On the molecular diversity of dimethylsulphoniopropionate catabolism by marine bacteria

Emily Kate Fowler

A thesis submitted to the University of East Anglia for the Degree of Doctor of Philosophy

April 2015

Declaration

I declare that the work in this thesis has not been previously submitted for a degree at the University of East Anglia or any other university, and all research has been carried out by myself, unless otherwise stated.

Signed

Emily Kate Fowler

Abstract

Dimethylsulphoniopropionate (DMSP) is the most abundant organic sulphur molecule in the oceans. Its breakdown by marine organisms is important for the global cycling of sulphur, and as a nutrient source for microbial life. In recent years, the molecular basis of DMSP catabolism by marine bacteria has begun to be unravelled, through the discovery of six different DMSP lyases and a DMSP demethylase, as well as downstream pathways. From these studies, it is becoming evident that there is great diversity in the way bacteria breakdown this important molecule. The work presented here further explores and expands our knowledge of this diversity. I have identified a novel DMSP lyase (DddK), which catalyses the cleavage of DMSP into acrylate and dimethyl sulphide (DMS) in the DMS-producing *Candidatus* Pelgaibacter ubique HTCC1062 - one of the most prolific bacteria on this planet. I have also shown that the γ-proteobacterium *Oceanimonas doudoroffii*, which has long been a study species for DMSP catabolism, has no fewer than three functional DMSP lyases - DddD, DddP1 and DddP2 - this being the first example of a species outside of the α-proteobacteria having multiple lyases. Additionally, I have presented a thorough bioinformatics analysis of the occurrence and synteny of genes associated with DMSP catabolism within sequenced members of the abundant Roseobacter clade, revealing some interesting patterns which warrant further experimental investigation. Finally, I have shown that the model marine Roseobacter species *Ruegeria pomeroyi* DSS-3 is able to use DMSP-derived acrylate as a sole carbon source via a fatty acid biosynthesis route, linked to propionate catabolism.

List of Contents

List of Figures

List of Tables

Acknowledgements

I would like to thank my supervisors Dr Jonathan Todd and Prof Andy Johnston, their knowledge, enthusiasm and support during this PhD has been truly invaluable. Special thanks also go to Dr Andy Curson, whose expertise combined with a seemingly infinite amount of patience has really made all this possible.

I must also thank all of the members of lab 1.29, past and present, for all their help and friendship throughout my time there, including Simone Payne, Ana Bermejo-Martinez, Dr Ornella Carríon-Fonseca, Dr Matt Sullivan, Dr Mark Kirkwood, Dr Rob Green, Pam Wells, Isabel Johnston and Ollie Burns. I feel very fortunate to have been part of such a hardworking and friendly team, and I hope there will be many more Food Days to come.

Additionally, I must thank Prof Tracey Chapman, for her remarkable patience, positivity and support while I have been writing this, and all the members of the Chapman lab, including Dr Damian Smith, Dr Janet Mason, Dr Wayne Rostrant, Dr Phil Leftwich, Dr Irina Mohorianu, Will Nash, Naomi Clarke, Michael Bolton and Liz Duxbury, for making the transition to a new group a wonderful experience.

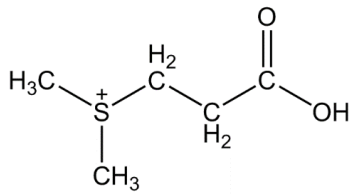
Finally, I thank Mio for all his support with my research, especially for being there during late nights in the lab. Most importantly, I must thank my family, Mum, Dad and David for their never-ending encouragement and belief in me. I hope I've made you proud.

Chapter 1

Introduction

1.1 Dimethylsulphoniopropionate

Dimethylsulphoniopropionate (DMSP) is a tertiary sulphonium zwitterion (**Figure 1.1**) produced in marine environments, mostly by photosynthetic eukaryotic organisms. It is a hugely abundant compound, with estimates of its production reaching 1 billion tonnes annually; indeed, it is the greatest single source of bio-organic sulphur in the world's oceans (Kettle and Andreae, 2000). In a biogeochemical sense, DMSP is also important as a precursor to the gas dimethyl sulphide (DMS), which is the major source of sea-to-land organic sulphur flux. Importantly, DMSP also acts as a valuable source of carbon for marine micro-organisms. It has been shown that DMSP accounts for 1-13% of bacterial carbon demand in surface waters, making it one of the most important single substrates for bacterioplankton identified so far (Kiene *et al*., 2000).



Dimethylsulphoniopropionate

Figure 1.1 Chemical structure of DMSP

The production of DMSP has been confirmed in many species of marine phytoplankton and macro-algae and a few angiosperms, which mostly reside in or near marine environments. The only animals known to produce DMSP are corals of the genus *Acropora* (Raina *et al*., 2013). Several different functions of DMSP in these organisms have been proposed, including that of an osmoprotectant, an antioxidant and an anti-stress molecule, and these will be described in detail later.

1.2 Biosynthesis of DMSP

Organisms that can synthesise DMSP are taxonomically diverse. They include species of uni- and multi-cellular algae, angiosperms, corals and phototrophic bacteria. The common link between all of these organisms is their habitat. With only a few exceptions, all are marine-based life-forms.

1.2.1 Algae

The most important producers of DMSP are phytoplankton including the Dinophyceae (Dinoflagellates), Prymnesiophyceae (including Coccolithophorids) and Chrysophyceae and Bacilloariophyceae (Diatoms). Within the Dinoflagellates, intracellular levels of DMSP vary greatly between species (Caruana and Malin, 2014) from 0.003 mM in *Pfiesteria piscicida* to a remarkable 7 M in *Symbiodinium* sp., a symbiont of corals (Broadbent *et al*., 2002). Of all studies to date, the median intracellular concentration of DMSP in dinoflagellates is 167 mM (Caruana and Malin, 2014).

The most studied species of the Prymnesiophyceae class are *Phaeocystis* sp. and the coccolithophore *Emiliania huxleyi*, due to their ability to form massive algal blooms, which have been associated with an increase in the production of dimethyl sulphide (Gibson *et al*., 1990; Levasseur *et al*., 1996). In *Phaeocystis* sp., intracellular DMSP levels can accumulate to 150 mM (Stefels and Boekel, 1993), while *E. huxleyi* also has high values, ranging from 50 to 250 mM (Steinke *et al*., 1998).

Some species of multicellular macro-alga have also been reported to produce DMSP, including Chlorophytes *Ulva lactuca* (Greene, 1962; Van Alstyne *et al*., 2007) and *Ulva* (previously *Enteromorpha*) *intestinalis* (Gage *et al*., 1997). The red alga *Polysiphonia fastigiata* also produces DMSP, and is the organism which led to the discovery of DMSP as a precursor of DMS (Challenger and Simpson, 1948).

1.2.2 Angiosperms

DMSP is also produced in a few angiosperms, that reside in marine environments. These include grasses of the genera *Spartina* (salt marsh grass) (Larher *et al*., 1977; Dacey *et al*., 1987) and the dicotyledon *Wollastonia biflora*, known colloquially as the beach sunflower (Hanson *et al*., 1994). Four species of *Spartina* have been confirmed to produce DMSP, namely *S. alterniflora, S. maritime, S. anglica* (a hybrid of the first two species), and *S. foliosa* (Otte *et al*., 2004), but interestingly, others (e.g. *S. cynosuroides* and *S. patens*) do not despite living alongside some of the DMSP-producing species (Otte and Morris, 1994). DMSP can accumulate to high

concentrations in producing plants (up to 250 μmol g^{-1} dry weight) indicating that it plays an important role in these organisms (Otte *et al*., 2004).

1.2.3 Corals

High concentrations of DMSP and DMS have been shown to be associated with coral reefs (Broadbent *et al*., 2002; Broadbent and Jones, 2004), but until recently, it was assumed that the producer of DMSP in this environment was exclusively *Symbiodinium*, an intracellular dinoflagellate symbiont of coral. It has now been shown that two species of coral (*Acropora millepora* and *A. tenuis*) are able to produce DMSP in the absence of any algal symbionts (Raina *et al*., 2013). This exciting research revealed that DMSP production is not restricted to photosynthetic organisms, as previously thought.

1.2.4 Cyanobacteria

DMSP has been measured in some species of marine unicellular and filamentous cyanobacteria, although concentrations were very low compared to those found in marine algae (Vogt *et al*., 1998). It is thought therefore that marine cyanobacteria are relatively minor producers of DMSP.

1.3 Pathways of DMSP Biosynthesis

Considering the importance of DMSP, and the vast amount of research into its presence and function in marine organisms, surprisingly little work has been done on elucidating the molecular and genetic mechanisms behind its production. As yet, not a single gene involved in DMSP biosynthesis has been confirmed in any organism. However, four different pathways to DMSP synthesis have been proposed in the angiosperms *Wollastonia biflora*, *Spartina alterniflora*, the macroalgae *Ulva intestinalis* and the dinoflagellate *Crypthecodinium cohnii*. These studies showed that in all cases the starting material in the dedicated pathways for DMSP synthesis is methionine, but the way in which this occurs varies between the different organisms (see **Figure 1.2**).

1.3.1 DMSP biosynthesis in *Wollastonia biflora*

As shown by Hanson *et al*. (1994), the first step in the synthesis of DMSP from methionine in *W. biflora* is the S-adenosyl methionine-dependent methylation of methionine to form S-methyl-methionine (SMM). The methyltransferase responsible for this step has been purified and shown to be a homo-tetramer of 115 kDa subunits (James *et al*., 1995a). The resultant SMM is likely converted to DMSP-aldehyde by successive transamination and decarboxylation steps, although

no intermediates in this step have been identified (James *et al*., 1995b). These steps could be carried out by one enzyme with dual functionality, or a closely coupled transamination-decarboxylase complex (Rhodes *et al*., 1997). Finally, the DMSP aldehyde is oxidised to DMSP (**Figure 1.2**) by an NAD-dependent dehydrogenase, which has also been purified from *W. biflora*. Interestingly, the DMSP-aldehyde dehydrogenase activity was recovered from the chloroplast stromal fraction, whereas the SMM:methionine S-methyltransferase activity was found in the cytosolic fraction, suggesting SMM is produced in the cytosol, before transportation to the chloroplast for conversion to DMSP (Trossat *et al*., 1996a, b; Trossat *et al*., 1998).

1.3.2 DMSP biosynthesis in *Spartina alterniflora*

A different, though related, pathway was identified in *S. alterniflora*. It differs from the *W. biflora* pathway in that a 3-dimethylsulphoniopropylamine (DMSP-amine) intermediate is produced from SMM (Kocsis *et al*., 1998). The enzymes catalysing the SMM → DMSP-amine → DMSP-aldehyde route are predicted to be an S-methyl-methionine decarboxylase, and a DMSP-amine oxidase (Kocsis and Hanson, 2000) (**Figure 1.2**).

1.3.3 DMSP biosynthesis in *Ulva intestinalis*

A third and entirely distinct pathway was discovered in the macroalga *U. intestinalis*. The intermediate 4-methylthio-2-oxobutyrate (MTOB) is produced from methionine via a transamination step. MTOB is then reduced to 4-methylthio-2-hydroxybutyrate (MTHB) in an NADPH-dependent reaction, and MTHB is methylated to form 4-dimethylsulphonio-2-hydroxybutyrate (DMSHB). Finally, DMSHB is oxidatively decarboxylated to DMSP (Gage *et al*., 1997) (**Figure 1.2**). Enzymes catalysing the first three steps of this pathway were partially purified and characterised as a 2-oxoglutarate-dependent aminotransferase, an NADPH-linked reductase and an S-adenosylmethionine-dependent methyltransferase (Summers *et al*., 1998).

1.3.4 DMSP biosynthesis in *Fragilariopsis cylindrus*

More recently, a proteomics study was carried out using the sea-ice diatom *Fragilariopsis cylindrus* (Lyon *et al*., 2011). The study found that intracellular DMSP concentration increased under hyper-saline conditions, along with a number of proteins which were identified by mass spectrometry. Five of these enzymes were predicted to be involved in the DMSP synthesis pathway, as they fitted in with the existing *Ulva intestinalis* model. These were an aminotransferase, a reductase, an S-adenosylmethionine-dependent methyltransferase and two decarboxylases. Interestingly, three of the enzymes had a chloroplast targeting sequence motif. The presence of these motifs hints at the possibility that, like in *Wollastonia* (see above), at least

part of DMSP synthesis takes place in the chloroplast. Importantly, this study provided the first candidate genes for a DMSP synthesis pathway, which are currently under investigation.

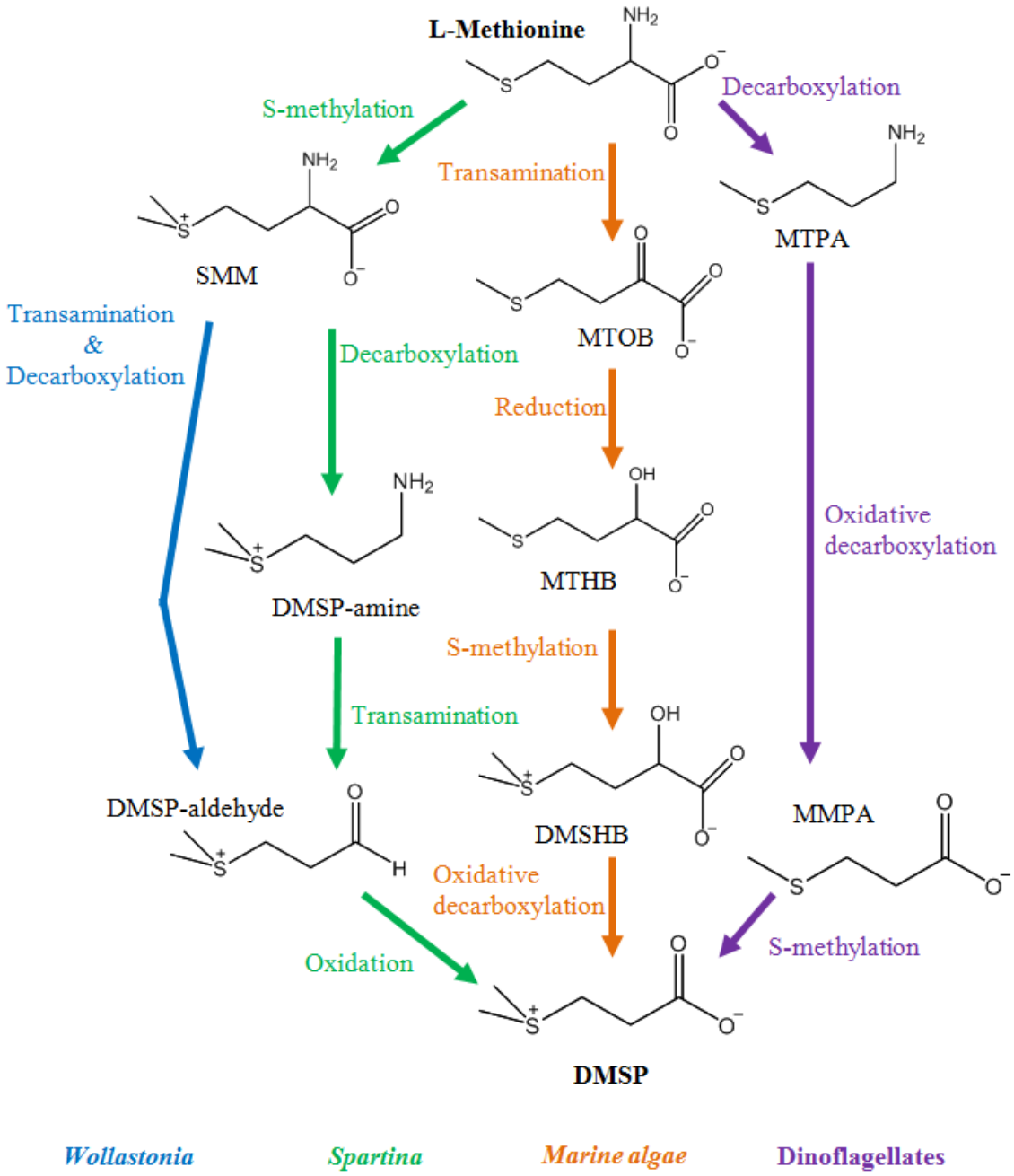


Figure 1.2 DMSP biosynthesis pathways in angiosperms, algae and dinoflagellates. Proposed DMSP biosynthesis pathways for *Wollastonia biflora* (blue arrows), *Spartina alterniflora* (green arrows), marine alga *Ulva intestinalis* and the diatom *Fragilariopsis cylindrus* (orange arrows) and the dinoflagellate *Crypthecodinium cohnii* (purple arrows). The chemical reactions are labelled for each pathway. SMM, S-methylmethionine; DMSP,

dimethylsulphoniopropionate; MTOB, 4-methyl-2-oxobutyrate; MTHB, 4-methyl-2-hydroxybutyrate; DMSHB, 4-dimethylsulfonio-2-hydroxybutyrate; MTPA, methanethiolpropanamine; MMPA, methylmercaptopropionate.

1.3.5 DMSP biosynthesis in dinoflagellates

The DMSP biosynthesis pathway described for the dinoflagellate *Crypthecodinium cohnii* is significantly different from those of the angiosperms and algae. Although the starting point is L-methionine, this is then decarboxylated to methanethiolpropanamine (MTPA) and subsequently converted to methylmercaptopropionate (MMPA) through oxidative decarboxylation. Finally, MMPA is methylated to produce DMSP (Kitaguchi *et al*., 1999) (**Figure 1.2**).

1.4 Functions of DMSP

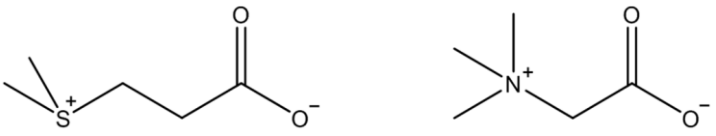
Although DMSP is abundant in the marine environment, its exact role in the organisms which produce it is not known. Several different functions have been proposed, based on correlative evidence and these are presented below.

1.4.1 DMSP as an osmoprotectant

In environments of high, or fluctuating salinity, there is a need for organisms to produce osmotically active solutes, which are compatible with metabolism. Since DMSP is almost exclusively produced in such environments, an attractive explanation is that DMSP plays a role in osmotic balance. Indeed, DMSP is a sulphonium analogue of the well-known compatible solute, glycine betaine (see **Figure 1.3**). Both glycine betaine and DMSP have been shown directly to enhance the salinity tolerance of *E. coli* at nanomolar levels, likely due to the presence of a high affinity osmoporter, ProU which could transport DMSP and glycine betaine (Cosquer *et al.*, 1999).

However, studies carried out in DMSP-producing organisms are less convincing. In some cases it was found that intracellular DMSP concentrations increased with salinity, for example in the coccolithophore *Hymenomonas carterae* (Vairavamurthy *et al*., 1985), and the diatom *Cylindrotheca closterium* (Van Bergeijk *et al*., 2003). Dacey *et al*. (1987) also noted a positive correlation between sediment salinity and the concentration of DMSP in the leaves of *Spartina alterniflora*. However, other studies have reported no effect of salinity on DMSP concentration

in this plant (Otte and Morris, 1994; Colmer *et al*., 1996). The best evidence to date that DMSP acts as an osmoprotectant is the study carried out in *F. cylindrus*, which showed an 85% increase in intracellular DMSP concentration (from ~15 mM to ~28 mM) when the diatom was gradually shifted from low to high salinity (Lyon *et al*., 2011).



DMSP Glycine betaine

Figure 1.3 Structures of DMSP and glycine betaine

1.4.2 DMSP as an antioxidant

Another possible function of DMSP is as a scavenger of hydroxyl radicals and other reactive oxygen species (ROS). Studies in marine unicellular algae have shown that cellular DMSP concentrations and DMSP lysis increase in response to a range of oxidative stressors. For example, in *E. huxleyi*, UV radiation, CO_2 limitation or exposure to high levels of Cu^{2+} and H_2O_2 all resulted in elevated levels of DMSP or DMS. The same was true for the coastal diatom *Skeletonema costatum* under iron-limiting conditions, and *T. pseudonana* in both iron and CO_2 limiting conditions (Sunda *et al*., 2002).

In *S. alterniflora*, applied oxidative stress did not result in elevated DMSP levels, but it did cause an increase in DMSP oxidation to dimethylsulphoxide (DMSO), which can scavenge ROS (Husband *et al*., 2010).

1.4.3 DMSP as a herbivore grazing deterrent

There is some evidence to suggest that DMSP or its cleavage products, DMS and acrylate, act as grazing deterrents. When presented with five different strains of *E. huxleyi*, all with varying levels of DMSP lyase activity, protozoan grazers avoided those strains with high DMSP lyase activity (Wolfe *et al*., 1997). However, it was noted in a separate study that the addition of DMSP reduced grazing on *E. huxleyi* by protozoa, but that the breakdown products (DMS and acrylate – see below) had no effect (Strom *et al*., 2003). In a study on different macroalgae (both Chlorophytes and Rhodophytes) DMSP was seen to act as a feeding attractant to species of sea urchin, whereas acrylic acid deterred feeding. Conversely, an isopod was not deterred by acrylic acid (Van Alstyne *et al*., 2001). From these studies it seems that the effectiveness of DMSP as a grazing deterrent may depend on both the DMSP producer and the herbivore species.

1.4.4 DMSP as a cryoprotectant

The observation that concentrations of DMSP in some Chlorophyceaen species from Antarctic regions are much higher than Chlorophyceaen species from temperate environments has led to the suggestion that DMSP acts as a cryoprotectant (Karsten *et al*., 1990). Indeed, DMSP was found to stabilize the enzyme phosphofructokinase at low temperatures (Nishiguchi and Somero, 1992). Later, it was shown for *Acrosiphonia arcta* that its malate dehydrogenase and lactate dehydrogenase activities were stabilised by DMSP at low temperatures and during freeze-thawing, respectively, and lactate dehydrogenase activity was even increased by the addition of DMSP (Karsten *et al*., 1996).

1.5 Environmental fate of DMSP

It has been estimated that 1 billion tonnes of DMSP are produced every year in the world's oceans (Kettle and Andreae, 2000). Measurements of the typical concentration of DMSP in seawater are patchy, but are usually in the range of 1-2 nM, which can increase dramatically to several micro-molar during algal blooms (van Duyl *et al*., 1998) or around coral reefs (Broadbent and Jones, 2004). This is due to the disruption of the cells of DMSP-producing organisms, for example by viral lysis, herbivore grazing, or senescence (Wolfe *et al*., 1994; Bratbak *et al*., 1996), at which point DMSP is released into the surrounding environment.

An important study by Kiene *et al*. (2000) used ^{35}S-tracer studies to examine the biochemical fate of DMSP in samples of oceanic and coastal waters. The investigation revealed that DMSP is rapidly degraded in seawater, and that sulphur from this molecule enters three major pools: particulates, dissolved non-volatile degradation products (DNVS) and volatiles. An average of

9% ended up as volatiles, and the majority of this was methanethiol (MeSH). Further investigation revealed that the sulphur from MeSH was incorporated into cellular macromolecules. A longer-lived volatile was dimethyl sulphide (DMS), which had a relatively slow turnover compared to DMSP and MeSH.

These two volatiles have served as indicators of different pathways of DMSP degradation in marine organisms. Thus the production of methanethiol from DMSP is attributed to the DMSP demethylation pathway, whereas DMS production may indicate the presence of a DMSP cleavage pathway. However, it is important to consider that DMS can also be produced from MeSH via a methylation step, so the mere production of these volatiles from DMSP is not truly enough to confirm the presence of either pathway.

In recent years, much work has been carried out on the molecular basis of DMSP-dependent DMS and MeSH production, mostly by the UEA lab, and Mary Ann Moran's group at the University of Georgia. Although these studies focussed on DMSP degradation in bacteria, other organisms also can degrade DMSP. For example, the coccolithophore *E. huxleyi* and the green alga *Enteromorpha clathrata* both have been shown to produce DMS from DMSP (Franklin *et al*., 2010; Steinke and Kirst, 1996). However, much more is known about the genetics of DMSP catabolism in bacteria, and while eukaryotic organisms are the major producers of DMSP, it is thought that bacteria are largely responsible for the further breakdown of this molecule (Kiene, 1992). For this reason, the following presentation on DMSP degradation will focus on pathways found in bacteria.

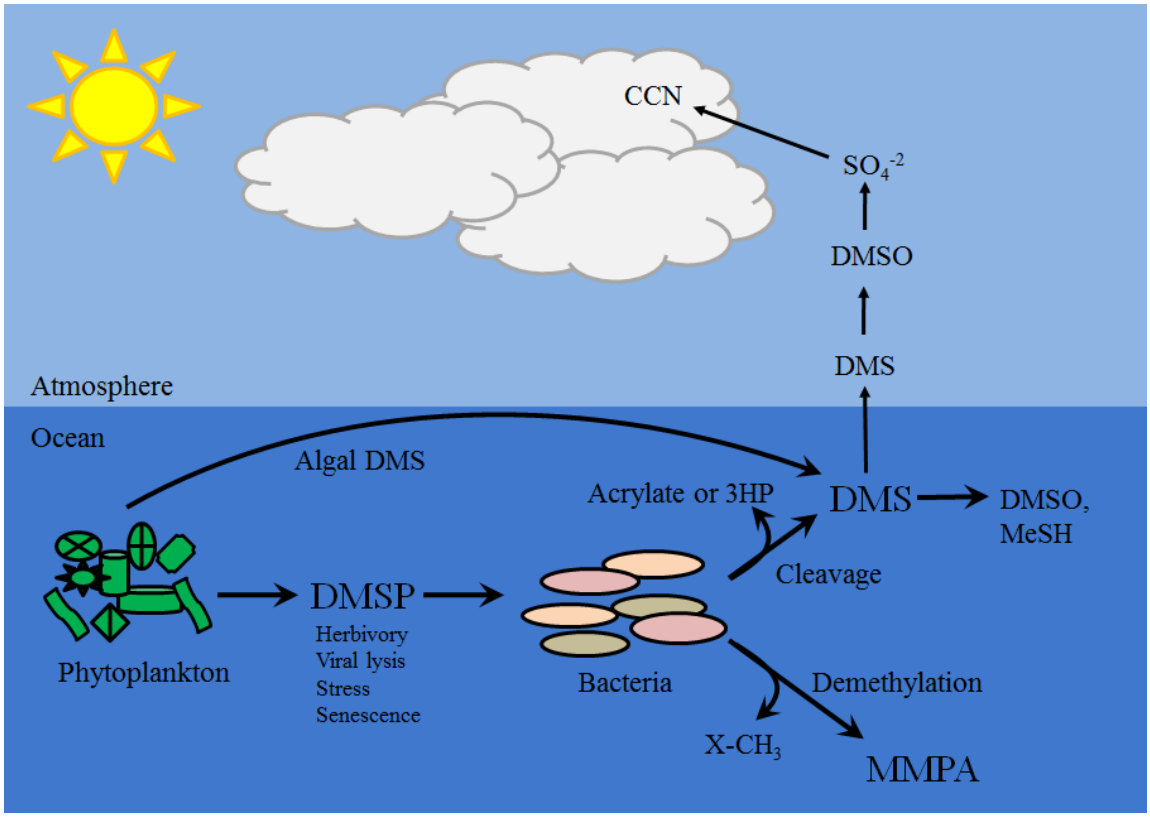


Figure 1.4 Simplified transformations of DMSP and DMS. Once produced by marine eukaryotic organisms, DMSP can be catabolised by some producing organisms to DMS, or it is released into the environment upon rupture of the cells, for example by herbivory, viral lysis, stress or senescence. DMSP can then be taken up by bacteria and catabolised either to produce DMS and acrylate or 3-hydroxypropionate (3-HP), or demethylated to methylmercaptopropionate (MMPA). DMS can be further catabolised by bacteria to dimethylsulphoxide (DMSO) or methanethiol (MeSH), or released into the atmosphere, where its oxidation products form cloud condensation nuclei (CCN). In the form of precipitation, the sulphur from DMS is returned to land, thus contributing to the global sulphur cycle.

1.6 DMSP demethylation

One pathway of DMSP degradation occurs via the initial removal of a methyl group from DMSP, resulting in the production of methylmercaptoproionate (MMPA). It is thought that this route accounts for ca. 70% of the global degradation of DMSP (Kiene *et al*., 2000). Importantly, the resultant MMPA can be further catabolised to methanethiol (MeSH) which is a major source of sulphur for bacterial protein synthesis (Kiene *et al*., 1999). The production of MMPA and MeSH from DMSP was discovered over 25 years ago, in studies of anoxic marine sediment (Kiene *et al*., 1988), but, despite its importance, the exact pathway and its genetics were not uncovered until very recently.

1.6.1 DmdA – discovery of the DMSP demethylase and the corresponding gene

The gene encoding the first enzymatic step in DMSP demethylation, and indeed *any* DMSP catabolic pathway, was discovered in 2006 in *Ruegeria pomeroyi* DSS-3, an α-proteobacterium in the Roseobacter clade of abundant marine bacteria (see below). As shown by González *et al*. (1999), this strain can produce MeSH as one of the end-products of the DMSP demethylation catabolic pathway. Using a colorimetric screen, Howard *et al*. (2006) obtained a mutant, from a transposon-based mutant library, which failed to make MeSH when the cells were grown in the presence of DMSP. The mutation was mapped to a single gene, namely SPO1913, which was termed *dmdA*.

The *dmdA* gene encodes a glycine cleavage-T family protein and, using the purified enzyme, was shown directly to demethylate DMSP to MMPA, using tetrahydrofolate (THF) as a methyl acceptor. Although it appears to have a strict substrate specificity for DMSP, it has a surprisingly high K_m (5.4 mM). However, *R. pomeroyi* can accumulate intracellular DMSP concentrations as high as 70 mM from an external concentration of just 1 mM (Reisch *et al*., 2008).

1.7 Downstream steps in the DMSP demethylation pathway

Having identified the initial enzyme, and gene, in the DMSP demethylation pathway, Moran's group began to investigate the subsequent degradation of MMPA. In an attempt to identify the genes and corresponding enzymes that catalyse the rest of the demethylation pathway, they focussed on a hypothetical demethiolation pathway, whereby MMPA would be degraded via a coenzyme A dependent cycle of fatty acid β-oxidation (Bentley and Chasteen, 2004).

1.7.1 The MMPA demethiolation pathway enzymes DmdB, DmdC, and DmdD

Since the focus was on a pathway involving coenzyme A intermediates, enzymes with MMPA-CoA ligase activity were purified from cell extracts of *R. pomeroyi* (Reisch *et al*., 2011). One of the four remaining enzymes following purification was identified as a medium-chain fatty acid CoA ligase. The gene encoding this enzyme, SPO2045, was cloned into an expression vector and confirmed to have MMPA-CoA ligase activity in *E. coli*. This gene was designated *dmdB*. Interestingly, a DmdB$^-$ mutant strain of *R. pomeroyi* still retained a reduced level of MMPA-CoA ligase activity, which was attributed to the presence of a second *dmdB* in DSS-3 (SPO0677) (see **Chapter 5**).

The remaining steps of the pathway were identified using incubations of *R. pomeroyi* crude cell extracts with MMPA-CoA which resulted in the production of methylthioacryloyl-CoA (MTA-CoA), and separate incubations with MTA-CoA, which released MeSH and free CoA. The enzyme catalysing the release of MeSH from MTA-CoA was purified and identified as an enoyl-CoA hydratase encoded by SPO3805. In the genome, SPO3805 is located immediately upstream of SPO3804 which was cloned and shown to encode the enzyme responsible for the production of MTA-CoA from MMPA-CoA. The genes were named *dmdC* (SPO3804) and *dmdD* (SPO3805) (Reisch *et al*., 2011). Thus one pathway of MeSH production from DMSP in *R. pomeroyi* occurs via MMPA-CoA and MTA-CoA, summarised in **Figure 1.4**.

Interestingly, although mutations were made in all of the downstream *dmd* genes, the mutants were not assayed for production of MMPA-dependent MeSH production, so the contribution of each of these genes to the production of MeSH from MMPA is not known. Nor is the overall flux through this pathway, compared to other, very different ways in which this strain – and other Roseobacters – can catabolise DMSP (see below).

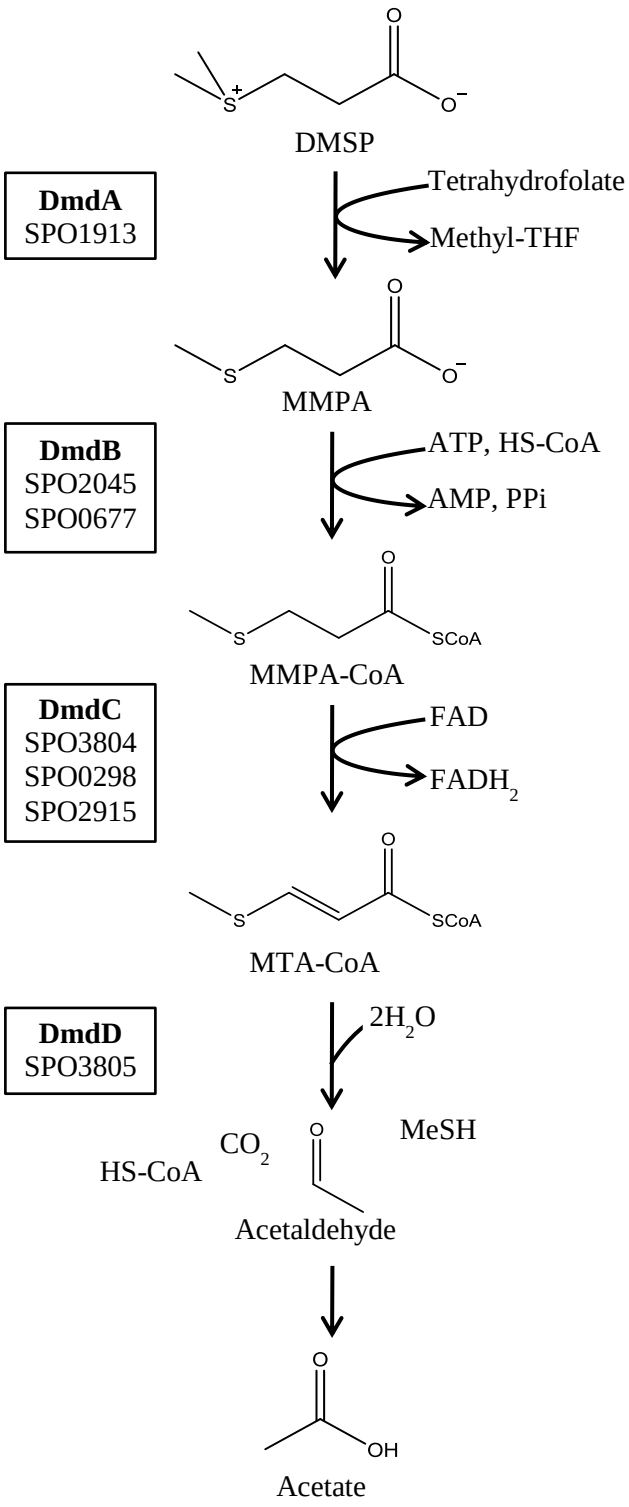


Figure 1.5 DMSP demethylation pathway in *Ruegeria pomeroyi* DSS-3. Enzymes and genes involved at each step are indicated in boxes. DMSP is demethylated to MMPA in a tetrahydrofolate-dependent manner by DmdA. MMPA is then converted to MMPA-CoA in an ATP-dependent reaction mediated by DmdB, and MMPA-CoA is dehydrogenated by DmdC to produce MTA-CoA. MTA-CoA is hydrated to acetaldehyde by DmdD, releasing MeSH, free

CoA and CO_2 in the process. Acetaldehyde can then be oxidised to acetic acid. Figure adapted from Reisch *et al*., 2011.

1.8 Distribution of DmdA, DmdB, DmdC and DmdD

Close homologues of DmdA are found in most (though not all – see below) of the genome-sequenced strains and species of the Roseobacters, a clade of the α-proteobacterial Rhodobacterales family. It is also seen in two other hugely abundant clades of marine-alphas, namely *Puniceispirillum*, and the even more numerous *Pelagibacter* SAR11s. Indeed, in the latter group, the corresponding gene was cloned and shown to make a functional enzyme, though, like that of *R. pomeroyi*, it had a surprisingly high K_m (13.2 mM) in *in vitro* assays. DmdA is also present in some marine γ-proteobacteria, for example strain HTCC2080 and the sponge symbiont *Thioalkalivibrio* sp. HK1 (**Figure 1.6**).

This widespread distribution of the *dmdA* gene in two of the most abundant taxa of bacteria *anywhere*, underpins the finding that *dmdA* homologues are so very frequently encountered in marine metagenomic data bases. Most notably, in the metagenomic Global Ocean Sampling (GOS) data (Rusch *et al*., 2007), sufficient numbers of *dmdA* homologues were found for almost 60% of sampled cells to contain this gene (Howard *et al*., 2008).

In contrast to DmdA, homologues to DmdB and DmdC are widespread in nature, in marine and terrestrial environments alike. However, a selection of DmdC enzymes from MeSH-producing strains have been verified as having MMPA-CoA dehydrogenase activity, and these peptides form a sub-clade from other acyl-CoA dehydrogenases (Reisch *et al*., 2011). These included DmdC from *Pelagibacter ubique*, *Burkholderia thailandensis*, *Pseudomonas aeruginosa* and *Ruegeria lacuscaerulensis*, plus two further homologues to DmdC in *R. pomeroyi* (SPO0298 and SPO2915). Similarly, DmdB from *P. ubique* and both copies of DmdB from *R. pomeroyi* have been verified as functional and they also form a sub-clade from other homologous CoA-ligases (Reisch *et al*., 2011).

Unlike DmdC and DmdB, homologues to DmdD are rare, even within those bacteria which produce MeSH from MMPA. For example, the closest homolog to DmdD in *P. ubique* (24% identity) did not possess MTA-CoA hydratase activity. However, the *dmdD*-negative strain *Ruegeria lacuscaerulensis*, which possesses *dmdB* and *dmdC*, was shown to have DmdD activity, suggesting a non-orthologous isofunctional enzyme may have replaced DmdD, at least in this bacterium (Reisch *et al*., 2011).

The relative abundances of DmdB, DmdC and DmdD homologues are reflected in metagenomic data. Analysis of the GOS database returned over 6000 homologues to *dmdB* and *dmdC*, indicating they may be present in 61% of sampled cells, whereas only 16 homologues were found for *dmdD* (Reisch *et al*., 2011).

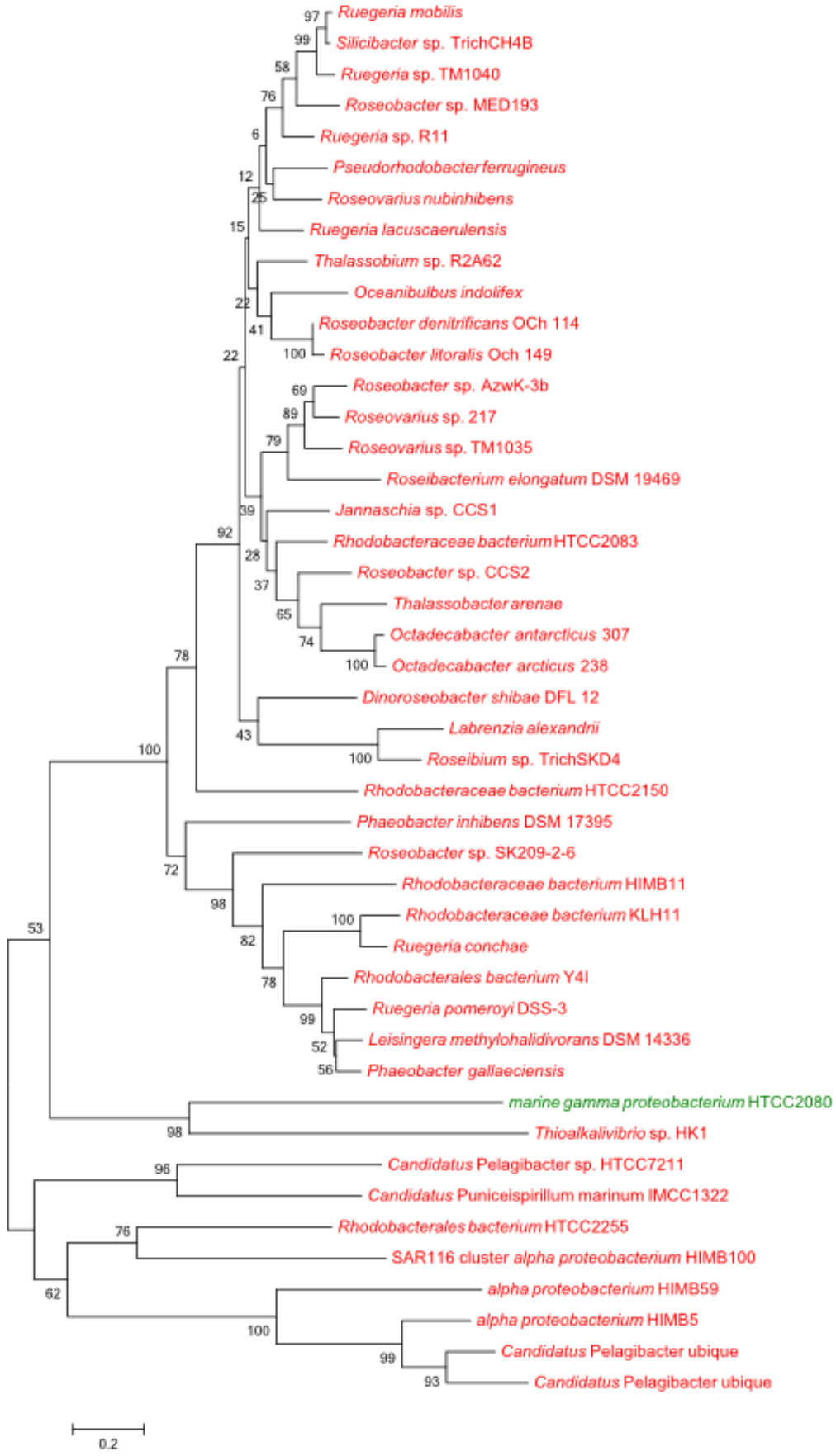

Ruegeria mobilis
Silicibacter sp. TrichCH4B
Ruegeria sp. TM1040
Roseobacter sp. MED193
Ruegeria sp. R11
Pseudorhodobacter ferrugineus
Roseovarius nubinhibens
Ruegeria lacuscaerulensis
Thalassobium sp. R2A62
Oceanibulbus indolifex
Roseobacter denitrificans OCh 114
Roseobacter litoralis Och 149
Roseobacter sp. AzwK-3b
Roseovarius sp. 217
Roseovarius sp. TM1035
Roseibacterium elongatum DSM 19469
Jannaschia sp. CCS1
Rhodobacteraceae bacterium HTCC2083
Roseobacter sp. CCS2
Thalassobacter arenae
Octadecabacter antarcticus 307
Octadecabacter arcticus 238
Dinoroseobacter shibae DFL 12
Labrenzia alexandrii
Roseibium sp. TrichSKD4
Rhodobacteraceae bacterium HTCC2150
Phaeobacter inhibens DSM 17395
Roseobacter sp. SK209-2-6
Rhodobacteraceae bacterium HIMB11
Rhodobacteraceae bacterium KLH11
Ruegeria conchae
Rhodobacterales bacterium Y4I
Ruegeria pomeroyi DSS-3
Leisingera methylohalidivorans DSM 14336
Phaeobacter gallaeciensis
marine gamma proteobacterium HTCC2080
Thioalkalivibrio sp. HK1
Candidatus Pelagibacter sp. HTCC7211
Candidatus Puniceispirillum marinum IMCC1322
Rhodobacterales bacterium HTCC2255
SAR116 cluster alpha proteobacterium HIMB100
alpha proteobacterium HIMB59
alpha proteobacterium HIMB5
Candidatus Pelagibacter ubique
Candidatus Pelagibacter ubique
0.2

Figure 1.6 Phylogenetic tree of DmdA polypeptides. Protein sequences of DmdA homologues were aligned with MEGA 6 and used to estimate an unrooted phylogenetic tree using *LG model*, *Gamma-Distributed*. The scale bar indicates number of substitutions per site. Bootstrap values of 500 replicates are given at the base of each branch pair. Species names are coloured according to their taxonomic status: α-proteobacteria (red); γ-proteobacteria (green).

1.9 Alternative pathways of MMPA degradation

Prior to the study by Reisch *et al*. (2011) described above, it was thought that MMPA produced from DMSP was either demethylated further to 3-mercaptopropionate (3-MPA) or demethiolated to produce MeSH. The latter route could occur via a simple elimination reaction, or reductive cleavage step, yielding acrylate or propionate respectively as the secondary catabolite (Taylor and Gilchrist, 1991). Evidence for the sequential demethylation of DMSP, first to MMPA and then to 3-MPA was produced in early studies of DMSP degradation in anoxic coastal marine sediment slurries. These slurries were incubated with DMSP, which caused an increase in the concentration of both methanethiol, and 3-MPA. Both molecules were presumed to be derived from MMPA, as addition of MMPA to the slurries had the same outcome (Kiene and Taylor, 1988). The production of 3-MPA from DMSP and MMPA was later shown in aerobic bacterial isolates, again accompanied by a production in MeSH (Taylor and Gilchrist, 1991). However, another study found that an aerobic methylotrophic bacterial strain named BIS-6 could grow on DMSP and MMPA producing 3-MPA, but never MeSH, in the process (Visscher and Taylor, 1994). Therefore, MMPA is not always degraded to MeSH and an organism may have the double demethylation pathway without any alternative volatile-producing route. As yet, no enzymes or genes have been identified which play a role in MMPA demethylation, or in the direct cleavage of MMPA to produce MeSH and much work remains to be done to determine if this pathway exists in any organism.

1.10 DMSP cleavage

The initially described ways in which DMSP could be catabolised involved a so-called "cleavage" reaction, in which the substrate was split into DMS and a C3 compound. This latter product was generally assumed (and sometimes confirmed) to be acrylate (though there is at least one instance in which 3-hydroxy propionate {3HP} is the C3 catabolite – see below).

In retrospect, the first hint of this process goes back to 1935 (Haas) who noted the red seaweed *Polysiphonia fastigiata* produced a 'penetrating, sickly odour' upon drying. Over a decade later, Challenger and Simpson (1948) identified the odorous gas arising from *P. fastigiata* as DMS, and showed that the gas originated from DMSP. Challenger and Simpson also revealed for the first time that the second product of DMSP 'cleavage' is acrylic acid.

One of the first DMSP catabolic reactions to be described was in a different species of red seaweed, *Polysiphonia lanosa* (Cantoni and Anderson, 1956). Extracts from *P. lanosa* were shown to cleave DMSP into DMS and acrylate, with high enzymatic activity. However, a problem with these early studies on DMSP lyase activity in seaweed extracts is that the seaweed samples were never confirmed to be axenic. Regardless, later studies working with axenic cultures have confirmed that some algal producers of DMSP can indeed catabolise it to DMS. For example, several strains of the coccolithophore *Emiliania huxleyi* (Steinke *et al*., 1998; Steinke *et al*., 2007), the Prymnesiophyte *Phaeocystis* (Stefels and van Boekel., 1993; Mohapatra *et al*., 2013) and the dinoflagellate species *Heterocapsa triquetra* and *Scrippsiella trochoidea* (Niki *et al*., 2000) have all been shown to possess DMSP lyase activity. DMSP lyase activity has also been demonstrated in extracts of green algae of the *Ulva* (previously *Enteromorpha*) genus (Steinke and Kirst, 1996), and a DMSP lyase enzyme has been purified from *Ulva curvata* (de Souza and Yoch, 1996a). Despite this, algal DMSP lyase enzymes remain poorly understood, and not a single gene encoding an algal DMSP lyase has been identified.

On the contrary, bacterial DMSP lyases have been well characterised and a diverse range of bacteria are known to degrade DMSP to DMS. The overwhelming majority of these bacteria inhabit marine environments, therefore it is somewhat ironic that the first bacterial isolate shown to produce DMS from DMSP was found in a freshwater river sediment. That isolate, a strain of *Clostridium propionicum* grows anaerobically on DMSP by fermenting it to DMS, propionate, acrylate, acetate, CO_2 and a proton (Wagner and Stadtman, 1962). The first marine bacterial DMSP degrader to be identified was also an anaerobe, isolated from intertidal sediments. This strain was named *Desulfovibrio acrylicus* on the basis of its ability to use the acrylate derived from DMSP, as a terminal electron acceptor (van der Maarel *et al*., 1996b).

Early studies of aerobic DMSP-cleaving bacteria were conducted mainly using four different strains: the β-proteobacterium *Alcaligenes faecalis*, isolated from the surface of a salt marsh (de Souza and Yoch, 1995a, 1995b); the γ-proteobacterium *Oceanimonas* (previously *Pseudomonas*) *doudoroffii*, isolated from oceanic waters (de Souza and Yoch, 1995a; de Souza and Yoch, 1996b); and two Roseobacter-related isolates, strain LFR (Ledyard and Dacey, 1994; Ledyard *et al*., 1993) and *Sagittula stellata* (González *et al*., 1999). All of these strains were shown to produce DMS from DMSP, and detailed molecular work on the DMSP lyases responsible for this phenotype were carried out in *A. faecalis* and *O. doudoroffii* (see below). Since these early studies, many more strains have been investigated for their ability to degrade DMSP to DMS, including representatives from all classes of proteobacteria and it has become clear that bacteria have a variety of ways to produce DMS from DMSP.

1.11 Molecular genetics reveal diversity of bacterial DMSP lyases

The first indications that different bacteria used different classes of enzymes to cleave DMSP, releasing DMS as one product, came from work in Yoch's laboratory. A comparison of the properties of DMSP lyases purified from *A. faecalis*, and *O. doudoroffii* revealed that these enzymes had different optimum conditions. For example, the *A. faecalis* lyase had two pH optimum peaks, at 6.5 and 8.8, whereas *O. doudoroffii* lyase had a single peak at pH 8.8. The K_m values for DMSP of the two lyases were also quite different, at 2 mM and <20 μM, for *A. faecalis* and *O. doudoroffii*, respectively (de Souza and Yoch, 1995b). Inhibitor studies also showed that the lyase activity was in a different cellular location in each strain, and it was predicted that *A. faecalis* had an extracellular DMSP lyase, whereas the activity in *O. doudoroffii* was likely cytosolic (de Souza and Yoch, 1996b). Recent work, all of it from our laboratory at UEA, has confirmed, and indeed extended the appreciation that there is a remarkable diversity of different enzymes, in different microbes and different sub-cellular compartments all of which can act on DMSP, releasing DMS as a product. Therefore, the term "DMSP lyase" is only used as useful shorthand and should not be viewed as a description of a particular polypeptide family.

To date, no fewer than six different DMSP lyases, in four wholly distinct polypeptide families have been described. These lyases were identified using the same general approach, as follows. Bacteria which produced DMS from DMSP were obtained, either directly from the environment, or from other laboratories. Then, cosmid libraries were made using the genomic material of the DMS-producing strains. The libraries were mobilised into suitable host species which provided a null-background in which to screen for DMSP-dependent DMS production, or for growth on DMSP as a sole source of carbon. To screen for DMS production, individual colonies of the

library-containing host were grown in the presence of DMSP in sealed vials. The headspace of the vials was then assayed for DMS production using gas chromatography. Those cosmids which conferred DMSP-dependent DMS production (Ddd^+) on the host were isolated, and the gene causing the Ddd^+ phenotype identified through sub-cloning from the cosmid. The following section describes our knowledge on six different "Ddd" enzymes and the corresponding *ddd* genes identified in this manner: *dddD, dddL, dddP, dddQ, dddY* and *dddW*.

1.12 DddY

Despite the fact that the DMSP lyase gene *dddY* was the fifth such gene to be discovered, it is a good place to start since it already had something of a history before it was discovered in 2011 (Curson *et al*.). As mentioned above, some of the early work on DMSP lyases was conducted in the salt marsh β-proteobacterium *Alcaligenes faecalis* M3A, by Yoch's group. They had shown that this strain could grow on DMSP, releasing DMS (Ansede *et al*., 1999) and had purified the DMSP lyase that was responsible for this phenotype. Remarkably, they even managed to obtain a partial N-terminal amino acid sequence for the purified lyase (de Souza and Yoch, 1996b). In addition, they had evidence that this DMSP lyase might be associated with the cell surface, unlike the other, cytoplasmic, enzymes that cleaved DMSP in other bacteria (Ansede *et al*., 1999; see below). Some years after these studies, an analysis of the genetic basis of DMSP catabolism in *A. faecalis* M3A revealed that the DMSP lyase was encoded by the *dddY* gene, a finding which fully supported the biochemical data from the Yoch lab.

The *dddY* gene was discovered by searching for the DMSP catabolism genes that allowed *A. faecalis* to use DMSP as a sole carbon source. A genomic library of *A. faecalis* was mobilised into *Pseudomonas putida*, chosen because it has many sigma factors and so may be able to express introduced "foreign" genes. The transconjugants were screened for their ability to grow on DMSP as a sole source of carbon (Curson *et al*., 2011). Two such transconjugants were selected and, upon sequencing, found to contain a cluster of eight genes, five of which were homologues of other *ddd* genes already linked to DMSP catabolism in other bacteria (see below), plus a novel gene, termed *dddY*. The *dddY* gene was cloned and expressed in *E. coli*, where it was shown to break DMSP down to DMS and acrylate, and significantly that the DMSP lyase activity was much higher in the periplasmic fraction (Curson *et al*., 2011).

Reassuringly, the deduced peptide sequence of DddY is predicted to encode a leader sequence which would guide it to the bacterial periplasm, consistent with earlier findings that suggested the DMSP lyase in *A. faecalis* is associated with the cell surface (Ansede *et al*., 1999). Even more reassuringly, if the leader sequence of DddY is cleaved at the predicted site (21 amino

acids from the N-terminus) then the resultant peptide has an N-terminal sequence exactly matching the sequence found for the purified lyase by Yoch (de Souza and Yoch, 1996b). Thus, it was clear that the *dddY* gene encoded the DMSP lyase described earlier by Yoch.

The DddY polypeptide is predicted to have a molecular weight of 45.5 kDa, similar to the 48 kDa protein purified from *A. faecalis* by de Souza and Yoch (1995a). It has no similarity whatsoever to any polypeptide of known function, or any predicted domain features.

1.12.1 Distribution of DddY

Homologues of DddY are found in species of β-, γ-, δ-, and ε-proteobacteria, as well as one species of Flavobacterium–*Gramella portivictoriae* (**Figure 1.7**). It is the only DMSP lyase, to date, which is not found in the deduced proteome of any sequenced α-proteobacteria. Almost all strains with *dddY* have been isolated from marine environments, like other DMSP-lyase containing bacteria (see below), although *Acinetobacter bereziniae* was reportedly isolated from a hospital environment (Nemec *et al*., 2010). Interestingly though, the isolates are rarely from open water samples. Many of the strains, such as *G. portivictoriae*, *Ferrimonas balearica*, *F. futtsuensis*, *Shewanella piezotoleans*, *S. fidelis* were isolated from marine sediment (Lau *et al*., 2014; Rosselló-Mora *et al*., 1995; Nakagawa *et al*., 2006; Xiao *et al*., 2007; Ivanova *et al*., 2003a). Others, like *Candidatus Accumulibacter* and *Acinetobacter baylyi* originated from activated sludge (Albertsen *et al*., 2012; Carr *et al*., 2003). Some strains were even isolated from sea creatures, such as *Ferrimonas kyonanensis* which was isolated from the alimentary tract of a little neck clam (Nakagawa *et al*., 2006), *F. senticii* which was found in the mucus of a puffer fish (Campbell *et al*., 2007), *Shewanella waksmanii* which was cultured from a marine worm (Ivanova *et al*., 2003b). Interestingly, *Arcobacter nitrofigilis*, the only strain with two *dddY* homologues, was isolated from salt marsh sediment associated with *Spartina alterniflora* (McClung and Patriquin., 1980). Given that *A. faecalis* was also isolated from such an environment, there is at least one example of bacteria from two different classes containing the same gene, living in similar habitats. Excitingly, this strain of *Arcobacter* was also confirmed to have a Ddd^+ phenotype (Curson *et al*., 2011), although neither *dddY* gene from *A. nitrofigilis* has been ratified as functional.

Strikingly, *dddY* is found in many different species of the γ-proteobacterium *Shewanella* (**Figure 1.7**), but not all. Two strains, *S. putrefaciens* and *S. halifaxensis* have been verified as having Ddd^+ activity, and *dddY* from the former strain has been cloned and shown to confer a Ddd^+ to *E. coli* (Curson *et al*., 2011). Significantly, *Shewanella oneidensis*, which does not have a homologue of *dddY*, did not make DMS from DMSP (Curson *et al*., 2011).

It is intriguing that there is a distinct lack of *dddY*-strains from open water environments, and that many of the environments that *dddY*-containing strains are isolated from are anoxic or microaerobic in nature. In addition, *dddY* often appears nearby to genes encoding membrane-bound cytochromes (Curson *et al*., 2011). These observations hint at the possibility of DddY being involved in anaerobic respiration. Connected to this theory, *Desulfovibrio acrylicus* uses acrylate as an electron acceptor in anaerobic respiration and also has a DddY homologue (van der Maarel *et al*., 1996a; van der Maarel *et al*., 1996b).

