine conversion to dimethylsulfoniopropionate: evidence for an unusual transamination reaction. *Plant Physiology*, *115*(4), 1541-1548.

Rudulier, D. L., Pocard, J.-A., Boncompagni, E., & Poggi, M. (1996). Osmoregulation in bacteria and transport of onium compounds. In *Biological and Environmental Chemistry of DMSP and Related Sulfonium Compounds* (pp. 253-263). Springer.

Sambrook, J., Fritsch, E. F., & Maniatis, T. (1989). *Molecular cloning: a laboratory manual*. Cold spring harbor laboratory press.

Sanchez, K. J., Chen, C.-L., Russell, L. M., Betha, R., Liu, J., Price, D. J., Massoli, P., Ziemba, L. D., Crosbie, E. C., & Moore, R. H. (2018). Substantial seasonal contribution of observed biogenic sulfate particles to cloud condensation nuclei. *Scientific reports*, *8*(1), 1-14.

Schäfer, A., Tauch, A., Jäger, W., Kalinowski, J., Thierbach, G., & 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*(1), 69-73.

Schäfer, H., Myronova, N., & Boden, R. (2010). Microbial degradation of dimethylsulphide and related C1-sulphur compounds: organisms and pathways controlling fluxes of sulphur in the biosphere. *Journal of experimental botany*, *61*(2), 315-334.

Scholz, A., Stahl, J., de Berardinis, V., Müller, V., & Averhoff, B. (2016). Osmotic stress response in Acinetobacter baylyi: identification of a glycine–betaine biosynthesis pathway and regulation of osmoadaptive choline uptake and glycine–betaine synthesis through a choline-responsive BetI repressor. *Environmental microbiology reports*, *8*(2), 316-322.

Serra, A. L., Mariscotti, J. F., Barra, J. L., Lucchesi, G. I., Domenech, C. E., & Lisa, A. T. (2002). Glycine betaine transmethylase mutant of Pseudomonas aeruginosa. *Journal of bacteriology*, *184*(15), 4301-4303.

Sheehan, V. M., Sleator, R. D., Fitzgerald, G. F., & Hill, C. (2006). Heterologous expression of BetL, a betaine uptake system, enhances the stress tolerance of Lactobacillus salivarius UCC118. *Applied and environmental microbiology*, *72*(3), 2170-2177.

Sheets, E. B., & Rhodes, D. (1996). Determination of DMSP and other onium compounds in Tetraselmis subcordiformis by plasma desorption mass spectrometry. In *Biological and environmental chemistry of DMSP and related sulfonium compounds* (pp. 55-63). Springer.

Shikuma, N. J., Davis, K. R., Fong, J. N., & Yildiz, F. H. (2013). The transcriptional regulator, CosR, controls compatible solute biosynthesis and transport, motility and biofilm formation in Vibrio cholerae. *Environmental microbiology*, *15*(5), 1387-1399.

Sievert, S. M., Kiene, R. P., & Schulz-Vogt, H. N. (2007). The sulfur cycle. *Oceanography*, *20*(2), 117-123.

Simó, R., & Pedrós-Alió, C. (1999). Role of vertical mixing in controlling the oceanic production of dimethyl sulphide. *Nature*, *402*(6760), 396-399.

Simon, R., Priefer, U., & Pühler, A. (1983). A broad host range mobilization system for in vivo genetic engineering: transposon mutagenesis in gram negative bacteria. *Bio/technology*, *1*(9), 784-791.

Slama, I., Abdelly, C., Bouchereau, A., Flowers, T., & Savouré, A. (2015). Diversity, distribution and roles of osmoprotective compounds accumulated in halophytes under abiotic stress. *Annals of Botany*, *115*(3), 433-447. https://doi.org/10.1093/aob/mcu239

Smiddy, M., Sleator, R. D., Patterson, M. F., Hill, C., & Kelly, A. L. (2004). Role for compatible solutes glycine betaine and L-carnitine in listerial barotolerance. *Applied and environmental microbiology*, *70*(12), 7555-7557.

Smirnoff, N., & Cumbes, Q. J. (1989). Hydroxyl radical scavenging activity of compatible solutes. *Phytochemistry*, *28*(4), 1057-1060. https://doi.org/https://doi.org/10.1016/0031-9422(89)80182-7

Smith, L. T., Pocard, J.-A., Bernard, T., & Le Rudulier, D. (1988). Osmotic control of glycine betaine biosynthesis and degradation in Rhizobium meliloti. *Journal of bacteriology*, *170*(7), 3142-3149.

Stefels, J. (2000). Physiological aspects of the production and conversion of DMSP in marine algae and higher plants. *Journal of Sea Research*, *43*(3-4), 183-197.