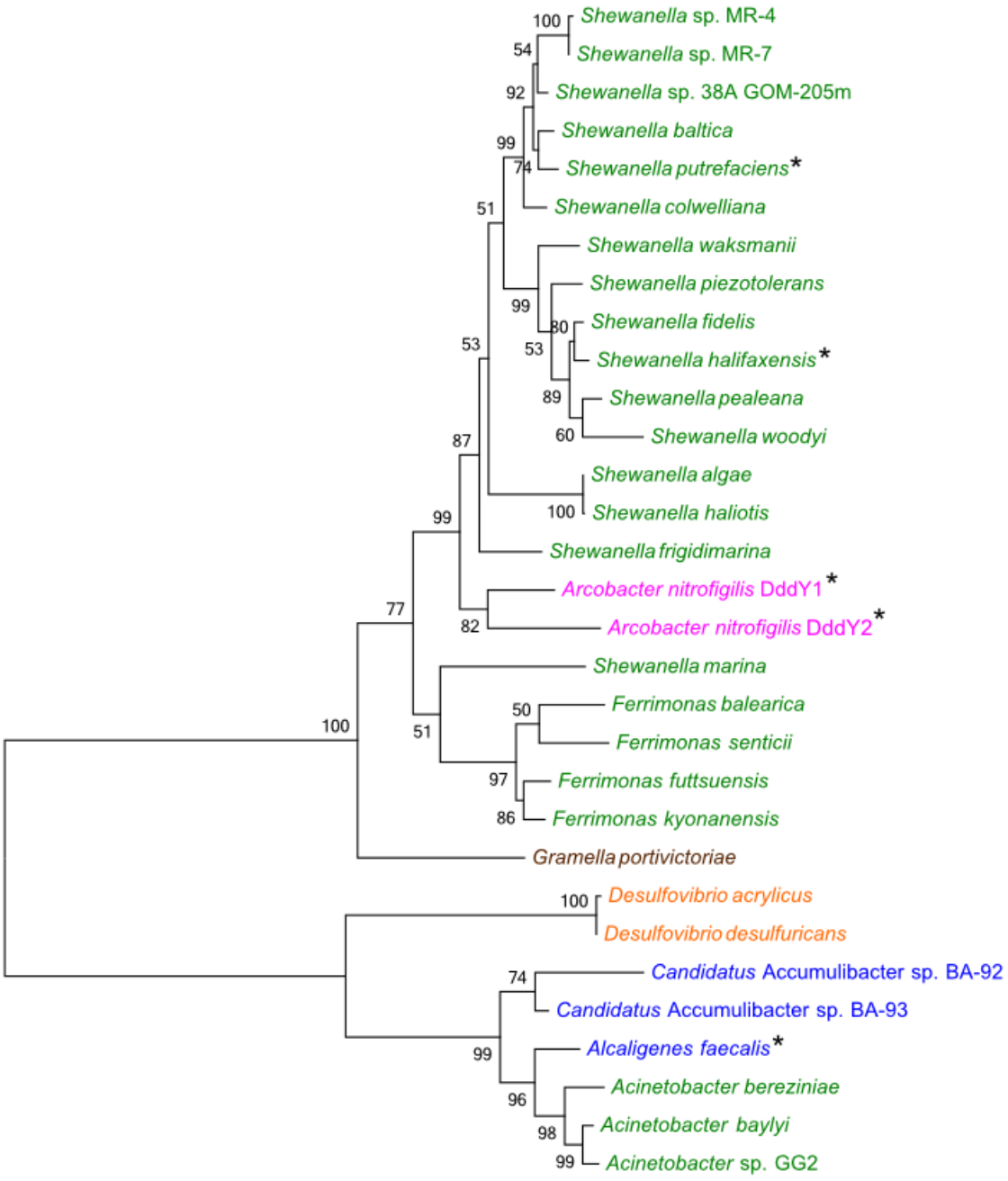


Figure 1.7 Phylogenetic tree of DddY polypeptides. Protein sequences of DddY homologues were aligned with MEGA 6 and used to estimate an unrooted phylogenetic tree using LG model, gamma distributed with invariant sites. The scale bar indicates number of substitutions per site. Bootstrap values of 500 replicates are given at the base of each branch pair. Species names are coloured according to their taxonomic status: γ-proteobacteria (green); β-proteobacteria (blue); ε-proteobacteria (pink) and δ-proteobacteria (orange); Flavobacteriales (brown). *Strains confirmed as producing DMS from DMSP.

It is worth re-iterating that *dddY* was discovered in a cosmid which conferred on *P. putida* the ability to use DMSP as a sole carbon source, something that *A. faecalis* does itself. The cosmid contained an operon of genes, of which *dddY* was a part, which have been shown to encode a pathway of DMSP degradation to acetaldehyde. These genes, *dddA*, *dddC*, *acuN* and *acuK* will be presented in detail later (see **Section 1.16**), but it is important to mention it here because this cluster of acrylate catabolism genes do appear next to another, completely different DMSP lyase gene called *dddD*, and bacterial strains containing *dddD* are also noted for their ability to use DMSP as a sole carbon source.

1.13 DddD

The *dddD* gene was the first DMSP lyase gene to be discovered. Thus it was slightly ironic that the encoded DddD enzyme was actually not a DMSP lyase in the "classical" sense. Instead of cleaving DMSP to DMS and acrylate, DddD converts DMSP to DMS and 3-hydroxypropionate (3HP) (see below).

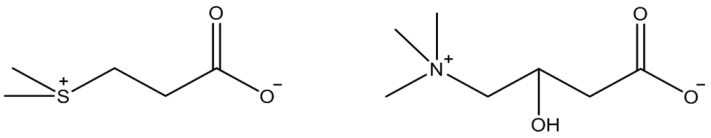
1.13.1 Discovery of the *dddD* gene

DddD was discovered in the γ-proteobacterium *Marinomonas* sp. MWYL-1. This strain was isolated from the rhizosphere of the salt marsh grass *Spartina anglica*, and was selected on the basis of its ability to grow well with DMSP as a sole carbon source, producing DMS in the process. The *dddD* gene was identified using the method described above, by mobilising a genomic library of MWYL-1 into *E. coli*. Since *E. coli* does not produce DMS from DMSP, it provided a null background in which to screen MWYL-1 library fosmids. A fosmid that conferred a Ddd^+ phenotype on *E. coli* was extracted and sequenced.

By sub-cloning fragments from the cosmid, it was found that a single gene, termed *dddD*, was sufficient to confer a Ddd^+ phenotype to *E. coli*, as long as an active promoter that functioned in

this host background was supplied in the cloning vector. And, in agreement with this, an insertional mutation in *dddD* completely abolished the ability of the mutant to make DMS and to grow on DMSP (Todd *et al*., 2007).

Earlier studies of DMSP lyase activity had shown (or predicted) that the production of DMS was via a simple cleavage step producing acrylate as the secondary catabolite. Therefore it was somewhat surprising that the sequence of DddD placed it in a family of type III acyl-CoA transferases. Its closest homologue (26% identity) with known function is *E. coli* CaiB – a γ-butyrobetainyl-CoA: carnitine CoA-transferase that mediates the addition of acyl-CoA to carnitine (an amino acid with structural similarity to DMSP, see **Figure 1.8**). Interestingly, the CaiB protein of *E. coli* is a homodimer of two separate CaiB polypeptides (Elssner *et al*., 2001), whereas the ~93 kDa DddD polypeptide contains two CaiB domains separated by a linker region, suggesting DddD acts as a form of "intra-molecular" dimer.



DMSP Carnitine

Figure 1.8 DMSP and its structural analogue, carnitine

In *E. coli*, CaiB mediates the transfer of CoA molecules from *L*-carnitinyl-CoA and crotonobetainyl-CoA to *L*-carnitine (**Figure 1.9**; Elssner *et al*., 2001). The similarity of DddD to CaiB, led to the prediction that DddD could act as a CoA transferase, in addition to a lyase (Todd *et al*., 2010a). Further studies on another γ-proteobacterium, *Halomonas* HTNK1, that also catabolises DMSP via a DddD enzyme, revealed that DddD is distinct from the other DMSP lyases. HPLC and ^{13}C-NMR analysis were employed to show directly that 3HP was a catabolite of DddD-mediated DMSP degradation, not acrylate as seen for the other lyases (Todd *et al*., 2010a). Unfortunately, the anticipated CoA intermediates were not seen in this work, but this was rectified in later studies on purified DddD from *Marinomonas* MWLY1 (Alcolombri *et al*., 2014). The authors confirmed that DddD acts as a CoA-transferase and lyase, converting DMSP and acetyl-CoA to DMS, acetate and 3HP-CoA (**Figure 1.10**). Although acetyl-CoA seemed to be the major CoA donor, they also showed that DddD can use the 3HP-CoA intermediate at a slower rate, releasing free 3HP in the process. This situation is analogous to CaiB, which uses the *L*-carnitinyl-CoA intermediate generated from *L*-carnitine.

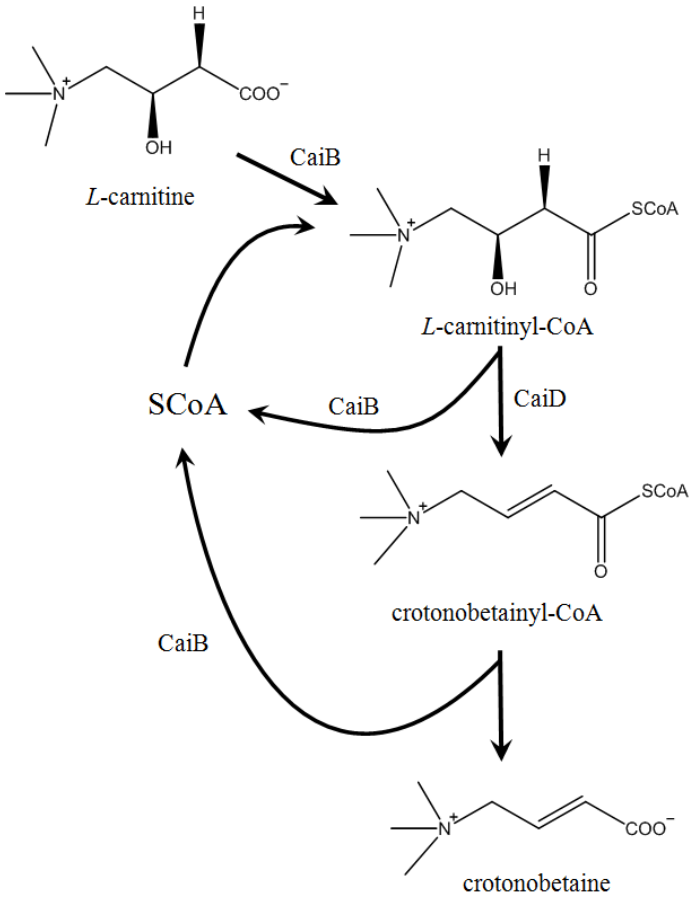


Figure 1.9 CaiB-mediated transfer of coenzyme A to carnitine. The enoyl-CoA hydratase (CaiD) dehydrates carnitinyl-CoA to crotonobetainyl-CoA, which is subsequently converted to crotonobetaine. Both *L*-carnitinyl-CoA and crotonobetainyl-CoA can serve as CoA donors for the action of CaiB, which transfers a CoA to the carboxyl group of *L*-carnitine to form carnitinyl-CoA. Adapted from Elssner *et al.*, 2001.

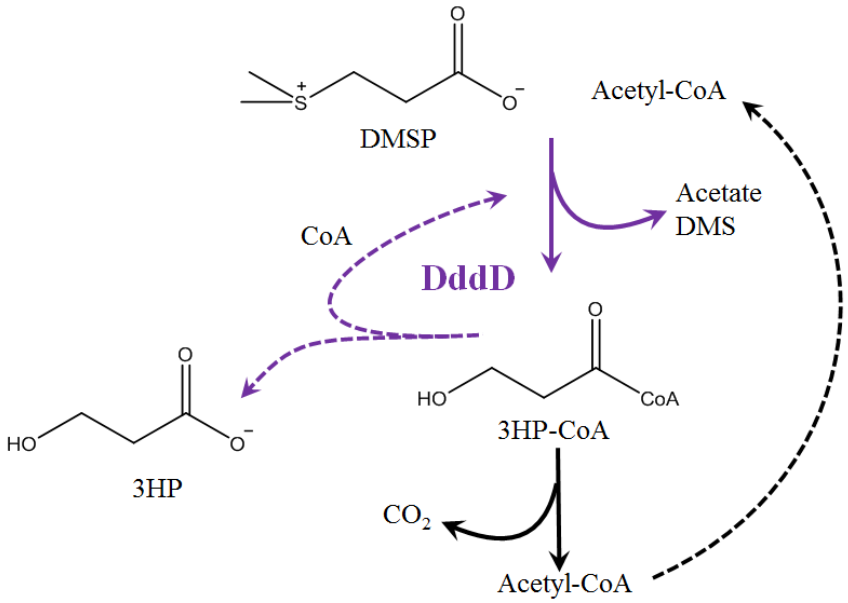


Figure 1.10 Predicted action of the DMSP-CoA transferase DddD. Reactions mediated by DddD are shown in purple. DddD acts as a CoA transferase and lyase, releasing 3-hydroxypropionate-CoA (3HP-CoA) and DMS from DMSP, using acetyl-CoA as a CoA donor. 3HP-CoA itself may act as a CoA donor in a slower reaction, yielding 3HP. Alternatively, 3HP-CoA is eventually converted to acetyl-CoA which is recycled as a substrate for DddD (Alcolombri *et al*., 2014).

1.13.2 Distribution of DddD in different organisms

Close homologues (42-76% identity and containing the CaiB-CaiB duplex domains) of DddD are found in phylogenetically and ecologically diverse groups of bacteria, including species of α- β- and γ-proteobacteria (**Figure 1.11**). Crucially, most of these bacteria reside in marine environments, and would likely be exposed to DMSP. For example, DddD homologues are found in several species of γ-proteobacteria, which were isolated from algae and salt marsh grasses, *Enterovibrio* spp., isolated from the gut of Turbot larvae (Thompson *et al*., 2002), β-proteobacterium MOLA814 which was isolated from the Canadian Beaufort Sea (Courties *et al*., 2013) and a deep sea worm symbiont, *Osedax symbiont* Rs2. Importantly, DddD is often found in those bacteria capable of using DMSP as a sole carbon source. In addition to *Marinomonas* MWYL-1, DddD is also found in *Halomonas* HTNK1, *Pseudomonas* J465 and *Psychrobacter* J466 which were all isolated on the basis of growth on DMSP as a sole carbon source (Todd *et al*., 2010a; Curson *et al*., 2010). In each of these isolates, the *dddD* gene is found in a cluster of genes involved in DMSP transport and the catabolism of 3HP, and in the case of *Halomonas*, acrylate. The genes involved in downstream DMSP catabolism will be presented in more detail in **Section 1.16**.

DddD homologues are also found sporadically amongst members of the Roseobacter clade (see **Chapter 5** for a more detailed discussion), and, intriguingly, also in some species of *Rhizobium* and *Burkholderia* (a rhizophore bacterium) which are known to colonise a wide range of legume and other plant hosts. Significantly, DddD-containing *Sinorhizobium fredii* NGR234 and *Burkholderia cepacia* AMMD both produced DMS from DMSP, whereas a selection of other *Rhizobium* and *Burkholderia* strains that lacked *dddD* did not (Todd *et al*., 2007). In addition, DddD enzymes from NGR234 and *B. cepacia* were cloned and expressed in *E. coli* where they conferred a Ddd^+ phenotype (Todd *et al*., 2007). This was unexpected, since there are very few terrestrial examples of DMSP-catabolising bacteria, but excitingly it might mean that these DddD-containing bacteria have unknown DMSP-producing angiosperm hosts.

Interestingly, DddD is not found uniformly among all members of any single genus. This was the case for the *Burkholderia* and *Rhizobium* mentioned above. For example, several closely related species of *Burkholderia* have had their genomes sequenced, but only some species contain a homologue of DddD (*Burkholderia* sp. WSM14176, *Burkholderia vietnamiensis*, *Burkholderia ambifaria*, and *Burkholderia phymatum*). Similarly, within the many sequenced strains of Rhizobiaceae, *Sinorhizobium fredii* NGR234, *Rhizobium leguminosarum* and *Agrobacterium tumefaciens* are among those containing DddD, but other members of this family do not contain a *dddD* gene.

This type of distribution is consistent with the *dddD* gene having undergone several rounds of horizontal gene transfer. Intriguingly, there is even a reasonably convincing DddD homologue (22% identical to *Marinomonas* MWYL-1 DddD) in a eukaryote, the coccolithophore *Emiliania huxleyi*, so such transfer may even span the boundaries of different domains.

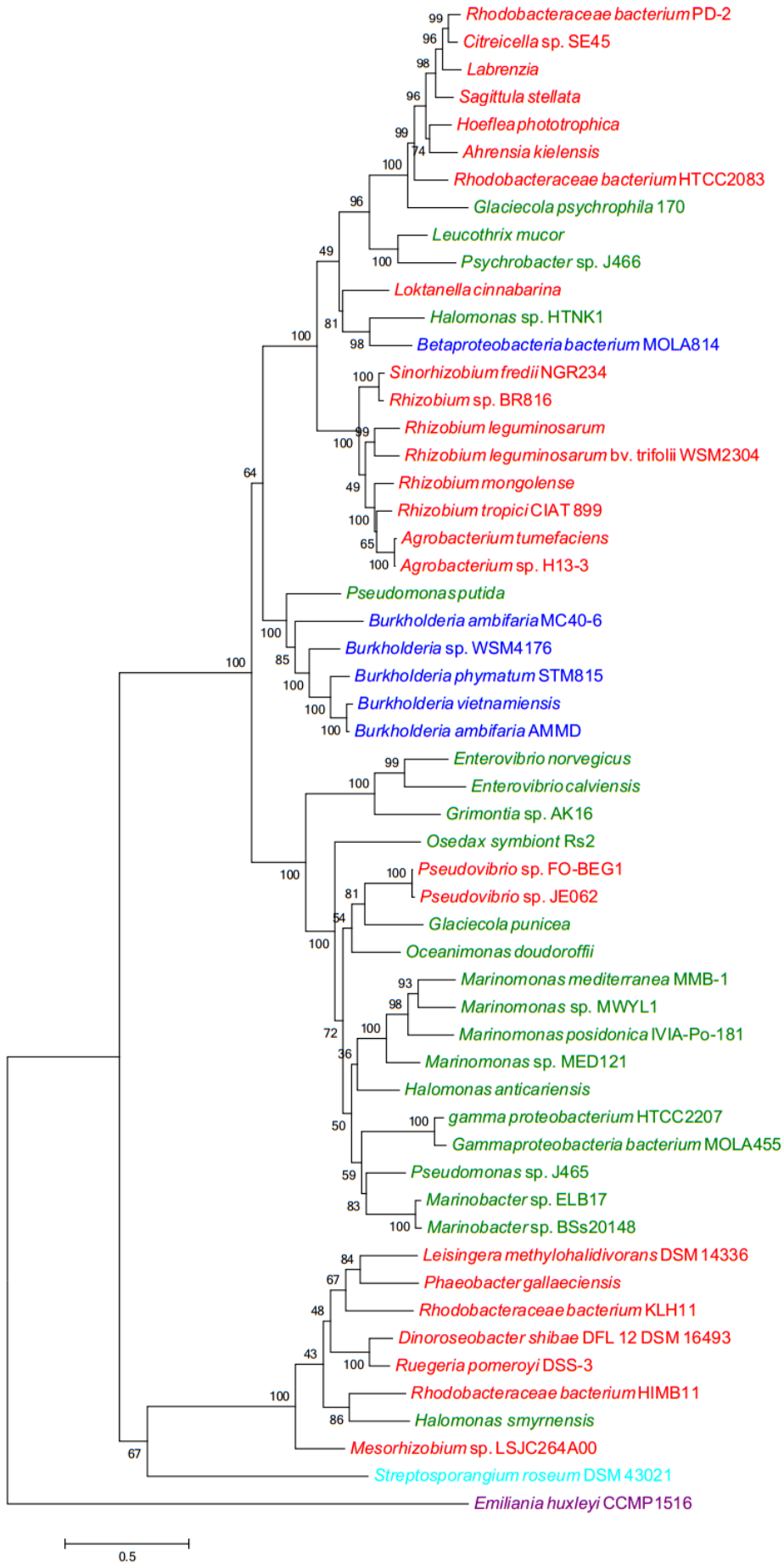

Rhodobacteraceae bacterium PD-2
Citreicella sp. SE45
Labrenzia
Sagittula stellata
Hoeflea phototrophica
Ahrensia kielensis
Rhodobacteraceae bacterium HTCC2083
Glaciecola psychrophila 170
Leucothrix mucor
Psychrobacter sp. J466
Loktanella cinnabarina
Halomonas sp. HTNK1
Betaproteobacteria bacterium MOLA814
Sinorhizobium fredii NGR234
Rhizobium sp. BR816
Rhizobium leguminosarum
Rhizobium leguminosarum bv. trifolii WSM2304
Rhizobium mongolense
Rhizobium tropici CIAT 899
Agrobacterium tumefaciens
Agrobacterium sp. H13-3
Pseudomonas putida
Burkholderia ambifaria MC40-6
Burkholderia sp. WSM4176
Burkholderia phymatum STM815
Burkholderia vietnamiensis
Burkholderia ambifaria AMMD
Enterovibrio norvegicus
Enterovibrio calviensis
Grimontia sp. AK16
Osedax symbiont Rs2
Pseudovibrio sp. FO-BEG1
Pseudovibrio sp. JE062
Glaciecola punicea
Oceanimonas doudoroffii
Marinomonas mediterranea MMB-1
Marinomonas sp. MWYL1
Marinomonas posidonica IVIA-Po-181
Marinomonas sp. MED121
Halomonas anticariensis
gamma proteobacterium HTCC2207
Gammaproteobacteria bacterium MOLA455
Pseudomonas sp. J465
Marinobacter sp. ELB17
Marinobacter sp. BSs20148
Leisingera methylohalidivorans DSM 14336
Phaeobacter gallaeciensis
Rhodobacteraceae bacterium KLH11
Dinoroseobacter shibae DFL 12 DSM 16493
Ruegeria pomeroyi DSS-3
Rhodobacteraceae bacterium HIMB11
Halomonas smyrnensis
Mesorhizobium sp. LSJC264A00
Streptosporangium roseum DSM 43021
Emiliania huxleyi CCMP1516
0.5

Figure 1.11 Phylogenetic tree of DddD polypeptides. Protein sequences of DddD homologues were aligned with MEGA 6 and used to estimate an unrooted phylogenetic tree using LG model, gamma distributed with invariant sites. The scale bar indicates number of substitutions per site. Bootstrap values of 500 replicates are given at the base of each branch pair. Species names are coloured according to their taxonomic status: α-proteobacteria (red); γ-proteobacteria (green); β-proteobacteria (blue); Actinobacteria (light blue) and Prymesiophycaea (purple).

1.14 DddP

The DddY and DddD lyases were discovered in bacteria which reside in marine sediments surrounding *Spartina* plants. In contrast, another, completely different DMSP lyase that cleaves DMSP to DMS and acrylate was discovered in the α-proteobacterium *Roseovarius nubinhibens* ISM, a member of the abundant Roseobacter clade, which mostly reside in the open ocean (see **Section 1.20.1**). It was noted that *R. nubinhibens* had a Ddd^+ phenotype, but its genome did not encode a homologue of any known DMSP lyase at the time. Again, the lyase was identified through the screening of a genomic library of *R. nubinhibens* for any cosmids that conferred a Ddd^+ phenotype to the heterologous host, which this time was the α-proteobacterium *Rhizobium leguminosarum*. The newly described DMSP lyase gene was termed *dddP* (Todd *et al*., 2009).

1.14.1 DddP is a member of the metallopeptidase family

DddP is a ~50 kDa polypeptide in the PepPXaa-Pro aminopeptidase metalloenzyme family. As expected from its name, members of this family generally cleave peptides, but there are some which act on non-peptide substrates. For example, the creatinase of *Pseudomonas putida* catalyses the cleavage of creatine and water to urea and sarcosine (Bazan *et al*., 1994). Therefore, DddP is unusual but not unique in cleaving a non-peptide. Enzymes of the metallopeptidase family contain an active site with a binuclear metal centre, and require metal cofactors such as cobalt, manganese, zinc, iron or nickel (Bazan *et al*., 1994; Schiffmann *et al*., 2006). In accordance with this, DddP polypeptides have five conserved amino acids predicted to form the metal binding sites in the active site of metallopeptidases (Schiffmann *et al*., 2006; Todd *et al*., 2009), and site directed mutations in those residues abolished DMSP lyase activity (Kirkwood *et al*., 2010a). Very recent structural studies using X-ray crystallography of DddP from another Roseobacter, *Roseobacter denitrificans*, revealed that DddP does indeed have a metallopeptidase-like fold, and furthermore it binds Fe^{2+} in its active site (Hehemann *et al*., 2014). Therefore, DddP is not a peptidase, but it does require iron as a metal co-factor, a characteristic of the metallopeptidase family.

The *R. nubinhibens* DddP protein was expressed and purified from *E. coli*, and shown to be a homodimeric protein, with a K_m of 14 mM for DMSP, and a V_{max} of 0.31 nmol DMS min^{-1} μg $protein^{-1}$. Although this is a relatively high K_m value, it is comparable to values obtained for DmdA in *R. pomeroyi* and *Candidatus* P. ubique, and suggests that *R. nubinhibens* may also accumulate DMSP to high internal concentrations (Kirkwood *et al*., 2010a).

Significantly, a mutation in *dddP* of *R. nubinhibens* severely reduced, but did not abolish, the ability of this bacterium to make DMS from DMSP. This prompted a search for a second DMSP lyase in this strain, which will be discussed later (**see Section 1.15.5**).

1.14.2 Distribution of DddP

Homologues of DddP are somewhat more abundant than other DMSP lyases, particularly so in the α-proteobacteria (**Figure 1.12**). Of the 42 sequenced Roseobacter strains, 22 have a homologue of DddP. DddP homologues are also found in members of the abundant SAR11 and SAR116 clades (see **Chapter 2**) and some γ-proteobacteria, including *Oceanimonas doudoroffi* which has two copies (see **Chapter 3**). Remarkably, homologues of DddP are also found in some species of fungi, notably within the *Aspergillus* and *Fusarium* genera. Some of these DddP-containing fungi were shown directly to produce DMS when grown in the presence of DMSP (Todd *et al*., 2009). Significantly, other closely related species lacking DddP were also tested; these did not possess DMSP lyase activity. To verify that fungi contained functional copies of DddP, the genes encoding this lyase were amplified from *Fusarium graminearum* cc19 and *Fusarium culmorum* Fu42, cloned into an expression vector and expressed in *E. coli*. Both copies of *dddP* conferred a Ddd^+ phenotype on *E. coli* (Todd *et al*., 2009). So, these fungi likely acquired *dddP* from bacteria through inter-domain HGT, and the fact that fungal DddPs are intron-less supports this idea. Although exciting, the finding that fungi were able to make DMS from DMSP was not unprecedented. In 1998, Yoch's group isolated the fungus *Fusarium lateritium* from seawater and salt marsh due to its ability to grow on DMSP. Using studies *in vivo*, they found that *F. lateritium* had DMSP lyase activity with a K_m of 1.2 mM and a V_{max} of 34.7 μmol min^{-1} mg $protein^{-1}$ (Bacic and Yoch, 1998). It may be that DddP confers a selective advantage to fungi which form close associations with DMSP-producing plants and other organisms. Bacic and Yoch hypothesised DMSP-lyase containing fungi could play an important role in the degradation of DMSP producers, such as macroalgae and salt marsh grasses (Bacic and Yoch, 1998). Indeed an opportunistic pathogen of corals, *Aspergillus sydowii*, also has a functional copy of DddP (Kirkwood *et al*., 2010b).

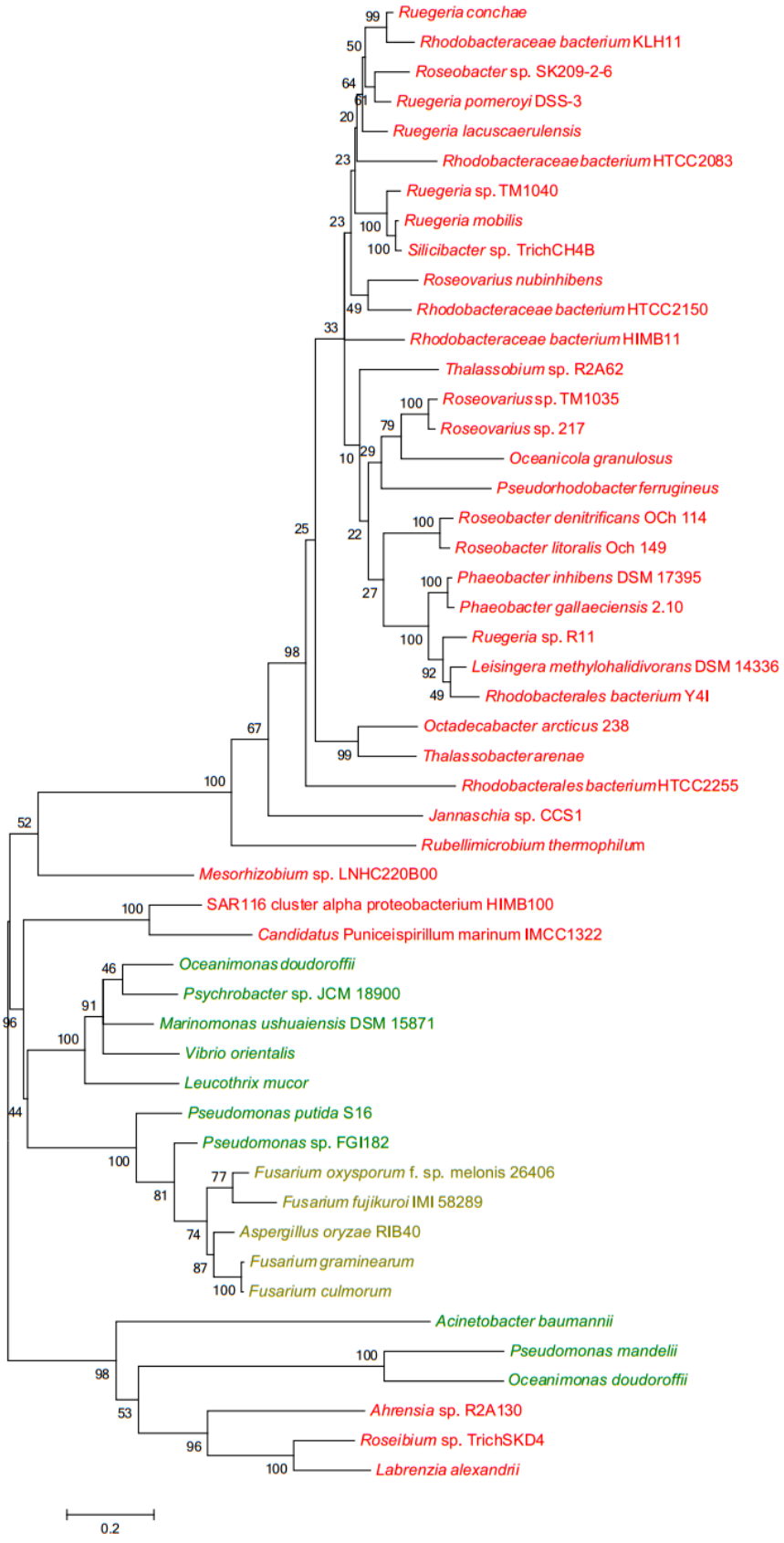

Ruegeria conchae
Rhodobacteraceae bacterium KLH11
Roseobacter sp. SK209-2-6
Ruegeria pomeroyi DSS-3
Ruegeria lacuscaerulensis
Rhodobacteraceae bacterium HTCC2083
Ruegeria sp. TM1040
Ruegeria mobilis
Silicibacter sp. TrichCH4B
Roseovarius nubinhibens
Rhodobacteraceae bacterium HTCC2150
Rhodobacteraceae bacterium HIMB11
Thalassobium sp. R2A62
Roseovarius sp. TM1035
Roseovarius sp. 217
Oceanicola granulosus
Pseudorhodobacter ferrugineus
Roseobacter denitrificans OCh 114
Roseobacter litoralis Och 149
Phaeobacter inhibens DSM 17395
Phaeobacter gallaeciensis 2.10
Ruegeria sp. R11
Leisingera methylohalidivorans DSM 14336
Rhodobacterales bacterium Y4I
Octadecabacter arcticus 238
Thalassobacter arenae
Rhodobacterales bacterium HTCC2255
Jannaschia sp. CCS1
Rubellimicrobium thermophilum
Mesorhizobium sp. LNHC220B00
SAR116 cluster alpha proteobacterium HIMB100
Candidatus Puniceispirillum marinum IMCC1322
Oceanimonas doudoroffii
Psychrobacter sp. JCM 18900
Marinomonas ushuaiensis DSM 15871
Vibrio orientalis
Leucothrix mucor
Pseudomonas putida S16
Pseudomonas sp. FGI182
Fusarium oxysporum f. sp. melonis 26406
Fusarium fujikuroi IMI 58289
Aspergillus oryzae RIB40
Fusarium graminearum
Fusarium culmorum
Acinetobacter baumannii
Pseudomonas mandelii
Oceanimonas doudoroffii
Ahrensia sp. R2A130
Roseibium sp. TrichSKD4
Labrenzia alexandrii
0.2

Figure 1.12 Phylogenetic tree of DddP polypeptides. Protein sequences of DddP homologues were aligned with MEGA 6 and used to estimate an unrooted phylogenetic tree using LG model, gamma distributed with invariant sites. The scale bar indicates number of substitutions per site. Bootstrap values of 500 replicates are given at the base of each branch pair. Species names are coloured according to their taxonomic status: α-proteobacteria (red); γ-proteobacteria (green) and Ascomycota (olive).

1.15 The Cupin DMSP lyases, DddL, DddQ and DddW.

Whereas DddY, DddD and DddP are of completely different polypeptide families to one another, the remaining three DMSP lyases share a common domain. These small lyases, termed DddL (~26 kDa), DddQ (~22 kDa) and DddW (~16 kDa), all have a conserved C-terminal β-barrel, known as a cupin ('*cupa*' is Latin for small barrel) domain. Cupin domains usually bind transition metals, and are found in a diverse range of polypeptides that are equally diverse in function (Dunwell *et al*., 2004).

1.15.1 DddL

The first of the cupin-type DMSP lyases, named DddL, was discovered in the marine α-proteobacterium *Sulfitobacter* EE-36. This strain was known to have a Ddd^+ phenotype (González *et al*., 1999), but a search of its genome sequence did not reveal any known DMSP lyase. Therefore, it was supposed that *Sulfitobacter* EE-36 must use a different DMSP degradation pathway.

As for *dddP* and *dddD*, the *dddL* gene was identified through the screening of a *Sulfitobacter* EE-36 genomic cosmid library for any cosmids that conferred DMSP-dependent DMS production, to a "null" bacterial recipient, which, this time, was a strain of *Rhizobium leguminosarum.* One such cosmid was obtained and a single gene was shown to be required and sufficient for conferring the Ddd^+ phenotype to *Rhizobium*. A *dddL* insertional mutation in the genome of *Sulfitobacter* itself completely abolished its Ddd^+ phenotype. Furthermore, when DddL was cloned alone, under the control of a constitutive promoter in a plasmid vector and introduced into *E. coli*, the resulting strain generated DMS from DMSP and also formed equimolar amounts of acrylate, as assayed by HPLC (Curson *et al*., 2008).

1.15.2 Distribution of DddL

In contrast to DddD, DddY and DddP, homologues of DddL are mainly restricted to one taxonomic branch of the α-proteobacteria, namely the Rhodobacterales (**Figure 1.13**). This bacterial family includes all the Roseobacters, as well as few other marine genera, including the much-studied *Rhodobacter*. Significantly, the "terrestrial" genera of Rhodobacterales, including *Paracoccus* spp., do not have a DddL homologue. Outside of the Rhodobacterales, DddL is only found sporadically in two species of *Marinobacter* (γ-proteobacteria) and an actinobacterium *Serinicoccus marinus*, thus it would appear that the *dddL* gene has undergone only limited HGT, in terms of taxonomic distance, and the environments inhabited by these bacteria that contain it. All bacteria containing DddL were isolated from marine, open water environments. In addition to *Sulfitobacter* sp. EE-36, the DddL from *R. sphaeroides* 2.4.1 and *Dinoroseobacter shibae* DFL-12 have also been confirmed as functional, although interestingly *D. shibae* did not make DMS from DMSP under laboratory conditions (Curson *et al*., 2012; discussed further in **Chapter 5**).

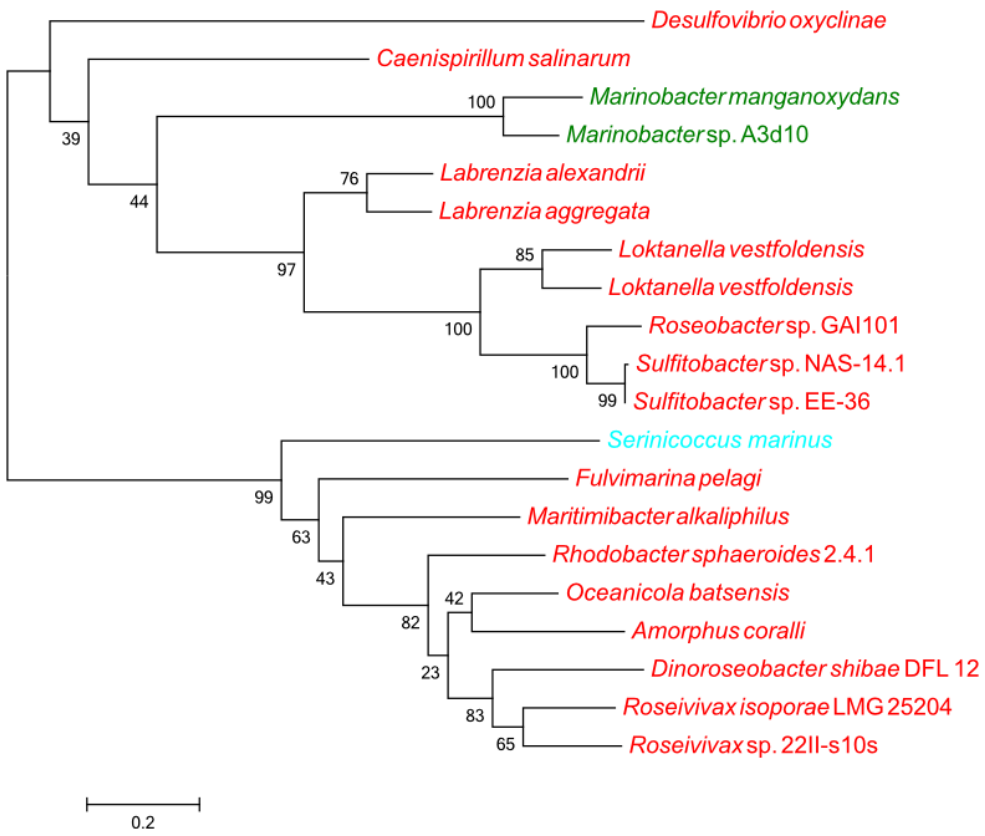


Figure 1.13 Phylogenetic tree of DddL polypeptides. Protein sequences of DddL homologues were aligned with MEGA 6 and used to estimate an unrooted phylogenetic tree using LG model, gamma-distributed. The scale bar indicates number of substitutions per site. Bootstrap values of 500 replicates are given at the base of each branch pair. Species names are coloured according to their taxonomic status: α-proteobacteria (red); γ-proteobacteria (green) and Actinobacteria (light blue).

1.15.3 DddW

Another cupin-type DMSP lyase, DddW, was initially discovered in *R. pomeroyi* DSS-3, but in a different way compared to the previously described *ddd* genes (Todd *et al*., 2012a). In this case, it was noted in a microarray survey of this strain that the expression of one gene, SPO0453, was massively induced (~37-fold) in cells that had been pre-grown in the presence of the DMSP substrate, compared to succinate-grown controls. Furthermore, the product of this gene had a predicted cupin domain near its C-terminus. Therefore, SPO0453 was PCR-amplified from *R. pomeroyi* genomic DNA, and when the PCR product was cloned into an expression vector it was found to confer a Ddd^+ phenotype to *E. coli*, and so was renamed *dddW*.

1.15.4 Distribution of DddW

Similarly to DddL, DddW is also found in the Roseobacter clade, but this enzyme is the least abundant DMSP lyase, as, to date, it is only found in two species - *R. pomeroyi* and *Roseobacter* sp. MED193.

1.15.5 DddQ

The finding that a *dddP*$^-$ mutant strain of *R. nubinhibens* still retained significant levels of DMSP lyase activity (see above) prompted a search for a second, as yet unknown, DMSP lyase in this strain. The *R. nubinhibens* genomic library was therefore further screened for cosmids that conferred a Ddd^+ phenotype to *Rhizobium* but which lack the *dddP* gene. One such cosmid was identified and was shown to contain a gene cluster, in a single transcriptional unit which was confirmed to be responsible. Within this predicted operon, two adjacent genes termed *dddQ1* and *dddQ2* were of interest. The gene products were 39% identical to each other, and when cloned individually, each conferred DMSP lyase activity to *E. coli*. The production of acrylate from DMSP by each of these enzymes was confirmed by NMR and HPLC analysis, so these too are "conventional" DMSP lyases (Todd *et al*., 2010b).

1.15.6 Distribution of DddQ

To date, homologues of DddQ are exclusive to the α-proteobacteria, where, like the other cupin lyases, they are mostly found in members of the Roseobacter clade. There are also DddQ homologues in *alpha proteobacterium* HIMB5 and HIMB100, members of the abundant marine SAR11 and SAR116 clades, respectively (see **Chapter 2**).

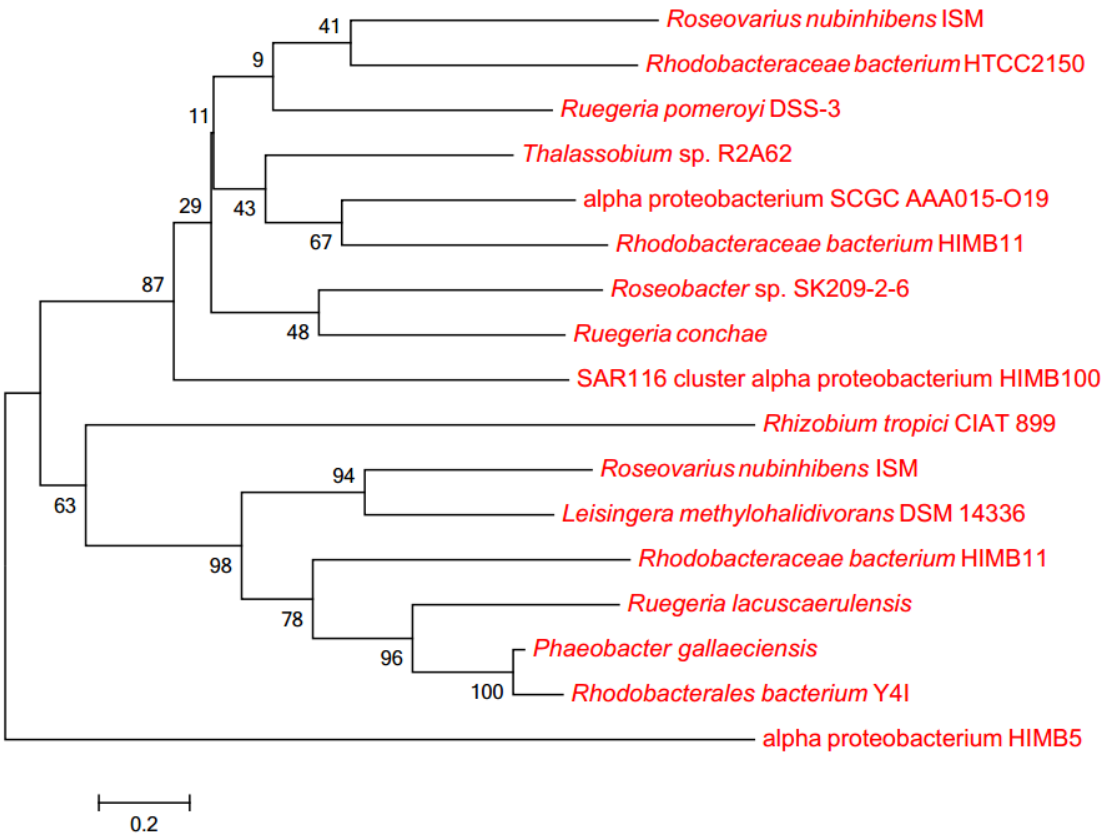


Figure 1.14 Phylogenetic tree of DddQ polypeptides. Protein sequences of DddQ homologues were aligned with MEGA 6 and used to estimate an unrooted phylogenetic tree using LG model, gamma distributed with invariant sites. The scale bar indicates number of substitutions per site. Bootstrap values of 500 replicates are given at the base of each branch pair.

1.15.7 Structure and mechanism of DddQ

Recently, the crystal structure of DddQ from *Ruegeria lacuscaerulenesis* ITI_1157 (which is 45% identical to DddQ2 of *R. nubinhibens*) was solved, providing insight into the catalytic mechanism of DMSP cleavage by this (and perhaps other) cupin-containing lyases (Li *et al*., 2014). DddQ consists of five α-helices and eight anti-parallel β-sheets which form the cupin β-barrel. The β-barrel surrounds the substrate-binding pocket, and this is covered by two loops that permit access to the pocket via a gating mechanism.

The presence of a metal ion is characteristic of cupin-superfamily proteins. The metal is usually iron, but copper, zinc, cobalt, manganese and nickel ions have also been found in cupin active sites. DddQ itself is a zinc metalloenzyme, containing a Zn^{2+} ion in the active site.

The authors also proposed a molecular mechanism for the production of DMS from DMSP. In the absence of DMSP, four amino acid residues in the active site (His125, Glu129, His163, and Tyr131) form coordination bonds with Zn^{2+}. Once DMSP enters, the oxygen atom of its carboxyl group forms a bond with Zn^{2+}, displacing the Tyr131 residue. Then, C_2 of DMSP interacts with the O^- of Tyr131, forming a carbanion which attacks C_3 of DMSP and weakens the S-C_3 bond. The proton of C_2-H binds the O^- of Tyr131, the S-C_3 bond is broken and a C_2=C_3 double bond is formed, resulting in DMS and acrylate which are released from the active site.

Importantly, the four amino acid residues that bind Zn^{2+} in the active site are highly conserved in DddQ homologues, and also in the other, cupin-containing DMSP lyases - DddL and DddW. It is therefore likely that a similar mechanism of DMSP cleavage occurs in all of these enzymes.

1.16 Fate of DMSP cleavage products

As described above, the DMSP is degraded by the "classical" DMSP lyases to produce acrylate, or by DddD to release 3HP. The production of either acrylate or 3HP is coupled with the production of DMS. In this section, I will present our current understanding of the fate of DMSP catabolites, starting with the volatile gas, DMS.

1.16.1 DMS is an environmentally important gas

The importance of DMS to the global sulphur cycle was realised when Lovelock *et al.* showed that it was the major volatile responsible for the transfer of sulphur from the sea to land. This role had previously been assigned to hydrogen sulphide, despite the low atmospheric concentration of this gas (Lovelock *et al*., 1972). Fifteen years later, an additional role was suggested for DMS, which proposed the biological regulation of climate through the production of this gas. The theory was named the CLAW hypothesis after the first letter of each of the author's names – Charlson, Lovelock, Andreae and Warren. They proposed that in the atmosphere, DMS oxidation products are rapidly converted to cloud condensing nuclei (CCN), thereby increasing cloud cover over the oceans and reflecting more UV radiation, which cools the climate. This in turn has an effect on the speciation and size of phytoplankton blooms, and therefore the amount of DMS released to atmosphere, forming an overall negative feedback loop (Charlson *et al*., 1987).

The CLAW hypothesis certainly stimulated a great deal of research into the production and emission of DMS. However, in 2011 a paper was published which challenged the hypothesis. This paper was based on two decades of research, and concluded that only very large emissions of DMS would have any significant effect on cloud cover. It highlighted the importance of non-DMS sources of CCN, such as sea salts and organics, which are much greater contributors to cloud formation than DMS (Quinn and Bates, 2011).

While DMS may not play a significant role in climate regulation, it certainly is important in other ways. It does indeed have a major role in global sulphur cycling, and has been calculated to contribute to a global sea to air flux of 28 Tg of sulphur per annum, which is approximately 50% of the global biogenic sulphur input into the atmosphere (Andreae, 1990; Bates *et al*., 1992; Lana *et al*., 2011). A very different, though significant role for DMS is that it is a potent chemo-attractant for several vertebrates including seabirds (Nevitt and Bonadonna, 2005; Cunningham *et al*., 2008; Amo *et al*., 2013), seals (Kowalewsky *et al*., 2006), fish (DeBose *et al*., 2008) and turtles (Endres and Lohmann, 2012), and invertebrates such as the copepod *Temora longicornis* (Steinke *et al*., 2006).

Despite the large contribution of DMS to the global sea to air flux of organic sulphur, around 90% of DMS made as a result of DMSP cleavage never reaches the atmosphere. This is because it is used by microbes as a source of energy, carbon or sulphur.

Some strains of bacteria can use DMS as a sole source of carbon (see Schäfer *et al*., 2010), including species of *Thiobacillus, Hyphomicrobium* and *Methylophaga*. This is thought to occur by one of two pathways – either via a DMS monooxygenase or a methyltransferase (De Bont *et al*., 1981; Visscher and Taylor, 1993), both resulting in the initial production of methanethiol. Methanethiol produced via the DMS monooxygenase pathway is then further degraded to formaldehyde, hydrogen peroxide and sulphide by a methanethiol oxidase. Formaldehyde is either directly assimilated, or oxidised to CO_2, and sulphide is converted to sulphite, and then sulphate. Hydrogen peroxide is reduced to water and oxygen. Although not much is known about the biochemistry and molecular basis of most of this pathway, the DMS monooxygenase from *H. sulfonivorans* has been purified and characterised, and the genes encoding this enzyme identified (Boden *et al*., 2011). Methanethiol oxidase has also been purified from several species, such as *Hyphomicrobium* EG (Suylen *et al*., 1987) and *Thiobacillus thioparus* (Gould and Kanagawa, 1992).

The alternative pathway, whereby the initial conversion of DMS to methanethiol is via a DMS methyltransferase, was proposed for *Methylophaga thiooxidans* sp. nov. In this pathway, the sulphur from DMS is incorporated into tetrathionate, rather than sulphate. The tetrathionate can be used as an energy source by chemolithoautotrophic and photosynthetic bacteria (Boden *et al*., 2010). A DMS methyltransferase step was also suggested for the initial step of DMS-degradation in *Thiobacillus* ASN-1 (Visscher and Taylor, 1993).However, a DMS methyltransferase enzyme or gene from any species remains to be identified.

A diverse range of bacteria can oxidise DMS to DMSO. In phototrophic bacteria, this provides electron donors for carbon dioxide fixation. The conversion of DMS to DMSO has also been seen in heterotrophic bacteria, for example *Delftia acidvorans* and the Roseobacter *Sagittula stellata* E-37 (Zeyer *et al*., 1987; Zhang *et al*., 1991; Juliette *et al*., 1993; González *et al*., 1997; Fuse *et al*., 1998; Sorokin *et al*., 2000). DMS and DMSO can also be used as a sulphur source. For example, a strain of *Marinobacter* was shown to assimilate sulphur from DMS in a light-dependent manner (Fuse *et al*., 2000). Strains of *Acinetobacter* (Horinouchi *et al*., 1997) and *Rhodococcus* (Omori *et al*., 1995) have been shown to use DMS as a sulphur source via its conversion to DMSO.

1.16.2 Fate of acrylate and 3-hydroxypropionate

As noted above, many of those γ-proteobacteria that contain the "primary" DddD DMSP lyase grow very well on DMSP as sole carbon source, as does the *dddY*-containing β-proteobacterial *Alcaligenes faecalis*. Therefore these DMSP-catabolising bacteria must have pathways that allow them to assimilate carbon from the products of DMSP cleavage. Genetic studies have revealed the presence of several other "Ddd" proteins that are variously involved in other aspects of DMSP catabolism – these include those that are involved in the import of DMSP, and in its downstream catabolism via acrylate and 3HP. In many cases, these are encoded by genes that are closely linked to the "primary" *ddd* genes, notably *dddD* and *dddY*.

1.16.3 DMSP catabolism in *Halomonas* HTNK1

The γ-proteobacterium *Halomonas* HTNK1 is able to use DMSP as a sole source of carbon. In this organism *dddD* is part of a six-gene transcriptional unit known to be involved in DMSP catabolism (see **Figure 1.15**). Two of these genes, *dddA* and *dddC* encode a flavin-containing alcohol dehydrogenase and an aldehyde dehydrogenase, respectively. Another two genes, *acuN* and *acuK* encode proteins resembling a crotonobetainyl-CoA:carnitine CoA transferase (CaiB) and an enoyl-CoA hydratase (CaiD) characterised in *E. coli* (Elssner *et al*., 2001; see above). Biochemical studies in *Halomonas* HTNK1 revealed how these genes are involved in the catabolism of 3HP produced from DddD cleavage of DMSP, and also, rather unexpectedly, acrylate (Todd *et al*., 2010a).

In *Halomonas*, imported DMSP is converted to DMS and 3HP by DddD. Studies using recombinant *E. coli* expressing *Halomonas* genes were used to show that 3HP is further catabolised by the products of *dddA* and *dddC*. Thus, DddA was shown to convert 3HP to malonate semialdehyde, while DddC degrades malonate semialdehyde to acetyl-CoA and CO_2 (see **Figure 1.16**).

Previously the gene products of *acuN* and *acuK* had no known links with DMSP catabolism. These proteins resemble CaiB and CaiD which, in *E. coli*, work cooperatively to catabolise carnitine (Elssner *et al*., 2001). Like CaiB and CaiD, AcuN and AcuK also work in tandem, and when cloned together and expressed in *E. coli*, they were shown to break down acrylate to 3HP (Todd *et al*., 2010a). It was surprising to find genes involved in acrylate catabolism so closely linked to *dddD*, a DMSP lyase which does not produce acrylate. However, it was noted that *Halomonas* HTNK1 was also able to grow on acrylate as a sole source of carbon. Other *dddD*-containing bacteria, such as *Marinomonas* MWYL1, *Pseudomonas* J465 and *Psychrobacter* J466 do not use acrylate as a sole carbon source, and in accordance with this, those strains lack the *acuN* and *acuK* genes (**Figure 1.15**).

Therefore, it seems in *Halomonas* the pathways of DMSP and acrylate catabolism initially run in parallel, and then converge at a common intermediate – 3HP (**Figure 1.16**).

1.16.4 DMSP catabolism in *Alcaligenes faecalis*

The DddY-containing β-proteobacterium *Alcaligenes faecalis* also contains a gene cluster near *dddY* which closely resembles the six-gene operon of *Halomonas* HTNK1 (see **Figure 1.15**). Importantly, *A. faecalis* is also able to grow on DMSP and acrylate as sole carbon sources, and it has a copy of *dddA* and *dddC*, and both *acuN* and *acuK*. Assuming these genes encode enzymes with the same functions shown for the homologues in *Halomonas*, then *A. faecalis* would also catabolise acrylate to 3HP, and further to malonate semialdehyde and acetyl-CoA. However, a key difference is that the DddY-mediated cleavage of DMSP produces acrylate, so in this case the catabolism of DMSP and acrylate would occur sequentially, as shown in **Figure 1.16**.

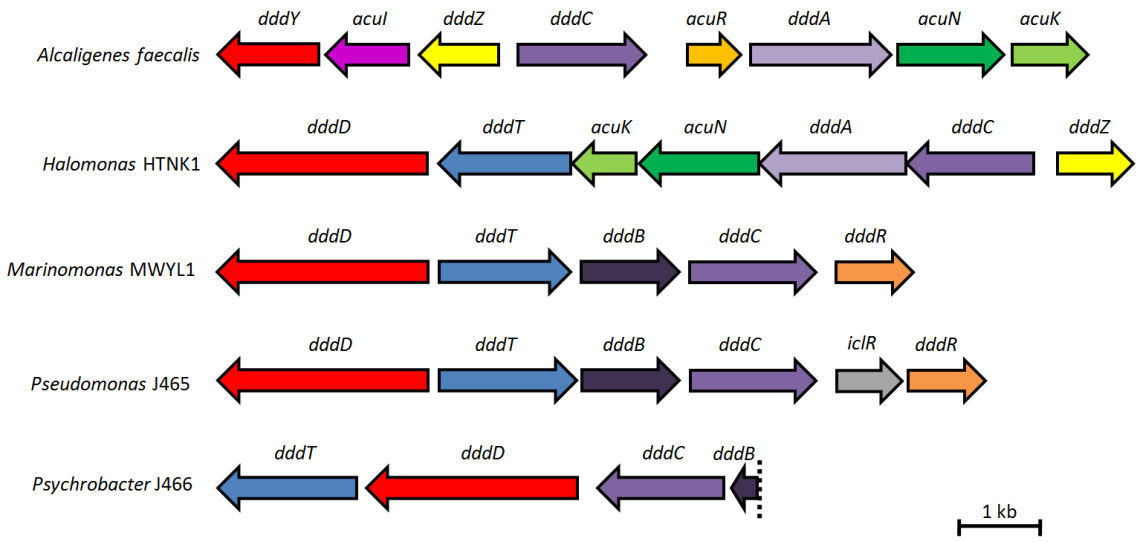


Figure 1.15 Arrangement of *ddd* genes in Ddd$^+$ strains. Genes identified as being involved in DMSP catabolism are shown for *Alcaligenes faecalis*, *Halomonas* HTNK1, *Marinomonas* MWYL1, *Pseudomonas* J465 and *Psychrobacter* J466. The dotted line indicates a contig boundary in the sequencing at *dddB* in *Psychrobacter* J466.

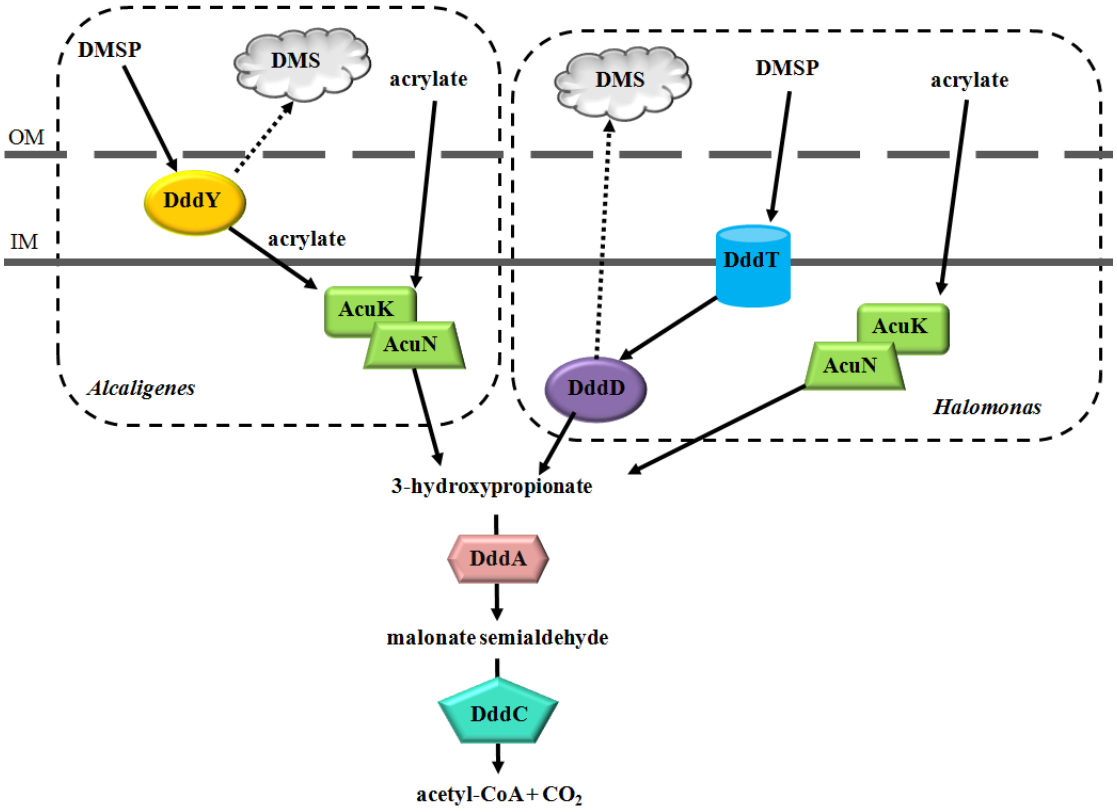


Figure 1.16 Pathways of DMSP and acrylate catabolism in *Halomonas* HTNK1 and *Alcaligenes faecalis*. In *Halomonas* HTNK1, DMSP must be imported by the transporter DddT, and then it is degraded by the cytoplasmic DddD to produce DMS and 3HP. Acrylate may also be imported from the environment and converted to 3HP by AcuN and AcuK. Contrastingly, *A. faecalis* has the periplasmic DMSP lyase, DddY and so there is no need for DMSP import across the inner membrane. Acrylate is produced by DMSP cleavage, or imported directly from the environment where it is acted on in the cytoplasm by AcuN and AcuK to produce 3HP. In both organisms, DddA converts 3HP to malonate semialdehyde, which is further catabolised to acetyl-CoA and CO_2 by DddC.

1.17 Regulation of DMSP catabolism

It is not surprising that the activities of DMSP lyase can be induced, sometimes markedly so, by prior growth of the bacteria in the presence of the DMSP substrate. And, indeed, this has been demonstrated for several of the *ddd* genes described above, in different bacteria. But again, there is variability, and no "one size fits all". Furthermore, there were some surprising features, the most striking of these being that induction of the *ddd* gene expression may be via the initial product (acrylate or 3HP) of the reaction rather than the DMSP itself.

In *Marinomonas* MWYL1 the *dddD* gene is transcribed divergently from the adjacent *dddTBCR* operon (**Figure 1.15**). The promoter-distal gene, *dddR*, encodes a LysR-type transcriptional regulator, which has been shown to positively regulate *dddD* in response to DMSP, enhancing its level of expression by at least 100-fold (Todd *et al*., 2007). Neither of these operons was affected by addition of either 3HP or acrylate, so this system most closely resembles the conventional type of LysR-type gene regulation, in which the substrate acts as the co-inducer. Typical induction of DMSP lyase expression by the substrate is also seen for the *dddQ* and *dddP* genes in *Ruegeria pomeroyi*, and *Roseovarius nubinhibens*. In *R. pomeroyi*, *dddW* is also induced greatly by DMSP (ca. 40-fold), and it is adjacent to a gene whose predicted product is a LysR-type transcriptional regulator, which has been shown to activate the expression of *dddW* in response to DMSP (Todd *et al*., 2012a).

As with the Ddd^+ bacteria described above, DMS production by both *Alcaligenes faecalis* and *Halomonas* HTNK1 is also inducible by DMSP but, unusually, it is enhanced more significantly by the initial products of DMSP catabolism. In both organisms, acrylate, and 3HP in the case of *Halomonas*, induces expression of the operon containing *dddY* or *dddD*. Indeed, it was shown in *Alcaligenes* that DMSP actually does not act as a direct co-inducer, and it must first be converted to the true inducer, acrylate (Curson *et al*., 2011).

A conceptually analogous situation was demonstrated in *Rhodobacter sphaeroides*, although both the type of regulatory gene (termed *acuR*, of the *tetR* family) and the DMSP lyase (*dddL*) differ from those above. These two genes are the promoter-proximal and promoter-distal genes in a three-gene operon (**Figure 1.17**) whose expression is markedly enhanced by either DMSP or acrylate. But, as in the case of *Alcaligenes*, the DMSP acts indirectly, and requires a conversion to acrylate, the *bona fide* co-inducer. Another unusual regulatory feature of the *acuR-acuI-dddL* operon in *R. sphaeroides* is that the mRNA transcript is leaderless, so lacks a 5'-untranslated region and ribosome binding site (RBS) upstream of *acuR*. As a result, *acuR* is translated at a lower efficiency than the downstream *acuI* gene (see **Section 1.19** below for description of the *acuI* gene product) which does have a good match to an RBS (Sullivan *et al*., 2011). This feature

allows *acuR* to regulate the expression of *acuI* and *dddL* in response to acrylate, while not being highly expressed itself.

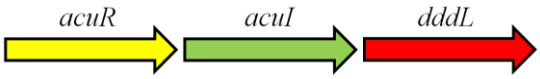


Figure 1.17 *acuR-acuI-dddL* operon in *Rhodobacter sphaeroides* 2.4.1. The DMSP lyase, *dddL* is distal in a three gene transcriptional unit, downstream of the regulatory gene *acuR*, and *acuI*, which encodes an acryloyl-CoA reductase.

1.18 DMSP transport

In most cases (DddY being the exception) DMSP must be transported into the cytoplasm before it is acted on by a DMSP lyase or demethylase. Two different families of proteins – the BCCT-type (betaine-carnitine-choline) and the ABC-type (ATP-binding cassette) transporters – are capable of importing betaines (Ziegler *et al*., 2010; Eitinger *et al*., 2011), and proteins belonging to these families have been directly shown to import DMSP, which is a sulphonium analogue of glycine betaine (Sun *et al*., 2012). Significantly, genes encoding BCCT-type transporters are found closely linked to *dddD* in some bacteria, for example the *dddD* genes of *Marinomonas* MWYL1, *Halomonas* HTNK1, *Pseudomonas* J465 and *Psychrobacter* J466 are adjacent to a gene, termed *dddT* (see **Figure 1.15**), whose predicted product encodes a BCCT transporter. The *dddT* genes of *Halomonas* HTNK1 and *Marinomonas* MWYL1 are both capable of transporting DMSP, as shown by expressing them individually in strain MKH13, an *E. coli* mutant defective in all three of its betaine uptake systems. Only the mutant strain with a *dddT* gene was able to transport DMSP, which corrected the hypersensitivity phenotype of MKH13 to NaCl (Sun *et al*., 2012).

The *dddD* gene of α- and β-proteobacteria is also adjacent to genes that are predicted to be involved in DMSP import, but in these cases this involves an ABC-type system. This was seen in such diverse bacteria as *Burkholderia ambifara*, *Rhizobium* sp. NGR234, *Rhodobacterales* bacterium KLH11 and *Hoeflea phototropica*. To confirm the role of these genes in DMSP transport, those of *B. ambifara* were cloned and were found to correct the defect of *E. coli* MKH13, though only partially and not as effectively as the cloned *dddT* genes, above.

In bacteria containing DddL, DddP, DddQ, DddW (mainly the Roseobacters) the corresponding DMSP lyase genes are normally in single gene units, and are not near those that are predicted to encode transporters. However, it is clear that the Roseobacters do have transporters that efficiently import DMSP. Thus, Sun *et al*. (2012) identified two different clusters of genes that encoded the ABC-type transporters in the Roseobacter *Roseovarius nubinhibens* ISM, and another ABC-type cluster in *Sulfitobacter* sp. EE-36. When cloned, the genes that encode these transporters corrected the osmotic sensitivity of the *E. coli* MKH13 strain, as described above. In contrast to the situation with the *dddD* genes (above) these ABC transport genes were not linked to those involved in other aspects of DMSP catabolism, *dddL* in the case of *Sulfitobacter* and *dddP*, two copies of *dddQ* and also the DMSP demethylase gene *dmdA* in the case of *R. nubinhibens*.

Lastly, the DddY lyase, found in *Alcaligenes faecalis*, is so far the only DMSP lyase located in the periplasm rather than the cytoplasm. This precludes the need for *A. faecalis* to import DMSP, and consistent with this the cluster of *ddd* genes near *dddY* in this organism lacks the copy of *dddT* which is present in DddD-containing bacteria with similar *ddd* clusters (see **Figure 1.15**).

The diversity in DMSP transport systems is interesting, as is the fact that some bacteria appear to have multiple transporters involved in DMSP uptake. In addition to the two different ABC transport systems of *Roseovarius nubinhibens* shown to be functional DMSP transporters, *Ruegeria pomeroyi* also has no less than five bioinformatically predicted BCCT-transporters (Moran *et al*., 2004), although none was experimentally ratified. Thus, it appears that these bacteria may have multiple ways of importing DMSP as part of their general betaine uptake system(s) rather than a dedicated transport system.

1.19 AcuI – an extremely abundant enzyme involved in acrylate detoxification

In the course of the studies on the growing list of *ddd* genes, there was one constant, amid all the diversity described above. Nearly all the clusters of *ddd* genes involved in the initial transport, regulation and catabolism of DMSP, is a gene termed *acuI*. For example, *acuI* is found next to *dddL* in *Rhodobacter sphaeroides* 2.4.1, *dddY* in *Alcaligenes faecalis* and *dddD* in *Halomonas* HTNK1. Additionally, in the Roseobacter clade, there is a version of *acuI* that lies immediately downstream of the *dmdA* DMSP demethylase gene (see **Chapter 5**).

Biochemical studies on the AcuI gene product of *Rhodobacter sphaeroides* 2.4.1 (Schneider *et al*., 2012; **Figure 1.17**) showed that it was an acryloyl-CoA reductase, converting acryloyl-CoA

to propionyl-CoA. This provided a clear explanation of some of the phenotypes that had been associated with this gene in DMSP-catabolising bacteria. Most notably, AcuI$^-$ mutants of *Rhodobacter* were less efficient at catabolising acrylate, as measured by assaying labelled CO_2 in bacteria fed with 1-^{14}C-acrylate. More strikingly, these mutants were significantly more sensitive to the toxic effects of exogenously added acrylate (Sullivan *et al*., 2011). Since acryloyl-CoA is extremely cytotoxic (Herrmann *et al*., 2005) the role of AcuI may be to protect those bacteria that synthesise acrylate from self-inflicted damage, due to the subsequent formation of excess acryloyl-CoA.

Interestingly, close homologues of *acuI* are not confined to those bacteria that catabolise DMSP. Indeed, *E.coli* contains a gene, *yhdH*, previously of unknown function, whose product is 54% identical to the AcuI of *R. sphaeroides*. The purified YhdH protein has been shown to have acryloyl-CoA reductase activity *in vitro*, and YhdH$^-$ mutants are exquisitely sensitive to added acrylate in the medium (Todd *et al*., 2012b). There are many other bacteria with close homologues of AcuI and there is also evidence that other systems for acryloyl-CoA detoxification exist in those bacteria that lack *acuI* (Curson *et al*., 2014). Thus, AcuI and the detoxification of acryloyl-CoA may have wider relevance that stretches beyond marine environments, or the catabolism of DMSP.

1.20 The α-proteobacteria: Roseobacters and the SAR11 clade

There is one group of organisms which consistently emerge as key players in DMSP utilisation. These are members of the sub-phylum of α-proteobacteria, one of the largest and most diverse groups of Eubacteria. Among this sub-phylum there are two groups which are particularly abundant in the oceans, and play a key role in DMSP turnover – the Roseobacter clade and the SAR11 clade. Both clades have been a particular focus of this work, and will be introduced in detail in later chapters. However, this section provides a brief overview of each group.

1.20.1 The Roseobacter Clade

Due to their abundance, and physiological and geographical diversity, the group of α-proteobacteria known as the *Roseobacter* clade are the most intensely studied group of marine bacteria to date. Members of this clade are estimated to contribute up to 25% of the bacterial community in some marine environments. To date there are at least 20 different described genera of *Roseobacters*, many of which contain multiple species and strains, as well as hundreds of uncharacterised isolates and sequences. The clade forms a distinct cluster in the family of *Rhodobacteraceae* with *Roseobacter* members sharing at least 88% identity of the 16S rRNA

gene (Brinkhoff *et al*., 2008). Unlike other genera in the Rhodobacteraceae, the *Roseobacters* are notably absent from freshwater or terrestrial environments, but the marine environments they do reside in are diverse. They range from open seas (pelagic) to coastal and deep sea sediments, from the polar ice to tropical regions. They are often found living in close association with other marine organisms, including algal blooms (Buchan *et al*., 2005). Genes involved in DMSP cleavage and demethylation are particularly abundant amongst the Roseobacter clade, and several strains have been shown to degrade DMSP (see **Chapter 5**). One strain, *Ruegeria pomeroyi* DSS-3 has become something of a model organism of the Roseobacter clade and DMSP utilization by this strain has been studied extensively. One reason for this is that *R. pomeroyi* is capable of both DMSP cleavage and DMSP demethylation, and possesses a multitude of DMSP-related genes. This organism was the focus of part of this work, and so will be discussed in more detail in **Chapter 4**.

1.20.2 The SAR11 clade

The second of the two groups of α-proteobacteria known to be involved in DMSP degradation is known as the SAR11 clade, belonging to the Order Rickettsiales. It is predicted that members of this clade compose ca. 25% of the oceans' bacteria. Initially identified through culture independent techniques such as 16S rRNA sequencing, some SAR11 strains have now been cultivated (Giovannoni *et al*., 1990; Rappé *et al*., 2002). DMSP degradation in the SAR11 clade is the focus of **Chapter 2** and, as such, these important bacteria will be introduced in more detail then.

1.21 Aims and objectives

DMSP is a significant source of carbon and sulphur in marine food webs, and a precursor of the environmentally important gas DMS. In order to support environmental and ecological observations of the amounts and functions of this sulphur molecule, it is necessary to fully understand the genetics and molecular mechanisms underlying DMSP catabolism.

When this project started in 2010, it was already clear that molecular mechanisms used by bacteria to break down DMSP were diverse. In addition to the DMSP demethylase, DmdA, four DMSP lyase genes, *dddD*, *dddL*, *dddP* and *dddQ* had already been identified and published, while work on *dddY* and *dddW* was still in progress. These initial genetics studies, whilst addressing the long unanswered question of how bacteria are able to break down DMSP, had also opened up a new set of exciting questions. More work was required to understand why so many different lyases existed, why some bacteria have multiple DMSP-enzymes, and how the different pathways are regulated. In addition to the diversity in lyases, there was also a variance in how different bacteria use DMSP as a nutrient source, with only some species able to use DMSP or its breakdown products acrylate and 3HP as sole carbon sources.

My project was therefore to investigate this diversity in DMSP breakdown further, while addressing the following specific points:

- To perform a thorough bioinformatics analysis of the DMSP-related genes of the Roseobacter clade, particularly focussing on the synteny of DmdA and each of the DMSP lyases.
- To investigate how the model marine bacterium *Ruegeria pomeroyi* DSS-3 is able to assimilate carbon from DMSP and acrylate.
- To study the multiple DMSP lyases of *Oceanimonas doudoroffii*, in order to understand how each lyase gene is regulated.
- To identify the enzyme responsible for DMSP-dependent DMS production in the abundant and ubiquitous SAR11 strain HTCC1062.

Chapter 2

DMSP lyases of the ubiquitous *Candidatus* Pelagibacter ubique (SAR11) clade of marine bacteria

2.1 Introduction

As described in **Chapter 1**, the "SAR11" clade form an important, hugely abundant, group of marine α-proteobacteria (the nomenclature reflecting the fact that their existence was first demonstrated in the Sargasso Sea). The clade was discovered as part of a culture-independent study into the genetic diversity of the marine environment, through the phylogenetic analysis of bacterial 16S ribosomal RNA genes amplified from DNA extracted from the Sargasso Sea (Giovannoni *et al*., 1990).

The SAR11 clade was later shown to be hugely abundant and widespread in the oceans. Between 25 and 50% of bacterial ribosomal RNA genes in seawater belong to members of this clade (Morris *et al*., 2002), making them (probably) the most prolific group of microbes on the planet. There is also a SAR11 cluster, found at low abundances in freshwater lakes (Bahr *et al.*, 1996; Grote *et al*., 2012).

The SAR11 clade can be further divided into seven sub-clades (**Table 2.1**) (Grote *et al*., 2012). Genome sequences are available for a total of seven strains: five from sub-clade Ia (HTCC1062, HTCC7211, HTCC1002, HIMB083 and HIMB5); HIMB114 from sub-clade IIIa and HIMB59 from sub-clade V.

Table 2.1 Sub-clades of the SAR11 group

Sub-clade	Strain	Comments
Ia	HTCC1062*	>98% 16S rRNA identity to each other
	HTCC7211*	
	HTCC1002*	Most numerically dominant
	HTCC9565	Ubiquitous
	HIMB083*	
	HIMB5*	
Ib	SAR193	
	SAR11	
II	Arctic95B-1	
	SAR211	
IIIa	HIMB114*	88% 16S rRNA identity to HTCC1062
	OM155	
IIIb	S9D-28	Freshwater strains
	LD12	
IV	DQ009255	
V	DQ009262	Most distantly related sub-clade
	HIMB59*	82% 16S rRNA identity to HTCC1062

SAR11 strains are grouped into seven sub-clades. Complete genome sequences are available for seven strains, indicated by an asterisk.

Despite the ubiquity of the SAR11 clade, difficulties in cultivating these strains have hampered phenotypic studies. However cultures of some strains, including HTCC1002 and HTCC1062, have been obtained, with difficulty. The strains grow extremely slowly, reaching a maximum cell density of 2.5×10^5-3.5×10^6 cells per ml after 30 days incubation (Rappé *et al*., 2002).

So far, all cultivated strains of *Candidatus* Pelagibacter ubique have very small cells with a length of 0.37-0.89 μm and diameter of 0.12-0.2 μm (Rappé *et al*., 2002). Furthermore all genome-sequenced Pelagibacter ubique bacteria have tiny genomes at <1.5 Mb. Thus, for example HTCC1062, which is something of a model for this clade has a genome of just 1.31 Mb (Giovannoni *et al*., 2005).

Even with this small genome size, SAR11 cells have the great majority of basic functions that can be found in α-proteobacteria with much larger genomes, a feature which has been attributed to genome streamlining. These genomes contain very little redundancy or non-functional DNA, and the average intergenic space is a mere 3 base pairs (Giovannoni *et al*., 2005). However, SAR11 cells do have reduced metabolic capabilities. One such example with relevance to the work in this chapter is that of sulphur metabolism in SAR11, which has been studied in some

detail. It was found that despite the almost unlimited availability of sulphate in the oceans, SAR11 cannot use this as a sulphur source because they lack a complete assimilatory sulphate reduction pathway. Instead, they depend on reduced forms of sulphur, such as methionine, thiosulfate or DMSP (Tripp *et al*., 2008).

The importance of DMSP as a source of sulphur to SAR11 strains is reflected in their genetics. In 2008, Reisch *et al*. purified the DmdA enzyme (SAR11_0246) from strain HTCC1062 and showed it to be a functional DMSP demethylase. A BLASTp search of homologues to the SAR11_0246 sequence revealed that a further five SAR11 strains have a convincing homologue, with identities ranging from 41-48% (see **Table 2.2**). Strain HIMB114 has a very weak homologue with only 24% identity to the *Ruegeria pomeroyi* DmdA (SPO1913), so this is unlikely to be a functional DmdA enzyme.

Table 2.2 DmdA homologues amongst the SAR11 clade

Strain	Locus Tag	Identity to SAR11_0246	*E* value
HTCC1002	PU1002_05126	99%	0.0
HIMB083	Pelub83DRAFT_1008	78%	0.0
HIMB5	HIMB5_00000090	75%	0.0
HTCC7211	PB7211_770	70%	0.0
HIMB59	HIMB59_00001390	55%	$8e^{-149}$

Homologues were predicted using the peptide sequence of the functionally-verified SAR11_0246 as a query in a BLASTp search. Locus tags of homologues are presented, along with percentage sequence identity and *E* value (cut-off = $8e^{-149}$).

Although the purified SAR11_0246 protein was shown to have DMSP demethylase activity in 2008 (Reisch *et al*., 2008), given the challenges faced in growing cultures of SAR11 strains (as explained above), it has taken some time to verify that SAR11 strains containing DmdA do indeed demethylate DMSP. However, very recently an investigation into DMSP consumption and MeSH production by SAR11 strain HTCC1062 was carried out. This work, conducted by Stephen Giovannoni's group at the University of Oregon, showed that HTCC1062 cells removed DMSP over 18 hour incubations in artificial sea water, while simultaneously producing methanethiol. This showed for the first time in any SAR11 strain the likely presence of the DMSP demethylation pathway (S. Giovannoni, personal communication).

Surprisingly, the accumulation of methanethiol was only enough to account for 21% of the DMSP decrease. The predominant sulphur product released from DMSP was in fact DMS, accounting for 59% of DMSP decrease. Excitingly, this means that this SAR11 strain may possess both the DMSP demethylation and the cleavage pathway. This finding was especially

intriguing given that HTCC1062 was not known to possess any homologues of the known DMSP lyases. Therefore, the aim of this work was to identify the genetic basis of this Ddd^+ phenotype in HTCC1062, and also investigate the functionality of other DMSP lyase homologues in the SAR11 clade. To do this, homologues of *dddQ* from strain HIMB5, *dddP* from HTCC7211 and a candidate DMSP lyase gene from HTCC1062 were all synthesised and checked for DMSP-dependent DMS production.

2.2 Results

2.2.1 *Pelagibacter ubique* HTCC1062 gene SAR11_0394 encodes a cupin-containing polypeptide

Pelagibacter ubique HTCC1062 was shown to produce DMS as the major end product when grown in the presence of DMSP as the sole source of sulphur. To check for homologues to known DMSP lyases, the genome sequence of HTCC1062 (Giovannoni *et al*., 2005) was interrogated in a BLASTp search using DMSP lyase sequences as queries. The query sequences were: DddD of *Halomonas* HTNK1 (ACV84065); DddY of *Alcaligenes faecalis* M3A (ADT64689); DddL (EE36_11918) of *Sulfitobacter* sp. EE-36; DddQ (SPO1596), DddP (SPO2299) and DddW (SPO0453) of *Ruegeria pomeroyi*. The former five queries returned no hits with an E value below 0.002. However, one gene, with the tag SAR11_0394 had very weak homology to DddW with 37% identity over 64% coverage, with an E value of $2e^{-16}$. SAR11_0394 was predicted to encode a 130 amino acid polypeptide, with a putative C-terminal cupin domain.

As discussed in **Chapter 1**, three other DMSP lyases have cupin domains - DddL, DddW and DddQ. The structure and possible enzymatic mechanism of *Ruegeria lacuscaerulenesis* DddQ has been solved and four key conserved residues in the cupin domain were shown to be critical for the binding and cleavage of DMSP into DMS and acrylate. The cupin domain of the SAR11_0394 gene product also contains these four residues, as shown in the alignment in **Figure 2.1**. In addition, the computationally predicted tertiary structure of the SAR11_0394 polypeptide has a very similar cupin structure to the experimentally determined DddQ from *R. lacuscaerulenesis* (Li *et al*., 2014; **Figure 2.2**). Therefore, SAR11_0394 provided a good candidate for a novel DMSP lyase.

Figure 2.1 Sequence alignment of the cupin regions of DddL, DddQ, DddW and SAR11_0394 homologues. Completely conserved residues are highlighted in red, and highly conserved in yellow. Four residues shown to be key to DddQ cleavage of DMSP are indicated by asterisks (Li *et al*., 2014). Sequences 1-8 are DddL polypeptides from the following: 1, *Oceanicola batsensis* (OB2597_08014); 2, *Rhodobacter sphaeroides* 2.4.1 (RSP_1433); 3, *Amorphus coralli* (WP_026318838); 4, *Roseivivax isoporae* (RISW2_01295); 5, *Dinoroseobacter shibae* (Dshi_3313); 6, *Roseivivax* sp. 22II-s10s (AT08_14527); 7, *Maritimibacter alkaliphilus* (RB2654_07950); 8. *Fulvimarina pelagi* (FP2506_12684). Sequences 9-14 are DddQ polypeptides from the following: 9, *Ruegeria pomeroyi* (SPO1596); 10, *Roseovarius nubinhibens* (ISM_14090); 11. *Roseobacter* sp. SK209-2-6 (RSK20926_17292); 12, *Thalassobium* sp. R2A62 (TR2A62_3487); 13, *Pelagibacter ubique* HIMB5 (HIMB5_00000220); 14, *Ruegeria lacuscaerulenesis* (SL1157_0332). Sequences 15 and 16 are DddW polypeptides from *Ruegeria pomeroyi* (SPO0453) and *Roseobacter* sp. MED193 (MED193_09710), respectively. Sequences 17-19 are potential DMSP lyase polypeptides from the *Pelagibacter ubique* strains HTCC1062 (SAR11_0394) (17); HTCC1002 (PU1002_04381) (18) and HIMB5 (HIMB5_00004730) (19).

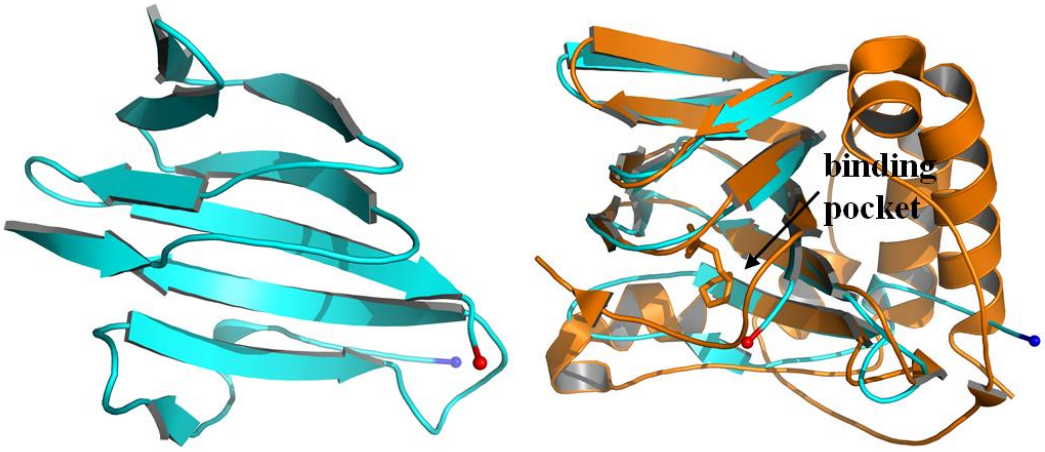


Figure 2.2 Predicted tertiary structure of the SAR11_0394 polypeptide. Computationally predicted tertiary structure of the SAR11_0394 polypeptide (left; Changjiang Dong, UEA). The same structure was superimposed onto the experimentally determined structure of DddQ from *Ruegeria lacuscaerulenesis* ITI_1157 (Li *et al*., 2014) using the homology modelling programme SWISS-MODEL (right; Bordoli *et al*., 2008). The substrate binding pocket is indicated.

2.2.2 *de novo* synthesis of SAR11_0394

The approach used to investigate the possible role of the gene in the Ddd^+ phenotype of strain HTCC1062, was to clone the gene in an expression vector and determine its phenotype, as was done for DmdA of this strain (Reisch *et al*., 2008). To do this, the SAR11_0394 sequence was optimized for codon usage in *E. coli* using the OPTIMIZER software (Puigbò *et al*., 2007), and that sequence was synthesised by *GenScript* USA inc. (Piscataway, New Jersey). To allow sub-cloning into protein over-expression plasmids, 5' and 3' extensions were incorporated, containing the restriction sites *Nde*I and *Bam*HI, respectively. The SAR11_0394 insert, initially cloned into the pUC57 plasmid, was sequence-verified by *GenScript*.

2.2.3 SAR11_0394 was sub-cloned into the expression plasmid pET16

A preparation of the pUC57 plasmid containing SAR11_0394 was digested with the restriction enzymes *Nde*I and *Bam*HI to release the synthesised SAR11_0394 insert. Following gel electrophoresis, a fragment of the expected size (700 bp) for SAR11_0394 was extracted from the gel and purified, and then ligated into the expression plasmid pET16b, to create pBIO2206. The pET16b plasmid contains an ampicillin resistance gene encoding β-lactamase, a *lacI*

repressor gene and the viral T7*lac*-promoter (*Novagen*). Thus, a gene cloned into the polylinker of pET16 will be transcribed from the T7*lac* promoter when transformed into a host expressing T7 RNA polymerase. In this case, pBIO2206 was transformed into *E. coli* strain BL21, which has a chromosomal copy of the T7 RNA polymerase gene, under the control of a *lacUV*5 promoter. This promoter is under the control of the *lacI* repressor, whose repression can be relieved by isopropyl β-D-1-thiogalactopyranoside (IPTG). Thus, in the presence of IPTG, the T7 RNA polymerase is expressed, along with the cloned gene under control of T7*lac* promoter. The pET16b plasmid also has a sequence upstream of the cloning site that encodes an N-terminal sequence encoding a string of ten histidine residues, such that cloned genes encode polypeptides with a His-tag, to facilitate their purification, as discussed below.

Following its construction and ratification, plasmid pBIO2206 was introduced into *E. coli* BL21 cells, selecting for ampicillin resistant transformants.

2.2.4 The SAR11_0394 encodes a DMSP lyase

2.2.4.1 DMSP-dependent DMS production

To establish if SAR11_0394 did encode a DMSP lyase, *E. coli* BL21 cells with pBIO2206, or with an 'empty' pET16b plasmid, were grown in LB media in the presence of 100 nM IPTG to induce expression of the SAR11_0394 gene product. Following overnight growth, the cells were washed and resuspended in M9 minimal media containing 10 mM succinate as a carbon source, and 5 mM DMSP, in a sealed vial for 1 hour before assaying by gas chromatography. The rate of DMS production was calculated as nmol DMS produced per hour, adjusted according to the total protein content of each vial, as measured by Bradford's assay. As shown in **Figure 2.3**, BL21 containing pBIO2206 produced significantly more DMS (0.89 nmol $hour^{-1}$ µg $protein^{-1}$) than the control cells (0.039 nmol $hour^{-1}$ µg $protein^{-1}$; Welch Two Sample t-test, t = -17.276, df = 2, p = 0.003). Thus, SAR11_0394 does encode a product with DMSP lyase activity and was re-named *dddK*.

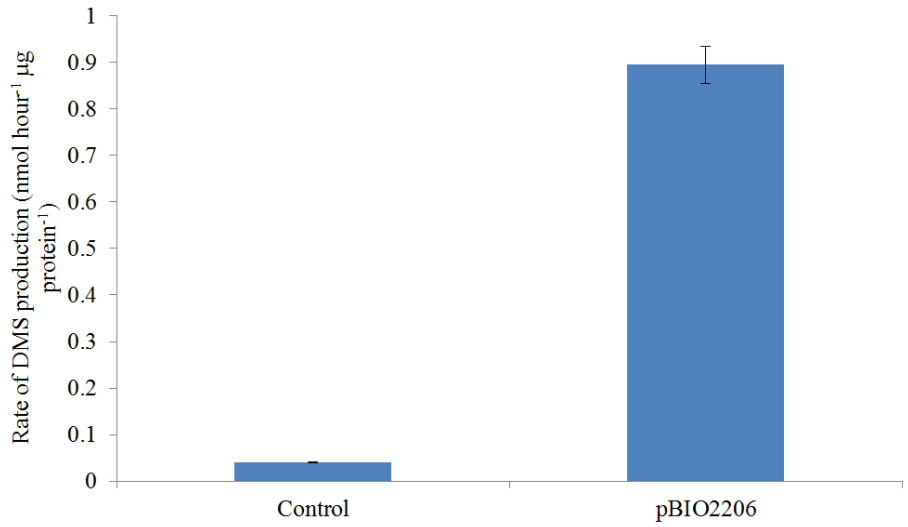


Figure 2.3 DMSP-dependent DMS production by *E. coli* expressing SAR11_0394. *E. coli* BL21 cells containing pBIO2206, or 'empty' pET16b (control) were grown in the presence of 100 nM IPTG to induce expression of SAR11_0394, then exposed to 5 mM DMSP for 1 hour in sealed vials and assayed by gas chromatography. Average rates of DMS production were calculated from triplicate samples, as nmol per hour, adjusted for total protein content. Error bars represent the standard error.

2.2.4.2. DMSP is cleaved by DddK to produce acrylate

Having shown that *E. coli*, containing the cloned SAR11_0394 gene generated DMS from the added DMSP substrate, NMR spectroscopy was used to identify the corresponding C3 catabolite. To do this, BL21 cells containing pBIO2206 were grown overnight in the presence of 100 nM IPTG as described above and then resuspended in deuterium oxide, and lysed by sonication. An aliquot of the soluble fraction was then incubated in the presence of 3 mM [3-^{13}C]-DMSP for 1 hour. A newly-formed ^{13}C catabolite was produced in each case as identified by NMR spectroscopy (carried out by Dr Yohan Chan in the School of Chemistry, UEA). As shown in **Figure 2.4** this ^{13}C catabolite exhibited a peak that did not exactly match the chemical shift of a reference sample of pure [3-^{13}C]-acrylate, but was identical to the DddK sample spiked with [3-^{13}C]-acrylate. The reason for the difference in acrylate spectra reflects a difference in pH between the pure solution and the mixed solutions (Y. Chan, personal communication). A negative control, in which the cell-free extract of wild type BL21 cells was used, did not yield acrylate, or any other product.

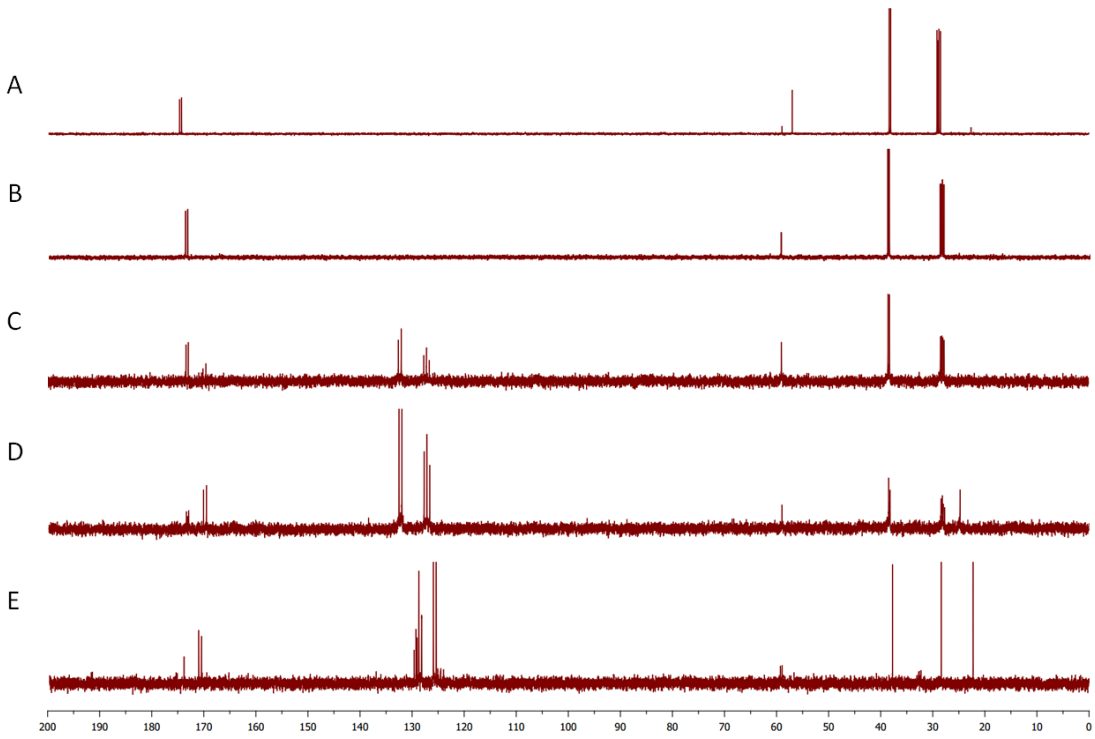


Figure 2.4 NMR spectra of *E. coli* containing *dddK*. NMR spectra showing (A) [3-^{13}C]-DMSP reference; (B) wild type *E. coli* fed with [3-^{13}C]-DMSP; (C) *E. coli* containing pBIO2206 fed with [3-^{13}C]-DMSP; (D) *E. coli* with pBIO2206 fed with [3-^{13}C]-DMSP, then spiked with [3-^{13}C]-acrylate; (E) [3-^{13}C]-acrylate reference.

2.2.5 Purification of DddK

Since the cell-free extract containing SAR11_0394 had DMSP lyase activity, the next step was to purify the polypeptide and determine the properties of this novel lyase.

2.2.5.1 Over-expression of DddK

To obtain significant amounts of DddK for purification and assaying, *E. coli* BL21 was used as the host strain a background for over-expression of DddK. This strain is not only designed to express genes cloned in pET16 at high level, but it has a mutation in the outer membrane protease VII which reduces proteolysis of expressed proteins. To over-express DddK, BL21 cells containing pBIO2206, were inoculated into LB containing ampicillin, incubated at 37°C until they reached mid-exponential phase (OD of 0.4), then 100 nM IPTG was added, prior to a further 4 hours incubation at 28°C, at which time the cells were harvested.

2.2.5.2 DddK is a soluble protein

The harvested *E. coli* cells were lysed and separated into insoluble and soluble fractions by centrifugation. The soluble fraction of the cells containing pBIO2206 was compared to the soluble fraction of wild type BL21 cells, using SDS-PAGE (**Figure 2.5**). The predicted size of His-tagged DddK is 15.8 kDa, and in accordance with this, a strong band was seen in the soluble fraction of pBIO2206 cells, at approximately 15 kDa. In contrast, the band was completely absent in BL21 cells lacking pBIO2206.

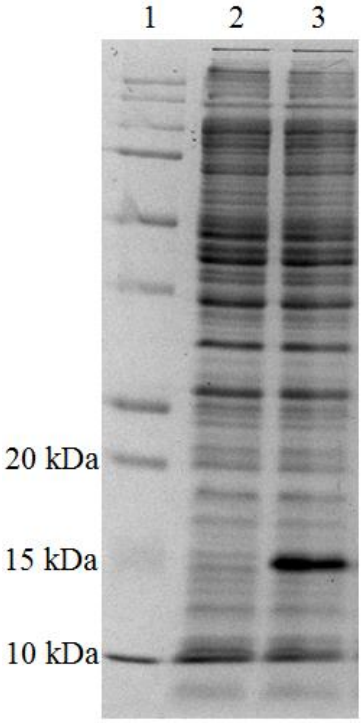


Figure 2.5 Expression of DddK polypeptide in *E. coli* containing the cloned pBIO2206. Polypeptides from the soluble fractions of *E. coli* BL21 cells were separated by SDS-PAGE on a 12% acrylamide gel, and stained with Coomassie Blue. Lane 1: Precision Plus protein standard. Lane 2: Fraction from wild type cells. Lane 3: Fraction from recombinant *E. coli* containing pBIO2206. A strongly staining band can be seen in Lane 3 at ca. 15 kDa, the approximate size of his-tagged SAR11_0394.

2.2.5.3 Partial purification of DddK polypeptide

In an attempt to purify the DddK polypeptide, a 50 ml culture of *E. coli* containing pBIO2206 was grown in LB at 28ºC in the presence of 100 nM IPTG. The cells were harvested, pelleted and re-suspended in 1.4 ml equilibration buffer, then lysed by sonication. The lysate was centrifuged at 13,000 RPM, and the soluble fraction was retained and was applied in two loads of 0.7 ml, to a *Qiagen* Ni-NTA spin column. Aliquots (5 µl) of the flow-through from each application were examined by SDS-PAGE (**Figure 2.6**, lanes 6 and 7). The flow-through contained a weakly staining band at ca. 15 kDa, suggesting that some of the His-tagged DddK had not bound to the Ni-NTA resin. The column was washed three times with wash buffer (**Figure 2.6**, lanes 8-10). The majority of remaining non-binding proteins were removed in the first wash. Again, a weakly stained band at 15 kDa was present in all three washes. Finally, the bound His-DddK was eluted twice using a buffer containing 300 mM of imidazole, which acts by competing for binding of the His-tag to the Ni ions, and thus displaces the bound polypeptide. As seen in **Figure 2.6** (lanes 11 and 12), both elutions contained a pronounced band at 15 kDa, the expected size of His-DddK. The total concentration of protein in eluate 1 and 2 was estimated using Bradford's assays, and shown to be 1.2 µg/µl and 0.74 µg/µl, respectively.

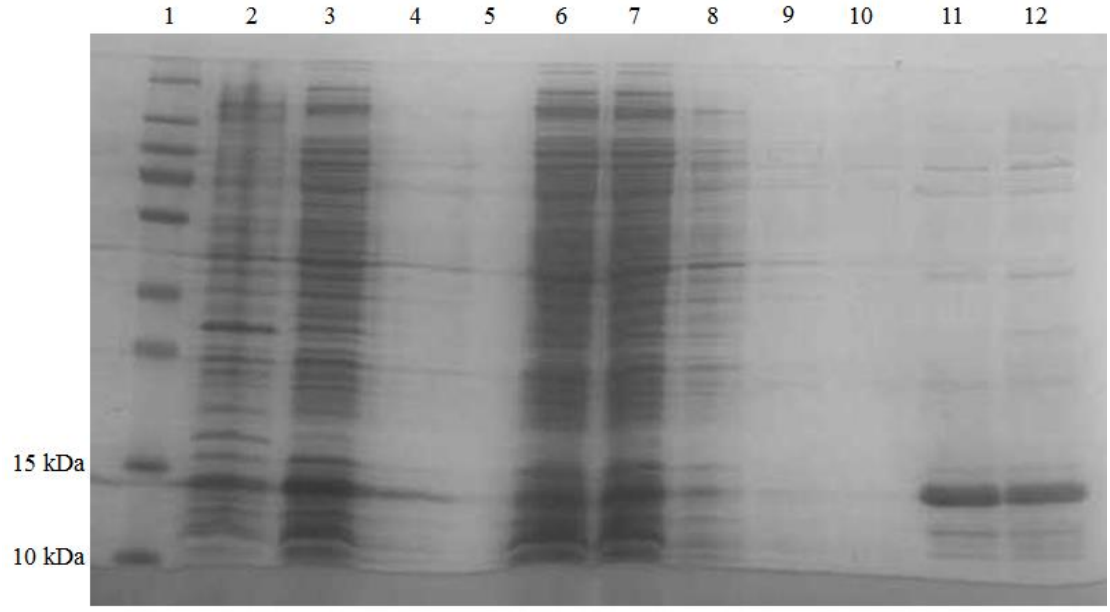


Figure 2.6 Partially purified *P. ubique* HTCC1062 DddK protein from *E. coli* cultures. Polypeptides were separated by SDS-PAGE on a 12% acrylamide gel and stained with Coomassie Blue. *E. coli* BL21 cells containing pBIO2206 were induced for DddK over-expression and separated into insoluble and soluble fractions following cell lysis. Lanes: 1, Precision-Plus protein standard (*Biorad*); 2, insoluble fraction; 3, soluble fraction; 4&5, left blank; 6-7, flow-through from first and second applications of soluble fraction onto an Ni-NTA spin column; 8-10, flow-through from three consecutive washes of the column; 11-12, Ni-NTA column eluate.

2.2.6 Enzyme kinetics of DddK

The partially pure DddK fraction obtained from the NI-NTA column was used for establishing the enzyme characteristics of DddK. To determine the K_m and V_{max} values, the initial rates of DMS production for different substrate concentrations (0-400 mM DMSP) were measured using 1.2 µg protein incubated at room temperature in sodium phosphate buffer (pH 8.0). The data are presented as a Michaelis-Menten curve in **Figure 2.7**. The V_{max} was 3.61 nmol DMS min^{-1}(µg protein)$^{-1}$, and the K_m was exceptionally high at ~82 mM DMSP.

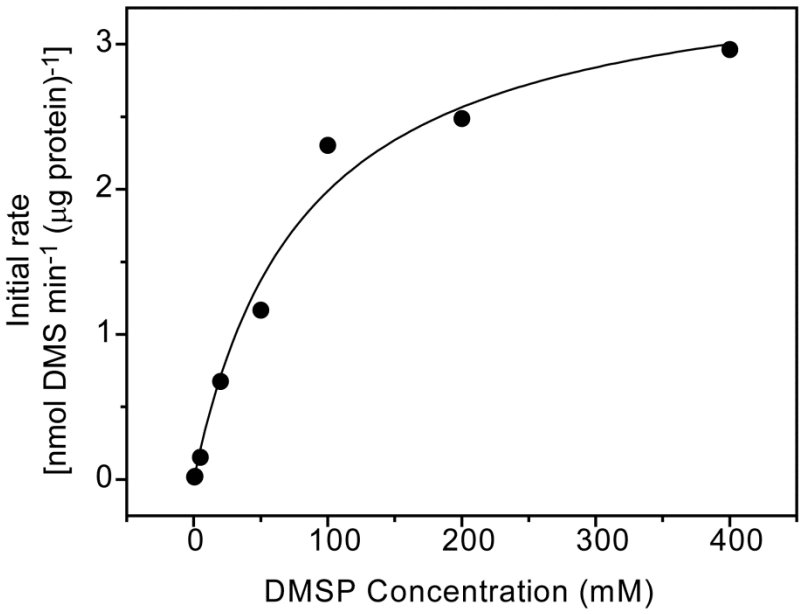


Figure 2.7 Kinetic analysis of DddK activity. Michaelis-Menten plot for the DMSP lyase activity of DddK. Data were fitted to the Michaelis-Menten equation using Origin software (version 8, Origin Labs). V_{max} was calculated as 3.61 ± 0.266 nmol DMS min^{-1}(µg protein)$^{-1}$, and K_m 81.87 ± 17.17 mM DMSP. DddK (1.2 µg) was in sodium phosphate buffer (pH 8.0.).

2.2.7 EDTA inhibits DddK activity

A characteristic of cupin superfamily proteins is that they contain a metal ion in their active site. For the majority of cupin enzymes, this metal ion is iron, but others have copper, zinc, cobalt, nickel or manganese (Dunwell *et al*., 2004). Indeed, the DMSP lyase DddQ has been shown to bind a Zn^{2+} ion in its active site (Li *et al*., 2014). Since DddK also contains a cupin-domain, it was of interest to see whether this enzyme requires a metal cofactor. To do this, the metal-chelating agent ethylenediaminetetra-acetic acid (EDTA) was used to determine if sequestering metals resulted in a decreased activity of DddK.

The EDTA experiments were carried out using the partially pure samples of DddK, in sodium phosphate buffer (pH 8.0). For this, 0.24 µg of DddK was incubated with or without 50 mM EDTA in 300 µl of buffer at 28ºC for 30 minutes. Then, 5 mM DMSP was added and the vials were sealed and incubated for a further 30 minutes, before assaying for DMS production by gas chromatography. The DMS produced was calculated as nmol min^{-1} and adjusted for protein content (**Figure 2.8**).

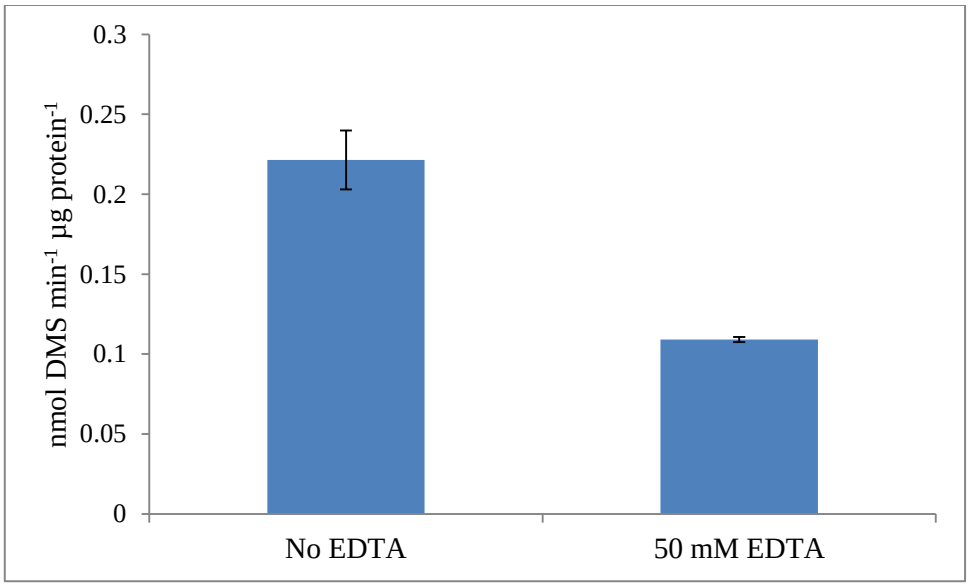


Figure 2.8 DddK activity in the presence and absence of EDTA. DddK aliquots were incubated in the presence or absence of EDTA, before the addition of 5 mM DMSP. Vials were sealed immediately, and samples were then assayed for DMS production by gas chromatography following 30 minutes incubation. Error bars represent the standard error from triplicate samples.

As shown in **Figure 2.8**, DddK in the absence of EDTA produced 0.22 nmol DMS min^{-1} µg $protein^{-1}$, but when 50 mM EDTA was present, this activity was significantly lower at 0.11 DMS min^{-1} µg $protein^{-1}$ (Welch Two Sample t-test, t = 4.946, df = 2, p = 0.037). It is therefore likely that DddK does have a metal co-factor, which is sequestered by EDTA. DddK was not, however, completely inhibited and this could be for a number of reasons. It may be that a 30 minute pre-incubation with EDTA was not enough time for the metal-chelating agent to bind all of the metal co-factors. The time-dependency of EDTA metalloenzyme inhibition was shown for a crayfish protease, which uses Zn^{2+} as a co-factor. In that case, the addition of 5 mM EDTA took 6 days to reduce the enzyme activity by 50%, a factor attributed to the tight binding of Zn^{2+} to the active site (Stöcker *et al*., 1988).

2.2.8 Homologues of DddK in the SAR11 clade

An investigation into the phylogenetic distribution of DddK revealed that this lyase is restricted to the SAR11 clade. Very close homologues (≥97% identical) are found in two other strains - HTCC1002 and HIMB5 (**Table 2.3**). These two strains and HTCC1062 belong to the SAR11 sub-clade Ia, which is the numerically dominant SAR11 sub-clade. The strains within this sub-clade are very closely related, with a 16S rRNA gene identity of ≥98% (Grote *et al*., 2012).

Table 2.3 Homologues of DddK in the SAR11 clade

Strain	Locus Tag	Identity to SAR11_0394	*E* value
HTCC1062	SAR11_0394	100%	$1e^{-92}$
HTCC1002	PU1002_04381	97%	$1e^{-89}$
HIMB5	HIMB5_00004730	73%	$3e^{-66}$

Homologues to DddK in the SAR11 clade were predicted using the peptide sequence of the functionally-verified SAR11_0394 as a query in a BLASTp search. Locus tags of homologues are presented, along with percentage sequence identity and *E* value (cut-off = $3e^{-66}$).

2.2.9 *dddK* synteny

In the HTCC1062 and HTCC1002 genomes, *dddK* is positioned downstream of *fabI*, *fabB* and *fabA* (see **Figure 2.9**) which are predicted to encode fatty acid biosynthesis pathway enzymes (Magnuson *et al*., 1993). Divergently transcribed from the *fab* genes in each case is a gene predicted to encode a Fur-family transcriptional regulator, a widely distributed family of proteins involved in the regulation of genes in response to iron availability (Andrews *et al*., 2006). Downstream of *dddK* in HTCC1062 is a gene whose product is predicted to be an S-adenosylmethionine-dependent methyltransferase. This is a large family of enzymes which catalyse the transfer of a methyl group from S-adenosylmethionine to a wide variety of acceptor substrates (Struck *et al*., 2012). For example, in DMSP biosynthesis, a SAM-methyltransferase is predicted to convert MTHB to DMSHB (Summers *et al*., 1998; see **Chapter 1**). In HTCC1002, *dddK* is upstream of a gene predicted to encode a polyribonucleotide nucleotidyl transferase, an enzyme involved in the degradation of mRNAs (Regnier *et al*., 1987).

Contrastingly, the *dddK* of HIMB5 is in a different genomic location to *dddK* of the other strains. It is divergently transcribed from a gene encoding a hypothetical protein, and downstream of a gene whose product falls into the YajQ-superfamily. YajQ proteins are involved in the temporal control of bacteriophage Phi6 gene transcription (Qiao *et al*., 2008). None of the genes surrounding *dddK* in any strain are predicted to be involved in DMSP degradation. The other cupin-type DMSP lyases are also usually found in single-gene transcripts or neighbouring genes of apparently no connection to DMSP degradation (see **Chapter 5**).

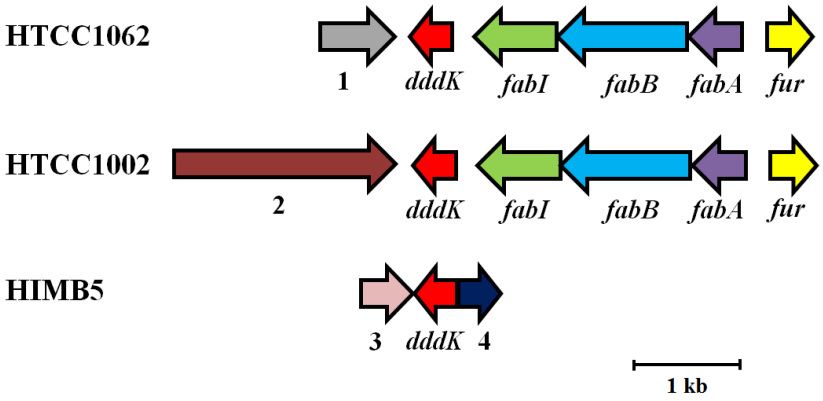


Figure 2.9 Genomic location of DddK. Arrangement of genes surrounding the *dddK* genes of HTCC1062, HTCC1002 and HIMB5 strains. Genes *fabI*, *fabB* and *fabA* encode the fatty acid biosynthesis pathway enzymes FabI, FabB and FabA, respectively. The *fur* gene encodes a Fur-family transcriptional regulator. Other genes are predicted to encode a SAM-dependent methyltransferase (1), polynucleotide phosphorylase (2), YajQ-superfamily protein (3) and a hypothetical protein (4).

2.2.10 Other DMSP lyases in the SAR11 clade

Of the seven sequenced strains of SAR11, four do not have homologues of DddK. To see if any other known DMSP lyases are present in the members of the SAR11 clade, each of the seven strains was searched using peptide sequences of DddD, DddY, DddP, DddL, DddQ and DddW, as described in **Section 2.2.1**. No homologues were found for DddD, DddY, DddL and DddW, but proteins with sequences similar to DddQ and DddP were present in some strains (**Table 2.4**). Homologues to DddP were seen in HTCC7211 (PB7211_1082), HIMB59 (HIMB59_00005110) and HIMB083 (Pelub83DRAFT_0483) with 48-51% identity to DddP2 from the γ-proteobacterium *Oceanimonas doudoroffii* (note: *O. doudoroffii* has two DddP enzymes, of which DddP2 is the most active {Curson *et al*., 2012 and see **Chapter 3**}). Interestingly, a single homologue of DddQ was seen in HIMB5 (HIMB5_00000220) with 28% sequence identity to the *R. pomeroyi* SPO1596 peptide, therefore this particular strain has homologues to two DMSP lyases – DddQ and DddK.

Table 2.4 Homologues of DddQ and DddP in the SAR11 clade

Strain	DddQ (SPO1596)	DddP (Od DddP2)
HTCC7211	-	PB7211_1082 48% $2e^{-130}$
HIMB5	HIMB5_00000220 28% $6e^{-29}$	-
HIMB59	-	HIMB59_00005110 51% $2e^{-139}$
HIMB083	-	Pelub083DRAFT_0483 49% $2e^{-134}$

Genomes of SAR11 clade were searched using BLASTp for homologues to DddQ (SPO1596) and *Oceanimonas doudoroffii* DddP2. Locus tags of the homologues are shown, with percentage identity and *E* value.

Despite the fairly low identity of the DddQ homologue, HIMB5_00000220, to the ratified DddQ from *Ruegeria pomeroyi* (SPO1596), the key residues essential for DddQ mediated DMSP cleavage are conserved, as shown in the sequence alignment of cupin-type lyases in **Figure 2.1**. This is also the case for the DddP-like homologues seen in strains HTCC7211, HIMB083 and HIMB59. A recent paper investigating the structure and mechanism of DddP from *Roseobacter denitrificans* showed that two Fe^{2+} ions bind the active site of DddP at six key residues – Asp297, Glu406, His371, Asp307, Asp295 and Glu421 (Hehemann *et al*., 2014). As shown in **Figure 2.10**, these six residues are conserved in PB7211_1082, HIMB59_00005110 and PelubDRAFT_0483, as well as functionally ratified DddPs from *R. pomeroyi*, *R. nubinhibens* and *R. denitrificans*. Two further residues, Asp377 and Tyr366 were suggested as candidates for the catalytic base in *R. denitrificans* DddP (Hehemann *et al*., 2014), and these are also conserved in all of the DddP-like sequences from SAR11 strains (**Figure 2.10**).

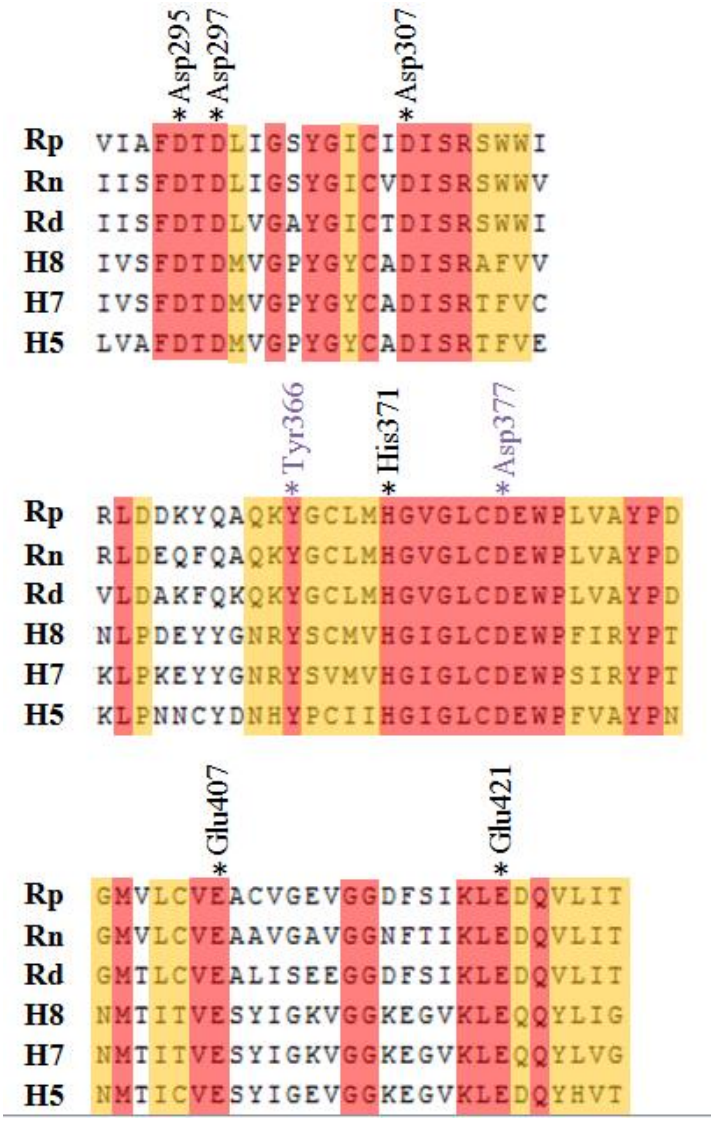


Figure 2.10 Alignment of conserved regions of DddP peptides. Completely conserved residues are highlighted in red, and residues conserved in the known DddPs, but different in the SAR11 DddPs are highlighted in yellow. Key residues predicted to be involved in DddP metal binding and catalytic activity are indicated by black or purple asterisks, respectively. Sequences are: *Ruegeria pomeroyi* SPO2299 (Rp); *Roseovarius nubinhibens* ISM_05385 (Rn); *Roseobacter denitrificans* RD1_2566 (Rd); *Pelagibacter ubique* HIMB083 Pelub83DRAFT_0483 (H8); *P. ubique* HTCC7211 PB7211_1082 (H7); *P. ubique* HIMB59 HIMB59_00005110 (H5).