Strand, T. A., Lale, R., Degnes, K. F., Lando, M., & Valla, S. (2014). A new and improved host-independent plasmid system for RK2-based conjugal transfer. *PloS one*, *9*(3), e90372. https://doi.org/10.1371/journal.pone.0090372

Studier, F. W., & Moffatt, B. A. (1986). Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. *Journal of molecular biology*, *189*(1), 113-130.

Summers, P. S., Nolte, K. D., Cooper, A. J., Borgeas, H., Leustek, T., Rhodes, D., & Hanson, A. D. (1998). Identification and stereospecificity of the first three enzymes of 3-dimethylsulfoniopropionate biosynthesis in a chlorophyte alga. *Plant Physiology*, *116*(1), 369-378.

Sun, J., Steindler, L., Thrash, J. C., Halsey, K. H., Smith, D. P., Carter, A. E., Landry, Z. C., & Giovannoni, S. J. (2011). One carbon metabolism in SAR11 pelagic marine bacteria. *PloS one*, *6*(8), e23973.

Sun, J., Todd, J. D., Thrash, J. C., Qian, Y., Qian, M. C., Temperton, B., Guo, J., Fowler, E. K., Aldrich, J. T., & Nicora, C. D. (2016). The abundant marine bacterium Pelagibacter

simultaneously catabolizes dimethylsulfoniopropionate to the gases dimethyl sulfide and methanethiol. *Nature microbiology*, *1*(8), 1-5.
Sun, L., Curson, A. R., Todd, J. D., & Johnston, A. W. (2012). Diversity of DMSP transport in marine bacteria, revealed by genetic analyses. *Biogeochemistry*, *110*(1), 121-130.
Sunagawa, S., Coelho, L. P., Chaffron, S., Kultima, J. R., Labadie, K., Salazar, G., Djahanschiri, B., Zeller, G., Mende, D. R., Alberti, A., Cornejo-Castillo, F. M., Costea, P. I., Cruaud, C., d'Ovidio, F., Engelen, S., Ferrera, I., Gasol, J. M., Guidi, L., Hildebrand, F., . . . Velayoudon, D. (2015). Structure and function of the global ocean microbiome. *Science*, *348*(6237), 1261359. https://doi.org/doi:10.1126/science.1261359
Sunda, W., Kieber, D., Kiene, R., & Huntsman, S. (2002). An antioxidant function for DMSP and DMS in marine algae. *Nature*, *418*(6895), 317-320.
Suzuki, H., Tamamura, R., Yajima, S., Kanno, M., & Suguro, M. (2005). Corynebacterium sp. U-96 contains a cluster of genes of enzymes for the catabolism of sarcosine to pyruvate. *Bioscience, biotechnology, and biochemistry*, *69*(5), 952-956.
Taalba, A., Xie, H., Scarratt, M., Bélanger, S., & Levasseur, M. (2013). Photooxidation of dimethylsulfide (DMS) in the Canadian Arctic. *Biogeosciences*, *10*(11), 6793-6806.
Tan, D., Crabb, W. M., Whitman, W. B., & Tong, L. (2013). Crystal structure of DmdD, a crotonase superfamily enzyme that catalyzes the hydration and hydrolysis of methylthioacryloyl-CoA. *PloS one*, *8*(5), e63870.
Tanaka, H., Esaki, N., & Soda, K. (1985). A versatile bacterial enzyme: L-methionine γ-lyase. *Enzyme and Microbial Technology*, *7*(11), 530-537.
Tang, K., & Liu, L. (2023). Bacteria are driving the ocean's organosulfur cycle. *Trends in Microbiology.*
Thume, K., Gebser, B., Chen, L., Meyer, N., Kieber, D. J., & Pohnert, G. (2018). The metabolite dimethylsulfoxonium propionate extends the marine organosulfur cycle. *Nature*, *563*(7731), 412-415.
Ticak, T., Kountz, D. J., Girosky, K. E., Krzycki, J. A., & Ferguson Jr, D. J. (2014). A nonpyrrolysine member of the widely distributed trimethylamine methyltransferase family is a glycine betaine methyltransferase. *Proceedings of the National Academy of Sciences*, *111*(43), E4668-E4676.
Timasheff, S. (2002). Protein Hydration, Thermodynamic Binding, and Preferential Hydration. *Biochemistry*, *41*, 13473-13482. https://doi.org/10.1021/bi020316e
Todd, J., Curson, A., Dupont, C., Nicholson, P., & Johnston, A. (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. *Environmental microbiology*, *11*(6), 1376-1385.
Todd, J. D., Curson, A. R., Nikolaidou-Katsaraidou, N., Brearley, C. A., Watmough, N. J., Chan, Y., Page, P. C., Sun, L., & Johnston, A. W. (2010). Molecular dissection of bacterial acrylate catabolism–unexpected links with dimethylsulfoniopropionate catabolism and dimethyl sulfide production. *Environmental microbiology*, *12*(2), 327-343.
Todd, J. D., Rogers, R., Li, Y. G., Wexler, M., Bond, P. L., Sun, L., Curson, A. R., Malin, G., Steinke, M., & Johnston, A. W. (2007). Structural and regulatory genes required to make the gas dimethyl sulfide in bacteria. *Science*, *315*(5812), 666-669.