2.2.11 Investigation of other DMSP lyases of the SAR11 clade

With the significant homology to known lyases, and the conservation of key residues thought to be involved in DMSP cleavage, the DddQ- and DddP-like proteins from SAR11 strains were proposed as good candidates for DMSP lyases. However, no work had previously shown these

putative *dddP* and *dddQ* genes from SAR11 members to encode functional DMSP lyases, so the next step was to ratify the function of these enzymes. Therefore, HIMB5_00000220 from strain HIMB5, and PB7211_1082 from HTCC7211 were chosen for further investigation.

2.2.11.1 Synthesis and cloning of HIMB5_00000220 and PB7211_1082

The DNA sequences of intact HIMB5_00000220 and PB7211_1082 genes, optimized for *E. coli* codon usage, were individually synthesised and cloned into pUC57 (*Eurofins MWG*) as described for SAR11_0394 above. For both genes, the 5' and 3' ends were modified to contain restriction sites *Nde*I and *Bam*HI, respectively. Using these restriction sites, the cloned DNA containing the PB7211_1082 and the HIMB5_00000220 genes were each excised from the pUC57-based recombinant plasmids and ligated into pET16b, creating plasmids pBIO2207 and pBIO2204, respectively. These two plasmids were each introduced into *E. coli* BL21 by transformation, and the resultant strains were examined for their ability to catabolise DMSP as follows.

2.2.11.2 Enzyme assays for HIMB5_00000220 and PB7211_1082

To determine whether the SAR11 *dddQ*-like and *dddP*-like genes encoded polypeptide products with DMSP lyase activity, *E. coli* BL21 cells containing pBIO2204 and pBIO2207 were grown overnight in LB containing ampicillin, and 100 nM IPTG to induce expression of their cloned putative *dddQ* or *dddP* genes respectively prior to assaying for DMSP-dependent DMS production as described above. As shown in **Figure 2.11**, *E. coli* containing pBIO2204 (*dddQ*-like gene, HIMB5_00000220) produced almost 6 nmol DMS $hour^{-1}$ μg $protein^{-1}$, approximately 150-fold greater than the control (Welch Two Sample t-test, $t = -176.2$, $df = 2$, $p < 0.001$). In contrast, the DMS produced by *E. coli* with pBIO2207 (PB7211_1082) was not significantly different to the negative control (Wilcoxon rank sum test, $W = 9$, $p = 0.1$).

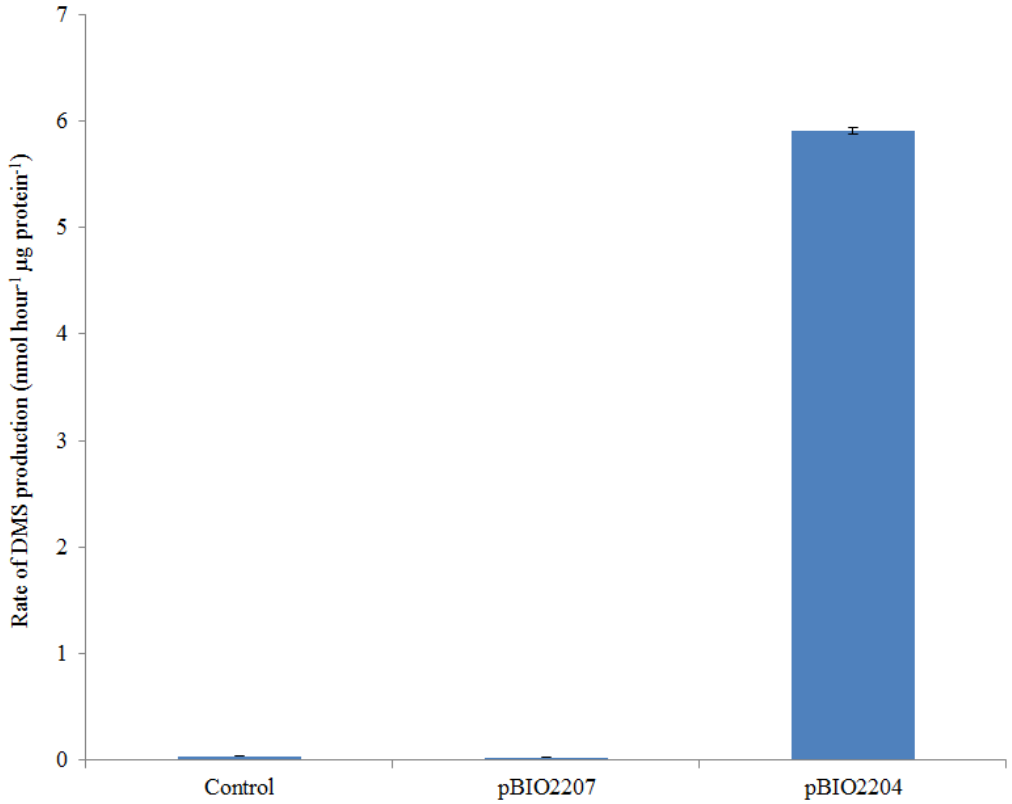


Figure 2.11 DMSP lyase activity in *E. coli* containing pBIO2204 and pBIO2207. *E. coli* BL21 cells containing pBIO2207 or pBIO2204, or empty pET16b (control), were grown in the presence of 100 nM IPTG to induce expression of *dddP*- and *dddQ*-like genes, then exposed to 5 mM DMSP for 1 hour in sealed vials and assayed for DMS production by gas chromatography. Average rates of DMS production were calculated from triplicate samples, as nmol per hour, and adjusted according to total protein content in each vial. Error bars represent the standard error of triplicate samples.

2.2.12 HIMB5_00000220 cleaves DMSP to produce acrylate

Since DddQ from HIMB5 conferred a Ddd^+ phenotype when expressed in *E. coli*, NMR spectroscopy was used to confirm the secondary catabolite produced was acrylate, as was the case in other DddQ enzymes, such as both DddQ homologues in *R. nubinhibens* ISM (Todd *et al*., 2010b). To do this, a strain of *E. coli* containing plasmid pBIO2204 (HIMB5_00000220 gene cloned in pET16b) was fed [3-C^{13}] DMSP, and cell-free extracts were obtained in the same way as described in **section 2.2.4.2** for DddK. The spectrum obtained for the DddQ extracts showed the appearance of C^{13}-acrylate (**Figure 2.12**), thus DddQ of SAR11 strain HIMB5 is able to cleave DMSP to DMS and acrylate.

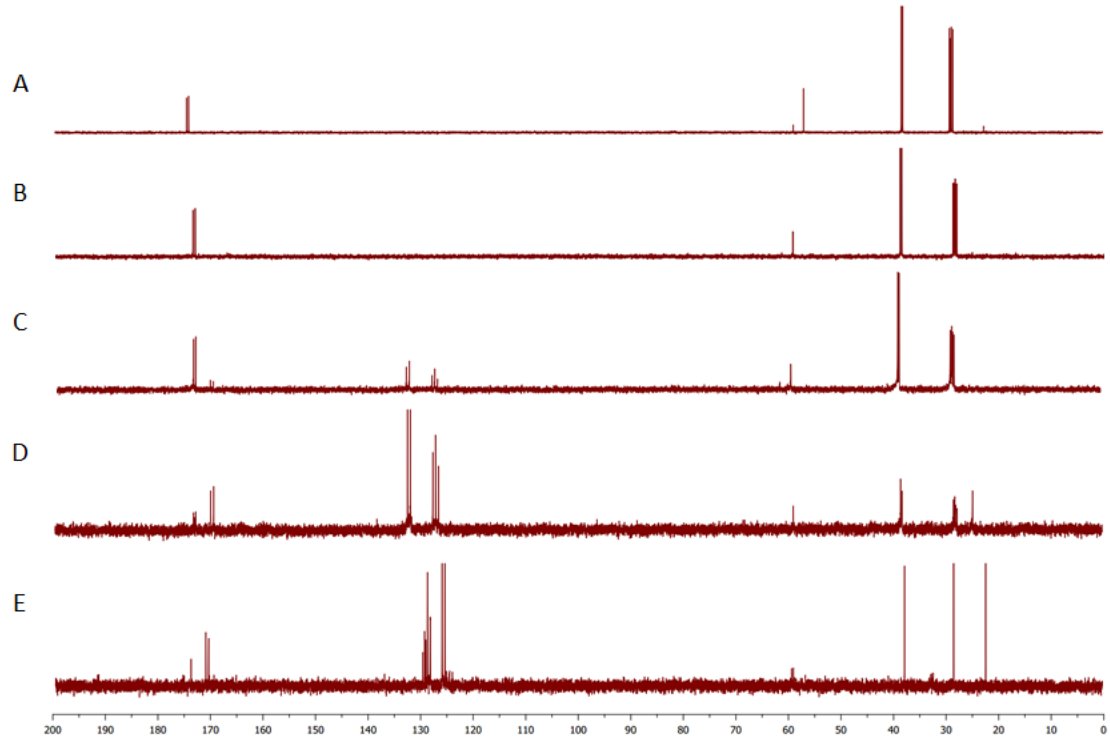


Figure 2.12 NMR spectra of *E. coli* containing DddQ. (A) [3-^{13}C]-DMSP reference; (B) wild type *E. coli* fed with [3-^{13}C]-DMSP; (C) *E. coli* containing pBIO2204 fed with [3-^{13}C]-DMSP; (D) *E. coli* with pBIO2206 fed with [3-^{13}C]-DMSP, then spiked with [3-^{13}C]-acrylate; (E) [3-^{13}C]-acrylate reference.

2.2.13 Enzyme kinetics for DddQ

Preliminary enzyme kinetics tests were carried out using cell lysates of *E. coli* BL21 containing pBIO2204 (HIMB5_00000220 gene in pET16b). To prepare the cells, *E. coli* with pBIO2204 was grown in 5 ml LB in the presence of 100 nM IPTG overnight. Of the resultant culture, 1 ml was pelleted and re-suspended in 1 ml sodium phosphate buffer (pH 8.0), then sonicated to lyse cells. The total protein concentration was measured as 0.2 µg/µl. To determine the K_m and V_{max} values, the initial rates of DMS production for different substrate concentrations (0-400 mM DMSP) were measured using 5 µl cell free extract (1.0 µg total protein), incubated at room temperature in 295 µl sodium phosphate buffer (pH 8.0). The data are presented in **Figure 2.13** as a Michaelis-Menten curve. As with DddK, DddQ of HIMB5 also had a high K_m of 131 mM for the DMSP substrate and its V_{max} value was 1.42 µmol DMS min^{-1} mg $protein^{-1}$.

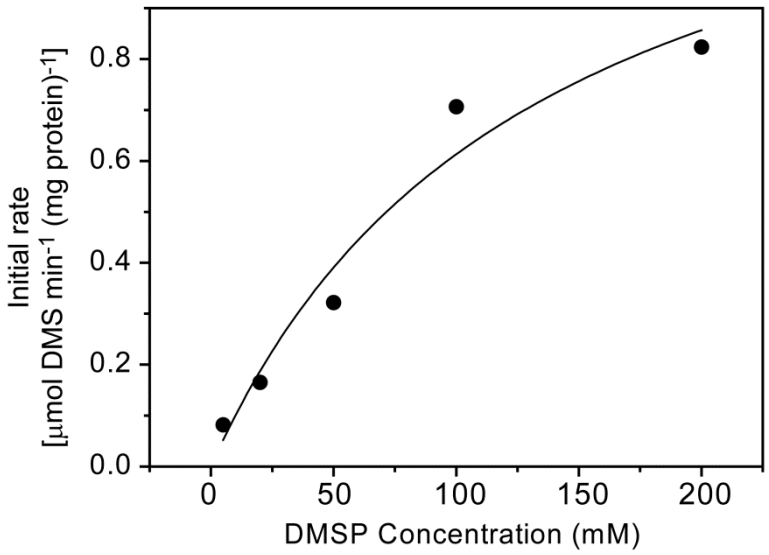


Figure 2.13 Kinetic analysis of DddQ activity. Michaelis-Menten plot for the DMSP lyase activity of *E. coli* cell extract containing pBIO2204 (cloned *dddQ* from SAR11 strain HIMB5). Data were fitted to the Michaelis-Menten equation using Origin software (version 8, Origin Labs). V_{max} was calculated as 1.42 ± 0.36 µmol DMS min^{-1}(mg protein)$^{-1}$, and K_m 131.9 ± 64 mM DMSP. DddQ was in sodium phosphate buffer (pH 8.0.).

2.3 Summary

This work, in collaboration with Stephen Giovannoni and colleagues at Oregon State University, is the first investigation into DMSP cleavage by members of the ubiquitous SAR11 clade of marine bacteria. While homologues of known DMSP lyases are present in some sequenced strains of SAR11 bacteria (see above, and discussed below), this was not the case for strain HTCC1062, pure cultures of which yielded both methanethiol and, more unexpectedly, DMS, when grown in the presence of DMSP. The work carried out here goes some way to identifying the genetic basis of this Ddd^+ phenotype, with the discovery of the novel DMSP lyase gene *dddK*.

2.3.1 DddK – a novel DMSP lyase

The initial interest in *dddK* as a potential DMSP lyase gene came from very limited sequence similarity between its encoded polypeptide and DddW. The similarity in sequence was attributable to both proteins containing a predicted cupin-barrel motif at their C-termini, a region of critical importance for the function of at least one cupin-containing DMSP lyase, namely DddQ. Furthermore, a computationally predicted structure of DddK is comparable to the empirically-determined structure of DddQ, with both enzymes possessing a similar cupin-barrel. This motif is certainly a recurring feature of DMSP lyases, with DddL also containing a cupin-barrel.

2.3.2 Functional DMSP lyases in the SAR11 clade

Currently there are complete genome sequences available for seven strains of *P. ubique*, and all but one have a homologue of DmdA, the exception being HIMB114. Interestingly, HIMB114 is also the only strain which does not have a homologue of any known DMSP lyase. The remaining six strains (HTCC1062, HTCC1002, HTCC7211, HIMB5, HIMB59 and HIMB083) have at least one convincing homologue of a DMSP lyase - DddK, DddQ or DddP. The work here shows that DddK and DddQ from strains HTCC1062 and HIMB5, respectively, are functional DMSP lyases which produce acrylate and DMS from DMSP. Conversely, this work could not confirm the functionality of a DddP homologue from HTCC7211 under the conditions used, despite the enzyme containing all key residues shown to be involved in the DddP cleavage of DMSP (Hehemann *et al.*, 2014). However, it is still possible that the enzyme works *in situ*, and it would be interesting to know whether strain HTCC7211 has a Ddd^+ phenotype.

2.3.3 DddK and DddQ have high K_m values for DMSP

The K_m values reported here for DddK and DddQ of 82 and 131 mM DMSP, respectively, seem, at first sight, to be very high. However, this seems to be a characteristic for enzymes that act on DMSP, as determined by *in vitro* assays. For example, DddP from *Roseovarius nubinhibens* had a K_m of 13.8 mM (Kirkwood *et al*., 2010a), while DmdA from *R. pomeroyi*, and indeed strain HTCC1062, had values of 5.4 mM and 13.2 mM, respectively (Reisch *et al*., 2008). These high K_m values make sense when one considers the intracellular DMSP concentrations in *R. pomeroyi* have been measured at 70 mM (Reisch *et al*., 2008). This particular characteristic of DMSP lyases will be discussed in more detail in Chapter 6.

2.3.4 Multiple DMSP lyases in SAR11 strains

Interestingly, despite having a small genome of just 1.34 Mbp, SAR11 strain HIMB5 has a functional DddQ enzyme, as well as a DddK homologue (which has since been synthesised and confirmed as functional by J. Todd {personal communication}). The retention of two DMSP lyases in such a streamlined genome suggests it is beneficial for this bacterium to have multiple DMSP enzymes. Indeed, there are many more examples of multiple DMSP lyases in species of the α-proteobacteria, particularly in the Roseobacter clade. This feature is much rarer outside of the α-proteobacteria, but it does exist, as will be explored in the next Chapter.

Chapter 3

Multiple DMSP lyases in the γ-proteobacterium *Oceanimonas doudoroffii*

3.1 Introduction

As mentioned in **Chapter 1**, the first purification and characterisation of any DMSP lyase was carried out by de Souza and Yoch in 1995(a), using the salt marsh isolate *Alcaligenes faecalis* M3A, and in a later publication they even obtained a short N-terminal sequence of the purified enzyme (de Souza and Yoch, 1996b). Almost eight years later, the UEA lab identified the corresponding gene in this strain, and termed it *dddY,* in recognition of Yoch's contributions. Thus, Curson *et al*. (2011) identified *dddY* by screening an *Alcaligenes* genomic library for recombinant cosmids that conferred a Ddd^+ phenotype when transferred to a strain of *Pseudomonas putida*. It was clear by analysing the deduced DddY gene product that this corresponded to the DMSP lyase that de Souza and Yoch had identified. Indeed, the DddY product was strongly predicted to encode a periplasmic protein whose transport through the cytoplasmic membrane involved a SecA-dependent cleavage of an N-terminal leader sequence. When this leader was removed *in silico*, the N-terminal sequence of the processed enzyme corresponded to that of the sequence that de Souza and Yoch found directly for the enzyme (de Souza and Yoch, 1996b).

In parallel to the earlier studies, de Souza and Yoch also purified and characterised another DMSP lyase from a marine species isolated from surface waters off the coast of Hawaii (de Souza and Yoch, 1995b). This was the γ-proteobacterium *Oceanimonas doudoroffii* (previously known as *Pseudomonas doudoroffii* {Brown *et al*., 2001; Baumann *et al*., 1972}). Their comparison of the DMSP lyases from *Alcaligenes* and *Oceanimonas* revealed some important similarities, but also some differences between these enzymes. Although their N-terminal sequences were very similar to each other, inhibitor studies using cyanide and *p*-chloromercuribenzoate suggested that they had different cellular locations in these two bacteria. In keeping with the periplasmic location of DddY, and therefore the absence of the need for DMSP transport, neither inhibitor affected DMS production by *Alcaligenes* cells. Conversely, both inhibitors stopped DMS production by *Oceanimonas in vivo* but not *in vitro*, suggesting the need for DMSP transport prior to its breakdown in *Oceanimonas*. Intriguingly, these observations pointed to the possibility of two versions of DddY - a periplasmic and a cytoplasmic form.

Work to find the *ddd* gene(s) in *O. doudoroffii* was all carried out by Andrew Curson in the UEA laboratory (Curson *et al*., 2012). Firstly, a near-complete (~98.5% coverage) genome sequence of *O. doudoroffii* was acquired, but, surprisingly, no gene corresponding to *dddY* could be found. Therefore, a cosmid library was constructed and this yielded three different cosmids, each of which conferred a Ddd^+ phenotype to *E. coli*. Upon closer examination, these cosmids were

found to encode homologues of previously known DMSP lyases; DddD in one case, and two somewhat different versions of DddP in the other two cosmids.

Following the identification of these three lyases by A. Curson, the aim of the work described belowwas to further investigate DMSP-dependent DMS production in *O. doudoroffii*, with a focus on the regulation of each of the *ddd* genes. All data presented below were generated through my own work, and the values obtained for promoter activities of each gene are my contribution to Curson *et al*, 2012 (see Appendix). I also present my data for DMS production by DddP1 and DddP2, and *Oceanimonas doudoroffii* itself. These DMSP lyase assays were also carried out separately by A. Curson for the published manuscript, and thus the DMS production values in this Chapter differ slightly from values given in Curson *et al*., 2012.

3.2 Results

3.2.1 *In silico* analysis of DMS-producing cosmids

Each of the three cosmids mentioned above contained a different region of *O. doudoroffii* genomic DNA, each of which was assumed to contain a functional *ddd* gene. In order to identify the genes in the cloned DNA, the sequence of the entire insert DNA of each of these was determined by 454 sequencing in the Department of Biochemistry, University of Oxford, UK (cosmid sequences were deposited at NCBI Genbank with accession numbers: JN541238; JN541239 and JN541240).

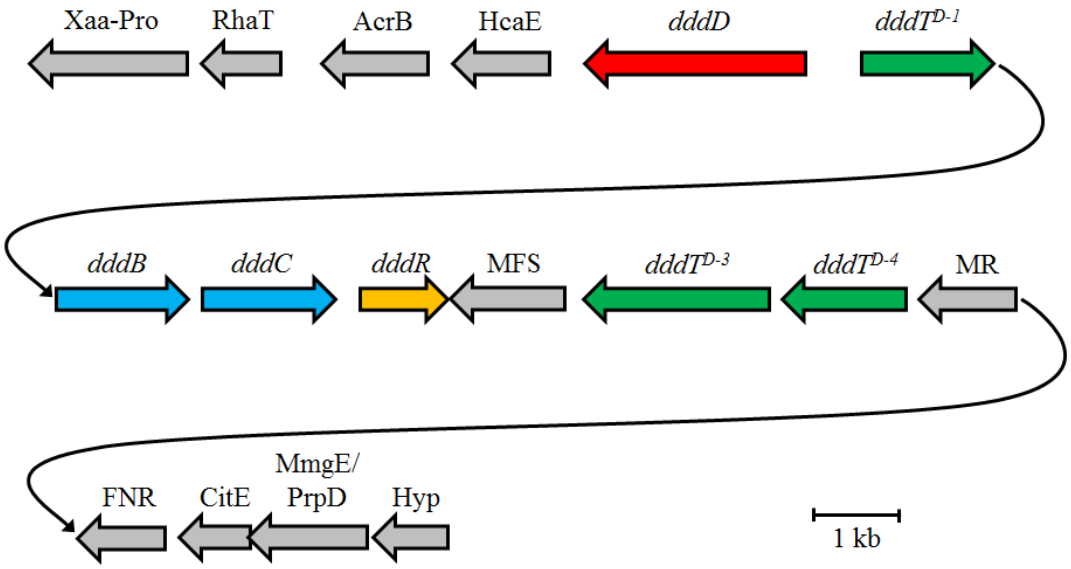


Figure 3.1 *Oceanimonas doudoroffii* genomic DNA insert in cosmid pBIO1932. Gene map showing 26,650 bp region of *O. doudoroffii* DNA inserted into pBIO1932. The arrows reflect gene orientation, and the names of each gene, or their predicted encoded polypeptide, are indicated above the arrows. A gene encoding a homologue of the DMSP lyase DddD is shown in red. Genes encoding potential DMSP transporters with similarity to DddT are shown as green arrows. Blue arrows represent genes encoding homologues of DMSP catabolism enzymes, DddB and DddC, and the yellow arrow shows a gene encoding a homologue of the LysR-type regulator DddR. Grey arrows represent genes with no predicted function in DMSP catabolism: Xaa-Pro – M24 metallopeptidase; RhaT – drug/metabolite transporter family; AcrB – cation/multidrug efflux pump; HcaE - oxygenase family polypeptide with C-terminal Rieske domain; MFS – major facilitator superfamily transporter; MR – mandelate racemase; FNR – Ferredoxin reductase; CitE – citrate lyase; MmgE/PrpD – MmgE_PrpD superfamily protein; Hyp – hypothetical protein.

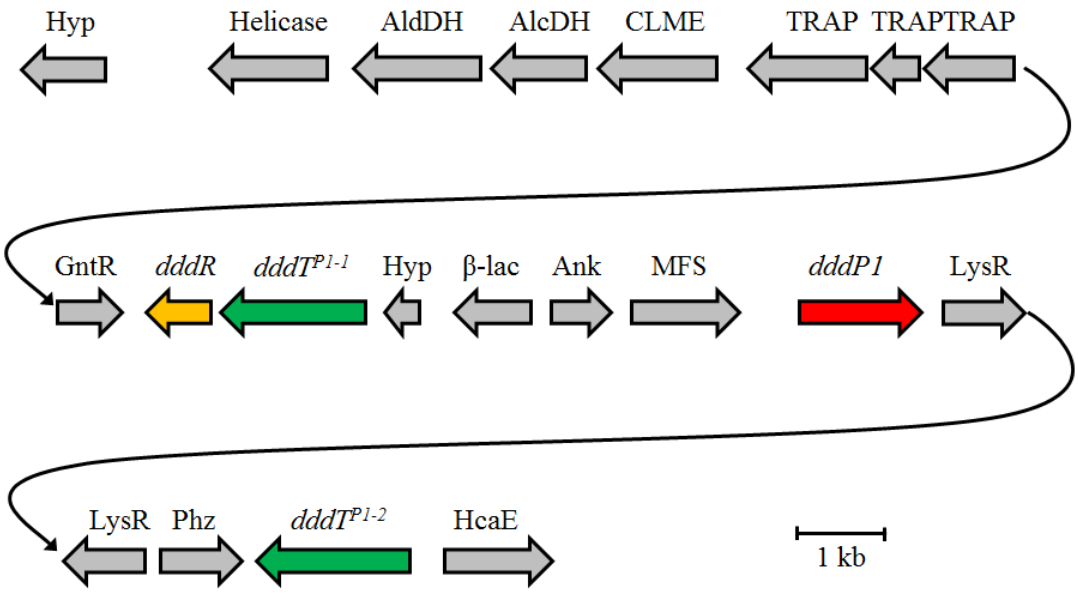


Figure 3.2 *Oceanimonas doudoroffii* genomic DNA insert in cosmid pBIO1930. Gene map showing 27,720 bp region of *O. doudoroffii* genomic DNA in pBIO1930. The arrows reflect gene orientation, and the names of each gene, or their predicted encoded polypeptide, are indicated above the arrows. The red arrow represents a gene encoding a homologue of the DMSP lyase DddP. The green arrows represent genes encoding putative BCCT transporters with homology to DddT. Grey arrows represent genes with no known connection to DMSP catabolism: Hyp – hypothetical protein; Helicase – DEAD/DEAH family helicase involved in

RNA unwinding; AldDH – aldehyde dehydrogenase; AlcDH – alcohol dehydrogenase; CLME - 3-carboxy-cis,cis-muconate cycloisomerase-like protein; TRAP – dicarboxylate transport system protein; GntR – GntR family transcriptional regulator; LysR – LysR family transcriptional regulator; β-lac – β-lactamase; Ank – ankyrin repeat-containing protein; MFS – major facilitator transporter; Phz – phenazine biosynthesis protein; HcaE - oxygenase family polypeptide with C-terminal Rieske domain.

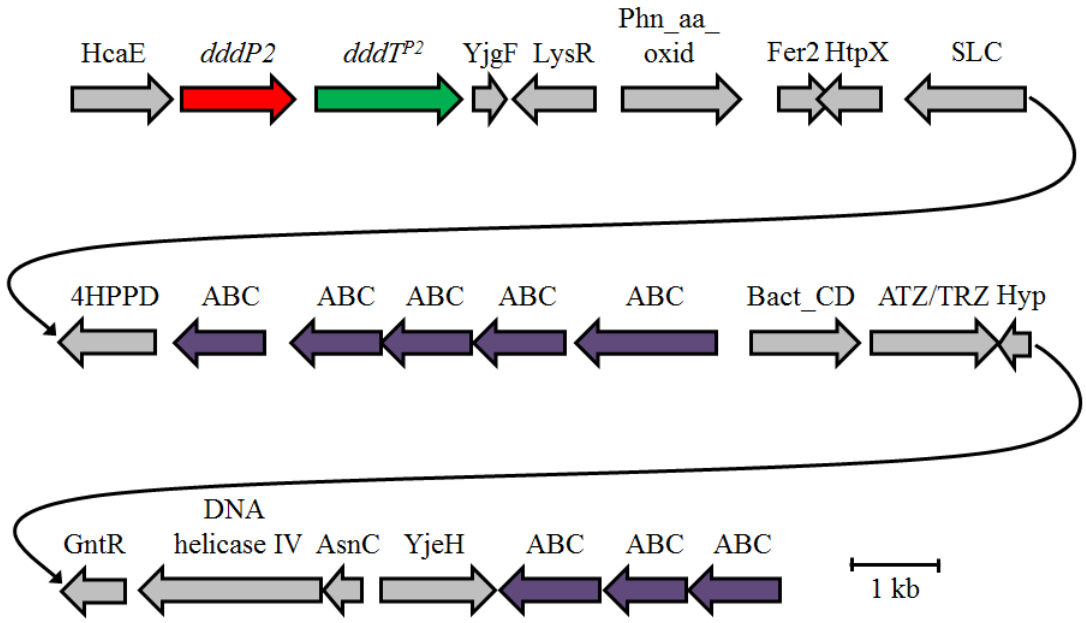


Figure 3.3 *Oceanimonas doudoroffii* genomic DNA insert in cosmid pBIO1931. Gene map of 28,830 bp region of *O. doudoroffii* genomic DNA in pBIO1931. The arrows reflect gene orientation, and the names of each gene, or their predicted encoded polypeptide, are indicated above the arrows. The red arrow represents a gene encoding a homologue to the DMSP lyase DddP. Green arrow shows a gene with homology to the BCCT-type transporter DddT. Purple arrows show genes encoding components of an ABC-transport system. Grey arrows represent genes encoding proteins not predicted to be involved in DMSP catabolism: HcaE - oxygenase family polypeptide with C-terminal Rieske domain; YjgF - YjgF/YER057c/UK114-like protein; LysR – LysR family transcriptional regulator; Phn_aa_oxid – FAD dependent oxidoreductase; Fer2 – oxidoreductase; HtpX – M48 family peptidase; SLC – sodium:neurotransmitter symporter; 4HPPD – 4-hydroxyphenylpyruvate dioxygenase; Bact_CD – N-isopropylammelide isopropylaminohydrolase; ATZ/TRZ – N-ethylammeline chlorohydrolase; Hyp – hypothetical protein; GntR – GntR-type transcriptional regulator; AsnC – AsnC-family transcriptional regulator; YjeH – inner membrane protein.

The inserted DNA in each cosmid was aligned with the genome sequence, and found to be contiguous (**Figures 3.1, 3.2 and 3.3**). A copy of *dddP* was present in both pBIO1930 and pBIO1931, whereas pBIO1932 had a homologue of *dddD*. These genes are discussed in detail in the following sections.

3.2.1.1 pBIO1930

The DddP homologue encoded by *dddP1* in pBIO1930 is 448 amino acids in length with a predicted molecular weight of 49.97 kDa. As discussed in **Chapter 1**, DddP polypeptides fall into different sub-groups (see **Figure 1.12**) – the Roseobacter-type DddPs, the fungal and γ-proteobacterial DddPs, and a third, smaller subgroup consisting of outliers from a few Rhodobacterales and Pseudomonadales. A BLASTp search with DddP1 revealed close homologues in species in the second subgroup of γ-proteobacteria and fungi. For example, DddP1 is 66% identical to the DddP of *Aspergillus oryzae* NIB40 and of *Fusarium graminearum* cc19, the genes of which have both been cloned and shown to be functional when expressed in *E. coli* (Todd *et al*., 2009). Close homologues of DddP1 are also present in the γ-proteobacteria *Psychrobacter* sp. JCM18900 (80% identical), *Marinomonas ushuaiensis* DSM15871 (82% identical) and various *Vibrio* spp. (79% identical). In terms of homology to the Roseobacter-type DddPs, DddP1 is 55% identical to DddP of the α-proteobacterium *Roseovarius nubinhibens* ISM, which has been shown to be functional (Todd *et al*., 2009). Near *dddP1*are two genes named $dddT^{P1\text{-}1}$ and $dddT^{P1\text{-}2}$ which are both predicted to encode BCCT-type transporters (**Figure 3.2**). As discussed in **Chapter 1**, a BCCT-type transporter from *Halomonas* sp. HTNK1, named DddT, was shown to transport DMSP (Sun *et al*., 2012). The BCCT-type transporters encoded by $dddT^{P1\text{-}1}$ and $dddT^{P1\text{-}2}$ are each 31% identical to DddT from *Halomonas* HTNK1, and 24% identical to each other. The phylogenetic relationship of $DddT^{P1\text{-}1}$ and $DddT^{P1\text{-}2}$ compared to other DddT peptides is shown in **Figure 3.4**, and will be discussed further in a later section.

Interestingly, there is a gene downstream of $dddT^{P1\text{-}1}$ which is predicted to encode a LysR-type transcriptional regulator, whose closest homologue in any sequenced organism is DddR from *Marinomonas* sp. MWYL1. The gene, annotated *dddR* in **Figure 3.2**, encodes a polypeptide with 34% sequence identity to *Marinomonas* DddR. As discussed in **Chapter 1**, DddR from *Marinomonas* is a positive regulator of the DMSP lyase gene *dddD* (Todd *et al*., 2007), and is found divergently transcribed from the *dddD* operon. There are other examples of close homologues to *Marinomonas* DddR being encoded by genes near *dddD*, for example in *Pseudomonas* sp. J465, but there are no other examples of *dddR* near to a *dddP* gene. It would

therefore be of interest to see if the *dddR* gene near *dddP1* in *O. doudoroffii* encodes a regulator involved in expression of DMSP lyase.

3.2.1.2 pBIO1931

The second of the *Oceanimonas* DddP homologues is 416 amino acids long, with a predicted molecular weight of 47.64 kDa. Designated DddP2, it is 51% identical to DddP1. In contrast to DddP1, DddP2 has close homologues amongst the smaller subgroup of outliers from the γ-proteobacteria *Pseudomonas mandelii* and *Acinetobacter baumannii*, with identities of around 70% (see **Figure 1.12**). Significantly, DddP2 is also 48% identical to PB7211_1082, the SAR11 DddP-like peptide which did not have DMSP lyase activity under laboratory conditions (see **Chapter 2**). However, DddP2 also has 43% identity to the *Roseovarius nubinhibens* DddP, which, contrastingly, has been confirmed as functional (Todd *et al*., 2009).

Similarly to *dddP1*, *dddP2* is also positioned near to, in this case 150 bp upstream of, yet another *dddT*-like gene annotated as $dddT^{P2}$ (**Figure 3.3**), which encodes a predicted BCCT-transporter with 32% identity to DddT from *Halomonas* sp. HTNK-1.

Also of note, 50 bp downstream of, and possibly co-transcribed with, *dddP2* is a gene encoding an iron-sulphur cluster polypeptide, HcaE. Although this protein has no known role in DMSP catabolism, *hcaE* is also found divergently transcribed from $dddT^{P1\text{-}2}$ (**Figure 3.2**), and 330 bp downstream of *dddD* (**Figure 3.1**).

3.2.1.3 pBIO1932

The DddD of *O. doudoroffii* is very similar to functional homologues found in *Halomonas*, *Marinomonas* and *Pseudomonas* spp. (56%, 71% and 75% identical, respectively). In keeping with DddD polypeptides, it also consists of two CaiB-like domains with an interlinking region. In *Oceanimonas,* the *dddD* gene is divergently transcribed from a *dddTBCR* gene cluster, which is also found in *Marinomonas* sp. MWYL1, and which encodes proteins involved in the transport and downstream catabolism of DMSP – in which DddB is an aldehyde dehydrogenase and DddC is an alcohol dehydrogenase. Interestingly, in addition to the transporter gene $dddT^{D\text{-}1}$ in this cluster, two further genes encoding putative betaine transporters are found near *dddD*.

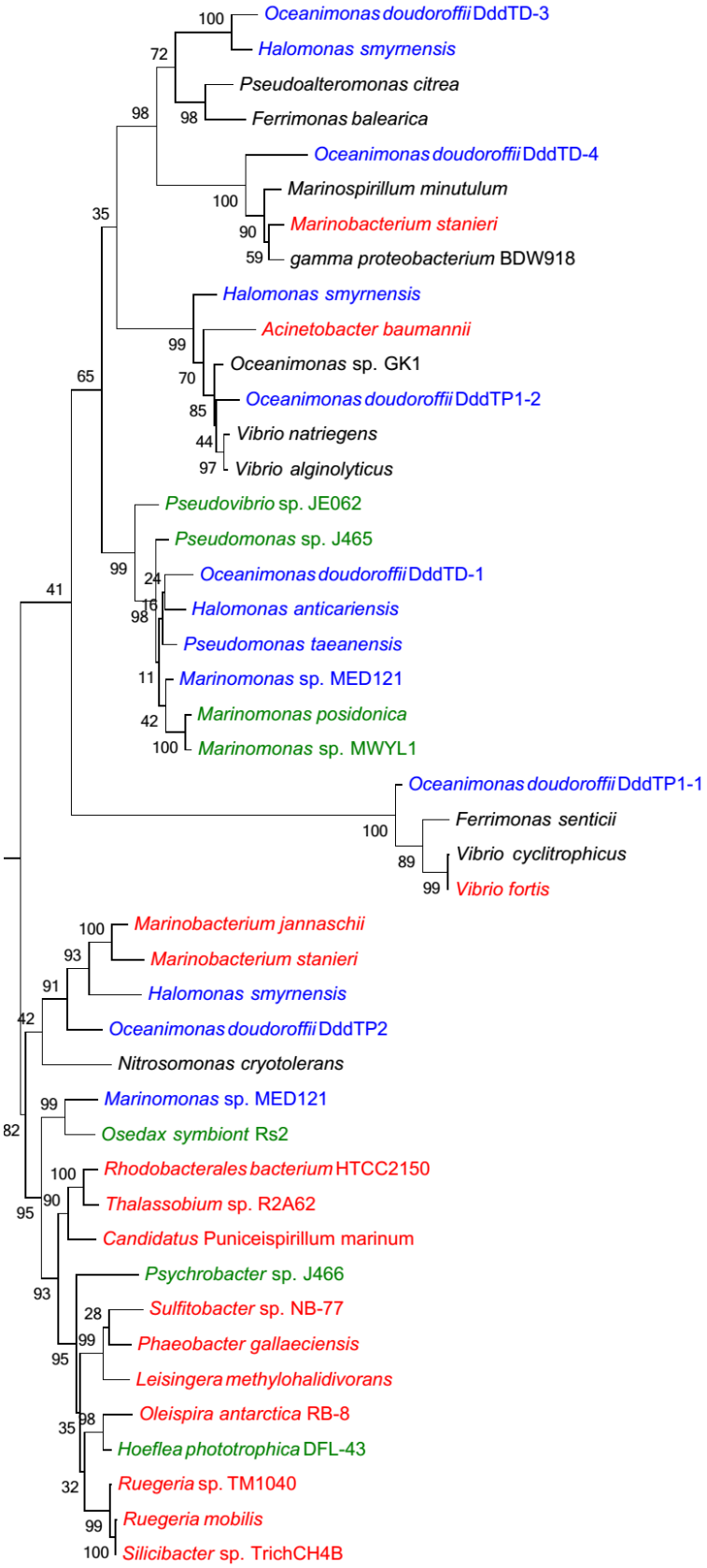

Oceanimonas doudoroffii DddTD-3
Halomonas smyrnensis
Pseudoalteromonas citrea
Ferrimonas balearica
Oceanimonas doudoroffii DddTD-4
Marinospirillum minutulum
Marinobacterium stanieri
gamma proteobacterium BDW918
Halomonas smyrnensis
Acinetobacter baumannii
Oceanimonas sp. GK1
Oceanimonas doudoroffii DddTP1-2
Vibrio natriegens
Vibrio alginolyticus
Pseudovibrio sp. JE062
Pseudomonas sp. J465
Oceanimonas doudoroffii DddTD-1
Halomonas anticariensis
Pseudomonas taeanensis
Marinomonas sp. MED121
Marinomonas posidonica
Marinomonas sp. MWYL1
Oceanimonas doudoroffii DddTP1-1
Ferrimonas senticii
Vibrio cyclitrophicus
Vibrio fortis
Marinobacterium jannaschii
Marinobacterium stanieri
Halomonas smyrnensis
Oceanimonas doudoroffii DddTP2
Nitrosomonas cryotolerans
Marinomonas sp. MED121
Osedax symbiont Rs2
Rhodobacterales bacterium HTCC2150
Thalassobium sp. R2A62
Candidatus Puniceispirillum marinum
Psychrobacter sp. J466
Sulfitobacter sp. NB-77
Phaeobacter gallaeciensis
Leisingera methylohalidivorans
Oleispira antarctica RB-8
Hoeflea phototrophica DFL-43
Ruegeria sp. TM1040
Ruegeria mobilis
Silicibacter sp. TrichCH4B

Figure 3.4 Phylogenetic relationship of selected DddT homologues. Protein sequences of DddT homologues were aligned with MEGA 6 and used to estimate an unrooted phylogenetic tree using LG model, gamma distributed with invariant sites. The scale bar indicates number of substitutions per site. Bootstrap values of 500 replicates are given at the base of each branch pair. Species in blue have a homologue of both DddD and DddP, species in red have only DddP and species in green have only DddD. Species names written in black do not have any known DMSP lyase.

3.2.1.4 Multiple DddT transporters in *O. doudoroffii*

As described in the previous three sections, genes encoding homologues of the BCCT-type DMSP transporter DddT are prevalent near to the DMSP lyase genes in *O. doudoroffii*. Thus, two *dddT* genes are closely linked to *dddP1*, another is immediately adjacent to *dddP2* and no fewer than three different genes are found near *dddD*. While the encoded polypeptides of these six genes do have homology to each other, they fit phylogenetically into different subgroups (**Figure 3.4**), and form some interesting patterns with the presence or absence of the lyases DddD or DddP in these organisms. Thus, DddT^{D-3}, DddT^{D-4}, DddT^{P1-1} and DddT^{P1-2} are most similar to polypeptides from organisms which mostly lack any DMSP lyase, although there are a few exceptions. For example, DddT^{D-3} and DddT^{P1-2} are each very similar (69% identical) to transporters from *Halomonas smyrnensis* which, like *O. doudoroffii*, also has *dddD* and *dddP* genes. However, *H. smyrnensis* has at least three different DddT homologues, and the third of these is more similar to DddT^{P2}. Some other exceptions are *Marinobacterium stanieri*, which has a DddT 62% similar to DddT^{D-4}, *Acinetobacter baumannii*, with a homologue 64% identical to DddT^{P1-2} and *Vibrio fortis* with a homologue 68% identical to DddT^{P1-1}. All three of these species have homologues to the *Roseovarius nubinihibens* DddP of between 33-55%.

In contrast, the closest homologues to DddT^{D-1} and DddT^{P2} are found in organisms which mostly have *dddP* or *dddD* genes, but again, these are divided into different groups. Thus the closest seven hits to DddT^{D-1} are all species of γ-proteobacteria with a copy of *dddD*, or as for *Halomonas anticariensis*, *Pseudomonas taeanensis* and *Marinomonas* sp. MED121, both *dddD* and *dddP*. Contrastingly, the closest eighteen homologues of DddT^{P2} are predominantly α-proteobacteria from the Roseobacter clade, which have a copy of *dddP*. Again, there are a few exceptions, such as the β-proteobacterium *Nitrosomonas cryotolerans* which doesn't have any known DMSP lyase, and a few γ-proteobacteria which have a copy of *dddP*, *dddD* or both genes.

It is very likely that DddT^{D-1} is involved in the transport of DMSP, given its location in the *dddTBCR* operon that is found in other *dddD*-containing bacteria, and its 70% sequence identity

to DddT of *Marinomonas* sp. MWYL-1 which has been shown to be a functional DMSP transporter (Sun *et al*., 2012). It is interesting that $DddT^{P2}$ should cluster with homologues from α-proteobacteria which have *dddP*. This is especially intriguing, given that the *dddP* genes in these organisms are not closely linked on the genome to any BCCT-transporter (see **Chapter 5**), yet in *O. doudoroffii*, *dddP2* and $dddT^{P2}$ are contiguous. So far, the only confirmed DMSP transporters in Roseobacter species are ABC-type transport systems from *Roseovarius nubinhibens* ISM and *Sulfitobacter* sp. EE-36 (Sun *et al*., 2012). However Roseobacter strains also have multiple BCCT-transporters (Moran *et al*., 2004) and it would be interesting to see if they use their $DddT^{P2}$ homologues to import DMSP.

3.2.2 Cloning *dddP1* and *P2* into the expression vector pET21a

Since the two DddP enzymes in *O. doudoroffii* are only 51% identical to each other, it was of interest to check if both of these enzymes were functional DMSP lyases. To do this, each intact *dddP* gene was PCR amplified from *O. doudoroffii* genomic DNA. Restriction sites *Nde*I and *Bam*H1 were integrated into the forward and reverse primers in each case. The purified PCR products of each *dddP* gene were digested with *Nde*I and *Bam*HI and ligated into the vector pET21a, cut with the same enzymes (work done by A. Curson). The pET21a vector works in the same way as pET16b (see **Chapter 2**), allowing for the expression of cloned genes under the T7 promoter. The recombinant pET21a plasmids containing either *dddP1* or *dddP2* were transformed into *E. coli* BL21 and transformants were selected on the basis of ampicillin resistance. Following sequence verification of cloned inserts, pET21a containing *dddP1* was designated pBIO1933 and pET21a with *dddP2* was renamed pBIO1934.

3.2.3 DddP1 and DddP2 confer a Ddd^+ phenotype on *E. coli* BL21

To assay DMSP-dependent DMS production by each DddP enzyme, BL21 transformants containing pBIO1933 and pBIO1934 were each inoculated to LB containing ampicillin and incubated at 37°C for 2 hours. Then, cultures were incubated for a further 4 hours at 30°C in the presence or absence of 0.2 mM IPTG. The cultures were then washed in M9 minimal media and added to GC vials containing 5 mM DMSP. The cells were assayed by gas chromatography after 2 hours incubation at room temperature. The amount of DMS produced was calculated in pmol per minute, and normalised to total protein content, as measured by Bradford's assay. The background level of DMS produced by BL21 cells with an empty vector was 0.012 pmol min^{-1} μg $protein^{-1}$. As shown in **Figure 3.5**, *E. coli* containing each of the cloned *dddP1* and *dddP2* genes, and whose expression was induced by addition of IPTG were effective in producing DMS form DMSP – by factors of some 30 and 560-fold above background respectively. As

expected, when IPTG was omitted, there was little sign of DMSP lyase activity with either of these *dddP* genes.

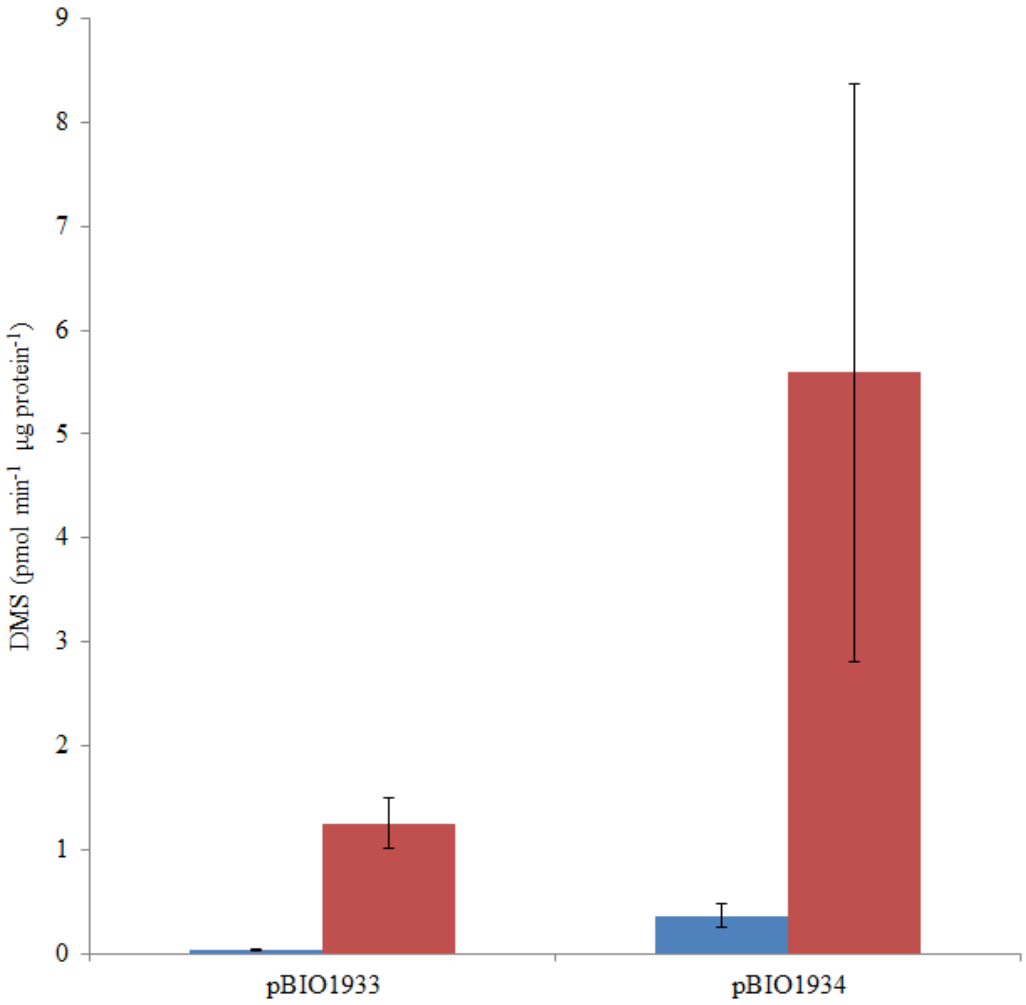


Figure 3.5 DMSP-dependent DMS production by *E. coli* BL21 containing pBIO1933 and pBIO1934. *E. coli* BL21 cells containing pBIO1933 or pBIO1934 were grown in the presence of 0.2 mM IPTG (red) to induce expression of the cloned *dddP1* and *dddP2* genes, or no inducer (blue). Cells were exposed to 5 mM DMSP for 2 hours prior to assaying DMS production by gas chromatography. Rates of DMS production were calculated in pmol DMS per minute and adjusted according to protein content. Assays were carried out in duplicate, and error bars show standard error.

3.2.4 Induction of DMS production in *Oceanimonas doudoroffii*.

To investigate the induction of the Ddd$^+$ phenotype, separate cultures of *O. doudoroffii* were grown overnight in M9 minimal media containing 10 mM succinate as the carbon source, with or without potential inducers, namely DMSP (5 mM), acrylate (1 mM) and 3HP (1 mM). Once in stationary phase, the cultures were washed in M9 minimal media to remove inducers, and an aliquot of each culture was transferred to a glass vial with 5 mM DMSP substrate, and incubated for 25 minutes prior to assaying by gas chromatography. The buffer-only control comprised M9 media plus 5 mM DMSP, but no cells.

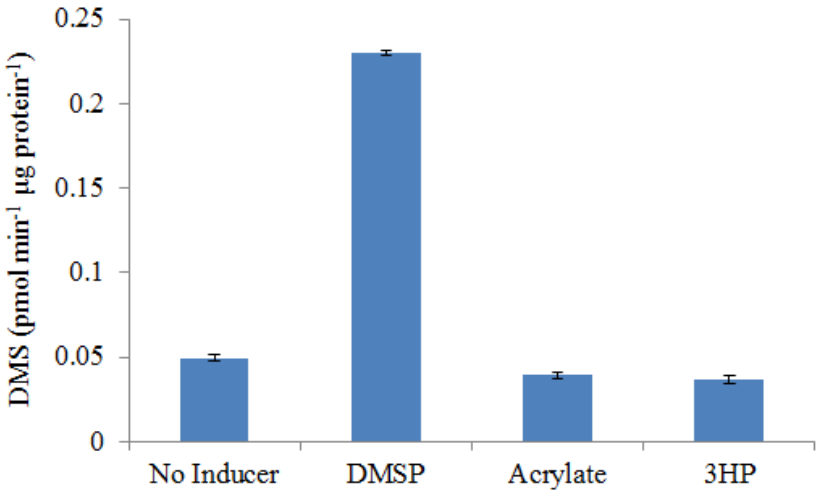


Figure 3.6 Rates of DMS production in *O. doudoroffii*. *O. doudoroffii* cultures were grown in the presence of DMSP, acrylate, 3HP or no inducer, then exposed to 5 mM DMSP substrate for 25 minutes before measuring the amounts of DMS. Average rates of DMS production from duplicate cultures were calculated in pmol per minute, per µg total protein in each vial. Standard error bars are shown.

As shown in **Figure 3.6**, DMS production in *O. doudoroffii* was enhanced ca. 4.6-fold when cells were pre-grown in the presence of DMSP, compared to when no inducer was present. Pre-growth in acrylate or 3HP did not increase DMS production above "no-inducer" levels.

3.2.5 Induction of DMS production in *O. doudoroffii* is likely due to enhanced transcription of *dddD*.

Having shown that DMS production by *O. doudoroffii* was enhanced by DMSP, it was of interest to see if this was due to the increased transcription of any of the three DMSP lyases, *dddP1*, *dddP2* or *dddD*. Transcription of the *dddT* genes immediately adjacent to *dddD* and *dddP2* was also investigated, to determine if DMSP transport might also be up-regulated by the presence of this substrate.

3.2.6 Construction of transcriptional *lacZ* fusions using the reporter plasmid pMP220

To investigate expression of the various *ddd* genes of *O. doudoroffii*, appropriate transcriptional *lacZ* fusions were constructed in the plasmid pMP220 (Spaink *et al*., 1987) (**Figure 3.7**). This plasmid contains an *E. coli lacZ* gene that lacks a promoter and operator, and has a multicloning site (MCS) positioned 5' of the promoter-less *lacZ* and a ribosomal binding site derived from the *E. coli* chloramphenicol acetyl transferase gene. There are no transcriptional terminators between the MCS and *lacZ*, so the reporter gene can be expressed from promoters cloned into the MCS. Assays of β-galactosidase activity can then be used to indicate the abundance of *lacZ* transcripts, and thus promoter activity. A critical feature of pMP220 is its wide host-range, enabling the determination of promoter activity in strains grown in different environmental conditions. In this case, pMP220 constructs in *E. coli* were individually mobilised into a rifampicin-resistant strain of *O. doudoroffii* (strain J495) via conjugal transfer using a patch cross. Since pMP220 confers tetracycline resistance upon its host, successful crosses into J495 were selected for using rifampicin and tetracycline.

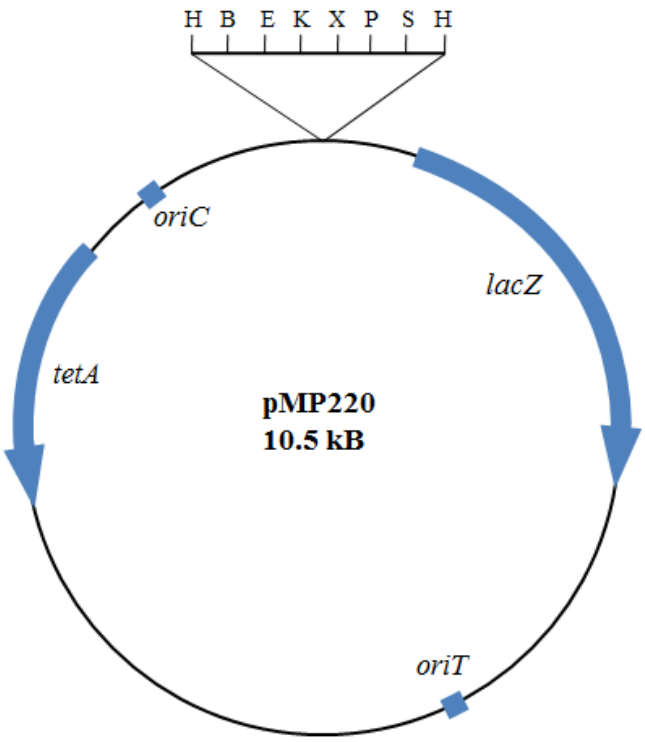


Figure 3.7 Schematic representation of pMP220. Map of pMP220 with *lacZ* reporter gene, tetracycline resistance gene (*tetA*), origin of replication (*oriC*) and origin of transfer (*oriT*). The multicloning site (MCS) is upstream of *lacZ* (H, *Hind*III; B, *Bgl*II; E, *Eco*RI; K, *Kpn*I; X, *Xba*I; P, *Pst*I; S, *Sph*I).

3.2.7 PCR amplification of *ddd* promoter regions

In order to clone the promoters upstream of each of the two *dddP* genes, *dddD*, $ddd T^{P2}$ and $ddd T^{D-1}$, PCR was used to amplify the upstream regions using *O. doudoroffii* genomic DNA as a template. To ensure that the promoter was cloned, in each case primers were designed to amplify the entire upstream intergenic space, with the 5' end originating in the upstream gene, and the 3' end extending into the start of the gene of interest (see **Figure 3.8, Table 3.1** and **Table 7.5**). The exception to this was *dddP2*, where the space between this and the upstream gene encoding HcaE is only 50 bp, so *dddP2* may be transcribed from the *hcaE* promoter. Therefore to ensure the *dddP2* promoter was cloned, primers were designed to amplify a 1500 bp fragment containing the *hcaE* promoter region, the intact *hcaE* gene and the intergenic space upstream of *dddP2*.

In all cases, primers were designed to contain restriction sites to enable directional cloning into pMP220. For *dddP2*, the forward and reverse primer contained an *Xba*I and a *Pst*I site,

respectively. For all other genes, the forward primer contained an *Eco*RI site, and the reverse primer an *Xba*I site. Following PCR, an aliquot of the reaction was checked for a product of the expected size. These were purified, digested using the restriction enzymes stated above, and separated on an agarose gel. Digested fragments were then extracted and individually ligated into pMP220, which had also been digested with the same restriction enzymes.