Tripp, H. J., Kitner, J. B., Schwalbach, M. S., Dacey, J. W., Wilhelm, L. J., & Giovannoni, S. J. (2008). SAR11 marine bacteria require exogenous reduced sulphur for growth. *Nature*, *452*(7188), 741-744.

Trossat, C., Nolte, K. D., & Hanson, A. D. (1996). Evidence that the pathway of dimethylsulfoniopropionate biosynthesis begins in the cytosol and ends in the chloroplast. *Plant Physiology*, *111*(4), 965-973.

Uchida, A., Ooguri, T., Ishida, T., Kitaguchi, H., & Ishida, Y. (1996). Biosynthesis of dimethylsulfoniopropionate in Crypthecodinium cohnii (Dinophyceae). In *Biological and environmental chemistry of DMSP and related sulfonium compounds* (pp. 97-107). Springer.

Van Diggelen, J., Rozema, J., Dickson, D., & Broekman, R. (1986). β-3-dimethylsulphoniopropionate, proline and quaternary ammonium compounds in Spartina anglica in relation to sodium chloride, nitrogen and sulphur. *New Phytologist*, *103*(3), 573-586.

Varaljay, V. A., Gifford, S. M., Wilson, S. T., Sharma, S., Karl, D. M., & Moran, M. A. (2012). Bacterial dimethylsulfoniopropionate degradation genes in the oligotrophic north pacific subtropical gyre. *Applied and environmental microbiology*, *78*(8), 2775-2782.

Vernette, C., Lecubin, J., Sánchez, P., Coordinators, T. O., Sunagawa, S., Delmont, T. O., Acinas, S. G., Pelletier, E., Hingamp, P., & Lescot, M. (2022). The Ocean Gene Atlas v2.0: online exploration of the biogeography and phylogeny of plankton genes. *Nucleic acids research*, *50*(W1), W516-W526. https://doi.org/10.1093/nar/gkac420

Villar, E., Vannier, T., Vernette, C., Lescot, M., Cuenca, M., Alexandre, A., Bachelerie, P., Rosnet, T., Pelletier, E., Sunagawa, S., & Hingamp, P. (2018). The Ocean Gene Atlas: exploring the biogeography of plankton genes online. *Nucleic acids research*, *46*(W1), W289-W295. https://doi.org/10.1093/nar/gky376

Visscher, P. T., Diaz, M. R., & Taylor, B. F. (1992). Enumeration of bacteria which cleave or demethylate dimethylsulfoniopropionate in the Caribbean Sea. *Marine ecology progress series. Oldendorf*, *89*(2), 293-296.

Visscher, P. T., & Taylor, B. F. (1993). A new mechanism for the aerobic catabolism of dimethyl sulfide. *Applied and environmental microbiology*, *59*(11), 3784-3789.

Waditee, R., Tanaka, Y., Aoki, K., Hibino, T., Jikuya, H., Takano, J., Takabe, T., & Takabe, T. (2003). Isolation and Functional Characterization ofN-Methyltransferases That Catalyze Betaine Synthesis from Glycine in a Halotolerant Photosynthetic Organism Aphanothece halophytica. *Journal of Biological Chemistry*, *278*(7), 4932-4942.

Wagner, M. A., & Jorns, M. S. (2000). Monomeric sarcosine oxidase: 2. Kinetic studies with sarcosine, alternate substrates, and a substrate analogue. *Biochemistry*, *39*(30), 8825-8829.

Wang, Z., Gerstein, M., & Snyder, M. (2009). RNA-Seq: a revolutionary tool for transcriptomics. *Nat Rev Genet*, *10*(1), 57-63. https://doi.org/10.1038/nrg2484

Wargo, M. J. (2013). Homeostasis and catabolism of choline and glycine betaine: lessons from Pseudomonas aeruginosa. *Applied and environmental microbiology*, *79*(7), 2112-2120.

Wargo, M. J., Szwergold, B., & Hogan, D. A. (2008). Identification of Two Gene Clusters and a Transcriptional Regulator Required for Pseudomonas aeruginosa Glycine Betaine Catabolism. *Journal of bacteriology*, *190*, 2690 - 2699.