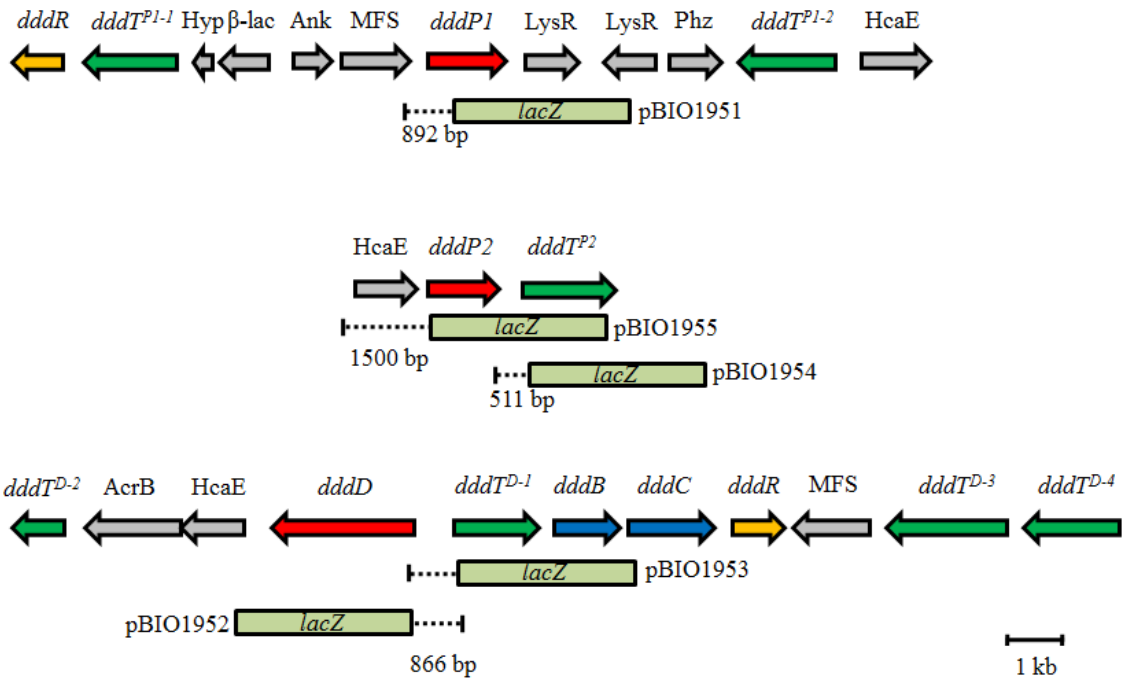


Figure 3.8 Transcriptional fusion constructs to *ddd* genes of *O. doudoroffii*. Transcriptional *lacZ* fusions were made for *dddP1*, *dddP2* (extending upstream HcaE), *dddT*^{P2}, *dddD* and *dddT*$^{D-1}$. The PCR-amplified DNA is shown as dotted lines, with sizes of each product displayed below. The names of the resultant plasmid constructs are also shown. Diagram is drawn to scale, as indicated by the scale bar.

Table 3.1 Primers used to amplify promoter regions of *O. doudoroffii ddd* genes

Gene	Forward Primer	Reverse Primer
dddP1	OdP1proEcoRFOR1	OdP1proXbaREV1
$ddd T^{P1}$	OdDddTproEcoFOR	OdDddTproXbaREV
dddP2	OdP2proFOR3	OdP2proXbaREV1
dddD	dddDproEcoFOR1	dddDproXbaREV1
$ddd T^{D-1}$	dddDTproEcoR1	dddDTproXbaRev1

Sequences of each primer are presented in Table 7.5.

3.2.8 Cloning promoter regions into pMP220

Following ligation and transformation, transformant colonies were picked to LB broth containing tetracycline and grown overnight, prior to restriction enzyme analysis. Plasmids containing an insert of the correct size were then verified by sequencing. The resultant *dddP1*-, *dddP2*-, $ddd T^{P2-}$, *dddD*- and $ddd T^{D-1}$-*lacZ* fusion plasmids were designated pBIO1951, pBIO1955, pBIO1954, pBIO1952 and pBIO1953, respectively.

3.2.9 β-galactosidase expression from *ddd* promoter fusions

To determine the activities of each of the cloned promoters, the reporter plasmids pBIO1951, pBIO1955, pBIO1954, pBIO1952, pBIO1953 and the pMP220 plasmid were individually mobilised into *O. doudoroffii* J495 via conjugal transfer. Transconjugants were selected for on LB agar containing rifampicin and tetracycline, and individual colonies were purified. Each transconjugant strain was grown overnight in LB broth with tetracycline. These cultures were washed to remove media and antibiotics and then used to inoculate M9 minimal media containing 10 mM succinate as a carbon source, plus each of the potential inducers DMSP, acrylate, 3HP or glycine betaine (GB, a structural analogue of DMSP), each at 2mM, or with no added inducer. Cultures were grown overnight, and cell densities were recorded before assaying for β-galactosidase activity. The enzyme activity of each transcriptional fusion strain was compared to background levels of β-galactosidase activity produced by J495 containing an "empty" pMP220 plasmid.

As shown in **Figures 3.9 and 3.10**, β-galactosidase activities from the *dddP2* promoter fusion (pBIO1955) and the *dddT*^{P2} fusion (pBIO1954), were no different to the background levels produced by the vector pMP220 itself, in any of the conditions tested. However, the *dddP1* fusion (pBIO1951) did have low level constitutive expression , which was significantly different to the negative control in the "no inducer", DMSP, 3HP and glycine betaine induction conditions (Pairwise t-tests, P <0.05). Additionally, promoter activity for pBIO1951 in the presence of acrylate, while not significantly different to the control, was still over 2-fold greater than pMP220 only. The expression of the *dddP1* fusion was not enhanced by any of the induction conditions (ANOVA, $F_{4,15}$, = 0.716, P = 0.594).

In contrast, the *dddD* fusion (pBIO1952) was very strikingly (~60-fold) induced by DMSP, but not by any of the other compounds that were tested (**Figure 3.11**; ANOVA, $F_{4,5}$ = 32.79, P <0.001, post-hoc Tukey HSD tests showed DMSP was significantly different to other inducers, P <0.01). The *dddT*$^{D-1}$ fusion (pBIO1953) was also significantly induced by DMSP (ANOVA, $F_{4,5}$ = 15.28, P <0.01, post-hoc Tukey HSD, P <0.05), and unlike *dddD*, the *dddT*$^{D-1}$ fusion was also expressed, albeit at low level, in the absence of any co-inducer (**Figure 3.10**; Pairwise t-test shows a significant difference between the pMP220 control and pBIO1953 in "no-inducer", 3HP and glycine betaine induction conditions, P <0.01. There is no significant difference for the acrylate condition, but this data set included a value of 0 for one point, which may bias the data). In contrast, expression of the *dddT*$^{P-1}$ fusion (pBIO1954) was not significantly different to the negative control in any condition (Pairwise t-test).

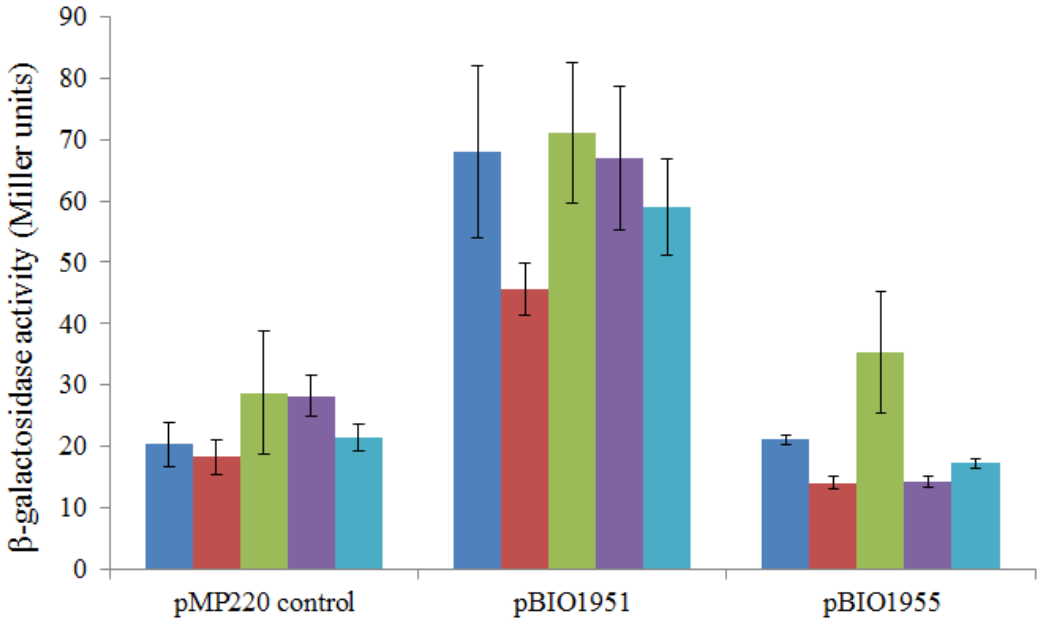


Figure 3.9 β-galactosidase activities of *dddP1*- and *dddP2-lacZ* fusion plasmids in *O. doudoroffii*. Activities of the *dddP1*- and *dddP2-lacZ* fusion plasmids (pBIO1951 and pBIO1955, respectively) and an "empty" pMP220 vector were measured in *O. doudoroffii*, following pre-growth in the presence of either DMSP (red), acrylate (green), 3HP (purple), or glycine betaine (light blue) compared to cells with no inducer (blue). Standard error bars represent data from two biological replicates, and two technical replicates.

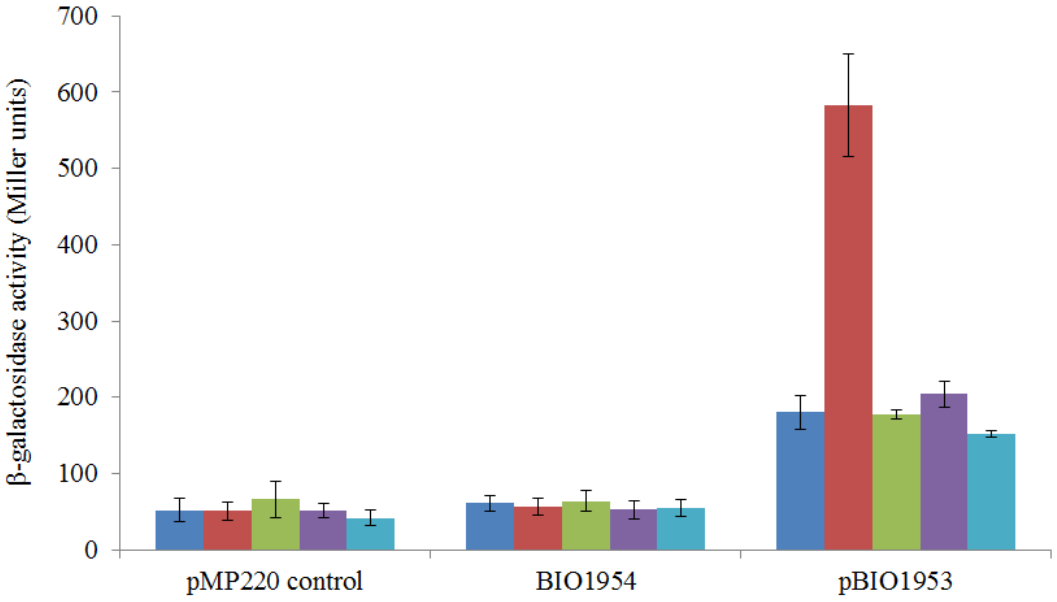


Figure 3.10 β-galactosidase activities of the *dddT^{P1}*- and *dddT^{D-1}-lacZ* fusion plasmids in *O. doudoroffii*. Activities of the *dddT^{P1}*- and *dddT^{D-1}-lacZ* fusion plasmids (pBIO1954 and pBIO1953, respectively) and an "empty" pMP220 vector were measured in *O. doudoroffii*, following pre-growth in the presence of either DMSP (red), acrylate (green), 3HP (purple), or glycine betaine (light blue) compared to cells with no inducer (blue). Standard error bars represent data from two biological replicates, and two technical replicates.

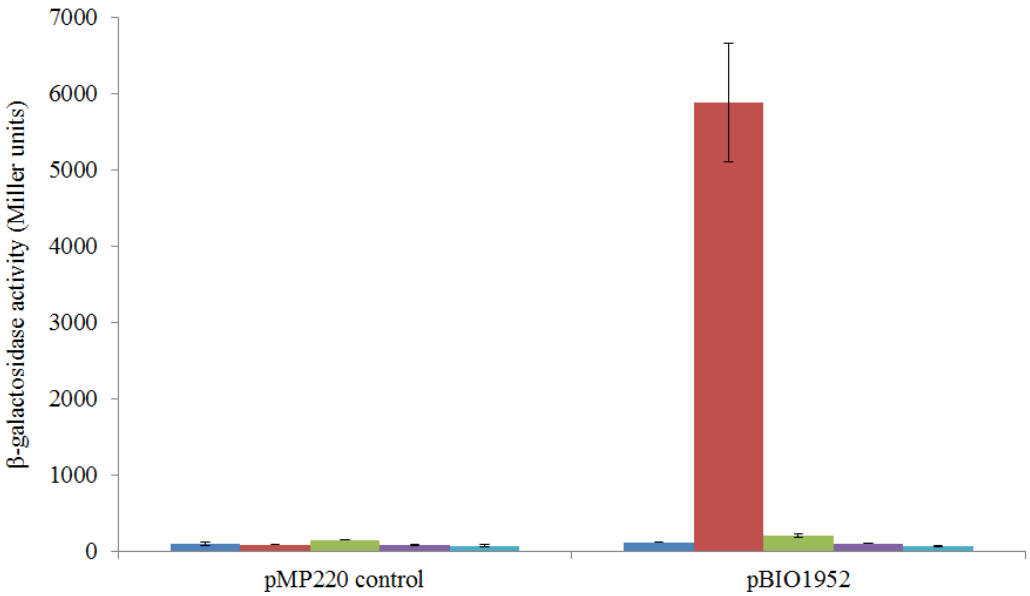


Figure 3.11 β-galactosidase activity of the *dddD-lacZ* fusion plasmid pBIO1952 in *O. doudoroffii*. β-galactosidase activities of the *dddD-lacZ* fusion plasmid (pBIO1952) and an "empty" pMP220 vector in *O. doudoroffii*, following pre-growth in the presence of either DMSP (red), acrylate (green), 3HP (purple), or glycine betaine (light blue) compared to cells with no inducer (blue). Standard error bars represent data from two biological replicates, and two technical replicates.

3.2.10 *O. doudoroffii* grows on DMSP as a sole carbon source

To test if *O. doudoroffii* could use DMSP or any of its breakdown products as sole sources of carbon, growth experiments were carried out on solid and liquid media. In each case, M9 minimal media was used. For growth on agar plates, the M9 media was supplemented with glucose, succinate, acrylate, 3HP or different concentrations of DMSP. A negative control, which lacked any added carbon source was used to rule out growth on the agar itself. After incubating at 28°C for 1 week, no growth was seen in the negative control, or on plates containing glucose, 3HP or acrylate. However, *O. doudoroffii* grew equally well on plates supplemented with 10 mM succinate and 5 mM DMSP. Weak growth was also seen on plates containing 0.5 mM DMSP (**Table 3.2**).

Table 3.2 Growth tests of *O. doudoroffii* on solid media with different sole carbon sources.

Carbon Source	Concentration	Growth
No Carbon	-	-
Succinate	10 mM	+++
Glucose	10 mM	-
DMSP	0.5 mM	+
DMSP	5 mM	+++
Acrylate	2 mM	-
3HP	2 mM	-

***O. doudoroffii* was streaked on M9 agar plates with or without different sole carbon sources. Level of growth represented by – (no growth), + (slight growth) or +++ (thick growth).**

Liquid growth tests were carried out in M9 minimal media supplemented with either 5 mM or 1 mM DMSP, acrylate, 3HP or glycine betaine. Again, negative controls contained no added carbon source. An overnight culture of *O. doudoroffii* grown in LB was washed, then diluted 1:100 into 5 ml M9 with or without the listed carbon sources. The cells were incubated at 28°C for 24 hours, when the OD_{600} was recorded (**Figure 3.12**). In accordance with the results seen using solid media, *O. doudoroffii* could grow with 5 mM DMSP as a sole carbon source, and showed slight growth when 1 mM DMSP was present, but did not grow on 3HP or acrylate. It was also noted that this strain could use the DMSP analogue glycine betaine as a sole source of carbon.

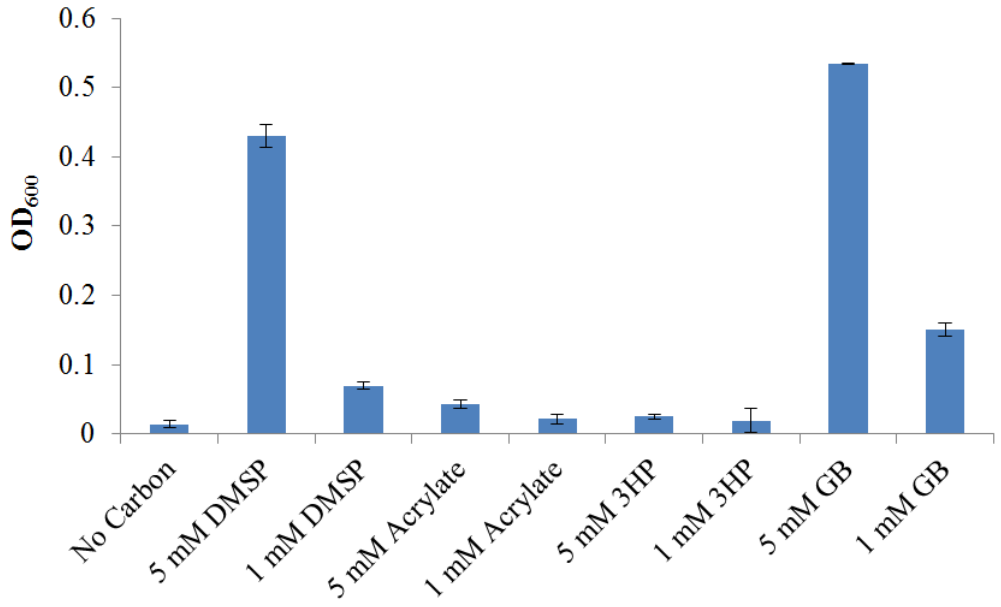


Figure 3.12 Growth of *O. doudoroffii* in liquid media supplemented with different sole carbon sources. Cultures were incubated at 28°C in M9 minimal media supplemented with 5 mM or 1 mM of DMSP, acrylate, 3HP or GB as sole carbon sources. Cell densities were measured after 24 hours as absorbance at 600 nm. Standard error bars represent data from duplicate growth tests.

3.3 Summary

3.3.1 *O. doudoroffii* has three functional DMSP lyases

Although the occurrence of multiple pathways for DMSP catabolism is quite common in the Roseobacter clade, and was also noted in SAR11 strain HIMB5 (see **Chapters 2 and 5**), *O. doudoroffii* is the first example of a γ-proteobacterium with multiple DMSP lyases. The presence of DddD is not unusual for this sub-phylum, but DddP is mostly restricted to the α-proteobacteria, along with some species of fungi (Todd *et al*., 2009). Furthermore, *O. doudoroffii* is the only organism known to have two copies of *dddP*. The work here shows that both of these DddP lyases are functional when cloned and expressed in *E. coli*, despite having limited sequence similarity, although DddP2 had greater lyase activity than DddP1. Indeed, DddP1 and DddP2 fall into two distinct phylogenetic groups, both quite distantly related to the DddP enzymes found in the Roseobacter clade, and each containing a miscellany of unrelated organisms. For example, DddP1 is most similar to homologues found in a few other γ-proteobacteria, including *Vibrio* and *Pseudomonas* spp. and some species of fungi. DddP2 is closely related to the DddP of *Acinetobacter baumannii* and *Pseudomonas mandelii*, and a few α-proteobacteria (see **Figure 1.12**). These include the SAR11 strains HTCC7211, HIMB59 and HIMB083 mentioned in the previous chapter, although the DddP-like homologue of HTCC7211 was not functional under laboratory conditions (**Chapter 2**). The sporadic occurrence of DddP in such unrelated species strongly indicates multiple incidences of horizontal gene transfer.

Currently, the only other example of a γ-proteobacterium with multiple DMSP lyases is the recently sequenced marine bacterium *Leucothrix mucor*, which has homologues to *dddP* and *dddD*. Thus the presence of multiple DMSP lyases is not confined to the Roseobacter clade, but occurs more widely, perhaps in other marine bacteria that may come into contact with high levels of DMSP.

3.3.2 DMSP-dependent DMS production is inducible by DMSP

In line with earlier findings (de Souza and Yoch, 1995b), the work here shows that DMS production in *O. doudoroffii* is inducible by pre-growth in the presence of DMSP. This form of positive regulation by the substrate of an enzyme in not unusual in bacteria; indeed the paradigm of the *E. coli lac* operon is regulated in this way (Jacob and Monod, 1961). DMSP degradation is also induced by DMSP in other DMS-producing species, for example in *Marinomonas* sp. MWYL-1 and *R. nubinhibens* ISM (See **Chapter 1**; Todd *et al*., 2007, 2009).

In *O. doudoroffii*, the work carried out here provides strong evidence that the induction of DMS production by DMSP is due to the much-enhanced transcription of *dddD* in the presence of this

substrate. In contrast, *dddP1* was expressed at low levels in all conditions tested, and *dddP2* was not expressed at detectable levels in any condition. Previous work had shown that in some members of the Roseobacter clade, the expression of *dddP* was inducible by DMSP, but only to a small degree. For example, *dddP* expression in *Roseovarius nubinhibens* is enhanced ca. 4-fold in response to DMSP (Todd *et al*., 2009), while microarray data from *Ruegeria pomeroyi* shows a 7-fold increase in *dddP* transcripts in DMSP-containing media (M Kirkwood, personal communication). It is not clear why the *dddP* genes of *O. doudoroffii* are not expressed in a similar way. However, given the relatively low factor of induction of *dddP* in *R. nubinhibens* and *R. pomeroyi* compared to other *ddd* genes (e.g. *dddD* in this study, which was induced 60-fold) it may be that the true inducer of *dddP* is yet to be discovered in all of these organisms.

3.3.3 *O. doudoroffii* can grow on DMSP as a sole carbon source

The work carried out here shows that *O. doudoroffii* can use DMSP as a sole source of carbon. However, it was unable to grow on the C3 catabolites acrylate and 3HP. This situation is similar to that of *Marinomonas* MWYL1, another γ-proteobacterium which can use DMSP, but not its C3 catabolites, as sole carbon sources. The reason behind this growth phenotype is likely to be the *dddTBCR* operon which is found adjacent to *dddD* in both *O. doudoroffii* and *Marinomonas* MWYL1, as well as other *dddD*-containing γ-proteobacteria which grow on DMSP. These ancillary *ddd* genes are described in detail in **Chapter 1**, but, in short, they provide the means to transport DMSP into the cytoplasm where it is converted to 3HP via DddD, which is then degraded to acetaldehyde by DddB and DddC. The transcriptional regulator DddR is probably required for the expression of *dddD*. In other Ddd^+ bacteria which also have the ability to use acrylate as a sole carbon source, for example *Halomonas* HTNK1 and *Alcaligenes faecalis*, additional genes known to be involved in acrylate catabolism, namely *acuN* and *acuK*, are also found nearby to the DMSP lyase gene. Consistent with its inability to grow on acrylate, *Oceanimonas* lacks these genes, but it does contain a copy of the acryloyl-CoA reductase AcuI (56% identical to *Rhodobacter sphaeroides* 2.4.1 AcuI), which may be present to deal with a toxic build-up of acryloyl-CoA resulting from the DddP-mediated cleavage of DMSP into acrylate.

The ability to use DMSP as a sole source of carbon certainly seems to be particularly associated with species of γ-proteobacteria, but it is not restricted to this class. Indeed, a few members of the Roseobacter clade are able to use DMSP, or sometimes acrylate, as carbon sources, despite the fact they lack the *dddTBCR* operon (González *et al*, 2003; Schäfer *et al*., 2005). Amongst these species is the model bacterium *Ruegeria pomeroyi* DSS-3, and the work presented in the next Chapter is an investigation into the pathway used by this organism to assimilate carbon from DMSP.

Chapter 4

DMSP and acrylate catabolism in *Ruegeria pomeroyi* DSS-3

4.1 Introduction

As described in the previous chapter, *Oceanimonas doudoroffii* can use DMSP as a sole carbon source, likely *via* the enzymes encoded by *dddD* and the *dddTBCR* operon (Curson *et al*., 2012). Operons with this same general organisation and function are often found in γ-proteobacteria containing DddD, and also in the DddY-containing β-proteobacterium *Alcaligenes faecalis* M3A. In both cases, the operon is linked to the primary *dddD* or *dddY* gene. In *Alcaligenes*, and also in *Halomonas* HTNK1, the cluster also contains genes for acrylate catabolism, namely *acuN* and *acuK*, which are responsible for the ability of these strains to use acrylate as a sole carbon source.

The ability to use DMSP and acrylate as sole carbon sources is much less common amongst the α-proteobacteria, which tend to possess either DddP, or the cupin-type DMSP lyases DddL, DddQ, and DddW. Furthermore, while homologues to DddA, DddC, AcuN and AcuK are present in α-proteobacteria, the corresponding genes are not clustered with the *ddd* genes that encode the lyases as is the case for *Alcaligenes* and *Halomonas*. Good homologues to these enzymes are actually widely distributed amongst marine and terrestrial bacteria, most with no connection at all to DMSP. Since all of these enzymes have highly conserved regions related to their generic functions as dehydrogenases or hydratases, it is difficult to pinpoint a cut-off point at which they no longer act on acrylate or 3HP.

Interestingly, a few species of the Roseobacter clade have been reported to grow on DMSP and acrylate as sole carbon sources. Amongst these is the model strain *Ruegeria pomeroyi* DSS-3, although previous reports of the growth of this strain on various sole carbon sources were inconsistent, even when these were from the same authors. Thus, in 1999, González *et al*. reported that it could use DMSP as a sole carbon source, but could not grow on acrylate. However, in a later study the same authors claimed that it used both DMSP and acrylate as sole carbon sources (González *et al*., 2003). Therefore one purpose of the work below was to establish if *R. pomeroyi* is indeed able to use these and other related carbon sources.

Like several members of the Roseobacter clade, *R. pomeroyi* can produce both MeSH and DMS from DMSP. A pathway that leads to MeSH production in *R. pomeroyi* was presented by Reisch *et al*. (2011), and is shown in detail in **Chapter 1** (see **Figure 1.4**). Briefly, DMSP is demethylated to MMPA by DmdA, which is then converted to MMPA-CoA by either of two versions of the MMPA-CoA ligase, termed DmdB and DmdB2, which is then converted to MTA-CoA by DmdC, of which there are at least three functional homologues in *R. pomeroyi*. The fourth step, mediated by DmdD, yields MeSH, CO_2 and acetaldehyde, which may contribute to central carbon metabolism via the ethylmalonyl-CoA pathway (Reisch *et al*., 2011). It is

interesting to note that mutations in *dmdB* and one of the *dmdC* genes (SPO3804) did not abolish growth on DMSP as a sole carbon source, whereas a mutation in *dmdD* made *R. pomeroyi* hypersensitive to the presence of DMSP. Thus, the results from that study did not rule out the possibility of an alternative route of DMSP carbon assimilation in *R. pomeroyi*.

In addition to DmdA, *R. pomeroyi* DSS-3 also possesses homologues of four DMSP lyase genes - *dddD, dddP*, *dddQ* and *dddW*. The products of the latter three have been shown to contribute to DMS production from DMSP in this strain. In addition to each cloned gene conferring a Ddd^+ phenotype to *E. coli*, mutations in *dddP*, *dddQ* or *dddW* all reduced the amount of DMSP-dependent DMS production by *R. pomeroyi* (Todd *et al*., 2010b; Todd *et al*., 2012a). Thus, DMS production decreased by 50% in each of the $dddP^-$ and $dddW^-$ mutant strains, and almost abolished in the $dddQ^-$ mutant, when compared to wild type (Todd *et al*., 2010b; Todd *et al*., 2012a). In contrast, a mutation in *dddD* did not decrease DMSP lyase activity in *R. pomeroyi*, suggesting that this gene does not encode a functional DMSP lyase in this strain (Todd *et al*., 2010b).

Thus, in *R. pomeroyi*, DMSP is broken down to acrylate by at least three different enzymes. At the start of this study it had not been established if this strain could metabolise the acrylate generated from DMSP, and, if so, it was certainly not known how. Although this bacterium, like virtually all bacteria, does have good homologues to AcuN, AcuK, DddA and DddC, the corresponding genes are spread throughout the genome, and are not found near those for any DMSP lyases, suggesting they are not involved in the same pathway as seen in the γ-proteobacteria. Given the lack of an acrylate catabolism gene cluster in *R. pomeroyi*, the main purpose of this study was to ascertain if and how *R. pomeroyi* is able to metabolise DMSP, via the downstream catabolite acrylate.

4.2 Results

4.2.1 *R. pomeroyi* growth on sole carbon sources

To test whether *R. pomeroyi* DSS-3 can use DMSP and other related sources of carbon in our laboratory conditions, its growth in minimal medium (MBM) containing different carbon sources was measured. In addition to DMSP, the carbon sources included the primary products of DMSP degradation, namely acrylate and MMPA, succinate as a positive control, and propionate. Propionate was tested as a possible downstream catabolite of acrylate.

To do this, starter cultures were grown overnight in rich media (1/2 YTSS), before adjusting to identical OD_{600} values of 1.0 and washing in MBM buffer. The washed cells were diluted 1:100 into 100 ml MBM containing either succinate, DMSP, or propionate (each at 5 mM), or MMPA or acrylate (both at 2 mM, due to their toxicity at higher concentrations) as sole carbon sources. The cultures were incubated at 28°C with shaking at 200 rpm and their OD_{600} were measured at intervals until cultures had reached stationary phase (20-80 hours, depending on the carbon source).

As shown in **Figure 4.1**, *R. pomeroyi* used all five compounds as a sole source of carbon, but with varying effectiveness. Succinate was by far the most effective, leading to a maximum OD_{600} of 0.6 by 20 hours. Growth on propionate was slightly slower, (maximum OD_{600} of 0.5 after 40 hours). The growth on DMSP and acrylate was very similar, and a maximum OD_{600} of 0.2 was reached by 40 hours. Finally, growth on MMPA was much slower, with a maximum OD_{600} of 0.2 at 80 hours.

Thus, *R. pomeroyi* is indeed able to use acrylate as a sole carbon source, and the following work was carried out to deduce a possible route of acrylate metabolism in this strain. An important observation was that *R. pomeroyi* also used propionate as a sole carbon source, a finding that was of interest since propionate, or its breakdown products, can be derived from acrylate-type molecules.

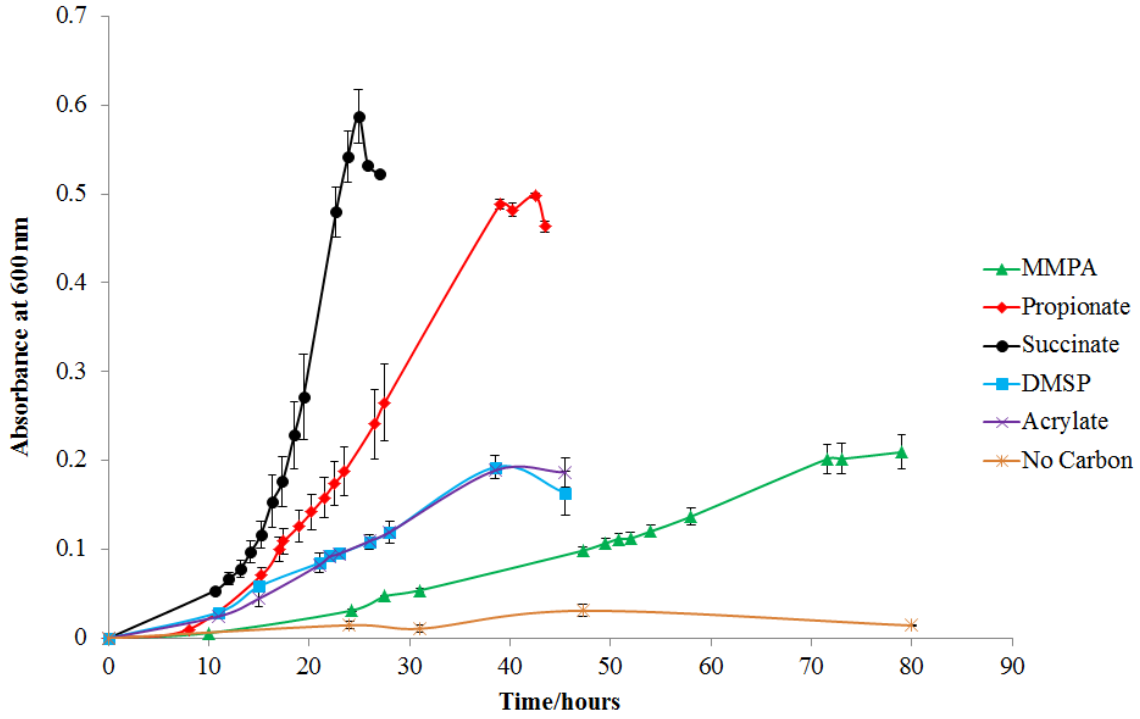


Figure 4.1 Growth of *Ruegeria pomeroyi* on sole carbon sources. Growth curves of *R. pomeroyi* were measured in MBM with either no carbon, or 5 mM succinate, 5 mM propionate, 5 mM DMSP, 2 mM MMPA, or 2 mM acrylate as sole carbon source. The OD_{600} values of the cultures were measured at regular intervals, until stationary phase was reached. Error bars represent the standard error of three biological replicates.

4.2.2 Strategy for identifying genes involved in DMSP/acrylate catabolism

The initial search for genes likely to be involved in DMSP and downstream acrylate catabolism took advantage of a set of microarray data on *R. pomeroyi* grown in the presence or absence of DMSP or acrylate, previously obtained by Mark Kirkwood. In total, the expression of 72 genes and 77 genes was enhanced >3-fold when DMSP or acrylate were present, respectively, compared to the succinate default medium. Of those, 19 genes were enhanced >3-fold in both DMSP and acrylate conditions (**Table 4.1**).

Table 4.1 Genes up-regulated >3-fold when *R. pomeroyi* was grown in the presence of DMSP and acrylate

Gene	DMSP	Acrylate	Putative protein product
SPO0363	3.31	6.393	Hypothetical
SPO0759	10.4	19.13	Hypothetical
SPO1094	**5.869**	**8.612**	**Propionyl-CoA carboxylase, beta-subunit**
SPO1095	4.992	10.89	Hypothetical
SPO1101	**3.264**	**6.052**	**Propionyl-CoA carboxylase, alpha-subunit**
SPO1105	**3.599**	**6.418**	**Methylmalonyl-CoA mutase**
SPO1177	6.563	8.574	Hypothetical
SPO1372	3.719	3.449	Hypothetical
SPO1809	5.369	14.35	Histidine kinase
SPO1810	3.203	7.9	Sodium:solute symporter
SPO1811	3.69	9.754	Hypothetical
SPO1912	3.375	5.88	GntR family transcriptional regulator
SPO1913	19.61	16.59	DmdA
SPO1914	**14.28**	**11.56**	**AcuI**
SPO2067	3.61	3.832	Hypothetical
SPO2203	4.412	3.192	Methylmalonate semialdehyde dehydrogenase
SPO2758	5.135	3.021	Hypothetical
SPO2792	3.553	3.333	Outer membrane protein
SPO2934	**3.26**	**3.879**	**Propionate-CoA ligase**

***R. pomeroyi* was grown in the presence of 10 mM succinate, with or without 5 mM DMSP or 2.5 mM acrylate. The fold change in expression of each gene (locus tag given in column one) in DMSP or acrylate, compared to the succinate-only control, was derived as the average of two biological replicates. The predicted protein encoded by each gene is listed in the right-hand column. Genes predicted to encode enzymes of the propionate catabolism pathway are shown in bold. Also in bold is SPO1914, a putative acryloyl-CoA reductase, which could catalyse the production of propionyl-CoA from acryloyl-CoA.**

In order to interpret the microarray data, advantage was taken of the KEGG (Kyoto Encyclopedia of Genes and Genomes) resource (http://www.genome.jp/kegg/). The KEGG

database project was started in 1995 as a reference resource for the biological interpretation of genomic sequencing data (Kanehisa *et al*., 2000; 2014). One feature of the resource is the KEGG pathway maps which represent networks of metabolism, and indicate the types of enzymes involved in each step. If the genes of a genome sequenced organism have been assigned KEGG numbers, one can display specific pathway maps for that organism. These maps are annotated with the locus tags of genes encoding proteins that are predicted to catalyse each step of the pathway. Since the genes from *R. pomeroyi* have been assigned KEGG numbers, I was able to generate the relevant KEGG maps. Another very useful feature of KEGG pathways is that the user is able to upload a set of transcriptome data, which is then incorporated into the metabolism maps. Thus, if at least one gene encoding a protein predicted to be involved in a certain step is differentially expressed, this step of the pathway will be coloured accordingly. Up-regulated genes are coloured from yellow to red, with red being a higher expression, and down-regulated genes are indicated in shades of green.

Thus, the microarray data for *R. pomeroyi* in the presence of acrylate or DMSP was uploaded onto KEGG for interpretation. Strikingly, genes predicted to be involved in propionate metabolism were all expressed at a higher level in the DMSP and acrylate conditions. These genes, and their fold-induction, are highlighted in bold in **Table 4.1**. They included SPO1094 and SPO1101, which are predicted to encode the subunits of propionyl-CoA carboxylase, and SPO1105 which is expected to encode a methylmalonyl-CoA mutase. Additionally, SPO0932 whose predicted product is a methylmalonyl-CoA epimerase, was also up-regulated, although to a lesser extent than the other genes (2.96-fold in DMSP, and 2.45-fold in acrylate). Significantly, the gene SPO1914 was also expressed at a higher level when DMSP and acrylate were present. This gene, predicted to catalyse the reduction of acryloyl-CoA to propionyl-CoA, will be discussed in more detail later. **Figure 4.2** shows the up-regulated genes and the biochemical pathway they are known to be involved in.

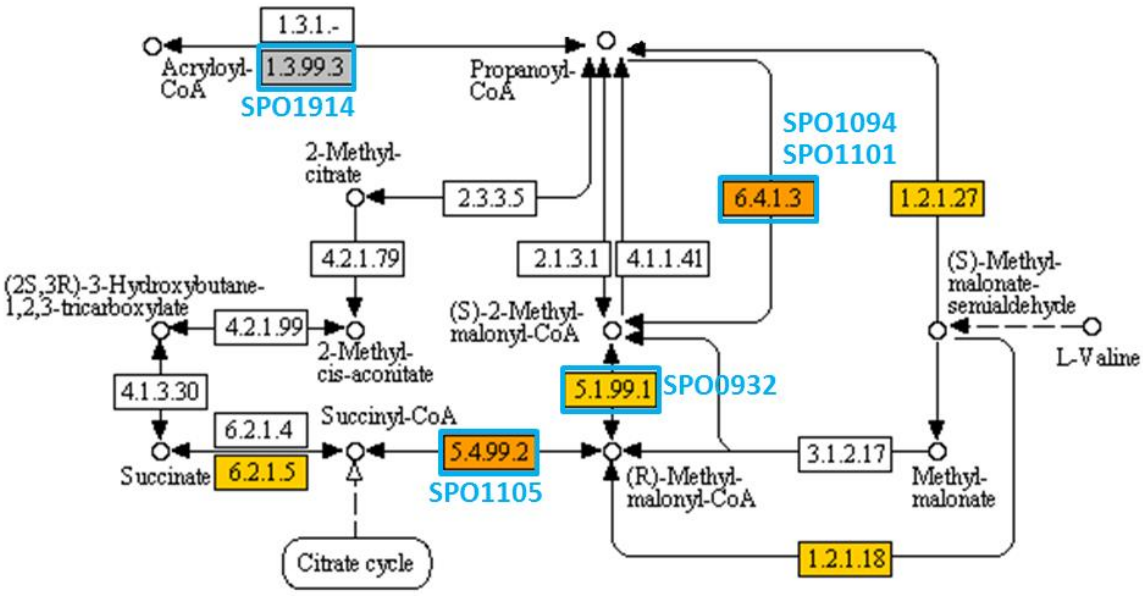


Figure 4.2 KEGG map showing possible route of acryloyl-CoA metabolism in *Ruegeria pomeroyi*. The boxes show the EC enzyme number for the corresponding step. Orange or yellow boxes show that at least one gene encoding that type of enzyme is up-regulated in the presence of acrylate, according to microarray data. The steps of interest are outlined in blue, with the locus tags of the corresponding acrylate-induced genes also in blue.

4.2.3 DMSP and acrylate-induced genes involved in propionate metabolism

The KEGG map predicts that four genes are involved in the conversion, in three steps, of propionyl-CoA to succinyl-CoA (**Figure 4.2)**. The first step converts propionyl-CoA to (S)-2-methylmalonyl-CoA, via an enzyme encoded by genes SPO1094 and SPO1101. Next, (S)-2-methylmalonyl-CoA is converted to the isoform (R)-2-methylmalonyl-CoA by an epimerase likely encoded by SPO0932. Finally, (R)-methylmalonyl-CoA is converted to succinyl-CoA by a mutase, encoded by SPO1105. The microarrays had shown that the expression of each of these genes was significantly enhanced in the presence of both DMSP and acrylate (**Table 4.1)** and the corresponding gene products are described in the following sections.

4.2.3.1 Propionyl-CoA carboxylase

Propionyl-CoA carboxylase (PCC) is a widely distributed, highly conserved, biotin-dependent enzyme. Similar to other carboxylases (e.g. pyruvate carboxylase and acetyl-CoA carboxylase), PCC uses the cofactor biotin to transfer carboxyl groups from the donor (in this case propionyl-CoA) to the recipient (methylmalonyl-CoA) (Attwood and Wallace, 2002). The crystal structure of the PCC complex from humans and bacteria, including (remarkably) *R. pomeroyi*, has been

solved and shown to be highly similar in all cases, being a hexamer of β-subunits, surrounded by 6 α-subunits (Huang *et al*., 2010). The α-subunit (PccA) contains the biotin carboxylase and biotin carboxylase carrier protein domains, and the β-subunit (PccB) has a carboxyltransferase domain. In *R. pomeroyi*, PccA is encoded by SPO1101 and PccB by SPO1094. The PCC holoenzyme in *R. pomeroyi* has been shown to use propionyl-CoA as the preferred substrate over acetyl-CoA (Huang *et al*., 2010).

In the microarray data, both SPO1094 and SPO1101 were induced by DMSP and acrylate when compared to succinate-only conditions. In the presence of DMSP, SPO1094 and SPO1101 were up-regulated approximately 6-fold and 3-fold, respectively. The genes were expressed to a higher level when acrylate was present, with SPO1094 up-regulated 8-fold and SPO1101 up 6-fold. The exact values are presented in **Table 4.1**.

4.2.3.2 Methylmalonyl-CoA epimerase

The resulting product, (S)-2-methylmalonyl-CoA, is then converted to the isoform (R)-2-methylmalonyl-CoA by an epimerase. Studies on this epimerase in bacteria have mostly been carried out in the Actinomycete *Propionibacterium shermanii* (McCarthy *et al*., 2001). The closest homologue (33% identical) in *R. pomeroyi* is encoded by SPO0932, which was up-regulated ~3-fold in the presence of DMSP and ~2.5-fold in acrylate.

4.2.3.3 Methylmalonyl-CoA mutase

The subsequent conversion of (R)-methylmalonyl-CoA to succinyl-CoA requires methylmalonyl-CoA mutase (MCM). Like PCC, this enzyme is also widely distributed in nature, but unlike PCC, the structure of MCM varies in different organisms. The mutase, like the methylmalonyl-CoA epimerase, has been extensively studied in *P. shermanii* (Francalanci *et al*., 1986). In this bacterium, two genes encode an α- and β-subunit of a heterodimeric MCM, but this is not the case for all bacteria. Others, like *Sinorhizobium meliloti* have a homodimeric MCM, only consisting of α-subunits (Miyamoto *et al*., 2003), and in *E. coli* the *sbm* gene encodes a functional single MCM subunit (Haller *et al*., 2000). Indeed, *R. pomeroyi* does not have a gene for the β-subunit, but does have SPO1105 which encodes a peptide with 62% sequence identity to the *P. shermanii* α-subunit. It is likely therefore that the MCM of *R. pomeroyi* is of the αα homodimer type, or functions as a single unit. In the microarray, SPO1105 was induced ~3.5-fold in DMSP and over 6-fold in acrylate.

The induction of genes predicted to be involved in propionyl-CoA metabolism in the presence of DMSP and acrylate was particularly interesting because of a recently described connection between acrylate and propionate metabolism in *Rhodobacter sphaeroides* 2.4.1. In this α-

proteobacterium, the *acuI* gene was shown to encode an acryloyl-CoA reductase, which reduces acryloyl-CoA to propionyl-CoA (Schneider *et al*., 2012; see **Chapter 1**). In *R. pomeroyi*, and indeed in most Roseobacters, *acuI* (SPO1914) is co-transcribed with the DMSP demethylase gene *dmdA* (see **Chapter 5** for more detail), and thus there is also a genetic link between acrylate and DMSP catabolism.

Since it was shown unambiguously that *R. pomeroyi* can use carbon from acrylate, and since some of the genes involved in propionate catabolism were induced by DMSP and acrylate, this suggested that acrylate might be fully catabolised via a propionate-linked pathway. To test this, individual mutant strains of *R. pomeroyi* with insertions in SPO1094, SPO1101 and SPO1105 were constructed.

4.2.4 Construction of mutant strains of *R. pomeroyi*

In order to examine the roles (if any) of the genes that were predicted to be involved in propionate catabolism, each of these was mutated and the phenotypes determined. To do this, strains of *R. pomeroyi* with insertional mutations in SPO1094, SPO1101, and SPO1105 were constructed using the suicide plasmid pBIO1879 (see **Figures 4.3**, **4.4** and **Chapter 7**). This plasmid, which is derived from pK19*mob* (Schäfer *et al*., 1994), contains an extra antibiotic resistance cassette (to spectinomycin), as well as kanamycin. Importantly, pBIO1879 can be mobilised into a wide range of host bacteria by conjugation, but only replicates in enteric bacteria, such as *E. coli* due to the PMB1 replicon. Therefore, the only way for antibiotic resistance to be maintained in a non-enteric host is via the integration by a single crossover of the entire plasmid into the host genome, provided that a region of the host's genome has previously been cloned into pBIO1879. Therefore to make a targeted insertional mutation, an internal fragment of a gene of interest is cloned into pBIO1879 and mobilised into the desired host strain, in this case *R. pomeroyi*. Transconjugants that are resistant to both spectinomycin and kanamycin should be due to the insertion of the recombinant plasmid into the gene of interest.

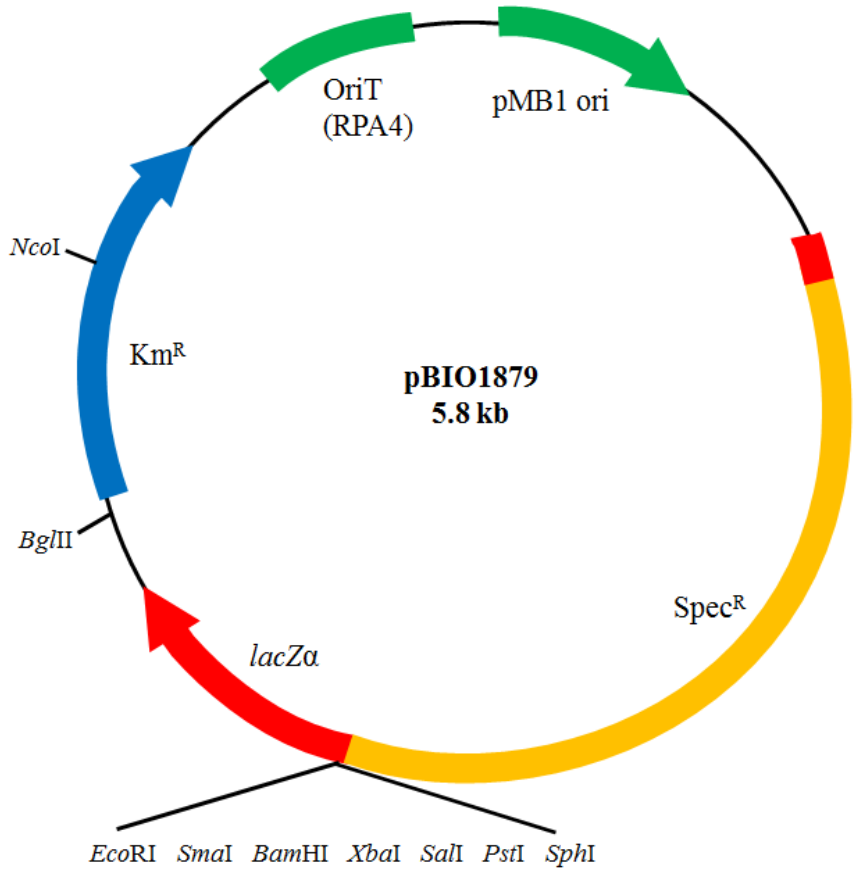


Figure 4.3 Schematic representation of pBIO1879. Map of pBIO1879 with kanamycin resistance cassette (Km^R), spectinomycin resistance cassette ($Spec^R$), the PMB1 replicon, and origin of transfer (*oriT*). The multicloning site (MCS) replaces 6-7 codons of the *lacZα* gene.

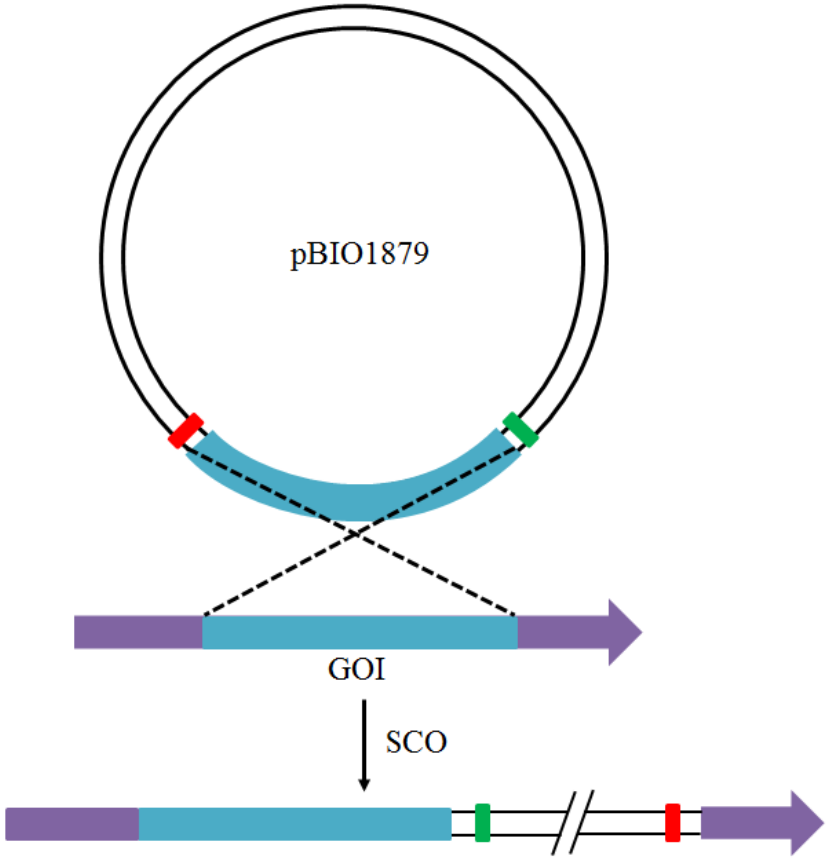


Figure 4.4 pBIO1879 mediated insertion mutagenesis. An internal fragment of the gene of interest (GOI) is cloned into the suicide vector pBIO1879 (indicated by the blue bar). The purple arrow represents the gene of interest and the position of the internal cloned region. Through a single crossover event (SCO) via homologous recombination, pBIO1879 inserts into the gene of interest at the site of the internal cloned fragment. The red and green bars show the orientation of inserted plasmid DNA. Any part of the cloned region may be the starting point for recombination.

4.2.4.1 Amplification of internal fragments of *R. pomeroyi* genes

To make the mutations, fragments internal to SPO1094, SPO1101, and SPO1105 were amplified from *R. pomeroyi* genomic DNA and each was cloned into pBIO1879. The positions of the primers and the sizes of the resultant fragments are shown in **Figure 4.5**, and primer sequences can be found in **Table 7.5**. Primers were designed so that the fragment was at least 175 bp away from each of the 5' and 3' terminals to ensure the gene was fully disrupted. In all cases, forward and reverse primers were designed to contain an *Eco*RI and *Xba*I site, respectively, to allow subsequent cloning into pBIO1879.

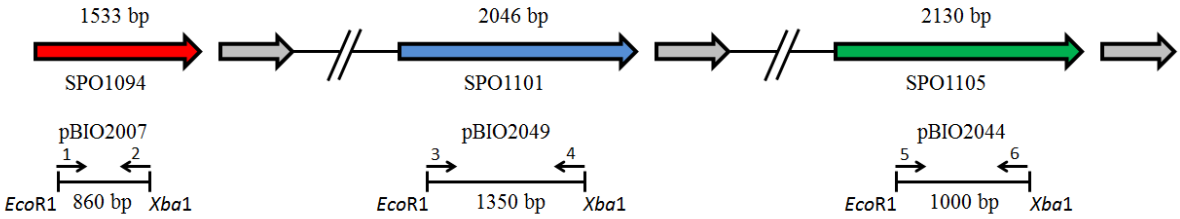


Figure 4.5 Internal fragments of SPO1094, SPO1101 and SPO1105 amplified and cloned into pBIO1879. Coloured arrows show SPO1094, SPO1101 and SPO1105 genes. The grey arrows represent the gene immediately downstream in each case. Broken black lines indicate where part of the genome has been omitted from the diagram. The sizes of each gene of interest are indicated above the arrows, in base pairs. Primers 1 and 2 (SPO1094PK19FOR and SPO1094PK19REV, respectively) were used to amplify an 860 bp fragment of SPO1094. Similarly, primers 3 and 4 (SPO1101PK19FOR and SPO1101PK19REV) were used to amplify a 1350 bp fragment of SPO1101, and primers 5 and 6 (SPO1105PK19FOR and SPO1105PK19REV) to amplify a 1000 bp fragment of SPO1105. *Eco*R1 and *Xba*I sites were incorporated into each forward and reverse primer, respectively, and used to clone the fragments individually into pBIO1879 (a pK19mob derivative with $spec^R$) to create pBIO2007 (SPO1094), pBIO2049 (SPO1101) and pBIO2044 (SPO1105).

4.2.4.2 Cloning the amplified gene fragments into pBIO1879

The PCR products of the internal gene fragments were purified, digested for 2 hours with *Eco*RI and *Xba*I and then separated by gel electrophoresis. They were then extracted from the gel and purified, and ligated to pBIO1879, which had also been digested with *Eco*RI and *Xba*I. The ligations were incubated at 4°C overnight, and then used to transform competent *E. coli* 803 cells, prior to plating onto LB agar containing kanamycin and spectinomycin, and overnight incubation at 37°C. Broth cultures were made from a selection of six transformant colonies and these were used for plasmid preparations. The extracted plasmids were analysed by restriction analysis and those with an insert of the correct size were verified by DNA sequencing. The plasmids were designated pBIO2007 (SPO1094), pBIO2049 (SPO1101), and pBIO2044 (SPO1105).

4.2.4.3 Mobilisation of pBIO2007, pBIO2049, and pBIO2044 into *R. pomeroyi*

The individual recombinant plasmids were each mobilised from *E. coli* 803 into a rifampicin-resistant strain of *R. pomeroyi* (J470) via tri-parental mating. The general method for tri-parental mating is described in **Chapters 2** and **7**, but in this case the procedure was slightly different. Since the pK19 plasmid must not only be transferred to the recipient, but also integrate into the recipient's genome, the conjugation efficiency should be as high as possible to increase the chances of this rare event occurring. Therefore, liquid cultures of overnight cultures of the three parent strains (the donor *E. coli* 803 containing the pBIO1879 derivatives, the helper pRK2013 strain and a Rif^R strain of *R. pomeroyi*) were mixed in a 2:1:1 ratio of recipient:helper:donor (1 ml: 0.5 ml:0.5 ml) then spread onto a membrane filter on 1/2YTSS agar plates. Following incubation for two nights at 28°C, the cells were washed off the filters in minimal media and then plated on 1/2YTSS agar plates containing rifampicin, kanamycin plus spectinomycin to select for *R. pomeroyi* containing the integrated pK19 plasmid.

For all three mutagenesis targets, single colonies arose on these selective media after 3 days incubation at 28°C. One colony per target gene was picked onto a fresh agar plate, and a culture of each strain was obtained. The genomic DNA was extracted from the resultant cultures and used to verify if pK19 had inserted into the desired gene.

4.2.5 Verification of *R. pomeroyi* mutants by Southern blotting

Potential mutant strains of *R. pomeroyi* were verified using Southern blotting. This technique is used for identifying specific DNA fragments, using a detectable probe which consists of a complementary DNA sequence labelled either radioactively, with a fluorescent dye, or with an enzyme which generates a visual signal when incubated with the correct substrate. To verify the presence of a genomic insertion in the desired region of DNA, wild type and mutant genomic DNA can be digested with specific restriction enzymes. A genomic insertion in the mutant strain will alter the pattern of fragments generated by restriction, by changing the distance between restriction sites, or introducing new target sequences for the chosen restriction enzymes. Then, a specific probe can be used to visualise and compare regions of interest between the wild type and potentially mutant strains. The following sections describe the probe design and restriction enzyme selection in this work.

4.2.5.1 Probe Design

To make the probes, primers were designed to amplify each intact gene of interest from wild type *R. pomeroyi* genomic DNA. The primers and the expected PCR product sizes are shown in **Table 4.2**. Each PCR product was separated using gel electrophoresis, and a band of the

expected size was excised and purified. In this case, the probes were then labelled with digoxigenin (DIG), a small hapten molecule that binds anti-DIG antibodies. Therefore, probe-target hybrids can be detected using anti-DIG antibodies in an enzyme-linked immunoassay. For this, the anti-DIG antibodies are conjugated with an alkaline phosphatase, which acts on two colourless substrates - 5-bromo-4-chloro-3'-indolyphosphate (BCIP) and nitro-blue tetrazolium (NBT) in a redox reaction, the result of which is the formation of a dark blue precipitate. Therefore, application of the NBT/BCIP substrate to a blot allows colorimetric detection of the positions of any anti-DIG antibodies bound to DIG-labelled probes, hybridised with target DNA fragments.

Table 4.2 Primers used to create Southern blot probes.

Gene	Forward Primer	Reverse Primer	PCR product size (bp)
SPO1094	SPO1094pS4FOR2	SPO1094pS4REV2	2117
SPO1101	SPO1101pS4FOR	SPO1101pS4REV	2585
SPO1105	SPO1105pRKFOR2	SPO1105pRKREv2	2422

4.2.5.2 Preparation and digestion of wild-type and mutant genomic DNA

Genomic DNA from wild-type *R. pomeroyi* and putative SPO1094^{-}, SPO1101^{-} and SPO1105^{-} mutant strains was prepared using isopropanol precipitation (full protocol given in **Chapter 7**). Approximately 1 µg of each preparation was digested with a restriction enzyme chosen specifically to provide the most diagnostic results (see **Figure 4.6** and **Table 4.3**). Thus, the probe designed for SPO1094 would label a ~7.6 kb *Nde*I fragment in wild type *R. pomeroyi*. Since there are no *NdeI* sites in pBIO2007, the successful insertion of this plasmid DNA into the genome would extend the fragment by ~7.9 kb (the size of pBIO2007), to ~15.5 kb. Similarly, the SPO1101 probe should bind a ~5.2 kb fragment in the wildtype genome, and the successful insertion of pBIO2049 would extend this to ~13.6 kb. The probe for SPO1105 would bind a ~4.6 kb wildtype *Nco*I fragment, but since pBIO2044 also contains one *Nco*I site, a successful insertion would split this region into two *Nco*I fragments, the combined size of which would be ~12.8 kb.

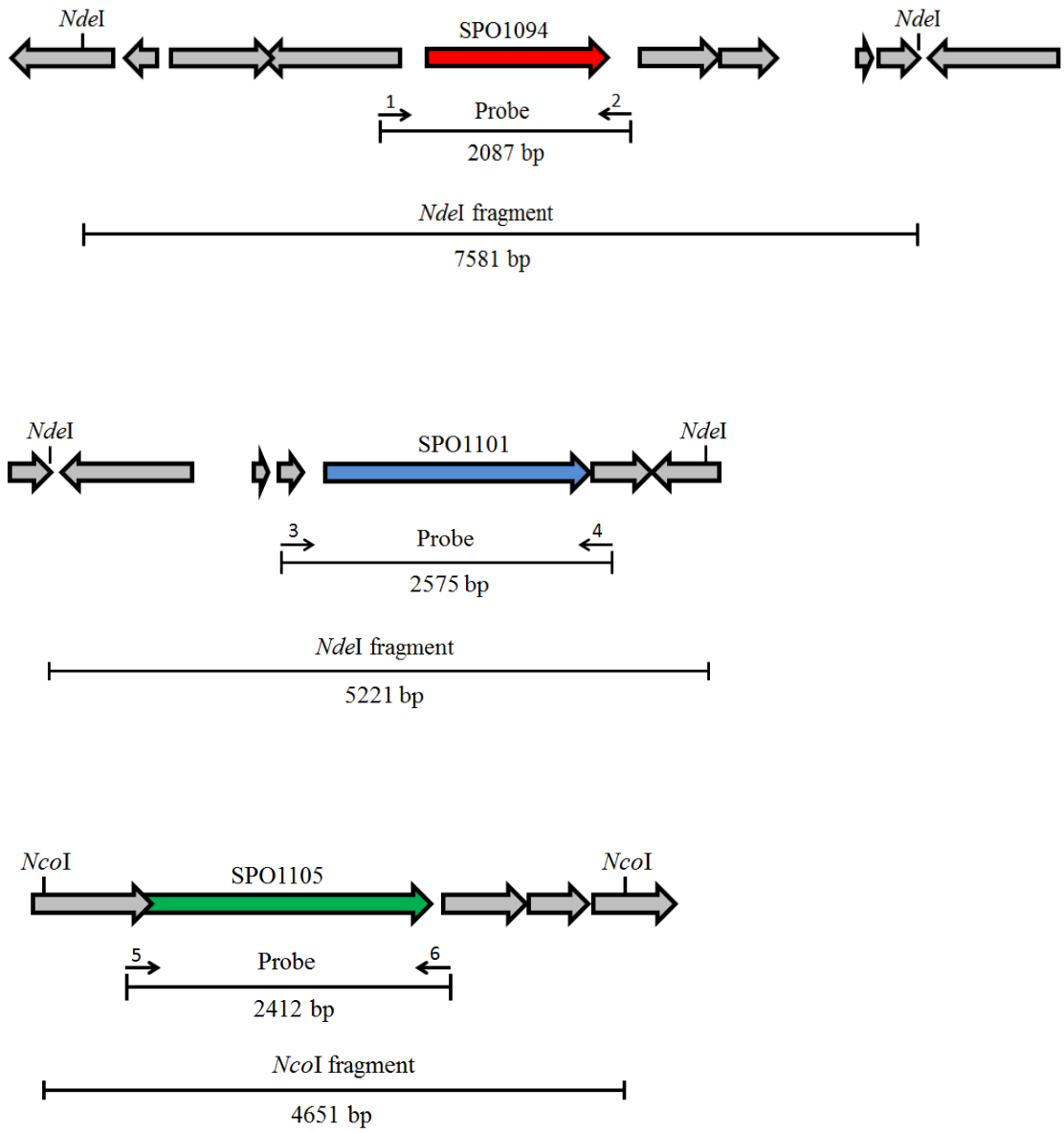


Figure 4.6 Probes used for Southern blotting of SPO1094⁻, SPO1101⁻ and SPO1105⁻ mutants. Probes for SPO1094, SPO1101 and SPO1105 were amplified using primers 1 and 2 (SPO1094pS4FOR2 and SPO1094pS4REV2), primers 3 and 4 (SPO1101pS4FOR and SPO1101pS4REV) and primers 5 and 6 (SPO1105pRKFOR2 and SPO1105pRKREV2), respectively. The sizes of each resulting probe are indicated in the diagram. The probes were designed to encompass the entire gene in each case. The probes for SPO1094 and SPO1101 bind to *Nde*I fragments of 7481 and 5221 bp in the wildtype genome, respectively. The probe for SPO1105 binds a 4651 bp *Nco*I fragment in the wildtype genome.

Table 4.3 Predicted restriction fragment sizes for wildtype *R. pomeroyi* and SPO1094⁻, SPO1101⁻, and SPO1105⁻ mutant strains

Gene	Restriction Enzyme	Predicted size of wildtype fragment (bp)	Size of plasmid insertion (bp)	Restriction sites on insertion?	Predicted size of mutant fragments (bp)
SPO1094	*Nde*I	7,581	7,887	No	15,468
SPO1101	*Nde*I	5,221	8,375	No	13,596
SPO1105	*Nco*I	4,651	8,212	Yes (1)	2 fragments totalling 12,863

Genomic DNA from wildtype *R. pomeroyi* and each of three mutant strains (SPO1094⁻, SPO1101⁻, and SPO1105⁻) was digested with either *Nde*I or *Nco*I (as indicated in column 2). The predicted size of the *Nde*I or *Nco*I fragment in which the gene of interest sits is given in column 3. The size of the plasmid insertion for each of the three mutant strains is given in column 4. There are no *Nde*I sites on either pBIO2007 or pBIO2049, so the plasmid insertion will simply extend the fragment size (column 6). In the case of the SPO1105⁻ mutant, the pBIO2044 insertion contains one *Nco*I site, thus the resulting region of probe-complementarity will be split into two fragments. The sizes of the two individual fragments cannot be predicted, since the plasmid may insert at any point along the region of homology. However, the combined size of the two fragments can be predicted (shown in column 6).

4.2.5.3 Southern blot procedure

The full details of the Southern blot protocol are given in **Chapter 7**. Briefly, the digested genomic DNA of *R. pomeroyi* wild type and the putative mutant strains were separated by gel electrophoresis, then blotted and probed with the appropriate DIG-labelled PCR product (**Chapter 7**). The resulting blots verified the successful insertion of plasmid DNA into the target gene in each case (**Figure 4.7**). However, for SPO1105⁻, an extra, unexplained band appeared at ~3.8 kb. Therefore, this mutant strain was double checked for a successful insertion using PCR. For this, genomic DNA from wildtype *R. pomeroyi* and the putative SPO1105⁻ mutant strain was used as a template in a PCR with SPO1105pRKREV2 (see **Table 7.5**) and universal M13F as primers. As expected, the wildtype DNA did not yield a PCR product, since pK19, and therefore the M13 site, is not present. In contrast, the SPO1105⁻ mutant strain DNA gave the expected product of 2130 bp, from the M13F primer site in the pK19 insertion, and the primer situated

just outside the SPO1105 gene. Thus the mutant strains were designated J559 (SPO1094$^-$), J561 (SPO1105$^-$) and J560 (SPO1101$^-$).

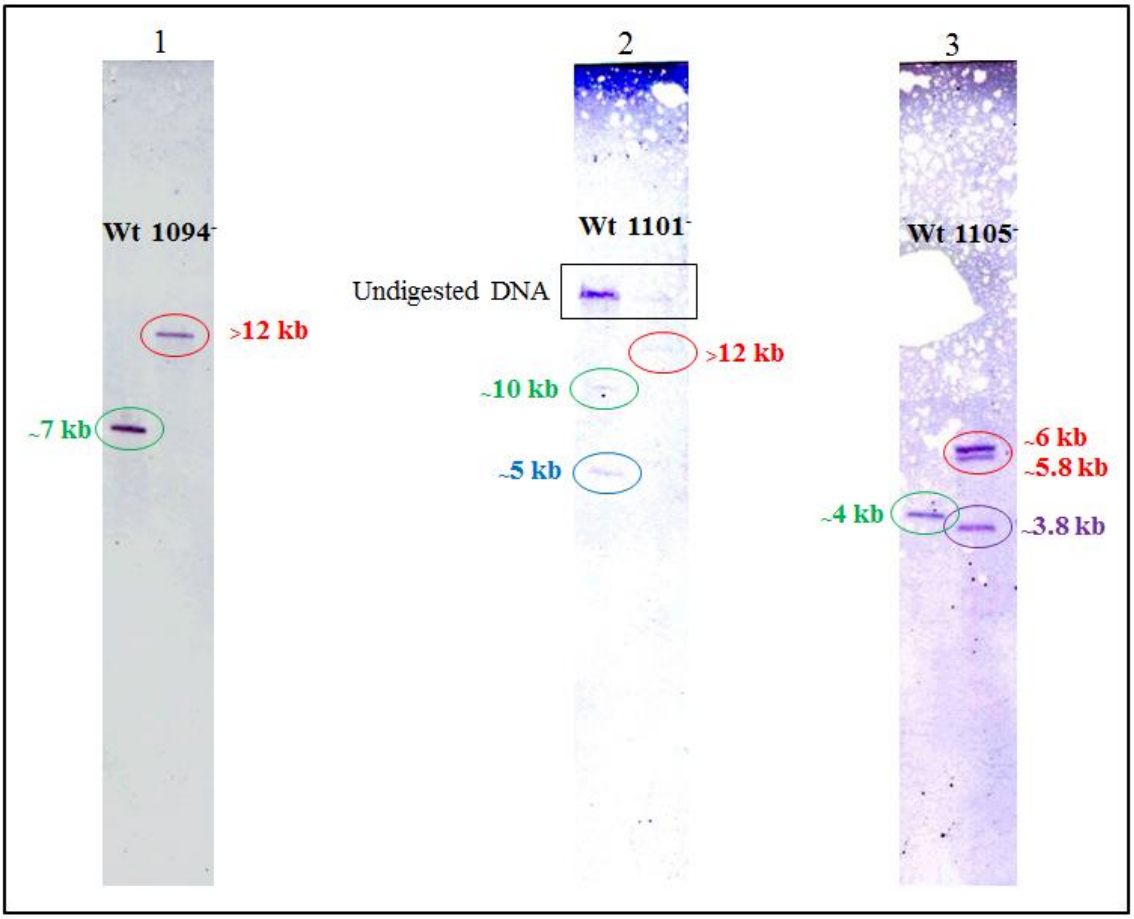


Figure 4.7 Verification of insertional mutations in SPO1094$^-$, SPO1101$^-$ and SPO1105$^-$ mutants by Southern Blots. Genomic DNA from wild type (Wt), 1094$^-$, 1101$^-$, and 1105$^-$ strains of *R. pomeroyi* was digested with *Nde*I (blots 1 and 2) or *Nco*I (blot 3) and probed with PCR product of intact SPO1094, SPO1101 or SPO1105, respectively. Approximate DNA lengths of the labelled bands are shown. In each case, the pattern was consistent with a plasmid insertion into the gene of interest in the mutant strain. Thus a successful insertion in SPO1094 was expected to expand the *Nde*I fragment size from ~7.5 kb to ~15 kb, consistent with blot 1 which shows a Wt band of ~7 kb and a larger band of >12 kb in the mutant strain (the DNA ladder used had an upper range of 12 kb, thus it is difficult to predict the sizes of fragments larger than this). Similarly, the insertion in SPO1101 was predicted to increase the *Nde*I fragment size from ~5.2 kb to ~13 kb, and blot 2 shows a Wt fragment of ~5 kb and a mutant fragment of >12 kb. There was also a band of ~10 kb present in the Wt lane, which may be explained by incomplete digestion of the genomic DNA. Consistent with this, there is also a

significant amount of undigested DNA in this lane which did not migrate from the well. Blot 3 shows the Wt DNA produced an *Nco*I fragment of ~4 kb, close to the predicted fragment size of ~4.6 kb. The SPO1105$^-$ mutant DNA produced three bands of ~6 kb, ~5.8 kb and ~3.8 kb. Only two fragments were expected in this case, totalling ~12.8 kb, which may correspond to the two larger bands in the blot. The ~3.8 kb band was not expected but colony PCR verified the strain as a SPO1105$^-$ mutant.

4.2.6 Phenotype of mutant strains

In their different ways, SPO1094, SPO1101 and SPO1105 were predicted to be involved in the growth of *R. pomeroyi* on DMSP, via acrylate. In light of this, the mutant strains were first tested for their growth in minimal MBM medium with either DMSP or acrylate as a sole carbon source. Media with succinate was used as positive control.

The wild type and all three mutant strains grew well on succinate, all reaching an OD_{600} of around 0.7-0.8 after 40 hours (**Figure 4.8**), which is significantly different to the "no carbon" control (ANOVA, $F_{4,10}$, <0.001). The wild type strain was also able to grow on DMSP, acrylate and propionate, as expected. Strikingly, however, the growth of mutant strains J559, J561 and J560 with DMSP, acrylate or propionate was not significantly different to the "no carbon" control (Tukey HSD, $P < 0.05$).

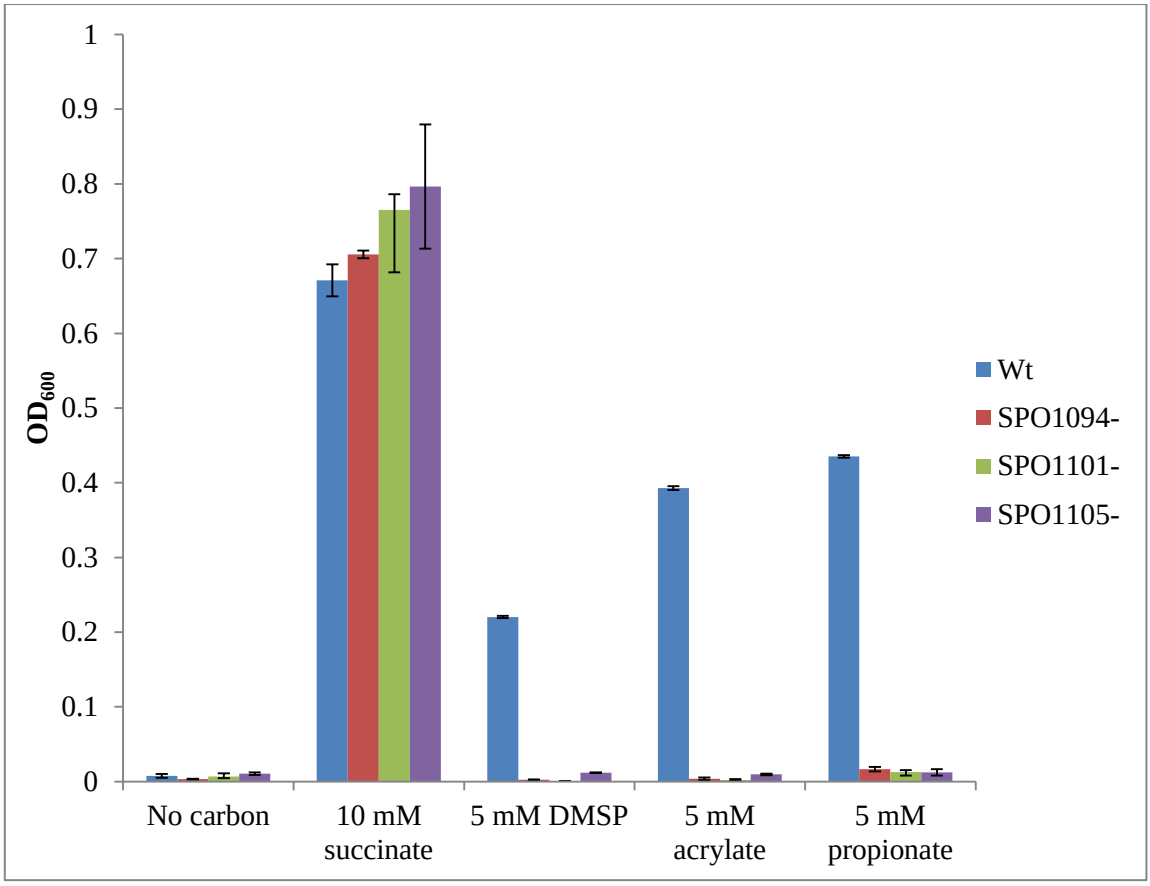


Figure 4.8 Growth of *R. pomeroyi* wild type and mutant strains on sole carbon sources. Growth tests were carried out in 5 ml MBM containing either 10 mM succinate or 5 mM DMSP, acrylate or propionate as the sole carbon source. The OD_{600} was recorded after 40 hours incubation at 28°C. The error bars represent the standard error between triplicate tests.

4.2.6.1 Mutant growth on succinate

The growth tests in **Figure 4.8** show the mutant strains reaching the same optical density as wild type *R. pomeroyi* after 40 hours when succinate was provided as a sole carbon source. Growth curves were also carried out to show that the mutants could grow at the same rate as the wild type on succinate. For this, overnight starter cultures of mutant and wild type strains were washed and inoculated to 100 ml MBM containing 10 mM succinate. Optical density readings were taken at regular intervals until the cultures had reached stationary phase. **Figure 4.9** shows that the mutant strains were comparable to the wild type strain for growth on succinate.

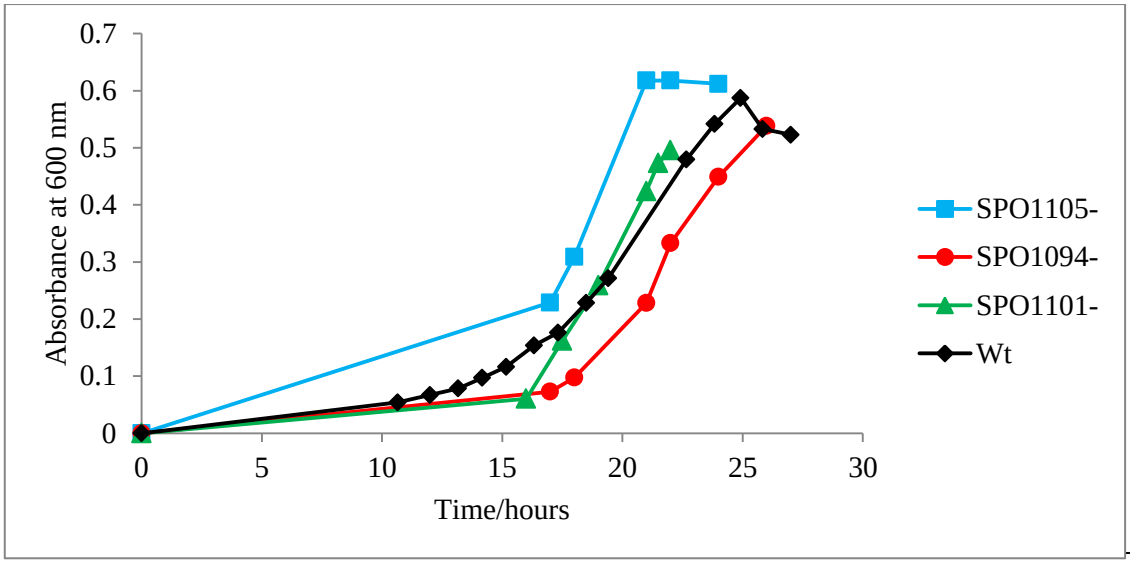


Figure 4.9 Growth curves of wild type and mutant strains of *R. pomeroyi* with succinate. Growth curves were carried out in 100 ml MBM with 10 mM succinate as a carbon source. Optical densities were recorded at regular intervals until cultures had reached stationary phase.

4.2.7 The role of AcuI in acrylate catabolism in *R. pomeroyi*

At this point it is appropriate to discuss the role of AcuI in DMSP catabolism. As mentioned earlier, the gene encoding this enzyme is in several cases co-transcribed with genes involved in DMSP breakdown. For example, in almost all *dmdA*-containing Roseobacters, *acuI* is co-transcribed with the DMSP demethylase gene *dmdA* (elaborated on in **Chapter 5**). In *Alcaligenes faecalis*, *acuI* is co-transcribed with *dddY* and in *Rhodobacter sphaeroides* strain 2.4.1 it is found in an operon with a regulatory gene (*acuR*) and *dddL*. Work carried out by Matthew Sullivan in *R. sphaeroides* showed that AcuI had a role in acrylate catabolism, since an AcuI$^-$ mutant is hypersensitive to the presence of acrylate in the medium, being inhibited for growth by 1 mM acrylate (the lowest concentration tested), some 10-fold less than that which is tolerated by wild type *R. sphaeroides* (M. Sullivan, personal communication). Furthermore, the AcuI$^-$ mutant was less effective than the wild type in generating $^{14}CO_2$, when fed with [1-^{14}C]-acrylate. Unfortunately, under the conditions used, no labelled intermediates which could be involved in acrylate catabolism were seen and its exact enzymatic function was not established. However, Schneider *et al.*, (2012) rectified this, through work exploring how *R. sphaeroides* is able to assimilate carbon from 3-hydroxypropionate (3HP). They showed that one route of 3HP metabolism was via its reductive conversion to propionyl-CoA via acryloyl-CoA. They noted that the *acuI* gene (RSP_1434) encodes a member of the medium chain dehydrogenase/reductase superfamily that also contains an acryloyl-CoA reductase in *Sulfolobus tokodaii* (Teufel *et al.*,

2009), and the acryloyl-CoA reductase domain of a propionyl-CoA synthase in *Chloroflexus aurantiacus* (Alber and Fuchs, 2002). This, coupled with the genomic association of *acuI* with genes involved in DMSP metabolism lead the authors to propose that AcuI was involved in the reductive conversion of 3HP to propionyl-CoA. Indeed, they showed that an *acuI*$^-$ mutant strain of *R. sphaeroides* was unable to grow on 3HP, and, furthermore, showed that 3HP and acrylate reduction activity was undetectable in the mutant strain. Thus, they suggested that AcuI of *R. sphaeroides* catalyses the reduction of acryloyl-CoA to propionyl-CoA (Schneider *et al*., 2012).

In *R. pomeroyi*, AcuI is encoded by the gene SPO1914, and as in most Roseobacters (see **Chapter 5**), it is co-transcribed with *dmdA*. Following the work by Schneider *et al*., it was predicted that this enzyme would also be involved in acrylate metabolism and so AcuI was investigated alongside SPO1094, SPO1101 and SPO1105. For this, I used the published AcuI$^-$ mutant strain J527 (Todd *et al*., 2012b), which, like the other mutants, has a pK19spec insertional mutation in the *acuI* gene.

4.2.7.1 AcuI$^-$ does not grow on DMSP or acrylate

The AcuI$^-$ mutant strain (J527) was tested for growth on DMSP and acrylate under the same conditions described above for the other mutants. After 40 hours incubation, the AcuI$^-$ mutant had reached a similar OD_{600} to the wild type when succinate and propionate were used as carbon sources, but could not grow using only DMSP or acrylate as carbon sources **Figure 4.10**). This was not surprising, since it has been shown that AcuI from *R. pomeroyi* is important in the detoxification of acrylate, or more accurately, the acryloyl-CoA derived from acrylate (Todd *et al*., 2012b). Several different bacterial strains have been shown to be hypersensitive to the presence of acrylate when a mutation is made in AcuI, including *E. coli* which has an AcuI homologue termed YhdH (Todd *et al*., 2012b). Therefore the absence of growth seen in the AcuI$^-$ *R. pomeroyi* strain may have been due to the accumulation of toxic acryloyl-CoA derived from the catabolism of acrylate or DMSP, as well as (or instead of) a failure to catabolise acrylate *per se.*

In light of these observations, it was of interest to test whether the poor growth of the SPO1094$^-$, SPO1101$^-$ and SPO1105$^-$ mutant strains on acrylate and DMSP was also due to a build-up of toxic intermediates.

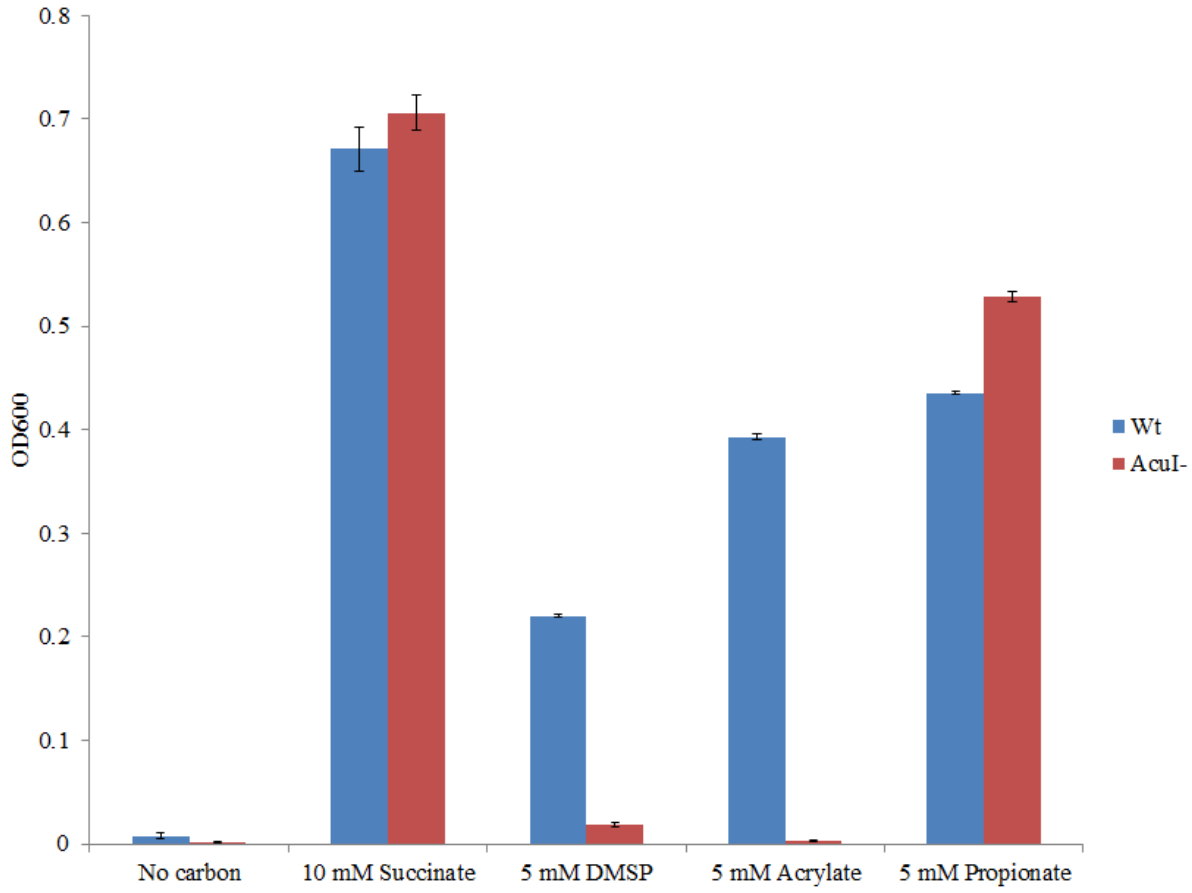


Figure 4.10 Growth of *R. pomeroyi* wild type and AcuI⁻ mutant on sole carbon sources. Wild type and AcuI mutant of *R. pomeroyi* were grown in 5 ml MBM containing either 10 mM succinate or 5 mM DMSP, acrylate or propionate as sole carbon source. The OD_{600} was recorded after 40 hours incubation at 28°C. Error bars represent standard error in triplicate tests.

4.2.8 Toxicity tests

To compare the wild type and SPO1094⁻, SPO1101⁻, SPO1105⁻ and AcuI⁻ mutant strains of *R. pomeroyi* for their sensitivity to acrylate, DMSP or propionate, 10 µl aliquots of overnight cultures were spotted onto MBM agar plates, all of which contained both succinate (10 mM) as a carbon source, plus eight different concentrations of DMSP, acrylate or propionate (ranging from 0 to 10 mM).

4.2.8.1 Hypersensitivity of mutant strains

The results of the spot tests are shown in **Figures 4.11 to 4.12**. The acrylate sensitivity tests were particularly striking. In this case the wild type strain failed to grow in the presence of 5 mM acrylate, whereas all of the mutants began to show signs of poor growth at 100 µM, and none of

the mutant strains could grow when 500 μM acrylate was present (**Figure 4.11**). These results were similar to the propionate sensitivity tests in that the 1094^{-}, 1101^{-} and 1105^{-} strains were also unable to grow at 100-500 μM propionate (**Figure 4.2**). However, an important difference in the propionate tests was that the AcuI^{-} mutant grew well in the presence of propionate. This result is logical, since neither AcuI, nor the acryloyl-CoA intermediate, are predicted to be involved in propionate catabolism.

In the DMSP sensitivity tests (**Figure 4.13**), the wild type grew well on all concentrations. However, the AcuI^{-} mutant showed signs of poorer growth at 2 mM, and failed to grow at 5 mM DMSP. The 1094^{-} and 1101^{-} strains showed slightly more resistance, with normal growth at 5 mM, but poor growth at 10 mM. Similarly, 1105^{-} grew well at 5 mM but failed to grow at all when 10 mM DMSP was present. These spot tests show the mutant strains have a slightly enhanced sensitivity to DMSP, which may explain why they could no longer use it as sole carbon source.

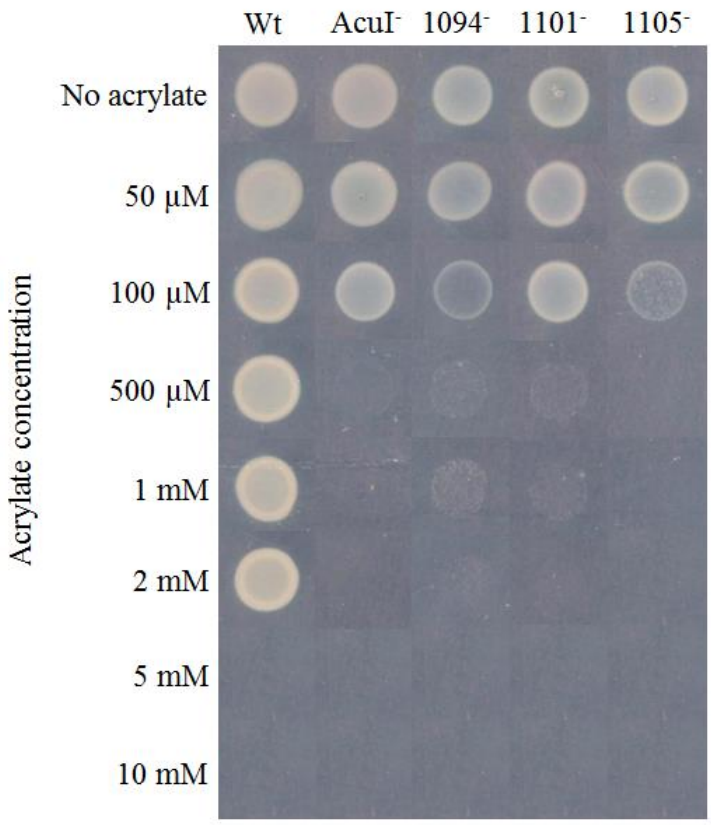


Figure 4.11 Effect of acrylate on growth of wild type and mutant strains of *R. pomeroyi*. Aliquots of 10 µl cultures (OD_{600} = 1.0) of wild type *R. pomeroyi* or AcuI$^-$, SPO1094$^-$, SPO1101$^-$ or SPO1105$^-$ mutant strains were spotted onto solid MBM medium with 10 mM succinate plus increasing concentrations of acrylate and were photographed after 3 days' incubation at 28°C.

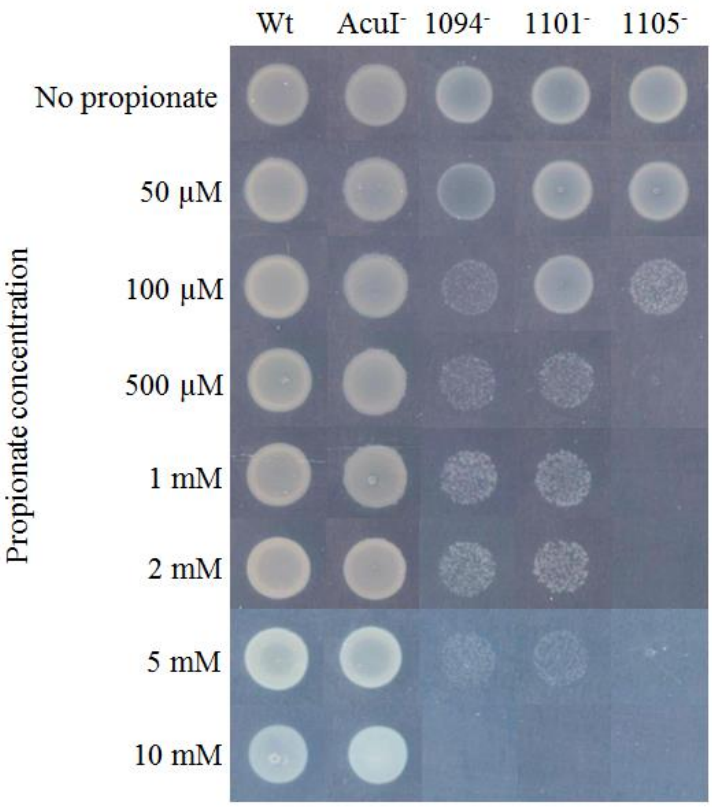


Figure 4.12 Effect of propionate on growth of wild type and mutant strains of *R. pomeroyi*. Aliquots of 10 µl cultures (OD_{600} = 1.0) of wild type *R. pomeroyi* or AcuI$^-$, SPO1094$^-$, SPO1101$^-$ or SPO1105$^-$ mutant strains were spotted onto solid MBM medium with 10 mM succinate plus increasing concentrations of propionate and were photographed after 3 days' incubation.

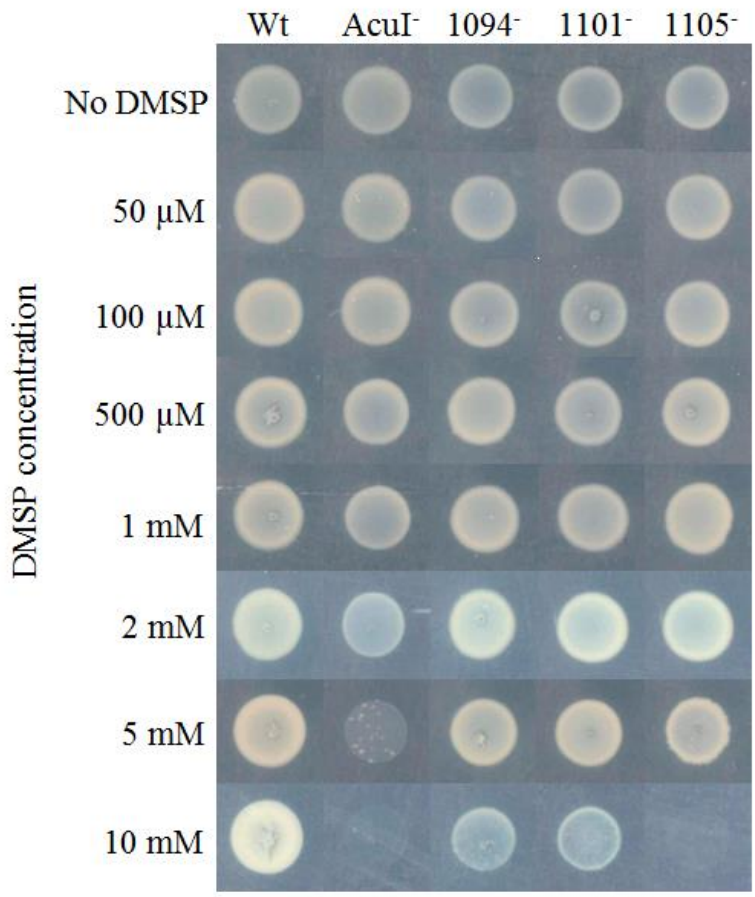


Figure 4.13 Effect of DMSP on growth of wild type and mutant strains of *R. pomeroyi*. Aliquots of 10 µl cultures (OD_{600} = 1.0) of wild type *R. pomeroyi* or AcuI$^-$, SPO1094$^-$, SPO1101$^-$ or SPO1105$^-$ mutant strains were spotted onto solid MBM medium with 10 mM succinate plus increasing concentrations of DMSP and were photographed after 3 day's incubation.

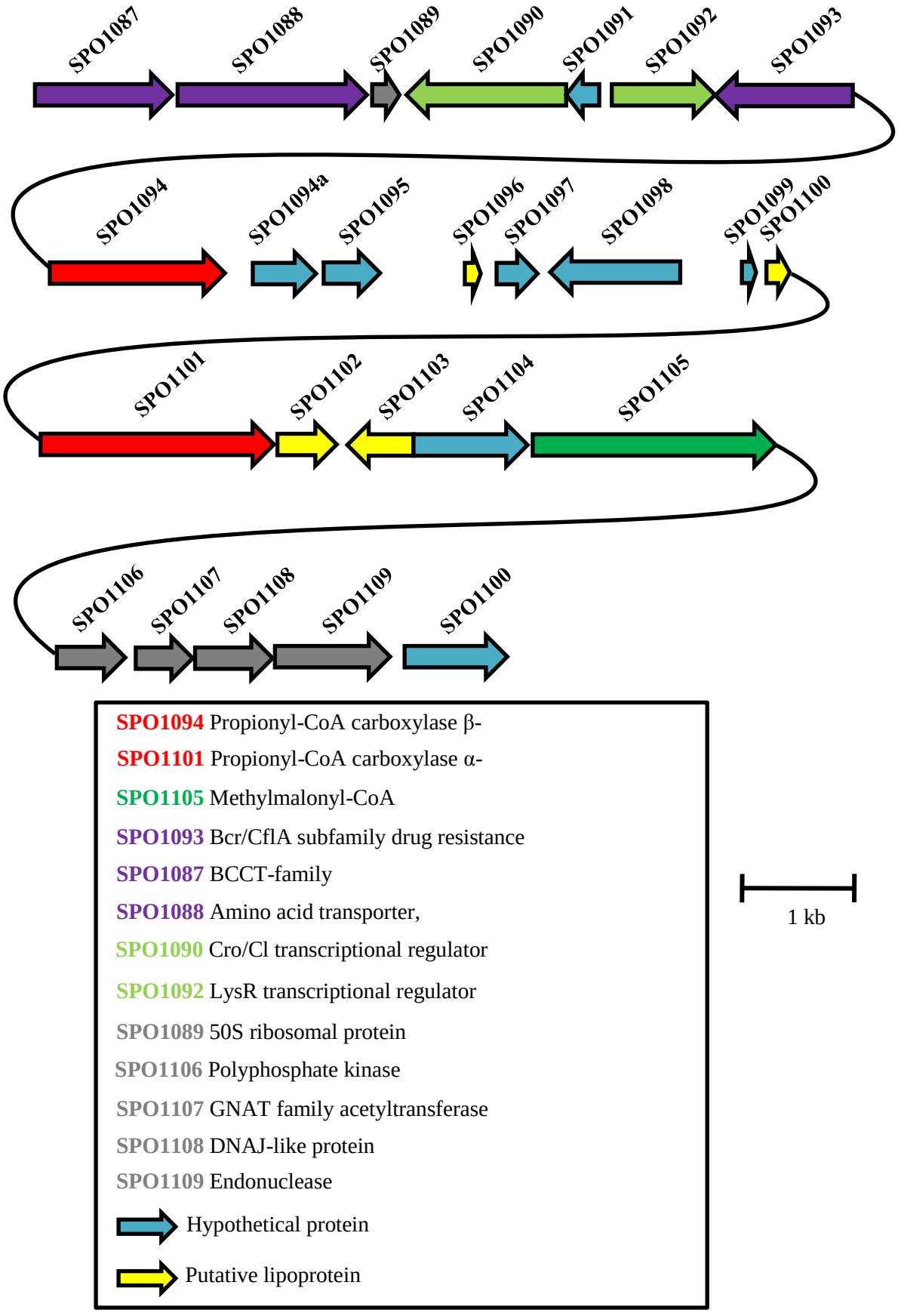
4.2.9 Arrangement of the propionate catabolism genes

In the *R. pomeroyi* genome, the SPO1094, SPO1101 and SPO1105 genes are closely linked, each separated by a few genes encoding hypothetical proteins or predicted lipoproteins (**Figure 4.14**). It was therefore feasible that this region may be contained within a single cosmid in a pre-existing genomic library of *R. pomeroyi* (made by Andrew Curson) since the average size of the DNA cloned in the vector pLAFR3 was ~25 kb.

In an attempt to isolate such a cosmid, 400 individual colonies of *E. coli* containing *R. pomeroyi* library cosmids were picked to membrane filters and a colony blot (see **Chapter 7**) was carried out, using P^{32}-labelled DNA that corresponded to the SPO1098 gene. Of 400 colonies, 20 were found to give a strong hybridising signal. The cosmid DNA was isolated from the corresponding

E. coli colonies and the termini of the cloned DNA were sequenced, using the universal M13forward and reverse primers. One cosmid had termini corresponding to SPO1087 and SPO1100, which are ~24 kb apart on the *R.* pomeroyi chromosome. Therefore, this cosmid was assumed to contain a contiguous region of cloned DNA, including the genes of interest. This was termed pBIO2037.

Figure 4.14 Map of genes involved in propionate catabolism in *R. pomeroyi*



4.2.10 Complementing mutant strains with pBIO2037

The pBIO2037 cosmid was mobilised into the SPO1094$^-$, SPO1101$^-$ and SPO1105$^-$ mutant strains by tri-parental mating, selecting for transconjugants on media containing the following antibiotics: rifampicin (to select for *R. pomeroyi* J470); spectinomycin and kanamycin (to retain the genomic insertion causing the mutation); and tetracycline (to select for the presence of tet^R pBIO2037). The complemented strains were then tested for their sensitivity to DMSP, acrylate, and propionate alongside the wild type and mutant strains. In all cases, the transconjugants restored a wild type phenotype with regard to their tolerance to the compounds tested above (**Figure 4.15**). Thus, the various sensitivity phenotypes of the mutants were due to the mutations in the individual genes in this region.

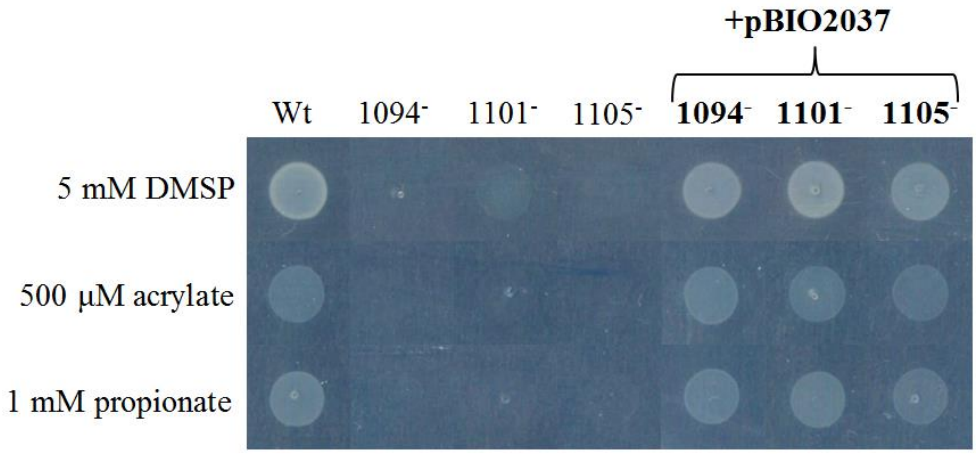


Figure 4.15 Spot tests showing complementation of mutant phenotypes by pBIO2037. Aliquots of 10 µl cultures (OD_{600} = 1.0) of wild type *R. pomeroyi* or AcuI$^-$, SPO1094$^-$, SPO1101$^-$ or SPO1105$^-$ mutant and the mutants containing cosmid pBIO2037 were spotted onto solid MBM medium with 10 mM succinate plus DMSP, acrylate or propionate, as indicated, and were photographed after 3 days' incubation.

4.2.11 The sensitivity phenotype

The observations that *R. pomeroyi* with mutations in the propionate metabolism pathway cause hypersensitivity to DMSP, acrylate and propionate lead to two important conclusions. The first concerns the nature of the toxicity itself. All of the mutations were in genes predicted to encode enzymes that degrade coenzyme A intermediates. It is known that a build-up of CoA molecules inhibits cell growth, due to the highly reactive nature of these intermediates (see below).

However, very few studies have explored the mechanism of toxicity of various CoA molecules, and very little is known on the exact causes. The work described here provides additional evidence for the toxic nature of acryloyl-CoA, propionyl-CoA and methylmalonyl-CoA, as presented in more detail in the next section.

The second, more important, conclusion is that at least some DMSP and acrylate carbon must be routed via the propionate metabolism pathway. However, it is not possible to say whether this is the major route of carbon assimilation from DMSP, since the lack of growth could also be caused by the build-up of toxic intermediates.

4.2.12 Introduction to CoA toxicity

It has long been known that short chain fatty acids such as acrylate and propionate inhibit the growth of microbes. Indeed, reports on the antimicrobial action of propionate can be traced as far back as 1913 (Kiesel, 1913), and in 1939 Hoffman *et al*. released a paper and a patent describing the use of propionate as a food additive to prevent mould formation. While propionate was initially recognised as having fungistatic properties, it was later noted that it was also inhibitory towards *Bacillus mesentericus,* a bacterium associated with bread spoilage (Chichester and Tanner, 1972).

The toxic properties of acrylate have also attracted much interest, although for different reasons. Acrylate and related compounds are commonly used as industrial chemicals in the manufacture of paints, plastics, and adhesives amongst many other things. The interest in acrylate toxicity thus stems less from an anti-microbial point of view, and more from a concern over the cytotoxicological effects it may have on animals and humans. Aside from its use in industry, acrylate is rarely seen in the environment, the major source being its production through the DMSP cleavage pathway in marine ecosystems. Related to this, some studies have demonstrated the antimicrobial effect of acrylate on microbes of the gastrointestinal tract of polar marine animals (Sieburth, 1961), and of seawater cultures (Slezak *et al*., 1994).

However, for both propionate and acrylate, there is evidence that their toxicity is in fact caused by accumulation of their respective downstream products acryloyl-CoA and propionyl-CoA. Several mechanisms of toxicity for these CoA molecules have been suggested, as follows.

4.2.12.1 Build-up of CoA intermediates sequesters free coenzyme A

One possibility is that the accumulation of any CoA-molecule depletes the pool of available CoA in the cell, thereby blocking other essential pathways that require this coenzyme. Many bacteria can synthesise the coenzyme A precursor pantothenate from the condensation of β-alanine and

pantoate, and then convert pantothenate into coenzyme A via several subsequent steps (Jackowski, 1996).

An early study conducted in 1952 showed that the presence of pantothenate enabled *Streptococcus faecalis* to tolerate higher levels of propionate (Hill, 1952). However, the reasons behind this observation were not as simple as replacement of sequestering CoA. A key study by King and Cheldelin in 1948 showed in the yeast *Saccharomyces cerevisiae* and the bacterium *Acetobacter suboxydans* that propionate inhibits the condensation of β-alanine and pantoic acid, thereby preventing the synthesis of pantothenate. This would also explain why addition of pantothenate relieved propionate sensitivity (Hill, 1952), although it also points to the toxicity being as a result of propionate, rather than propionyl-CoA.

4.2.12.2 Propionate inhibits acetate synthesis

Regardless of the means of the depletion of available coenzyme A, it seems likely that the main issue caused by a lack of CoA is the prevention of acetate synthesis. The Hill study showed that propionate sensitivity in *S. faecalis* could also be overcome by the addition of acetate (Hill, 1952). The formation of acetate from pyruvate is dependent on the presence of CoA, and so an absence of available CoA would inhibit this pathway.

Importantly, Maruyama and Kitamura (1985) showed that partially purified pyruvate dehydrogenase from *Rhodobacter sphaeroides* was inhibited by propionyl-CoA, but not by propionate. Also, a propionyl-CoA carboxylase deficient mutant was sensitive to propionate, but the sensitivity could be overcome by adding acetate to the media.

4.2.12.3 Acryloyl-CoA is a strong electrophile

Acryloyl-CoA is a very strong electrophile, which reacts readily with nucleophilic groups in molecules essential to the cell (Herrmann *et al*., 2005). Acrylates are particularly reactive with glutathionine, which is a strong nucleophile and an important antioxidant (Masip *et al*., 2006).

4.2.12.4 Propionyl CoA and methylmalonyl-CoA inhibit N-acetylglutamate synthetase

In rat liver mitochondria, propionyl-CoA, and, to a lesser degree, methylmalonyl-CoA is a potent inhibitor of N-acetylglutamate synthetase, competing with its natural substrate, glutamate (Coude *et al*., 1979). Some prokaryotes also have N-acetylglutamate synthetase enzymes, although there are no homologues of the *E. coli* enzyme in *R. pomeroyi*; perhaps propionyl-CoA acts as a competitive inhibitor for other glutamate enzymes.

4.2.12.5 *E. coli* is hypersensitive to acrylate under anaerobic conditions

Another study looked into the effect of acrylate on *E. coli* in aerobic and anaerobic conditions. They found that under anaerobic conditions, the strain could tolerate no more than 5 mM acrylate, whereas it could grow in the presence of up to 35 mM acrylate under aerobic conditions (Ayra *et al*., 2013). The authors suggested this could be due to the inhibition of pyruvate formate lyase by acrylate, or one of its derivatives, since this enzyme is required for anaerobic growth. In support of this, the authors noted a reduction in formate production from acrylate, in favour of lactate which does not require the pyruvate formate lyase (Ayra *et al*., 2013).

4.2.13 Attempts to relieve the hypersensitivity phenotype

Based on previous findings in the literature as described above, the effects on acrylate toxicity of adding acetate, pantothenate, glutathione, or glutamate to the media were determined.

4.2.13.1 Addition of acetate

Since a build-up of propionyl-CoA may inhibit acetate synthesis (see above), spot tests were done on MBM agar with 10 mM succinate plus different levels (1 – 10 mM) of acrylate and in the presence of 5 mM acetate incorporated into the medium. However, the acetate did not alleviate the sensitivity to acrylate in any case.

4.2.13.2 Addition of pantothenate

As described above, accumulated coenzyme A intermediates may either sequester, or somehow inhibit the synthesis of CoA. Since bacteria can synthesise CoA from the precursor pantothenate, spot tests were carried out as described above, but this time in the presence or absence of 10 mM pantothenate. As shown in **Figure 4.16**, the pantothenate did not rescue the mutant strains at any concentration of acrylate. Interestingly though, all strains, including the wild type, grew slightly better on succinate as a carbon source when pantothenate was added. This is not due to pantothenate acting as an additional carbon source, since *R. pomeroyi* did not grow when pantothenate was a sole carbon source.

Most bacteria can synthesise pantothenate from β-alanine and pantoate. Indeed, *E. coli* can produce 15 times as much pantothenate as is required for coenzyme A biosynthesis, and exports the excess from the cell (Jackowski and Rock, 1981). The genes and enzymes involved in pantothenate synthesis have been identified in *E. coli* and there are close homologues of all of these in *R. pomeroyi*; for example, SPO0103 is 45% identical to the *E. coli* pantothenate synthetase, PanC (Begley *et al*., 2001). Another PanC homologue in the α-proteobacterium

Rhizobium etli, shown to be essential for growth of this strain, is 49% identical to the SPO0103 gene product (Villaseñor *et al*., 2011).

It is therefore likely that *R. pomeroyi* can synthesise pantothenate. In addition the MBM media used for the spot tests also contains a low concentration of pantothenate as a vitamin supplement. However, the poorer growth of *R. pomeroyi* on succinate in the absence of extra pantothenate suggests that the bacterium cannot synthesise, or import enough of the molecule for optimal growth. Therefore, a greater concentration of pantothenate should be added to MBM for the growth of *R. pomeroyi* in the future.

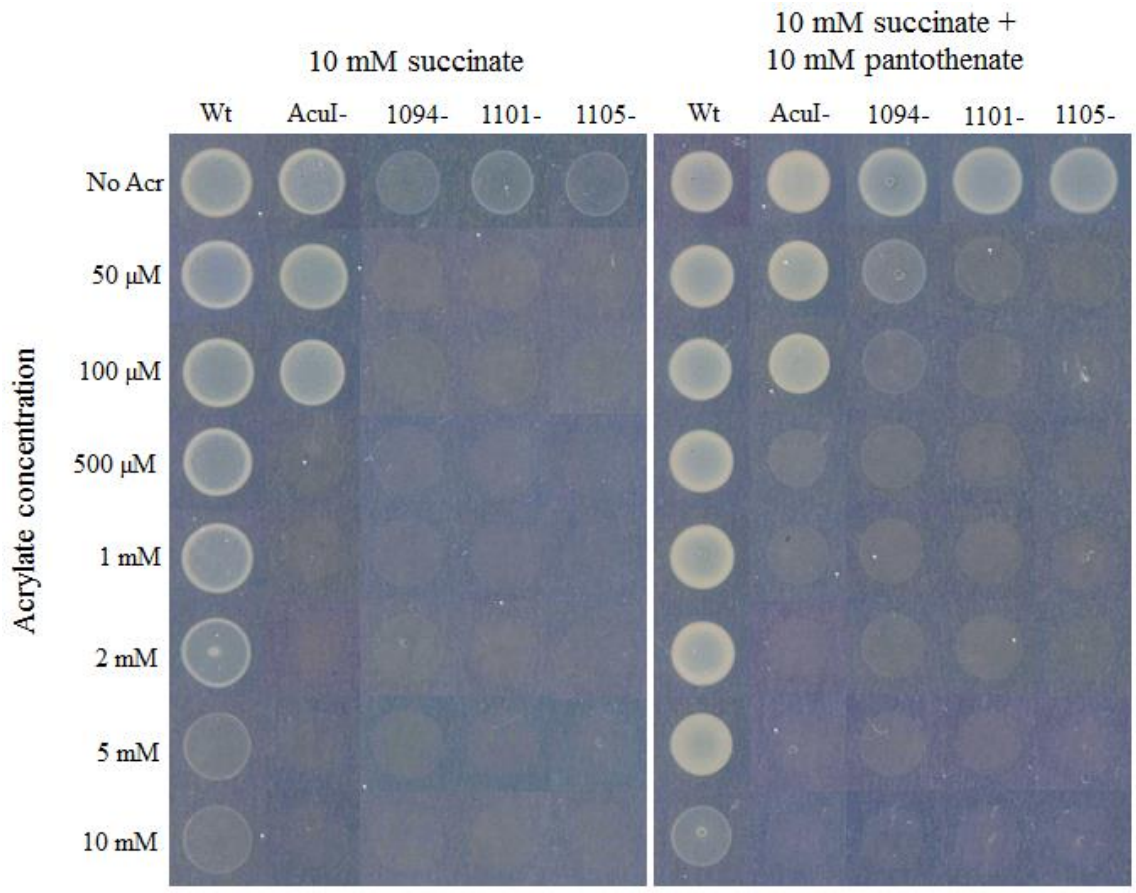


Figure 4.16 Effect of pantothenate on acrylate sensitivity in *R. pomeroyi* wild type and mutant strains. Aliquots of 10 µl cultures (OD_{600} = 1.0) of wild type *R. pomeroyi* or AcuI$^-$, SPO1094$^-$, SPO1101$^-$ or SPO1105$^-$ mutants were spotted onto solid MBM medium with 10 mM succinate, with or without 10 mM pantothenate, and varying concentrations of acrylate. Spots were photographed after 3 days' incubation.

4.2.13.3 Addition of glutathione and glutamate

Glutathione is one of the most abundant thiols produced by proteobacteria, with intracellular concentrations of 0.1 to 10 mM (Masip *et al.*, 2006). It is an important anti-stress molecule, with roles in osmo-adaptation, or coping with oxidative stress or low pH conditions (Masip *et al.*, 2006). Bacteria are able to synthesise glutathione from glutamate, which is combined with cysteine to yield γ-glutamylcysteine and glycine, a reaction which is catalysed by γ-glutamylcysteine synthase (GCS). The γ-glutamylcysteine produced is subsequently converted to glutathione by glutathione synthetase (GS). Since glutathione is a strong nucleophile, it is an important potential target of electrophilic attack by acryloyl-CoA. It was therefore of interest to

see if the addition of glutathione or its precursor, glutamate, in the media could alleviate the toxicity of acrylate to the *R. pomeroyi* mutants.

In contrast to acetate and pantothenate, the addition of glutathione and glutamate did have an effect on the hypersensitivity phenotype of the mutants. This was most dramatic when 2.5 mM glutathione was added to the media (**Figure 4.17**). In this case, the mutant strains almost regained a wild type phenotype, with signs of growth at 5-10 mM acrylate for all four mutant strains. However, the growth of the mutants was slightly poorer than wild type at acrylate concentrations of 500 µM and upwards. The presence of 0.25 mM glutamate was less effective at relieving sensitivity, and it did not have any effect on the growth of the AcuI$^-$ mutant (**Figure 4.18**). However, the 1094$^-$, 1101$^-$ and 1105$^-$ mutant strains had strong growth matching that of the wild type at 50 and 100 µM acrylate when glutamate was present.