Watkins, A. J., Roussel, E. G., Parkes, R. J., & Sass, H. (2014). Glycine betaine as a direct substrate for methanogens (Methanococcoides spp.). *Applied and environmental microbiology*, *80*(1), 289-293.

Welsh, D. T. (2000). Ecological significance of compatible solute accumulation by micro-organisms: from single cells to global climate. *FEMS microbiology reviews*, *24*(3), 263-290. https://doi.org/10.1111/j.1574-6976.2000.tb00542.x

Williams, B. T., Cowles, K., Bermejo Martínez, A., Curson, A. R., Zheng, Y., Liu, J., Newton-Payne, S., Hind, A. J., Li, C.-Y., & Rivera, P. P. L. (2019). Bacteria are important dimethylsulfoniopropionate producers in coastal sediments. *Nature microbiology*, *4*(11), 1815-1825.

Wolfe, G. V. (1996). Accumulation of dissolved DMSP by marine bacteria and its degradation via bacterivory. In *Biological and environmental chemistry of DMSP and related sulfonium compounds* (pp. 277-291). Springer.

Wood, J. M. (2015). Bacterial responses to osmotic challenges. *J Gen Physiol*, *145*(5), 381-388. https://doi.org/10.1085/jgp.201411296

Wood, W. B. (1966). Host specificity of DNA produced by Escherichia coli: bacterial mutations affecting the restriction and modification of DNA. *Journal of molecular biology*, *16*(1), 118-IN113.

Wördenweber, R., Rokitta, S. D., Heidenreich, E., Corona, K., Kirschhöfer, F., Fahl, K., Klocke, J. L., Kottke, T., Brenner-Weiß, G., & Rost, B. (2018). Phosphorus and nitrogen starvation reveal life-cycle specific responses in the metabolome of Emiliania huxleyi (Haptophyta). *Limnology and Oceanography*, *63*(1), 203-226.

Xu, T., Yu, M., Liu, J., Lin, H., Liang, J., & Zhang, X.-H. (2019). Role of RpoN from Labrenzia aggregata LZB033 (Rhodobacteraceae) in Formation of Flagella and Biofilms, Motility, and Environmental Adaptation. *Applied and environmental microbiology*, *85*(7), e02844-02818. https://doi.org/doi:10.1128/AEM.02844-18

Yancey, P. H. (2005). Organic osmolytes as compatible, metabolic and counteracting cytoprotectants in high osmolarity and other stresses. *Journal of experimental biology*, *208*(15), 2819-2830.

Yoo, E. J.-H., Feketeová, L., Khairallah, G. N., White, J. M., & Richard, A. (2011). Structure and unimolecular chemistry of protonated sulfur betaines,(CH 3) 2 S+(CH 2) n CO 2 H (n= 1 and 2). *Organic & Biomolecular Chemistry*, *9*(8), 2751-2759.

Zehr, J. P., Weitz, J. S., & Joint, I. (2017). How microbes survive in the open ocean. *Science*, *357*(6352), 646-647. https://doi.org/doi:10.1126/science.aan5764

Zheng, Y., Wang, J., Zhou, S., Zhang, Y., Liu, J., Xue, C.-X., Williams, B. T., Zhao, X., Zhao, L., Zhu, X.-Y., Sun, C., Zhang, H.-H., Xiao, T., Yang, G.-P., Todd, J. D., & Zhang, X.-H. (2020). Bacteria are important dimethylsulfoniopropionate producers in marine aphotic and high-pressure environments. *Nature Communications*, *11*(1), 4658. https://doi.org/10.1038/s41467-020-18434-4

Zhong, H., Sun, H., Liu, R., Zhan, Y., Huang, X., Ju, F., & Zhang, X. H. (2021). Comparative Genomic Analysis of Labrenzia aggregata (Alphaproteobacteria) Strains Isolated From the Mariana Trench: Insights Into the Metabolic Potentials and

Biogeochemical Functions. *Front Microbiol*, *12*, 770370. https://doi.org/10.3389/fmicb.2021.770370

Ziegler, C., Bremer, E., & Krämer, R. (2010). The BCCT family of carriers: from physiology to crystal structure. *Molecular Microbiology*, *78*(1), 13-34. https://doi.org/https://doi.org/10.1111/j.1365-2958.2010.07332.x

Zou, H., Chen, N., Shi, M., Xian, M., Song, Y., & Liu, J. (2016). The metabolism and biotechnological application of betaine in microorganism. *Applied Microbiology and Biotechnology*, *100*(9), 3865-3876. https://doi.org/10.1007/s00253-016-7462-3