The addition of glutathione to the media may have relieved the sensitivity phenotype by providing an additional source of this important molecule. It is known that bacteria can import glutathione; the *E. coli yliABCD* genes and the *Haemophilus influenzae dppBCDF* genes encode a dedicated ABC transporter for example (Suzuki *et al*., 2005; Vergauwen *et al*., 2010; Bachhawat *et al*., 2013). In both organisms, the genes are part of a single operon. In *H. influenzae*, DppB and DppC make up the transmembrane domains of the transporter, while DppD and DppF form a nucleotide binding domain. There are convincing homologues to the *H. influenzae* Dpp peptides in *R. pomeroyi*, although these are organised into two separate operons. Thus, DppB and DppC are 44% and 40% identical to SPO1544 and SPO1545, while DppD and DppF are 46% and 48% identical to SPO3777 and SPO3778, respectively.

The sensitivity phenotype was also somewhat relieved by glutamate, possibly as an indirect way of supplying more glutathione through its synthesis from glutamate (see above). *R. pomeroyi* has genes which are predicted to encode the enzymes involved in glutathione synthesis - GCS (SPO3626) and GS (SPO0401). Interestingly, microarray data for *R. pomeroyi* gathered by M. Kirkwood show that both of these genes are slightly up-regulated in the presence of DMSP in an *acuI*$^-$ mutant strain, compared to wild type strain (2.8-fold and 2.1-fold, respectively). Therefore it could be that a build-up of acryloyl-CoA derived from DMSP in the *acuI*$^-$ strain is depleting the glutathione store, which in turn is inducing the conversion of glutamate to glutathione.

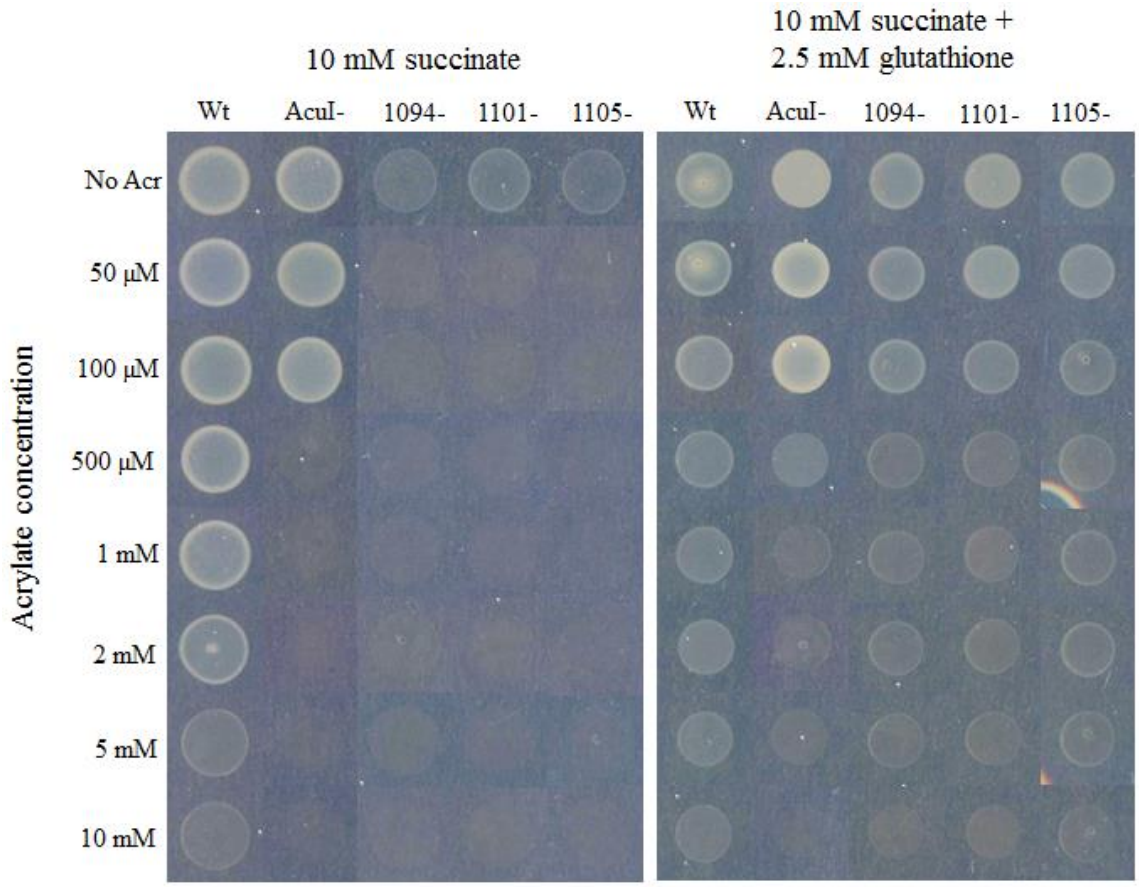


Figure 4.17 Effect of glutathione on acrylate sensitivity in *R. pomeroyi* wild type and mutant strains. Aliquots of 10 µl cultures (OD_{600} = 1.0) of wild type *R. pomeroyi* or AcuI$^-$, SPO1094$^-$, SPO1101$^-$ or SPO1105$^-$ mutants were spotted onto solid MBM medium with 10 mM succinate, with or without 2.5 mM glutathione, and varying concentrations of acrylate. Spots were photographed after 3 days' incubation.

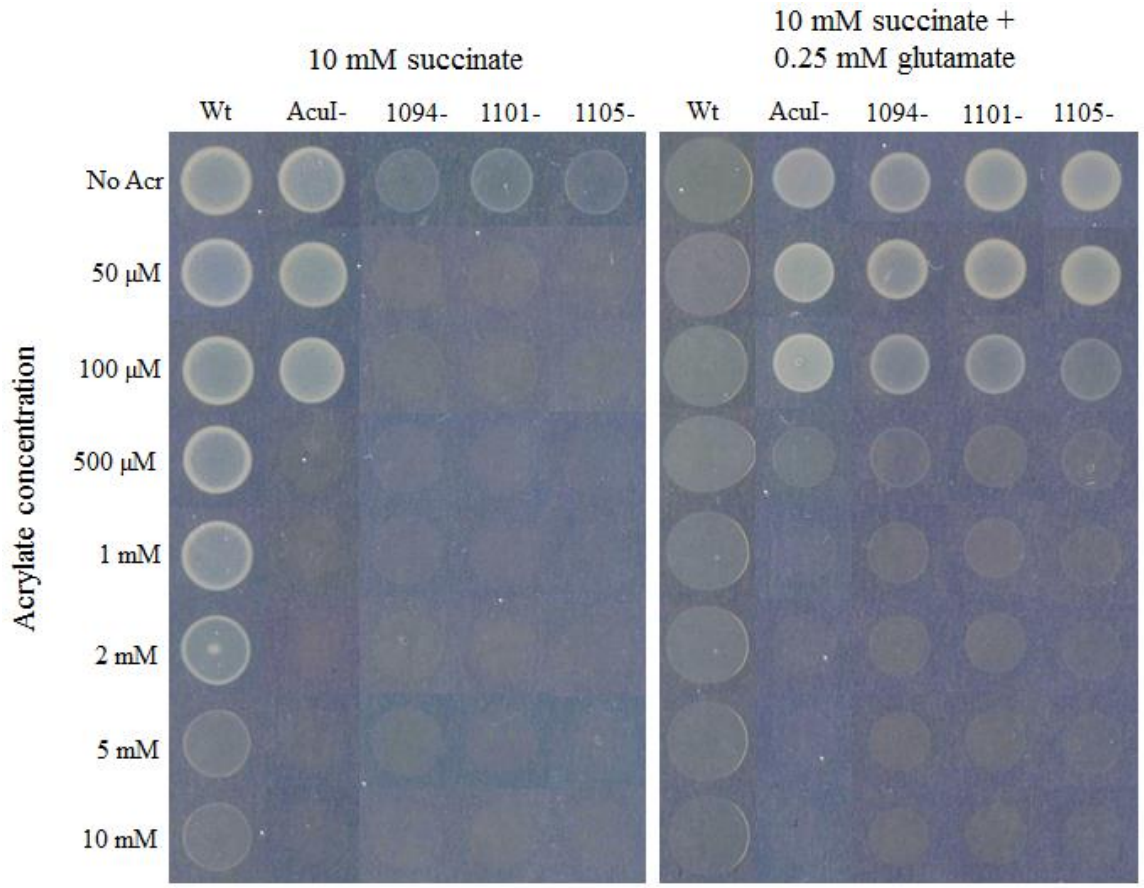


Figure 4.18 Effect of glutamate on acrylate sensitivity in *R. pomeroyi* wild type and mutant strains. Aliquots of 10 µl cultures (OD_{600} = 1.0) of wild type *R. pomeroyi* or AcuI$^-$, SPO1094$^-$, SPO1101$^-$ or SPO1105$^-$ mutants were spotted onto solid MBM medium with 10 mM succinate, with or without 0.25 mM glutamate, and varying concentrations of acrylate. Spots were photographed after 3 days' incubation.

4.2.14 Identifying the acryloyl-CoA ligase

If *R. pomeroyi* can indeed degrade acrylate via acryloyl-CoA and propionyl-CoA, there must be an enzyme which adds the CoA molecule to acrylate, presumably via an acryloyl-CoA ligase.

Inspection of the *R. pomeroyi* genome revealed 28 genes that are annotated as encoding CoA ligase enzymes. Of these, one (SPO2934) was up-regulated approximately 3-4 fold in the presence of both DMSP and acrylate as seen in the Kirkwood microarray data set and encodes a peptide with 45% sequence identity to the propionyl-CoA ligase of Enteric bacteria (Guo and Oliver, 2012).

Interestingly, there is a precedent for enzymes of the CoA-ligase family to display a low level of specificity for one substrate. For example, the propionyl-CoA ligases of *E. coli* and *S. typhimurium* can also use acetate as a substrate, albeit with a slightly lower efficiency (Guo and Oliver, 2012; Horswill and Escalante-Semerena, 1999). Furthermore, a study into the substrate specificity of acetyl-CoA ligase from yeast found that propionate and acrylate could each substitute for acetate in the reaction at 60 and 63% the rate of acetate, respectively (Patel and Walt, 1987). Significantly, the same enzyme in both *Ralstonia solanacearum* and *Salmonella choleraesuis* has previously been shown to possess propionyl- and acryloyl-CoA activity (Rajashekhara and Watanabe, 2004). The putative SPO2934 *R. pomeroyi* propionyl-CoA ligase was therefore investigated further to see if it could also act on acrylate as a substrate.

Initially, advantage was taken of the toxic nature of the predicted product, acryloyl-CoA. It was predicted that if the gene for a bona fide CoA ligase were cloned and over -expressed, this would lead to enhanced acrylate sensitivity due to accumulation of acryloyl-CoA.

Therefore, the SPO2934 gene was PCR amplified from the *R. pomeroyi* genome, using primers designed with *Xba*I and *Pst*I restriction sites, allowing the ligation of the PCR product into the vector pBluescript to form plasmid pBIO2094. This plasmid was mobilised into wild type *E. coli* strain K-12 and its acrylate sensitivity was compared to that of *E. coli* strain K-12 containing 'empty' pBluescript.

To do this, spot tests were carried out using overnight cultures of these two strains, each adjusted to an OD_{600} of ~1.0. Then, 10 µl aliquots were spotted onto LB agar plates containing different concentrations of acrylate and the plates were incubated overnight at 37°C. As seen in **Figure 4.19.** *E. coli* with the empty pBluescript vector could grow well on acrylate as high as 500 µM, but the presence of pBIO2094 prevented growth at concentrations greater than 30 µM.

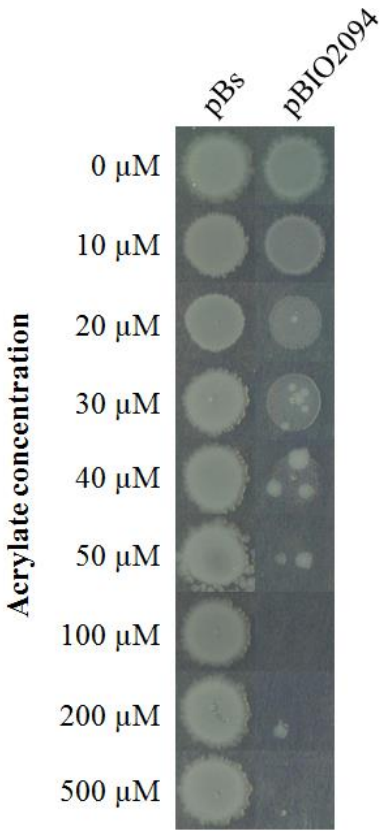


Figure 4.19 Effect of the cloned SPO2934 gene on the ability of *Escherichia coli* strain K-12 to grow in the presence of acrylate. Cultures of wild type *E. coli* strain K-12 with the cloned SPO2934 gene from *R. pomeroyi* DSS-3, in pBIO2094, or with the 'empty' plasmid pBluescript (PBs), were grown in LB medium to an OD_{600} of ~1.0. The cultures were washed in M9 salts and 10 µl aliquots were spotted onto LB agar plates, supplemented with varying levels of acrylate, as indicated. Plates were photographed after 1 nights' incubation at 37°C.

4.2.14.1 Constructing and characterising a SPO2934$^-$ mutant strain of *R. pomeroyi*

Having shown that the SPO2934 gene product likely acts on acrylate, a SPO2934$^-$ mutant strain of *R. pomeroyi* was then made using the pBIO1879 suicide vector as described above. To do this, a ~1000 bp region from within SPO2934 was amplified using primers with *Xba*I and *Pst*I restriction sites, digested and ligated to pBIO1879 cut with the same enzymes, creating pBIO2096. This recombinant plasmid was mobilised from *E. coli* into *R. pomeroyi* via tri-parental mating. Strains with a successful genomic insertion in SPO2934 were verified by Southern blot, as described previously in **Section 4.2.5**. Genomic DNA from wild type and putative SPO2934$^-$ strains were digested with *Nde*I. In wild type *R. pomeroyi*, the SPO2934 gene is positioned in a 9.6 Kb *Nde*I fragment (**Figure 4.20**), and a successful pBIO2096 insert would increase this to 16.4 Kb. This was the case for one strain (see **Figure 4.21**), so this was renamed J562.

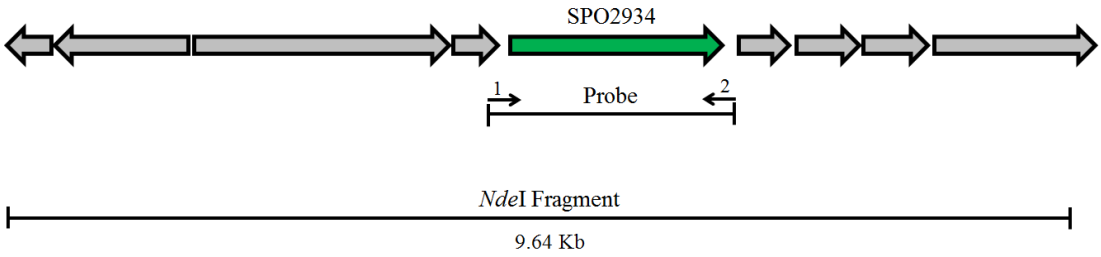


Figure 4.20 Probe used for Southern blotting of SPO2934$^-$ mutant. The probe for SPO2934 was amplified using primers 1 and 2 (SPO2934FOR and SPO2934REV), giving a PCR product, encompassing the entire gene, of ~2100 bp. The probe binds to an *Nde*I fragment of ~9.6 Kb in the wildtype genome.

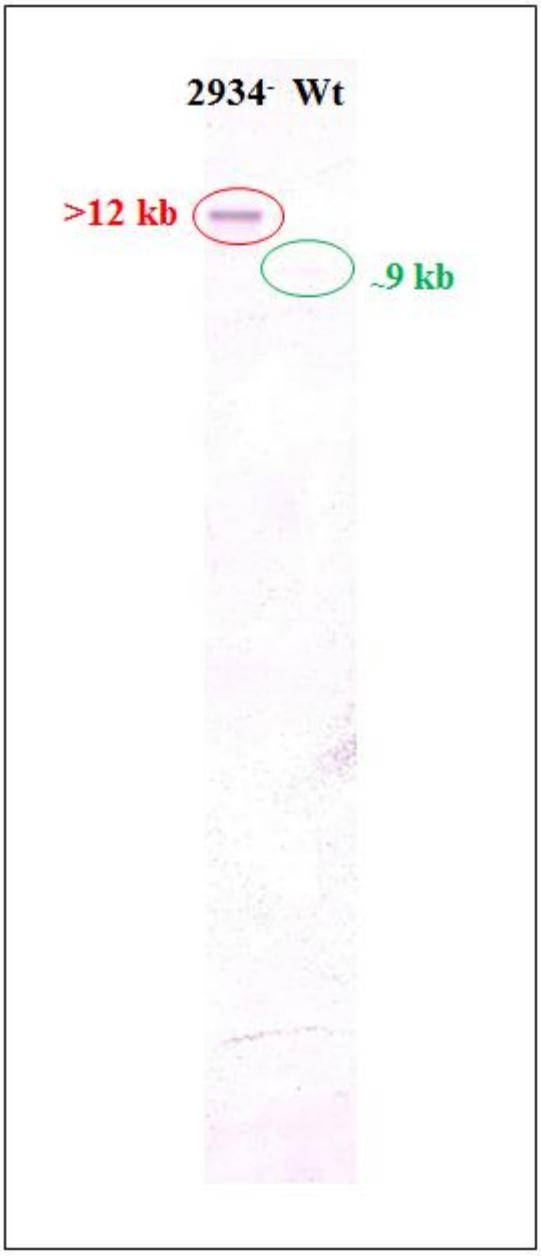


Figure 4.21 Southern blot of wild type *R. pomeroyi* and SPO2934⁻ mutant. Genomic DNA from wild type (Wt), and a putative 2934⁻ strain of *R. pomeroyi* was digested with *Nde*I and probed with PCR product of intact SPO2934. Approximate DNA lengths of the labelled bands are shown. A successful insertion in SPO2934 was expected to expand the *Nde*I fragment size from ~9.6 kb to ~16.4 kb, and the pattern of bands produced in this blot are consistent with this (Wt band of ~9 kb and a larger band of >12 kb in the mutant strain).

Since SPO2934 is predicted to convert acrylate to acryloyl-CoA, it was anticipated that a mutation in this gene might abolish growth on acrylate. However, this was not the case as the mutant grew just as well as the wild type in liquid MBM medium in which acrylate (5 mM) was

the sole carbon source. One possibility is that *R. pomeroyi* may have more than one enzyme that can add CoA to acrylate. This would be analogous to the situation in *S. typhimurium*, in which a mutation that abolished the *bona fide* propionyl-CoA ligase retained some enzymatic activity, which was in fact due to an acetyl-CoA ligase that could also use propionate as a substrate (Horswill and Escalante-Semerena, 1999).

4.2.14.2 Redundancy in acryloyl-CoA ligase activity in *R. pomeroyi*

As mentioned, the genome of *R. pomeroyi* has many genes whose products are annotated as acyl-CoA ligases, although SPO2934 is the only predicted acyl-CoA ligase gene to be induced in the presence of DMSP and acrylate. Two of these putative acyl-CoA ligase genes were chosen at random to check for their ability to use acrylate as a substrate. Thus, SPO2528 and SPO1014 were amplified from the *R. pomeroyi* genome using primers with *Xba*I and *Pst*I restriction sites, then individually cloned into pBluescript, creating plasmids pBIO2093 (SPO2528) and pBIO2095 (SPO1014) and mobilised into *E. coli* K-12 as described above.

The sensitivity spot tests were carried out as described above. As shown in **Figure 4.22**, both of these recombinant plasmids conferred some sensitivity to acrylate, but not to the same extent as did the cloned SPO2934. Thus, wild type *E. coli* grew well at acrylate concentrations in excess of 15 mM, but, with either pBIO2093 or pBIO2095 growth was compromised at acrylate concentrations of 5 – 7.5 mM. However, this was some 10-fold higher than the concentration that was tolerated by *E. coli* containing the cloned SPO2934 gene (see **Figure 4.19**).

Thus, it seems likely that both pBIO2093 (SPO2528) and pBIO2095 (SPO1014) encode CoA ligases that can act on acrylate as a substrate even though these may not be their "natural" substrate. If so, this could explain the lack of a growth phenotype seen in the SPO2934$^-$ mutant.

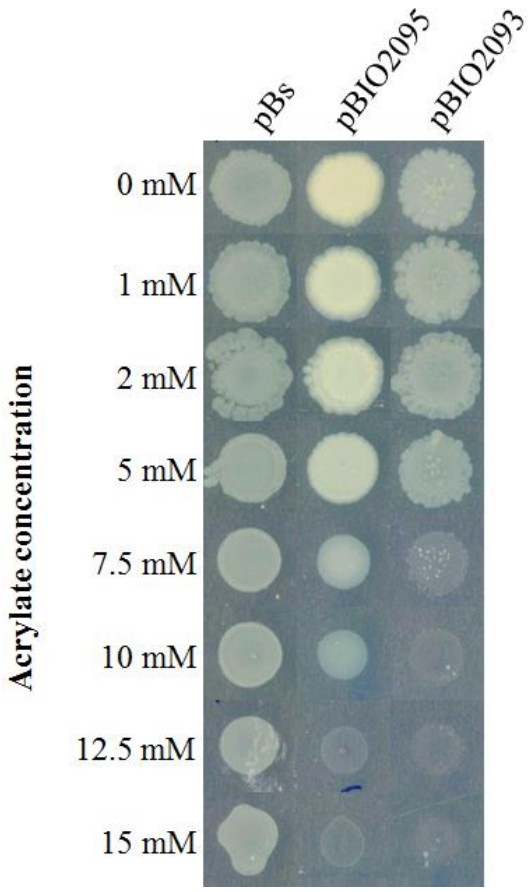


Figure 4.22 Effect of the cloned SPO2528 and SPO1014 genes on the ability of *Escherichia coli* strain K-12 to grow in the presence of acrylate. Cultures of wild type *E. coli* strain K-12 with the cloned SPO2528 and SPO1014 genes from *R. pomeroyi* DSS-3, in pBIO2093 and pBIO2095, respectively, or with the 'empty' plasmid pBluescript (PBs), were grown in LB medium to an OD_{600} of ~1.0. The cultures were washed in M9 salts and 10 µl aliquots were spotted onto LB agar plates, supplemented with varying levels of acrylate, as indicated. Plates were photographed after 1 night's incubation at 37°C.

4.3 Summary

The work in this chapter provides evidence for a pathway of DMSP catabolism through acrylate, acryloyl-CoA and propionyl-CoA in *R. pomeroyi*, as summarised in **Figure 4.23**. The gene SPO2934 was shown to encode a protein which, when expressed in *E. coli*, caused dramatic hypersensitivity to acrylate. This shows that SPO2934 likely encodes an acryloyl-CoA ligase in *R. pomeroyi*. Significantly, a SPO2934$^-$ mutant strain was still able to grow on acrylate and DMSP. In accordance with this, a redundancy in CoA-ligase activity was demonstrated, with at least two other genes in *R. pomeroyi* shown to encode enzymes with a lower level acryloyl-CoA ligase activity. Mutations were also made in the acryloyl-CoA reductase AcuI$^-$, the propionyl-CoA carboxylase PccA and PccB and the methylmalonyl-CoA mutase SPO1105. The hypersensitivity phenotype of each mutant strain to DMSP and acrylate suggests at least some of the carbon from DMSP and acrylate is routed via this pathway, although it does not rule out the presence of an alternative pathway of acrylate catabolism. The sensitivity phenotype was explored, and found to be partly relieved by the presence of glutamate or glutathione.

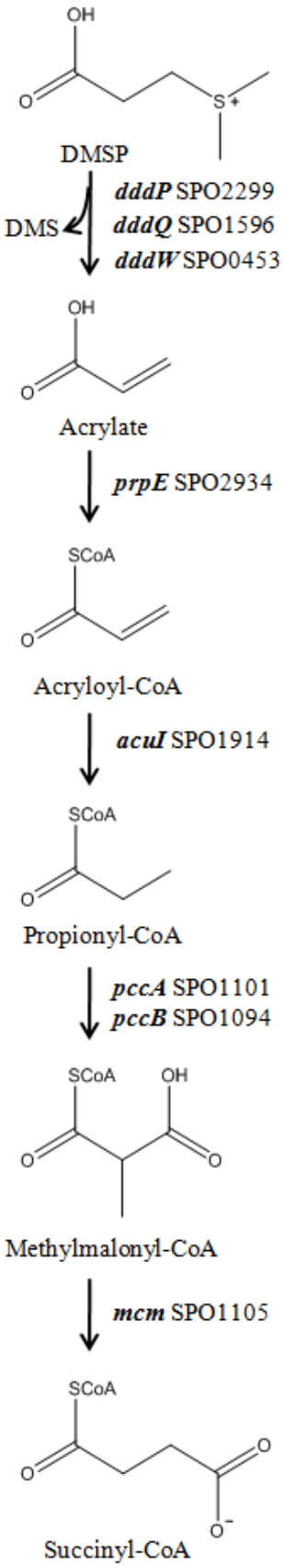


Figure 4.23 Proposed pathway of DMSP catabolism in *R. pomeroyi*. Proposed pathway of DMSP and acrylate catabolism in *R. pomeroyi*, consisting of DMSP lyases DddP (SPO2299),

DddQ (SPO1596) and DddW (SPO0453), acryloyl-CoA ligase PrpE (SPO2934), acryloyl-CoA reductase AcuI (SPO1914), propionyl-CoA carboxylase Pcc (SPO1094 and SPO1101), and methylmalonyl-CoA mutase (SPO1105).

Chapter 5

Bioinformatic Analysis of the Roseobacter Clade

5.1 Introduction

The characteristics of the important, DMSP-utilizing, Roseobacter clade of marine bacteria were briefly presented in **Chapter 1**. Using bioinformatic analyses, this chapter explores in detail the distribution and synteny of genes involved in DMSP catabolism in different strains and species of the Roseobacters.

5.1.1 The Roseobacter Clade

The identification of major (i.e. most abundant) clades of marine bacteria has arisen largely from culture-independent methods such as the amplification and sequencing of 16S rRNA genes from environmental samples. Most notable among these are the "SAR" bacteria and, of these, the remarkably abundant SAR11 and SAR116 bacteria. Many of those clades have very few or no cultivated representatives. In contrast, many members of the abundant Roseobacter clade are readily isolated and grown under laboratory conditions.

As reviewed by Buchan *et al*. (2005) and Wagner-Döbler and Biebl (2006), members of the Roseobacter clade are heterotrophic bacteria, widely distributed in marine environments. They have been isolated from a large variety of habitats, although they are most abundant in coastal waters and polar regions. Roseobacters are also well-represented amongst bacteria that inhabit the "phycosphere" - the area immediately surrounding phyto- and zoo-plankton. Additionally, increases in phytoplankton populations, such as during an algal bloom, have been shown to be accompanied by an increase in total bacterial numbers, including Roseobacter species (reviewed in Buchan *et al*., 2014).

The phycosphere is one environment where Roseobacter strains are exposed to high levels of DMSP, produced by many types of phytoplankton. A feature of many Roseobacters is their ability to catabolise DMSP by the cleavage or demethylation pathway. In many cases, both pathways are found in the same organism - a feature that is, so far, restricted to this clade and members of the SAR11 group (see **Chapter 2**).

Over the past decade, there has been a drive to increase the numbers of whole genome sequences for ecologically relevant marine bacteria, largely funded through the Gordon and Betty Moore Foundation. As a result, more than 50 genome sequences of Roseobacter strains are now available, all of which are listed on the 'Roseobase' website (http://www.roseobase.org/). These sequences provide a wealth of valuable data to be mined. In this chapter, the abundance and distribution of DMSP catabolism genes amongst members of the Roseobacter clade is explored, with particular attention to the genomic positioning of each of these genes.

5.2 Results

5.2.1 DMSP genes are abundant in the Roseobacter clade

According to the *Roseobase* website, there are currently 54 genomes from strains of Roseobacters available for BLAST searches. Of these listed strains, 42 have complete genomes, assigned to the Roseobacter group (taxonomic ID 31989) on the NCBI database. The *dmdA* and *ddd* gene products were used to interrogate the genomes of the Roseobacter group. For the purpose of identifying homologues, the query sequence used in each case was a sequence of a Roseobacter enzyme that has been shown to be functional. Thus, SPO1913 was used as the query for DmdA, ISM_05385 for DddP, EE36_11918 for DddL, ISM_14090 for DddQ, SPO0453 for DddW and SSE37_17628 for DddD. A summary of the findings is presented in **Table 5.1,** and more detailed results including locus tag and percentage identity for each individual gene product are reported in later tables.

Of the 42 Roseobacter strains queried, 35 were found to possess at least one homologue to a functional DmdA or Ddd protein sequence (**Table 5.1**). Of these, 22 strains possess genes encoding both the DmdA demethylase and at least one of the *ddd* genes for the cleavage pathway, with 10 having multiple DMSP lyase *ddd* genes.

In many cases, the genotype of the strain is reflected by its phenotype. For example, *Ruegeria pomeroyi*, *Roseovarius nubinhibens*, and *Ruegeria lacuscaerulensis* all have DmdA and at least one Ddd enzyme, and all produce both methanethiol and DMS from DMSP (González *et al*., 2003; Li *et al*., 2014). Similarly, *Roseobacter* sp. GAI-101, *Sagittula stellata* E-37 and *Sulfitobacter* EE-36, which all have at least one *ddd* DMSP lyase gene but no DMSP demethylase, have been found to produce DMS, but not methanethiol, from DMSP (González *et al*., 1999). There are, however, some unusual cases where the actual phenotype is not as expected. For example, *P. gallaeciensis* DSM17395 has both DmdA and DddP, but while it has been shown to demethylate DMSP with the eventual release of methanethiol, this strain did not have a Ddd^+ phenotype (Dickschat *et al*., 2010). Therefore, the DddP from this strain may not be active, at least under laboratory conditions, and no *in vitro* studies have been carried out on this enzyme. Another case is that of *Dinoroseobacter shibae* which possesses DmdA, plus a copy of *both* DddD *and* DddL. As with *P. gallaeciensis, D. shibae* has been shown to produce methanethiol from DMSP, indicating the presence of an active DMSP demethylation pathway (Dickschat *et al.*, 2010) but it also does not produce DMS from DMSP. The copy of DddD from *D. shibae* may not be functional, as has been shown for DddD from *R. pomeroyi* (see later), but DddL has been cloned and expressed in a heterologous host where it does possess DMSP lyase activity. It is worth noting the discrepancy between genotype and phenotype in *D. shibae* is not

restricted to DMSP catabolism genes. It is the only Roseobacter which, while possessing the necessary genes for using monomethylamine as a sole carbon or nitrogen source, is unable to grow on this substrate in laboratory conditions (Chen, 2012). Thus *D. shibae* may not behave in the lab as it does *in situ*.

Another, and potentially more exciting, example of an unexpected phenotype is seen in *Oceanibulbus indolifex* which does not have any *ddd* genes, but does make a small amount of DMS from DMSP (Dickschat *et al*., 2010). While it cannot be ruled out that this release of DMS from DMSP is by non-enzymatic means, it is also possible that there is another unidentified DMSP lyase gene in this strain.

These latter examples do serve as a reminder that genotype alone is not enough to determine the presence of biochemical pathways in bacteria, and, where possible, the phenotype of the organism, and ideally the enzyme function, should be confirmed.

A number of strains did not possess any convincing Ddd or DmdA homologues. These were: *Citreicella* sp. 357; *Loktanella hongkongensis* DSM 17492; *Loktanella vestfoldensis* DSM 16212; *Oceanicola* sp. S124; *Oceaniovalibus guishaninsula* JLT2003; *Pelagibaca bermudensis* HTCC2601 and *Wenxinia marina* DSM 24838. Of these strains, the Johnston lab had access to just one - *Pelagibaca bermudensis*.

Table 5.1 Homologues of *dmdA* and *ddd* products in sequenced Roseobacter strains.

Organism	**DmdA**	**DddP**	**DddL**	**DddQ**	**DddW**	**DddD**
Citreicella sp. SE45						✓
Citreicella sp. 357						
*Dinoroseobacter shibae*DFL12	✓		✓			✓
Jannaschia sp. CCS1	✓	✓				
Loktanella hongkongensis DSM17492						
Loktanella vestfoldensis DSM16212						
Loktanella vestfoldensis SKA53			✓			
Maritimibacter alkaliphilus HTCC2654			✓			
Oceanibulbus indolifex HEL45	✓					
Oceanicola batsensis HTCC2597			✓			
Oceanicola granulosus HTCC2516		✓				
Oceanicola sp. S124						
Oceaniovalibus guishaninsula JLT2003						

Octadecabacter arcticus 238	✓	✓				
Octadecabacter antarcticus 307	✓					
Pelagibaca bermudensis HTCC2601						
Phaeobacter gallaeciensis DSM17395	✓	✓				
Phaeobacter gallaeciensis 2.10		✓				
Rhodobacterales bacterium KLH11	✓	✓				✓
Rhodobacterales bacterium HTCC2083	✓	✓				✓
Rhodobacterales bacterium HTCC2150	✓	✓		✓		
Roseobacter denitrificans Och114	✓	✓				
Roseobacter litoralis Och149	✓	✓				
Roseobacter sp. AzwK-3b	✓					
Roseobacter sp. CCS2	✓	✓				
Roseobacter sp. GAI101			✓			
Roseobacter sp. MED193	✓				✓	
Roseobacter sp. SK209-2-6	✓	✓		✓		
Roseovarius nubinhibens ISM	✓	✓		✓✓		
Roseovarius sp. TM1035	✓	✓				
Roseovarius sp. 217	✓	✓				
Ruegeria lacuscaerulensis ITI-1157	✓	✓		✓		
Ruegeria pomeroyi DSS-3	✓	✓		✓	✓	✓
Ruegeria sp. TM1040	✓	✓				
Ruegeria sp. Trich CH4B	✓	✓				
Ruegeria sp. TW15	✓	✓		✓		
Ruegeria sp. R11	✓	✓				
Sagittula stellata E-37						✓
Sulfitobacter sp. EE36			✓			
Sulfitobacter NAS-14.1			✓			
Thalassiobium R2A62	✓	✓		✓		
Wenxinia marina DSM24838						

Ticks indicate the presence of the corresponding polypeptide in genome-sequences of Roseobacter strains. Two ticks indicate the presence of two DddQ homologues in *R. nubinhibens*. BLASTp searches used functionally verified Roseobacter Ddd and DmdA sequences as the query, and homologues were determined based on the following *E*-value cut-offs: DmdA, <e^{-149}; DddP, 0.0; DddL, <e^{-66}; DddQ, <e^{-28}; DddD, 0.0; DddW, <e^{-67}. Green-shaded boxes indicate that the corresponding polypeptide has been confirmed as a

functional DMSP lyase or demethylase, and the red-shaded box indicates a non-functional polypeptide.

5.2.1.1 *Pelagibaca bermudensis* does not make DMS or MeSH from DMSP

Since *P. bermudensis* lacks any known DMSP lyase or demethylase genes, it was of interest to see if this strain could make DMS or methanethiol from DMSP. To test this, cultures of *P. bermudensis* HTCC2601 were grown overnight in marine broth. The following day, the cultures were washed and resuspended in MBM. An aliquot of the culture was transferred to a vial with 5 mM DMSP and incubated at room temperature. The headspace of the vials was assayed for the presence of both DMS and of methanethiol by gas chromatography after 1 hour incubation, and then again after overnight incubation. In each case, *P. bermudensis* did not produce any DMS or methanethiol above background levels seen in DMSP + media-only controls. Thus this strain does not have a Ddd^+ or Dmd^+ phenotype in laboratory conditions, in accordance with the absence of known DMSP catabolism genes from its genome.

5.2.2 *dmdA*

DmdA of *R. pomeroyi* (locus tag: SPO1913) is a 364 amino acid DMSP demethylase, and was originally annotated as a glycine-cleavage T-family protein. When the protein sequence of the SPO1913 product was used in a BLASTp-search of the Roseobacter group on the NCBI database, a total of 24 other strains possessed a convincing homologue, with identities of at least 61%, and *E*-values of $3e^{-149}$ (The next-best hit was a peptide from *Roseobacter* sp. MED193 with a drop down to 41% identity, and an *E*-value of $2e^{-93}$), as shown in **Table 5.2**. The relatedness of the Roseobacter DmdA homologues is also shown as a phylogenetic tree in **Figure 5.1**. This shows that while the sequences are quite closely related, a cluster containing DmdA of *R. pomeroyi*, *Roseobacter* sp. SK209-2-6, *Rhodobacterales bacterium* KLH11 and *Ruegeria conchae* are slight outliers from the other strains. As mentioned above, several strains have been confirmed as able to produce methanethiol from DMSP (highlighted in blue in **Figure 5.1**), which is indicative of the presence of the DMSP demethylation pathway. However, only the SPO1913 gene product has been shown directly to be a functional DMSP demethylase (Howard *et al*., 2006). Significantly, and in contrast to the DMSP lyases, there is no evidence to suggest the existence of any alternative type of DMSP demethylase. Those strains lacking DmdA which have been tested did not produce methanethiol from DMSP (González *et al*., 1999)

Table 5.2 Homologues of DmdA (SPO1913) in the Roseobacter clade (continued on next page).

Organism	Locus tag	Identity to SPO1913	*E* value	Size (amino acids)
Dinoroseobacter shibae DFL 12	Dshi_2319	63%	3e-149	368
Jannaschia sp. CCS1	Jann_2379	61%	2e-165	367
Oceanibulbus indolifex HEL45	OIHEL45_13355	63%	8e-168	367
Octadecabacter arcticus 238	OA238_c20430	62%	6e-163	366
Octadecabacter antarcticus 307	OAN307_c32590	62%	3e-164	366
Phaeobacter gallaeciensis DSM17395 (BS107)	PGA1_262p01830	65%	1e-170	368
Rhodobacterales bacterium KLH11	RKLH11_1737	75%	0.0	367
Rhodobacterales bacterium HTCC2083	RB2083_1139	64%	2e-170	367
Rhodobacterales bacterium HTCC2150	RB2150_02909	65%	5e-176	368
Roseobacter denitrificans OCh 114	RD1_2288	64%	9e-168	367
Roseobacter litoralis Och 149	RLO149_c022350	63%	1e-166	367
Roseobacter sp. AzwK-3b	RAZWK3B_13009	64%	1e-164	369
Roseobacter sp. CCS2	RCCS2_18176	63%	1e-170	366
Roseobacter sp. MED193	MED193_02800	65%	1e-173	367
Roseobacter sp. SK209-2-6	RK20926_18022	68%	0.0	362

Organism	**Locus tag**	**Identity to SPO1913**	***E* value**	**Size (amino acids)**
Roseovarius nubinhibens ISM	ISM_00170	62%	3e-163	365
Roseovarius sp. TM1035	RTM1035_03430	64%	3e-167	370
Roseovarius sp. 217	ROS217_14471	64%	2e-161	369
Ruegeria lacuscaerulensis ITI-1157	SL1157_2967	67%	1e-177	367
Ruegeria pomeroyi DSS-3	SPO1913	100%	0.0	364
Ruegeria sp. TM1040	TM1040_1444	68%	1e-177	385
Ruegeria sp. Trich CH4B	SCH4B_2196	68%	3e-177	371
Ruegeria sp. TW15	N/A	76%	0.0	364
Ruegeria sp. R11	RR11_3440	65%	3e-175	367
Thalassiobium R2A62	TR2A62_0839	65%	6e-171	367

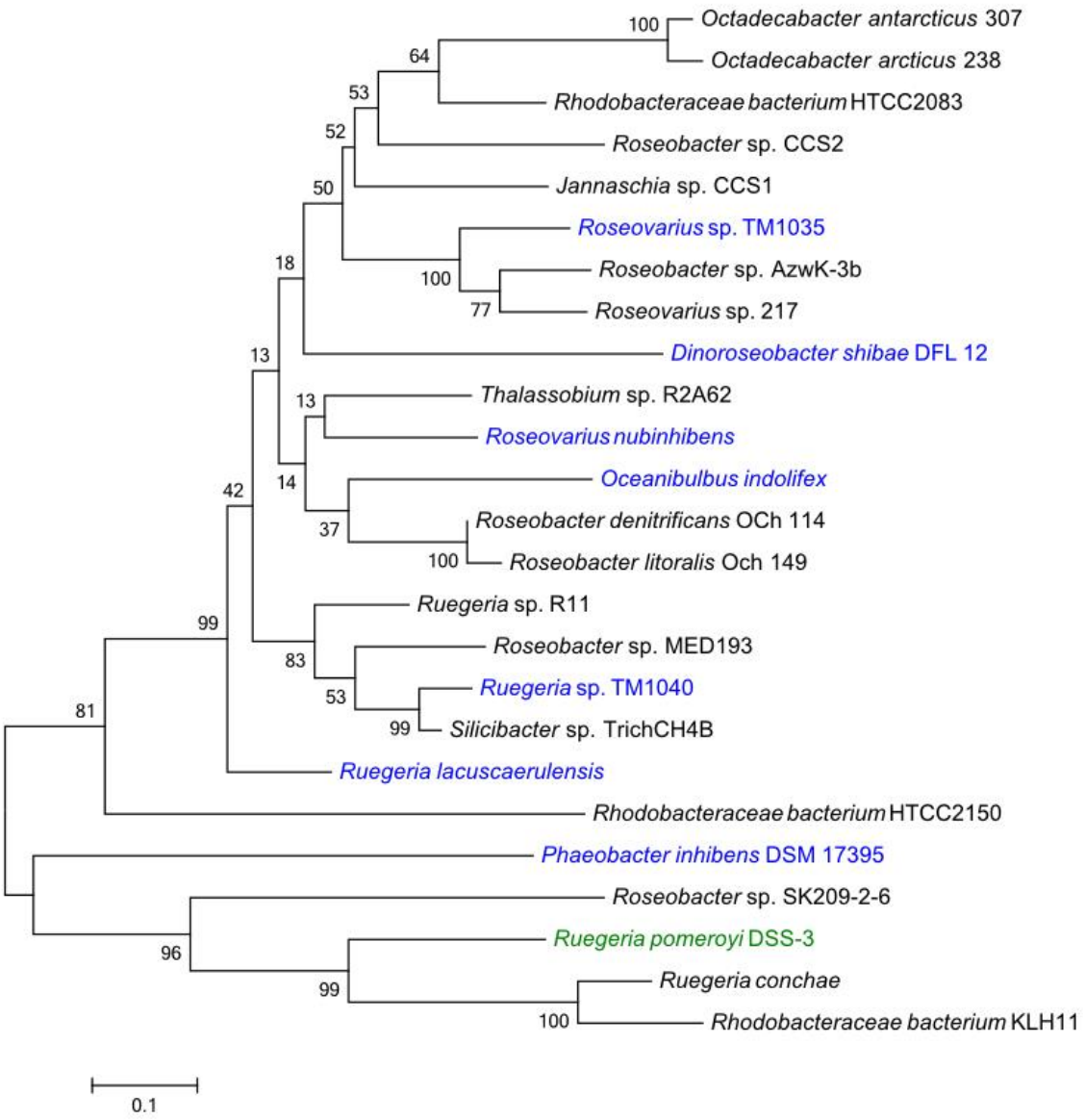


Figure 5.1 Phylogenetic relationship between DmdA homologues in Roseobacters. Protein sequences of DmdA homologues were aligned with MEGA 6 and used to estimate an unrooted phylogenetic tree using *LG model, Gamma-Distributed*. The scale bar indicates number of substitutions per site. Bootstrap values of 500 replicates are given at the base of each branch pair. Green text indicates the DmdA of that species has been confirmed as functional, blue indicates that organism has been shown experimentally to make MeSH from DMSP.

5.2.3 *dddP* genes are also abundant amongst Roseobacter strains

The product of the *Roseovarius nubinhibens dddP* gene (locus tag: ISM_05385) has been shown to possess DMSP cleavage activity using *in vitro* studies of the purified enzyme (Kirkwood *et al*., 2010a). Therefore, the sequence of this DddP peptide was used as the query in a BLASTp search for homologues amongst the Roseobacters. Hits to DddP were frequent, with 22 strains possessing homologues of at least 66% identity and *E*-values of 0.0. In the case of *Roseobacter* sp. CCS2, the CDS region predicted to encode DddP was annotated as two separate genes, both with 74% and 75% identity to each half of DddP. This annotation stems from the presence of translational stop codons in the middle of the gene in all possible reading frames, indicating either a sequencing error, or that this *dddP* gene may have acquired a mutation causing premature translation termination. No studies have been carried out to test the phenotype of strain CCS2, but since it does not contain any other known DMSP lyase, it would not be surprising if it lacked DMSP cleavage activity.

The *dddP* gene from *R. pomeroyi* (locus tag: SPO2299) has also been shown experimentally to encode a functional DddP lyase (Todd *et al*., 2010b), and both *Roseovarius* sp. TM1035 and *Phaeobacter inhibens* DSM17395, in which DddP is the only known DMSP lyase, have been shown to have Ddd^+ phenotypes. However, another strain, *Ruegeria* sp. TM1040, has been reported as being unable to produce DMS from DMSP under laboratory conditions (Miller and Belas, 2004), despite having a homologue of DddP which is closely related to that of *R. nubinhibens*. The same situation was also reported for *R. lacuscaerulensis* (Moran *et al*., 2012), although preliminary studies in our lab showed that this strain could make DMS from DMSP (unpublished observations), and it has at least been shown to possess a functional DddQ (Li *et al*., 2014). Therefore, its DddP may be expressed/functional only under a very specific set of conditions. One possibility is certain co-factors are required in the media for the enzyme to work. For example, recent work exploring the structure and mechanism of DddP from *Roseobacter denitrificans* found that the enzyme required iron as a metal co-factor (Hehemann *et al*., 2014).

Table 5.3 Homologues of DddP (ISM_05385) in the Roseobacter clade (continued on next page)

Organism	Locus tag	Identity to ISM_05385	*E* value	Size (amino acids)
Jannaschia sp. CCS1	Jann_1779	66%	0.0	443
Oceanicola granulosus HTCC2516	OG2516_03143	73%	0.0	444
Octadecabacter arcticus 238	OA238_c10540	74%	0.0	446
Phaeobacter gallaeciensis DSM17395 (BS107)	PGA1_c18750	82%	0.0	463
Phaeobacter gallaeciensis 2.10	PGA2_c17840	81%	0.0	450
Rhodobacterales bacterium KLH11	RKLH11_1853	83%	0.0	447
Rhodobacterales bacterium HTCC2083	B2083_2325	78%	0.0	446
Rhodobacterales bacterium HTCC2150	RB2150_0443	82%	0.0	446
Roseobacter denitrificans OCh 114	RD1_2566	77%	0.0	447
Roseobacter litoralis Och 149	RLO149_c019880	76%	0.0	447
Roseobacter sp. CCS2*	RCCS2_02043 RCCS2_02038	74% 75%	6e-147 2e-93	268 198
Roseobacter sp. SK209-2-6	RSK20926_21375	84%	0.0	447
Roseovarius nubinhibens ISM	ISM_05385	100%	0.0	446
Roseovarius sp. TM1035	RTM1035_11150	80%	0.0	447
Roseovarius sp. 217	ROS217_17567	80%	0.0	447

Organism	**Locus tag**	**Identity to ISM_05385**	***E* value**	**Size (amino acids)**
Ruegeria lacuscaerulensis ITI-1157	SL1157_2466	85%	0.0	447
Ruegeria pomeroyi DSS-3	SPO2299	84%	0.0	393
Ruegeria sp. TM1040	TM1040_1016	85%	0.0	447
Ruegeria sp. Trich CH4B	SCH4B_4481	83%	0.0	447
Ruegeria sp. TW15	N/A	83%	0.0	479
Ruegeria sp. R11	RR11_2557	81%	0.0	447
Thalassiobium R2A62	R2A62_0393	79%	0.0	446

***This organism has two *dddP* fragments, separated by a stop codon. BLASTp searches for DddP homologues used ISM_05385 as the query. Locus tags of convincing homologues (identities >66%, and *E*-values of 0.0) are presented, along with percentage identity to ISM_05385, *E*-values and size of predicted polypeptide.**

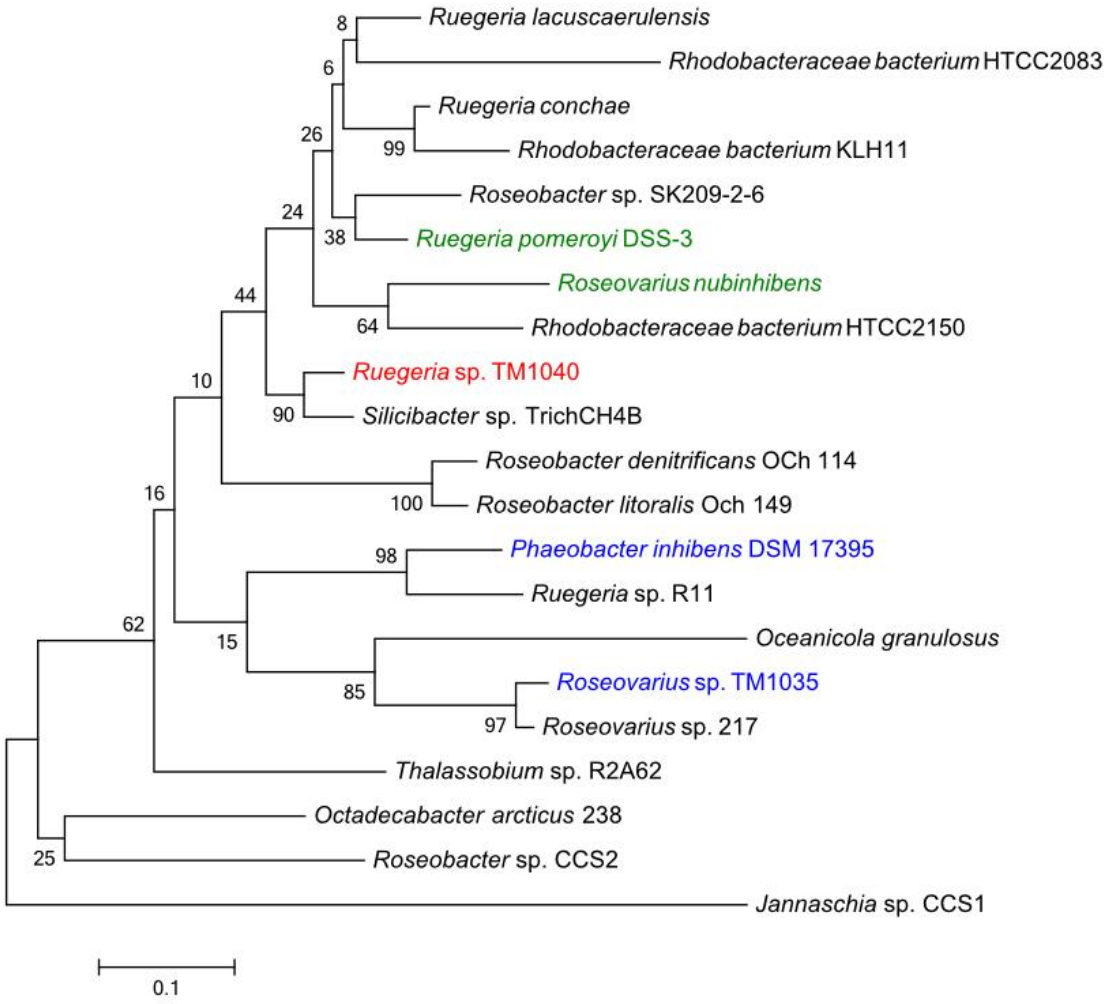


Figure 5.2 Relatedness of DddP sequences in Roseobacter strains. Protein sequences of DddP homologues were aligned with MEGA 6 and used to estimate an unrooted phylogenetic tree using LG model, gamma distributed with invariant sites. The scale bar indicates number of substitutions per site. Bootstrap values of 500 replicates are given at the base of each branch pair. Green text indicates the DddP of that species has been confirmed as functional, blue indicates that organism has been shown experimentally to make DMS from DMSP, and red indicates that this organism doesn't make DMS from DMSP.

5.2.4 Cupins: DddL, DddQ and DddW.

The *dddL*, *dddQ* and *dddW* genes encoding the cupin-type DMSP lyases are mostly restricted to the Roseobacter clade, although copies of DddQ, and the novel cupin-like lyase DddK, are also found amongst the SAR11 clade (see **Chapter 2**). Within the Roseobacters, homologues of DddL and DddQ were present in seven strains, and string matches to DddW were found in just two (**Tables 5.4** to **5.6** and **Figures 5.3** to **5.4**). However, poorly conserved homologues of DddW were present in four other species which will be discussed later.

DddL homologues fall into two distinct groups (**Figure 5.3**). One group contains homologues from *Sulfitobacter* sp. EE-36, *Sulfitobacter* NAS-14.1, *L. vestfoldensis* SKA-53 and *Roseobacter* GAI-101. Of these, the *dddL* from strain EE-36 has been cloned and shown to cleave DMSP when expressed in a heterologous host (Curson *et al*., 2008). Furthermore, a $DddL^-$ mutant strain of EE-36 lost its Ddd^+ phenotype. The remaining strains in this group have all been shown to have a Ddd^+ phenotype (González *et al*., 1999; Curson *et al*., 2008; Moran *et al*., 2012). Although no work has been done in these strains to confirm the functionality of their DddL peptides, they do not possess any other known DMSP lyases. The second phylogenetic group of DddL-like sequences contains *M. alkaliphilus*, *O. batsensis* and *D. shibae*, whose DddL sequences are more distantly related to those in Group I. Nevertheless, as mentioned above, DddL from *D. shibae* has been cloned and confirmed to have DMSP cleavage activity, despite the strain itself not possessing a Ddd^+ phenotype, so perhaps the Group II DddL enzymes are only expressed under specific conditions. However, another strain with this sub-type of DddL, *O. batsensis*, *has* been shown to produce DMS from DMSP under laboratory conditions (Curson *et al*., 2008), but the functionality of its DddL has not been investigated. Interestingly, with the exception of *Dinoroseobacter shibae*, strains that had a copy of *dddL* did not have any other DMSP genes. However, the significance of this is unknown.

Table 5.4 Homologues of DddL (EE36_11918) in the Roseobacter clade

Organism	Locus tag	Identity to EE36_11918	*E* value	Size (amino acids)
Dinoroseobacter shibae DFL 12	Dshi_3313	50%	8e-69	236
Loktanella vestfoldensis SKA53	SKA53_01756	74%	6e-119	244
Maritimibacter alkaliphilus HTCC2654	RB2654_07950	47%	3e-66	234
Oceanicola batsensis HTCC2597	OB2597_08014	49%	1e-67	235
Roseobacter sp. GAI101	RGAI101_3508	89%	3e-145	253
Sulfitobacter sp. EE-36	EE36_11918	100%	1e-165	223
Sulfitobacter NAS-14.1	NAS141_17149	99%	1e-164	223

BLASTp searches for DddL homologues used EE36_1191 as the query. Locus tags of convincing homologues (identities >47%, and *E*-values of <e^{-66}) are presented, along with percentage identity to EE36_1191, *E*-values and size of predicted polypeptide.

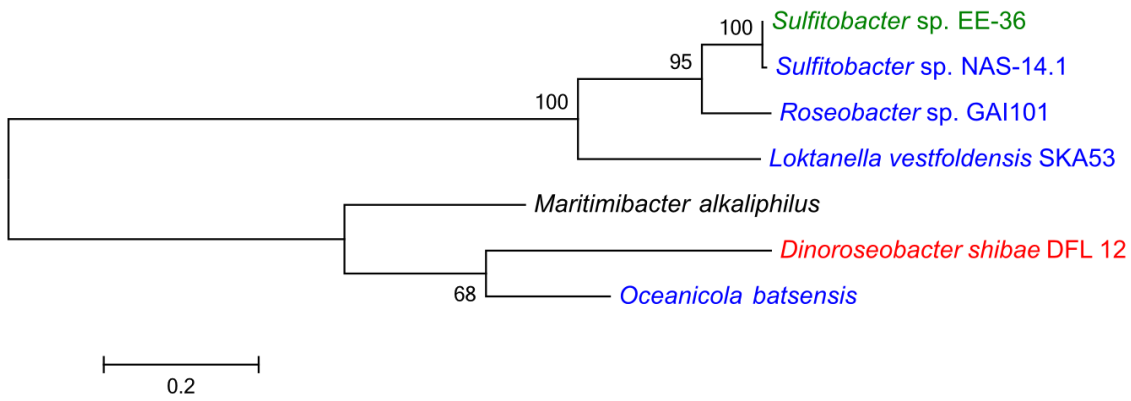


Figure 5.3 Relatedness of DddL sequences in Roseobacter strains. Protein sequences of DddL homologues were aligned with MEGA 6 and used to estimate an unrooted phylogenetic tree using LG model, gamma-distributed. The scale bar indicates number of substitutions per site. Bootstrap values of 500 replicates are given at the base of each branch pair. Green text indicates the DddL of that species has been confirmed as functional, blue indicates that organism has been shown experimentally to make DMS from DMSP, and red indicates that this organism doesn't make DMS from DMSP.

Table 5.5 Homologues of DddQ (ISM_14090) in the Roseobacter clade

Organism	Locus tag	Identity to ISM_14090	*E* value	Size (amino acids)
Rhodobacterales bacterium HTCC2150	RB2150_06543	44%	1e-50	203
Roseobacter sp. SK209-2-6	RSK20926_17292	37%	6e-37	197
Roseovarius nubinhibens ISM	ISM_14090	100%	2e-145	202
	ISM_14085	40%	8e-31	196
Ruegeria lacuscaerulensis ITI-1157	SL1157_0332	37%	3e-28	192
Ruegeria pomeroyi DSS-3	SPO1596	46%	6e-56	201
Ruegeria sp. TW15	N/A	39%	9e-36	200
Thalassiobium R2A62	TR2A62_3487	41%	2e-42	197

BLASTp searches for DddQ homologues used ISM_14090 as the query. Locus tags of convincing homologues (identities >37%, and *E*-values of $<e^{-28}$) are presented, along with percentage identity to ISM_14090, *E*-values and size of predicted polypeptide.

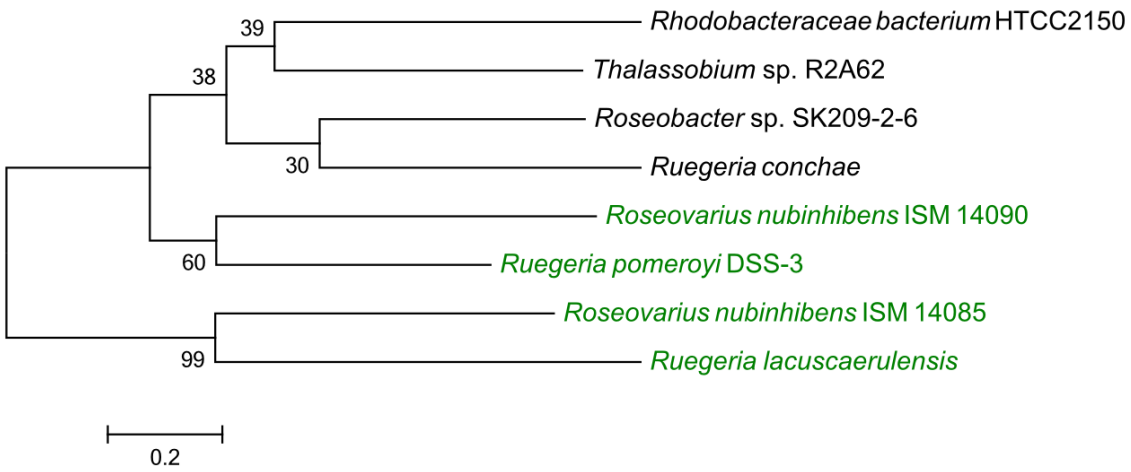


Figure 5.4 Relatedness of DddQ sequences in Roseobacter strains. Protein sequences of DddQ homologues were aligned with MEGA 6 and used to estimate an unrooted phylogenetic tree using LG model, gamma distributed with invariant sites. The scale bar indicates number of substitutions per site. Bootstrap values of 500 replicates are given at the base of each branch pair. Green text indicates the DddQ of that species has been confirmed as functional.

Compared to the other DMSP lyases, DddQ homologues have relatively low similarity to each other, but three *dddQ* genes, from strains *R. pomeroyi* and *R. nubinhibens* have been cloned and all have been shown to encode functional DMSP lyases (Todd *et al*., 2010b). Unusually, in *R. nubinhibens*, the two copies of *dddQ* are adjacent to each other. Initially it is tempting to attribute this to a gene duplication event, but the two genes have only 40% identity to one another, and so it may be that each was acquired separately.

Table 5.6 Homologues of DddW (SPO0453) in the Roseobacter clade

Organism	Locus tag	Identity to SPO0453	*E* value	Size (amino acids)
Ruegeria pomeroyi DSS-3	SPO0453	100%	1e-108	152
Roseobacter sp. MED193	MED193_09710	65%	2e-67	168
Rhodobacterales bacterium HTCC2083	RB2083_1887	54%	2e-37	148
Oceaniovalibus guishaninsula	OCGS_0874	53%	1e-34	139
Loktanella hongkongensis	Lokhon_01458	59%	3e-42	142
Citreicella sp. SE45	CSE45_0165	57%	9e-29	130

BLASTp searches for DddW homologues used SPO0453 as the query. Only one other close homologue of DddW was present, in *Roseobacter* sp. MED193. Poorly conserved homologues from four other Roseobacter strains are also presented.

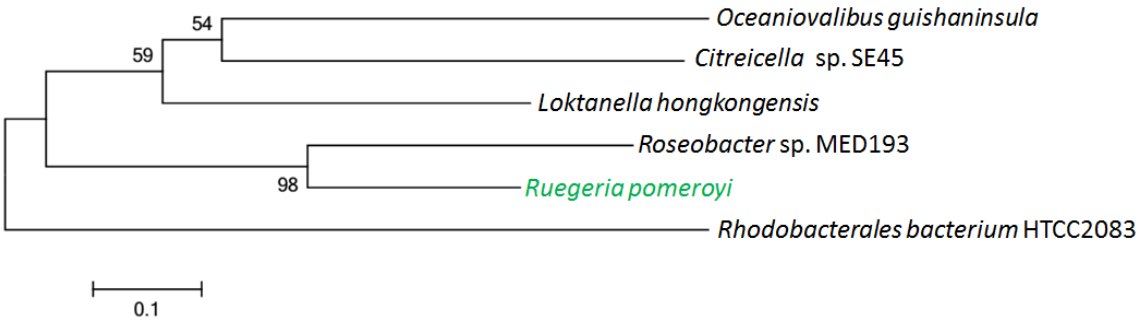


Figure 5.5 Relatedness of DddW sequences on Roseobacter strains. Protein sequences of DddW homologues were aligned with MEGA 6 and used to estimate an unrooted phylogenetic tree using WAG model, gamma distributed with invariant sites. The scale bar indicates number of substitutions per site. Bootstrap values of 500 replicates are given at the base of each branch pair. Green text indicates the DddW of *R. pomeroyi* has been confirmed as functional.

As mentioned above, only one other close homologue of *R. pomeroyi* DddW (SPO0453) is present in *Roseobacter* sp. MED193 (65% identity). However, there are also poorly conserved homologues in *Oceanivalibus guishaninsula*, *Citreicella* sp. SE45, *Loktanella hongkongensis* and *Rhodobacterales bacterium* HTCC2083 of between 53 and 59% identity to SPO0453 (**Table 5.6**). These peptides are vary slightly in length, from 130 to 148 amino acids, but are all shorter than SPO0453, and are more distantly related to SPO0453 than MED193_09710 (**Figure 5.5**). However, an alignment of DddW-like peptide sequences with SPO0453 revealed that some key residues in the cupin active site are conserved. For example, the two histidine residues indicated in **Figure 5.6** are completely conserved in all six peptides, and a glutamic acid and tyrosine residue are partly conserved. Other cupin DMSP lyases also have these four conserved residues, and in DddQ they were shown to form co-ordination bonds with a Zn^{2+} co-factor and a mutation in any of these residues resulted in dramatically reduced activity of DddQ (Li *et al*., 2014). Thus it would be interesting to see if the more poorly conserved DddW-like peptides do also function as DMSP lyases. Nevertheless, even counting the weaker homologues, DddW is among the rarest of the DMSP lyases in the deduced proteome of the Roseobacter clade, and indeed any metagenomic datasets.

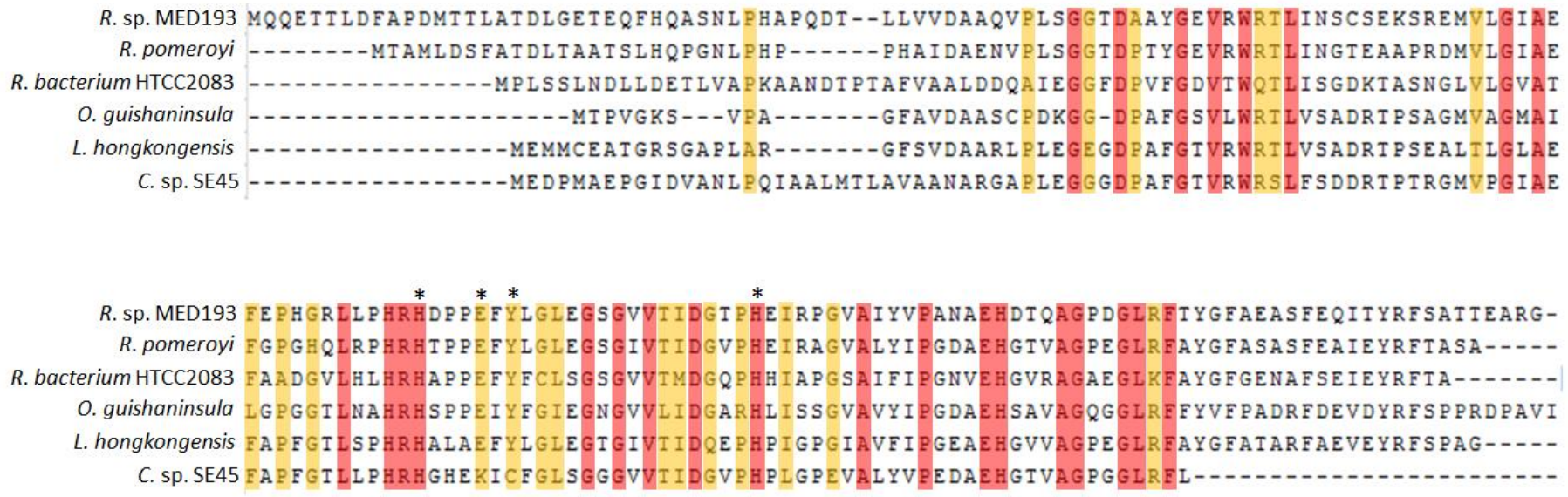


Figure 5.6 Alignment of DddW homologues in Roseobacter strains. Sequence alignment of DddW homologues. Completely conserved residues are highlighted in red, and highly conserved in yellow. Four residues shown to be key to DddQ cleavage of DMSP are indicated by asterisks (Li *et al*., 2014). Sequences are DddW-like polypeptides from the following: *Roseobacter* sp. MED-193 (MED193_09710); *Ruegeria pomeroyi* DSS-3 (SPO0453); *Rhodobacterales bacterium* HTCC2083 (RB2083_1887); *Oceanivalibus guishaninsula* (OCGS_0874); *Loktanella hongkongensis* (Lokhon_01558); and *Citreicella* sp. SE45 (CSE45_0165).

5.2.5 DddD

The gene SSE37_17628 from *Sagittula stellata* E-37 has been shown to encode a functional DddD enzyme, as determined through heterologous expression assays (Lei Sun, personal communication). Therefore, the peptide sequence of this enzyme was used to interrogate the NCBI Roseobacter group database for homologues. In total, five other strains had convincing homologues to the SSE37_17628 gene product, all with *E*-values of 0.0 (**Table 5.7**), and all had the CaiB-CaiB architecture (see **Chapter 1**). However, these were clearly divided into two distinct groups (termed "A" and "B"; **Figure 5.5**). Hits in *Citreicella* SE45 and *Rhodobacterales bacterium* HTCC2083 (in Group A) were very similar to SSE37_17628 (90% and 82% identical, respectively). Homologues in *D. shibae* DFL-12, *Rhodobacterales bacterium* KLH11 and *R. pomeroyi* DSS-3 (in Group B) only had identities of ~40%. DddD from *S. stellata*, which was shown to possess Ddd activity (Lei Sun, personal communication), is in Group A, but, in contrast, no DddD enzyme from Group B has been shown to be functional. *D. shibae*, as discussed above, does not have a Ddd^+ phenotype, and although *R. pomeroyi* can make DMS from DMSP via three different DMSP lyases, a mutation in *dddD* did not affect DMS production (Todd *et al*., 2010b). Thus it may be that only the DddD enzymes in Group A are functional, at least under laboratory conditions.

Table 5.7 Homologues of DddD (SSE37_17628) in the Roseobacter clade

Organism	Locus tag	Identity to SSE37_17628	*E* value	Size (amino acids)
Citreicella SE45	CSE45_4815	90%	0.0	836
Dinoroseobacter shibae DFL 12	Dshi_3632	41%	0.0	826
Rhodobacterales bacterium KLH11	RKLH11_3758	42%	0.0	824
Rhodobacterales bacterium HTCC2083	RB2083_930	82%	0.0	836
Ruegeria pomeroyi DSS-3	SPO1703	41%	0.0	826
Sagittula stellata E-37	SSE37_17628	100%	0.0	836

BLASTp searches for DddD homologues used SSE37_17628 as the query. Locus tags of convincing homologues (identities >41%, and *E*-values of 0.0) are presented, along with percentage identity to SSE37_17628, *E*-values and size of predicted polypeptide.

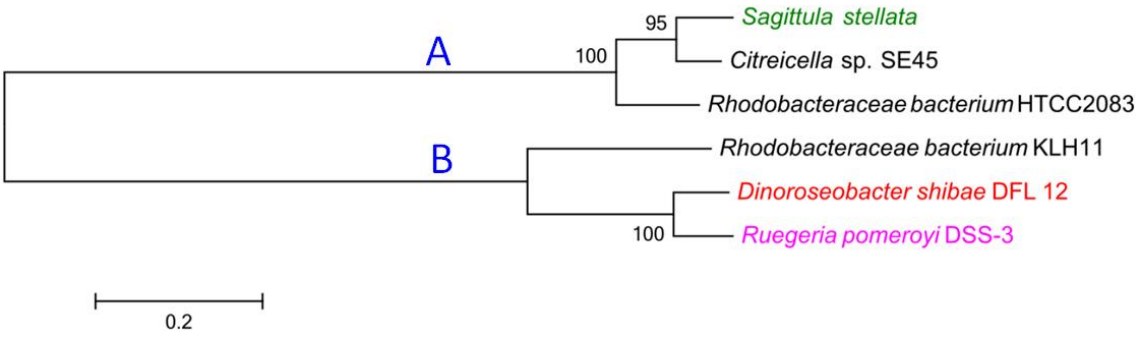


Figure 5.7 Relatedness of DddD sequences in Roseobacter strains. Protein sequences of DddD homologues were aligned with MEGA 6 and used to estimate an unrooted phylogenetic tree using LG model, gamma distributed with invariant sites. The scale bar indicates number of substitutions per site. Bootstrap values of 500 replicates are given at the base of each branch pair. Green text indicates the DddD of that species has been confirmed as functional, red indicates that this organism doesn't make DMS from DMSP, and purple indicates that *R. pomeroyi* does make DMS from DMSP, but a mutation in DddD did not reduce Ddd^+ activity.

5.2.6 Comparative Synteny of the *ddd* and *dmd* genes in different Roseobacters

As described above, genes encoding enzymes that act on DMSP are prevalent amongst the Roseobacters. However, no single gene is present in all of the strains, and conversely, there are many examples of very closely related strains that do not possess the same profile of "DMSP genes". Thus, almost certainly, these genes have been acquired by a series of HGT events, rather than vertically from a single common ancestor.

One purpose of exploring the arrangement of *dmdA* and the various *ddd* genes is to identify any re-occurring neighbouring genes. Since bacterial genomes are often arranged so that genes encoding proteins with related functions are clustered, the predicted functions of the genes near *dmdA* or *ddd* might be informative.

With this in mind, I examined the synteny of each of the primary DMSP genes found in the Roseobacter clade.

5.2.6.1 DmdA

Strikingly, the arrangement of genes surrounding *dmdA* is very similar in almost all cases (**Figure 5.8**). In all but one instance, *dmdA* is located upstream of *acuI*. As discussed in previous chapters, *acuI* encodes an acryloyl-CoA reductase, which converts acryloyl-CoA to propionyl-CoA. Therefore there is a clear functional link between *acuI* and DMSP cleavage to acrylate, and indeed, there are known cases where *acuI* abuts, and/or is co-transcribed with a DMSP lyase gene. For example, in *Rhodobacter sphaeroides*, *acuI* is in a three-gene operon with *dddL* and a regulator, *acuR* (Sullivan *et al*, 2011). In *Alcaligenes faecalis*, *acuI* is part of a cluster containing *dddY* and several other genes involved in DMSP catabolism (Curson *et al*, 2011). In fact, the Roseobacters, are an exception, in that their *acuI* is not near DMSP lyase genes, but instead is co-transcribed with the DMSP demethylase gene *dmdA*. Furthermore, it is known that the *dmdA-acuI* operon is up-regulated in the presence of DMSP and acrylate.

Another gene that is almost always present near *dmdA* is predicted to encode a GntR family transcriptional regulator. In *Ruegeria pomeroyi* and *Roseobacter* MED193, the divergently transcribed *gntR* is immediately upstream of *dmdA*. In other Roseobacters, *gntR* is separated from *dmdA* by a few other genes. This conserved close linkage suggested that *gntR* might be involved in the regulation of the *dmdA/acuI* genes Indeed, the *gntR* of *R. pomeroyi* (SPO1912) was shown to act as a transcriptional repressor of *acuI* expression, when cloned and expressed in the heterologous host *Rhizobium leguminosarum* containing an *acuI-lacZ* transcriptional fusion. This repression could not be relieved by DMSP, acrylate, MMPA, DMS or methanethiol, so although SPO1912 does negatively regulate the *dmdA* operon, the exact nature of this regulation is unclear (Mark Kirkwood, personal communication).

In most cases, there are two or three other genes between *gntR* and *dmdA*. Two of these encode "conserved hypothetical proteins" with domains of unknown function (DUF1326 and DUF2182). The "DUF1326" and "DUF2182" genes are almost exclusively found as a pair, not just in Roseobacter strains, but also in other bacteria, such as *Mesorhizobium* spp. Interestingly, in *Mesorhizobium* they are upstream of a gene encoding a Zn-dependent alcohol dehydrogenase of the MDR superfamily, but the protein product has no particular similarity to AcuI.

The third gene, found immediately upstream of *dmdA* encodes a putative DinB_2 superfamily (DNA-**d**amage **in**ducible) protein. DinB_2 superfamily proteins have very diverse sequences, but are united in their structures, and the presence of a conserved histidine triad motif, which may indicate metal-binding properties (Cooper *et al*., 2010). This histidine triad is conserved in the *dinB*-like gene near *dmdA* in each strain. The *dinB* genes of *Bacillus subtilis* and *E. coli* are induced in response to DNA damage by environmental stressors, and are under control of the SOS-repair system (Cheo *et al*., 1991; Wagner *et al*., 1999). The *dinB* gene of *E. coli* encodes

DNA polymerase IV. This polymerase plays a role in spontaneous mutagenesis, since it lacks proofreading ability, and has propensity to elongate misaligned templates, sometimes resulting in frameshift mutations (Friedberg *et al*., 2000). Other proteins of the DinB_2 superfamily have a different function, including YfiT from *Bacillus subtilis*, which is a bacillithiol *S*-transferase, and EF_3021, a glutathione *S*-transferase from *Enterococcus faecalis* (Newton *et al*., 2011). However, many DinB_2 superfamily proteins, including those in the Roseobacters, have no known function, and the significance of a *dinB*-like gene positioned immediately upstream of *dmdA* in many Roseobacters is not known.

In *Thalassobium* R2A62 and *R. bacterium* HTCC2150, there is a gene encoding a putative BCCT-type betaine transporter downstream of *acuI*, which are ~32% identical to DddT of *Halomonas* HTNK1, a confirmed functional DMSP transporter (Sun *et al*., 2012). It would be surprising if these were *not* involved in DMSP import.

As previously mentioned, in all but one strain, *dmdA* is always next to *acuI*, *gntR* and a sometimes a few other genes including *dinB*. But, uniquely, this is not the case for *Phaeobacter gallaeciensis* DSM 17395, in which, *dmdA* is the first gene in a four gene cluster containing the other genes known to encode DMSP demethylation enzymes: *dmdB2*, *dmdD* and *dmdC*. So in this instance all four genes encoding the complete demethylation pathway are contiguous, possibly in one operon. Although the demethylation pathway has been known since 2011 this is the first time this arrangement has been noted in any organism. In most other Roseobacter strains, *dmdC* and *dmdD* are found as an adjacent pair, but they are unlinked to *dmdA* and *dmdB*. In some Roseobacter strains, for which we have partial genome sequences, *dmdA* is next to *dmdB*, (e.g. in some strains of *Labrenzia*), but this was not the case for any of the Roseobacter strains investigated here. This provides exciting insights into the evolution of DMSP catabolism, and bacterial operons in general.

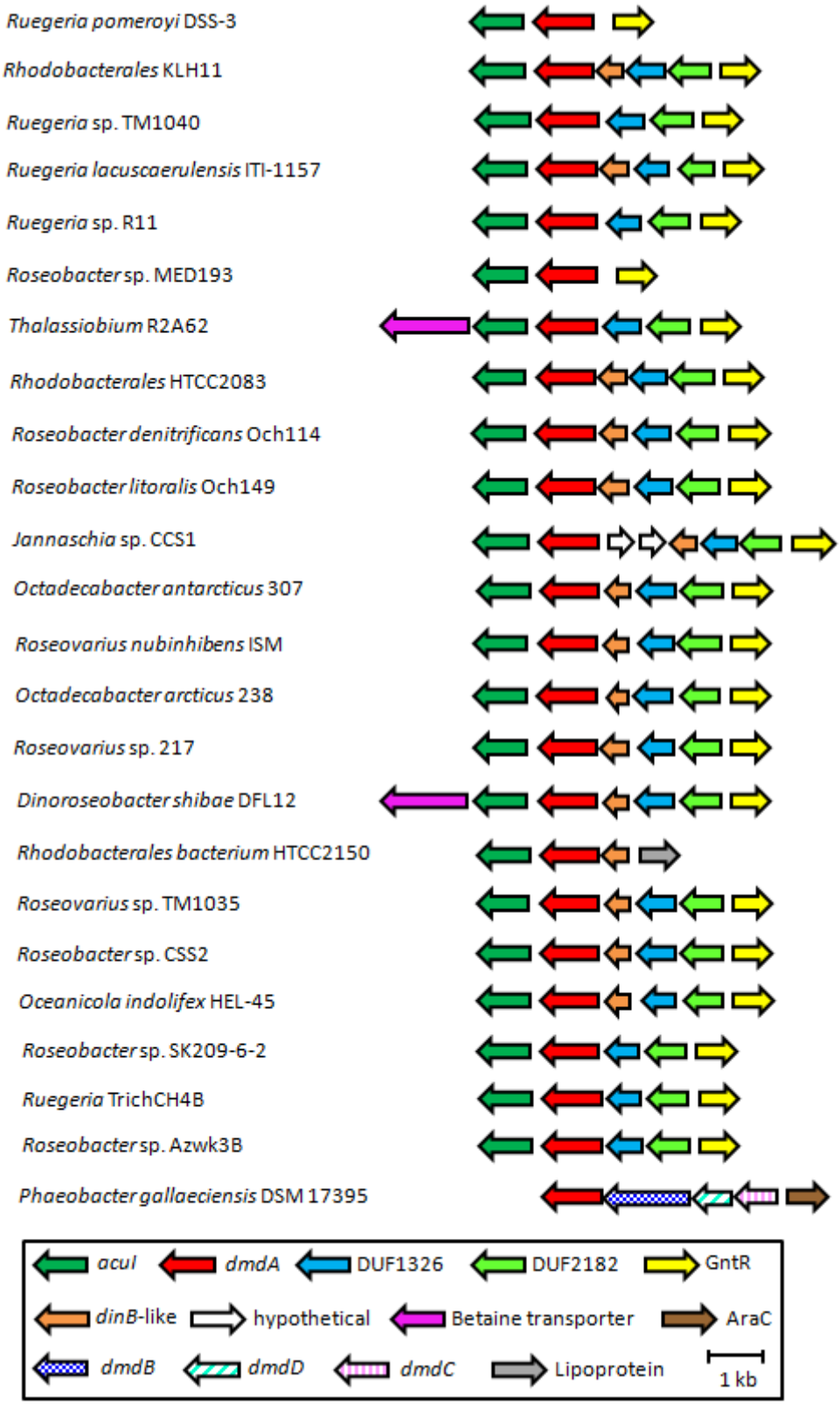


Figure 5.8 Genes near *dmdA* in different Roseobacter species. In almost all Roseobacter strains, *dmdA* is positioned upstream of *acuI*, which encodes an acryloyl-CoA reductase. Upstream of *dmdA* is often a *dinB*-like gene, whose product has no known function, two hypothetical proteins with conserved domains of unknown function (DUF1326 and DUF2182)

and a gene encoding a GntR-type transcriptional regulator. In two species, *D. shibae* and *Thalassobium* R2A62, there is a gene encoding a putative betaine transporter downstream of *acuI*. In *P. gallaeciensis*, *dmdA* is found in a cluster with other genes encoding the demethylation pathway: *dmdB*, which encodes the MMPA-CoA ligase DmdB; *dmdC* which encodes the MMPA-CoA dehydrogenase, DmdC and *dmdD* which encodes the MPA-CoA hydratase DmdD. This cluster is divergently transcribed from a gene whose predicted product is an AraC-type transcriptional regulator.

5.2.6.2 DddP

In most cases *dddP* is in a single gene operon (**Figure 5.9**), but it is close to the same, corresponding genes in several different strains. Thus, *dddP* is often transcribed divergently from a gene encoding the conserved hypothetical protein DUF3445. However, this gene is not restricted to strains containing *dddP*, and it has no predicted role in DMSP catabolism.

Another gene that is almost always found just downstream of *dddP* encodes a protein with a hydrolase superfamily domain and an osmotically inducible protein C (OsmC) domain. The OsmC superfamily contains proteins involved in defence against oxidative stress (Lesniak *et al*., 2003).

The other genes found nearby to *dddP* are of different types, but a few occur in more than one strain. These include a gene encoding a NUDIX hydrolase, which is a family of enzymes that hydrolyse a wide range of pyrophosphates (McLennan, 2006), and another gene encoding an endonuclease/exonuclease/phosphatase (EEP) domain protein, a diverse set of proteins that share the common catalytic feature of cleaving a phosphodiester bond (Dlakić, 2000).

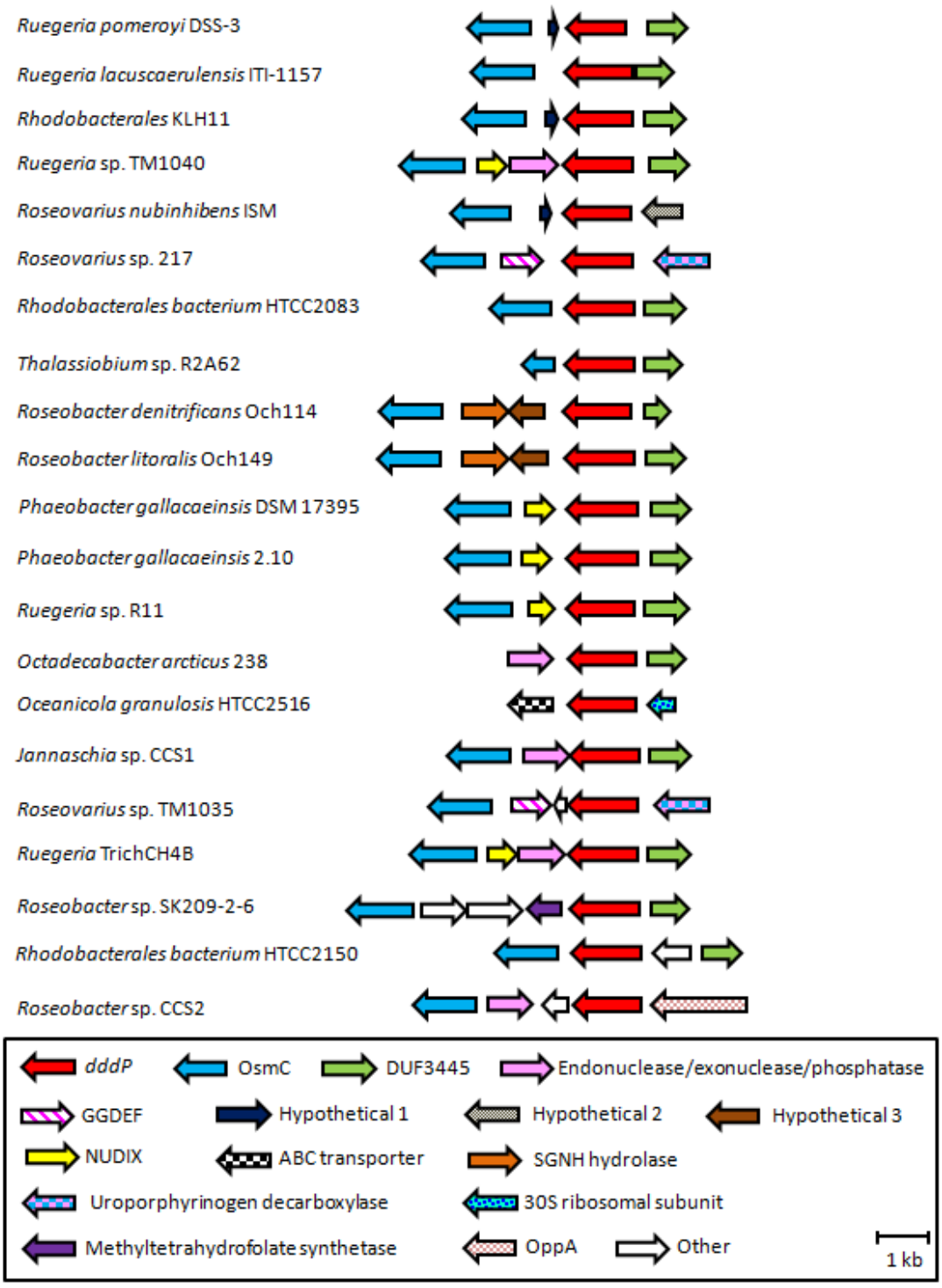


Figure 5.9 Genes near *dddP* in different Roseobacter species. The *dddP* genes of the Roseobacter species are usually found nearby a gene encoding an OsmC-like protein, and another whose product is a hypothetical protein with a conserved domain of unknown function (DUF3445). These three genes are often separated by an eclectic mix of different genes. The gene names, or predicted gene products, are indicated in the key: GGDEF – proteins of this family have diguanylate cyclase activity; SGNH hydrolase – a diverse family of lipases and esterases; OppA – a component of the ABC transport system. White arrows indicate genes

encoding various different proteins, which have no obvious link to DMSP catabolism, and do not appear near *dddP* in more than one species.

5.2.6.3 DddL

In most bacteria, the *dddL* gene is also found as a single gene transcriptional unit, and certainly within the Roseobacter clade, *dddL* is only found as a single gene, and the genes to either side of *dddL* have no known connection to DMSP catabolism (**Figure 5.10**. This is in contrast to another Rhodobacterales species, *Rhodobacter sphaeroides* 2.4.1, in which *dddL* is part of an operon with *acuR*, which encodes an acrylate-responsive repressor of the *dddL* operon, and the acryloyl-CoA reductase encoding *acuI* (see **Chapter 1**). In four of the *dddL*-containing Roseobacters, namely *Sulfitobacter* sp. EE-36, *Sulfitobacter* sp. NAS141, *Loktanellla vestfoldensis* and *Roseobacter* sp. GAI101, *dddL* is upstream of a gene encoding DNA topiosomerase, and in three of these cases the gene downstream of *dddL* encodes a protein from the alpha/beta hydrolase superfamily.

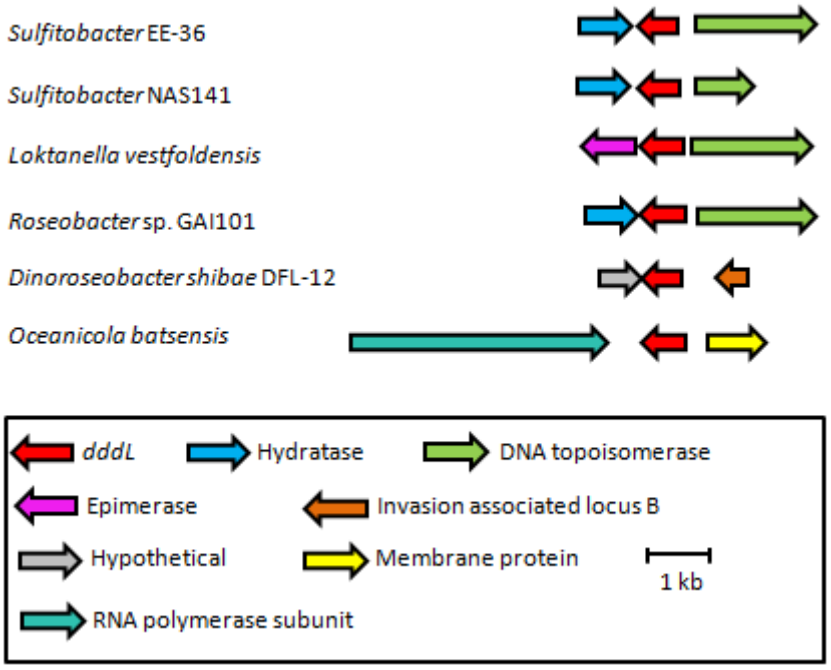


Figure 5.10 Genes near *dddL* in different Roseobacter species. In four cases, the *dddL* gene is divergently transcribed from a gene encoding a DNA topisomerase. In three of these strains, the other adjacent gene encodes a putative hydratase, and in the remaining strain (*L. vestfoldensis*) *dddP* is next to a gene whose predicted product is an epimerase. In the remaining two species, *dddL* is found near genes encoding a hypothetical protein, and a homologue of the virulence

factor, invasion-associated locus B (in the case of *D. shibae*), and genes whose predicted products are an RNA polymerase subunit and a membrane protein (in the case of *O. batsensis*).

5.2.6.4 DddQ

In the three strains that have been shown to possess a functional DddQ enzyme, namely *R. pomeroyi*, *R. nubinhibens* and *R. lacuscaerulensis*, the genes surrounding *dddQ* are conserved (**Figure 5.11**). However, there is no clear link between the products of these genes and DMSP catabolism. For example, in all three cases, *dddQ* is adjacent to a gene predicted to encode a Zn-dependent alcohol dehydrogenase, in the same super-family as, but distinct from, the AcuI, acryloyl-CoA reductase. Nearby are also genes encoding proteins annotated as being related to mandelate racemase or muconate-lactonizing enzyme. Thus, in *R. pomeroyi* and *R. lacuscaerulenesis*, there are two adjacent genes encoding distinct proteins belonging to the mandelate racemase family. However, in *R. nubinhibens*, there are four linked genes whose products all belong to this family. Interestingly, as mentioned above, *R. nubinhibens* also has multiple copies of *dddQ* in this region. Given the relatively low sequence identity of the two DddQ proteins, it is unlikely that this multiplicity arose during a gene duplication event (**Figure 5.11**). Similarly, the four mandelate racemase genes are also relatively diverse in sequence, and do not appear to be a simple duplication of the two homologues found in *R. pomeroyi* and *R. lacuscaerulensis* (see **Figures 5.12** and **5.13**). In *Pseudomonas putida*, mandelate racemase and muconate-lactonizing enzyme catalyse separate reactions necessary for aromatic acid catabolism (Neidhart *et al*., 1990), but it is not known if the homologues in the Roseobacters also have this function, or if they have any role in DMSP breakdown.

Another gene, encoding a putative glutamate semialdehyde aminomutase is found further downstream, again with no known connection to DMSP.

Upstream of *dddQ* in *R. lacuscaerulensis* and in *Thalassiobium* R2A62 is a gene encoding a putative glycine cleavage T family protein. Although in the same general family as the *dmdA* gene product, the product of this *dddQ*-linked gene is quite distinct from the DmdA DMSP demethylase, with the limited similarity restricted to the C-terminus.

In *Thalassiobium* R2A62, *Rhodobacterales bacterium* HTCC2150 and *Roseobacter* sp. SK209-2-6, there are no conserved genes of interest, just a sporadic collection of genes predicted to encode proteins with no obvious connection. However, it may be worth noting that a gene encoding a predicted Rieske-type protein is found just upstream of *dddQ* in *Roseobacter* sp. SK209-2-6. As discussed briefly in **Chapter 3**, Rieske-type genes (*hcaE*) are also found nearby

dddP1, *dddP2* and *dddD* of *O. doudoroffii*, but there is no known connection between these proteins and DMSP catabolism. Furthermore, the sequence similarity between the Rieske of *Roseobacter* sp. SK209-2-6 and those of *O. doudoroffii* is only 28-32%, and there is no indication that these form a specific DMSP-connected subgroup.

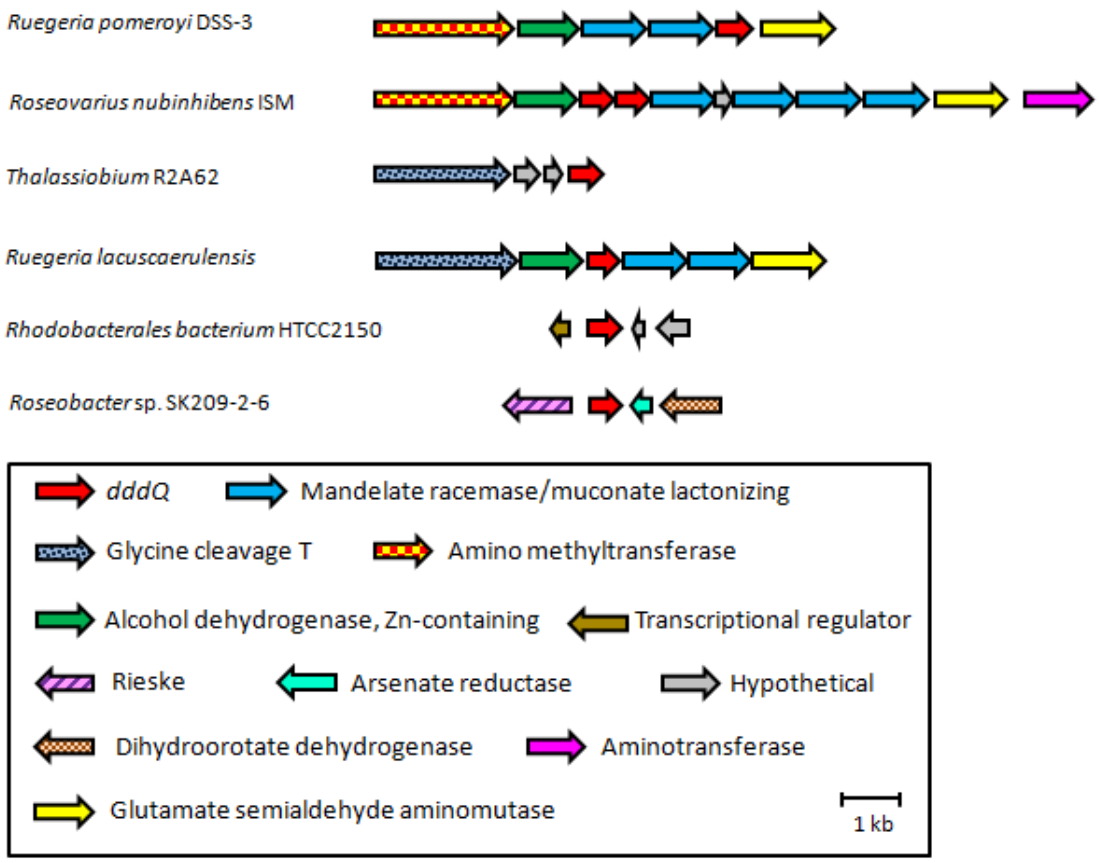


Figure 5.11 Genes near *dddQ* in different Roseobacter species. Gene maps showing *dddQ* region of different Roseobacter genomes. The gene names, or predicted gene products, are indicated in the key.

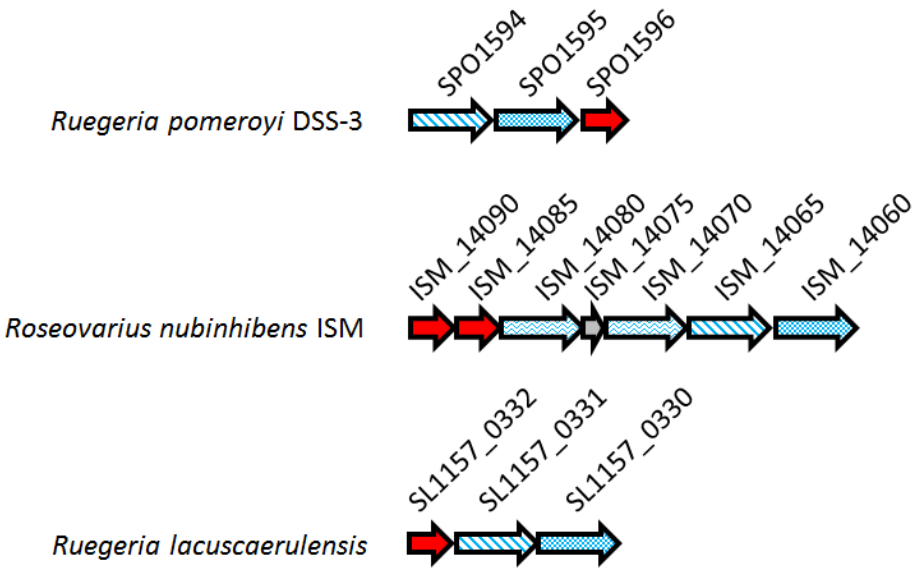


Figure 5.12 Gene clusters encoding mandelate racemase-like proteins in *dddQ* strains. Several genes encoding mandelate racemase-like proteins are found near *dddQ* in three Roseobacter species. These genes are represented by blue arrows. Arrows with the same pattern indicate the gene products are homologous. Thus, ISM_14080 and ISM_14070 are 35% identical, SPO1594, ISM_14065 and SL1157_0331 are 50-88% identical and SPO1595, ISM_14060 and SL1157_0330 76-80%. Red arrows represent *dddQ* genes.

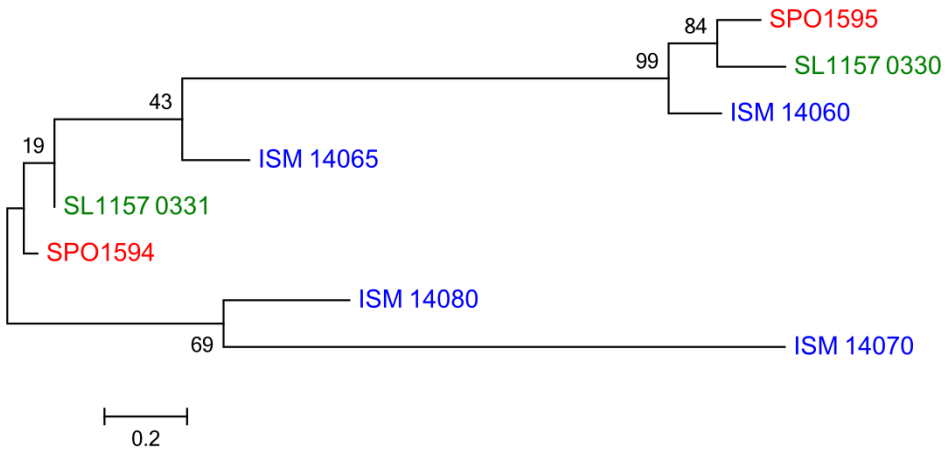


Figure 5.13 Relatedness of mandelate racemase-like proteins encoded near *dddQ*. Phylogenetic relationship of putative mandelate racemase-like proteins encoded by genes located near *dddQ* in *Ruegeria pomeroyi*, *Ruegeria lacuscaerulensis*, and *Roseovarius nubinhibens*. The proteins are annotated using the locus tag of the corresponding gene, and the prefix represents

the strain as follows: SPO – *R. pomeroyi* (red text); SL1157 – *R. lacuscaerulensis* (green text); ISM – *R. nubinhibens* (blue text). The tree was estimated using an LG + G model in MEGA6. The scale bar indicates number of substitutions per site. Bootstrap values of 500 replicates are given at the base of each branch pair.

5.2.6.5 DddW

As mentioned, a close homologue of the *R. pomeroyi* DddW has been found in one other species - *Roseobacter* sp. MED193. In both of these strains, the arrangement of genes either side of *dddW* is identical (**Figure 5.14**). The downstream gene is predicted to encode a conserved hypothetical protein, in the family of tellurium-resistance (TerB) proteins. However, the function of this family of proteins is not known, and there is no apparent connection to DMSP metabolism.

In a microarray analysis of *R. pomeroyi*, *dddW* was up-regulated (~37-fold) in the presence of DMSP (Todd *et al*., 2012a). Upstream of *dddW* is a LysR family transcriptional regulator (SPO0424), whose expression is up-regulated in the presence of DMSP (~2.6-fold), and given its proximity to *dddW*, was a candidate regulator of *dddW* expression. To show that SPO0424 is a regulator of *dddW*, Todd *et al*. made a *dddW* promoter fusion to *lacZ* in pBIO1878 and crossed this into *Rhizobium leguminosarum* 3841. Using β-galactosidase assays to measure level of *lacZ* expression, they showed that *dddW* is expressed at a constitutively low level in the heterologous *Rhizobium* host. However, when a plasmid containing the intact SPO0424 gene was also introduced into *Rhizobium*, the *dddW-lacZ* fusion was induced ~5-fold when grown in the presence of DMSP (Todd *et al*., 2012a). Thus, SPO0424 acts as a transcriptional activator of *dddW* and responds to DMSP. Additionally, Todd *et al*. showed that SPO0424 operates as a negative auto-regulator, which is typical of the LysR-type family (Maddocks and Oyston, 2008). This was confirmed by introducing a SPO0424-*lacZ* fusion into *R. leguminosarum*, which expressed β-galactosidase constitutively, but was repressed ~5-fold when an intact SPO0424 gene was also present (Todd *et al*., 2012a).

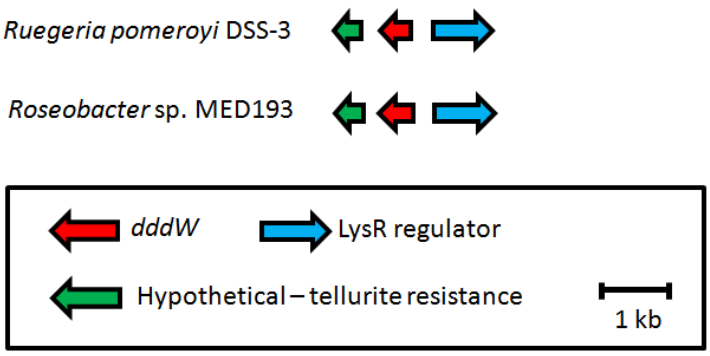


Figure 5.14 Genes near *dddW* in different Roseobacter species. The *dddW* genes of *R. pomeroyi* and *Roseobacter* sp. MED193 are found in the same genomic location, upstream of a gene encoding a conserved hypothetical protein, with a tellurite resistance domain, and divergently transcribed from a gene encoding a LysR-type regulator.

5.2.6.6 DddD

As discussed earlier, DddD enzymes of the Roseobacter group fall into two distinct clades - those of *Rhodobacterales* HTCC2083, *Citreicella* sp. SE45 and *S. stellata*, and those of *R. pomeroyi*, *D. shibae* and *Rhodobacterales* KLH11. Based on evidence so far, it seems possible that the former group of DddD enzymes do have Ddd^+ activity, but that the latter do not. Strikingly, the arrangement of genes surrounding DddD also differs between the two types (**Figure 5.15**). Group B (the non-functional group) *dddD* genes are immediately upstream of a gene predicted to encode a hypothetical protein with a glutamine amidotransferase domain. In the case of *R. pomeroyi* and *D. shibae*, *dddD* is divergently transcribed from a gene encoding an AraC-type transcriptional regulator. However, in Group A, from which at least one DddD has been shown to be functional, the genes are apparently co-transcribed with a gene encoding a close homologue of DddA. To recap, DddA is an alcohol dehydrogenase, which in *Halomonas* HTNK1 acts to convert DMSP-derived 3HP to malonate semialdehyde (Todd *et al*., 2010a). In the case of several γ-proteobacteria, *dddD* is part of an extended operon containing *dddA*, and also *dddC* and *dddT*. However, the Roseobacter strains do not have these extra DMSP catabolism genes, and it may be that that they have acquired only part of the full operon through horizontal gene transfer.

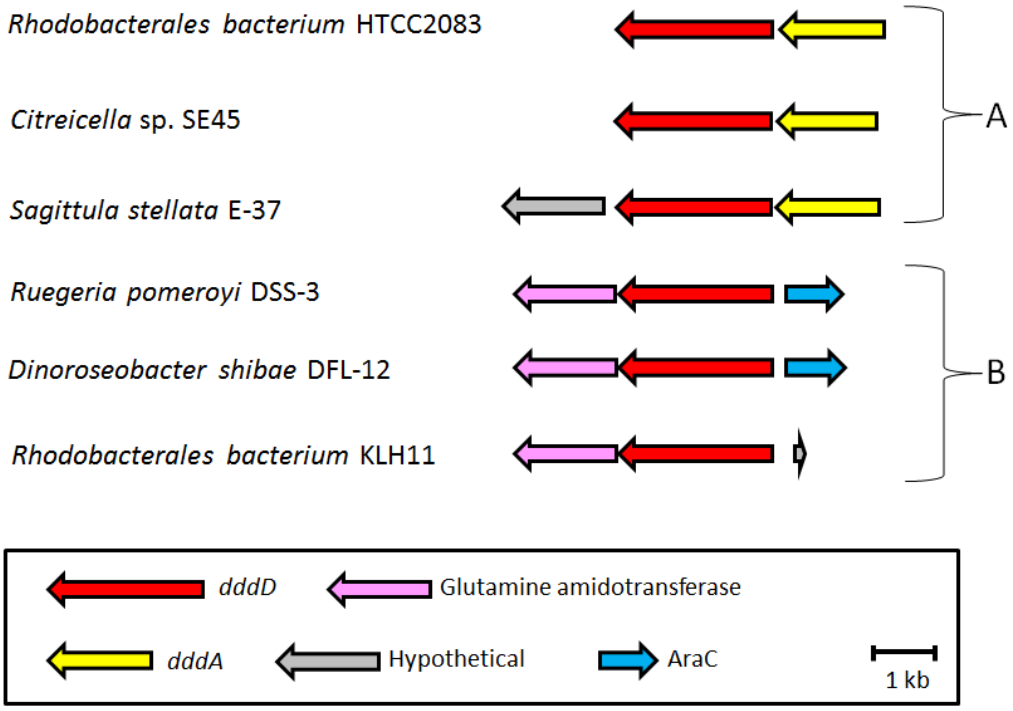


Figure 5.15 Genes near *dddD* in different Roseobacter species. Roseobacter *dddD* genes fall into two distinct Groups – A and B. Group A genes are positioned downstream of *dddA*, which encodes an alcohol dehydrogenase responsible for the degradation of malonate semialdehyde in the DMSP catabolic pathway in *Halomonas* sp. HTNK1 (Todd *et al*., 2010a). Group B *dddD* genes are transcribed upstream of a gene whose product is a putative glutamine amidotransferase. In the case of *R. pomeroyi* and *D. shibae*, *dddD* is divergently transcribed from a gene encoding an AraC-family transcriptional regulator.

5.2.7 Genes involved in the downstream catabolism of DMSP

5.2.7.1 The demethylation route

The demethylation route of DMSP catabolism is described in detail in **Chapters 1** and **4**. Briefly, following the conversion of DMSP to MMPA by DmdA, the MMPA is degraded via a series of CoA intermediates to acetaldehyde, CO_2 and methanethiol. The enzymes which catalyze this pathway are DmdB, DmdC and DmdD. In *R. pomeroyi* there are two forms of the MMPA-CoA ligase, DmdB (DmdB1 - SPO2045 and DmdB2 - SPO0677), and also three copies of the MMPA-CoA dehydrogenase DmdC (DmdC1 - SPO2804, DmdC2 - SPO0298 and DmdC3 - SPO2915). All of the Dmd enzymes from *R. pomeroyi* have been shown to have activity when cloned and expressed in *E. coli* (Reisch *et al*., 2011). To investigate the occurrence of these

enzymes amongst the Roseobacter clade, each Roseobacter species genome sequence was interrogated in a BLASTp search using the *R. pomeroyi* Dmd sequences as the queries.

All of the query sequences returned convincing homologues with sequence identities ranging from a minimum of 47% in the case of DmdB1 to a maximum of 86% as seen for DmdB2. However, the number of homologues for each enzyme varied considerably, with DmdC2 being present in all strains and DmdC1 only found in four cases. Similarly, almost all strains had a copy of DmdB2 whereas only 23 had DmdB1, and good homologues to DmdD were found in only eight of the 41 strains.

When these results are compared to the presence of DmdA amongst the Roseobacters, it is clear that while all DmdA-containing strains also have DmdB2 and DmdC2, many lack DmdD. However, bacteria that lack DmdD do not necessarily lack DmdD activity as shown for *Ruegeria lacuscaerulensis* (Reisch *et al*., 2011), suggesting there is a non-orthologous isofunctional enzyme in that organism.

Interestingly, where DmdC1 is present, the *dmdC1* gene is always in a two-gene operon with *dmdD*. Where DmdD is present without DmdC1, *dmdD* is either in a single gene unit (as in both the *Sulfitobacter* strains) or immediately adjacent to genes encoding proteins with no known function in DMSP catabolism. As mentioned earlier, the *P. gallaeciensis* DSM17395 *dmdA, dmdB2, dmdC* and *dmdD* genes are contiguous, the only known example of this. In addition, this strain has a second copy of both DmdC and DmdD, which are also encoded by adjacent genes, unlinked to the 4-gene *dmdABCD* cluster.

Table 5.8 Homologues of DMSP demethylation pathway enzymes in the Roseobacter clade (continued overleaf)

Organism	DmdB1	DmdB2	DmdC1	DmdC2	DmdC3	DmdD
Citreicella sp. SE45		CSE45_0327 57% 0.0		CSE45_3649 74% 0.0	CSE45_1290 71% 0.0	
Citreicella sp. 357	C357_19601 66% 0.0	C357_02029 56% 0.0		C357_14172 71% 0.0	C357_03630 63% 0.0	
Dinoroseobacter shibae DFL12		Dshi_0833 56% 0.0		Dshi_3085 70% 0.0	Dshi_0839 72% 0.0	
Jannaschia sp. CCS1		Jann_2979 56% 0.0		Jann_3894 64% 0.0	Jann_2970 81% 0.0	
Loktanella hongkongensis DSM17492				Lokhon_02077 66% 0.0	Lokhon_01092 72% 0.0	
Loktanella vestfoldensis SKA53		SKA53_08001 60% 0.0 SKA53_09349 53% 0.0		SKA53_14231 70% 0.0	SKA53_09324 78% 0.0	
Maritimibacter alkaliphilus HTCC2654	RB2654_10309 48% 6e-175	RB2654_21033 53% 0.0		RB2654_12329 72% 0.0	RB2654_20558 74% 0.0	
Oceanibulbus indolifex HEL45	OIHEL45_12740 71% 0.0	OIHEL45_09303 53% 0.0		OIHEL45_08725 73% 0.0	OIHEL45_09348 85% 0.0	
Oceanicola batsensis HTCC2597	OB2597_20456 53% 0.0 OB2597_11066 49% 2e-170	OB2597_06305 54% 0.0		OB2597_09869 70% 0.0	OB2597_06245 76% 0.0	
Oceanicola granulosus HTCC2516				OG2516_16671 68% 0.0	OG2516_13586 67% 0.0	
Oceanicola sp. S124	N/A 49% 2e-174			N/A 69% 0.0	N/A 74% 0.0	
Oceaniovalibus guishaninsula JLT2003	OCGS_2656 61% 0.0	OCGS_1155 52% 1e-175		OCGS_2069 69% 0.0	OCGS_1159 69% 0.0	

Organism	DmdB1	DmdB2	DmdC1	DmdC2	DmdC3	DmdD
Octadecabacter arcticus 238		OA238_c11160 58% 0.0 OA238_c13380 59% 0.0 OA238_c33110 56% 0.0		OA238_c37680 72% 0.0	OA238_c33030 70% 0.0	
Octadecabacter antarcticus 307		OAN307_c36980 53% 0.0		OAN307_c01790 71% 0.0	OAN307_c36890 70% 0.0	
Pelagibaca bermudensis HTCC2601		R2601_18288 57% 0.0		R2601_13860 75% 0.0	R2601_25771 73% 0.0	
Phaeobacter gallaeciensis DSM17395	PGA1_c35550 61% 0.0	PGA1_262p01840 51% 2e180 PGA1_c11990 54% 0.0	PGA1_262p01860 69% 0.0 PGA1_262p01550 69% 0.0	PGA1_c03150 77% 0.0	PGA1_c12080 89% 0.0	PGA1_262p01850 84% 5e-170 PGA1_262p01540 83% 3e-170
Phaeobacter gallaeciensis 2.10	PGA2_c33750 61% 0.0	PGA2_c11990 55% 0.0	PGA2_239p1730 69% 0.0	PGA2_c02730 77% 0.0	PGA2_c12070 89% 0.0	PGA2_239p1720 83% 3e-170
Rhodobacterales bacterium KLH11	RKLH11_2348 78% 0.0	RKLH11_325 85% 0.0		RKLH11_1667 80% 0.0	RKLH11_236 91% 0.0	
Rhodobacterales bacterium HTCC2083	RB2083_2298 69% 0.0	RB2083_4048 53% 0.0		RB2083_1403 73% 0.0	RB2083_3917 86% 0.0	RB2083_3634 75% 4e-145
Rhodobacterales bacterium HTCC2150		RB2150_15146 50% 5e-176		RB2150_07608 67% 0.0	RB2150_15181 76% 0.0	
Roseobacter denitrificans Och114		RD1_3974 54% 0.0		RD1_3417 74% 0.0	RD1_3969 87% 0.0	
Roseobacter litoralis Och149		RLO149_c004620 52% 0.0		RLO149_c027750 72% 0.0	RLO149_c004690 85% 0.0	
Roseobacter sp. AzwK-3b	RAZWK3B_10116 46% 1e-158	RAZWK3B_02755 60% 0.0 RAZWK3B_07664 55% 0.0		RAZWK3B_04470 72% 0.0	RAZWK3B_07694 72% 0.0	

Organism	DmdB1	DmdB2	DmdC1	DmdC2	DmdC3	DmdD
Roseobacter sp. CCS2		RCCS2_03809 51% 0.0		RCCS2_12239 72% 0.0	RCCS2_03779 79% 0.0	
Roseobacter sp. GAI101	RGAI101_1869 67% 0.0	RGAI101_340 58% 0.0 RGAI101_3142 54% 0.0		RGAI101_153 75% 0.0	RGAI101_1412 83% 0.0	RGAI101_2703 78% 4e-153
Roseobacter sp. MED193	MED193_13597 60% 0.0	MED193_06009 53% 0.0	MED193_17339 60% 0.0	MED193_10818 77% 0.0	MED193_05939 86% 0.0	MED193_17334 80% 3e-155
Roseobacter sp. SK209-2-6	RSK20926_0882 2 60% 0.0	RSK20926_17192 55% 0.0		RSK20926_04367 77% 0.0	RSK20926_17252 85% 0.0	
Roseovarius nubinhibens ISM		ISM_02075 53% 0.0		ISM_09776 71% 0.0	ISM_02040 80% 0.0	
Roseovarius sp. TM1035	RTM1035_15747 47% 8e-160	RTM1035_16647 61% 0.0 RTM1035_0530 55% 0.0		RTM1035_11850 74% 0.0	RTM1035_16902 82% 0.0	
Roseovarius sp. 217	ROS217_05134 47% 4e-160	ROS217_05929 60% 0.0 ROS217_23097 54% 0.0		ROS217_22022 74% 0.0	ROS217_11341 82% 0.0	
Ruegeria lacuscaerulensis ITI-1157	SL1157_2728 82% 0.0	SL1157_1815 86% 0.0		SL1157_0694 81% 0.0	SL1157_2180 83% 0.0	
Ruegeria pomeroyi DSS-3	SPO2045 100% 0.0	SPO0677 100% 0.0	SPO3804 100% 0.0	SPO0298 100% 0.0	SPO2915 100% 0.0	SPO3805 100% 0.0
Ruegeria sp. TM1040		TM1040_1170 57% 0.0 TM1040_1565 54% 0.0		TM1040_3059 78% 0.0	TM1040_1557 87% 0.0	
Ruegeria sp. Trich CH4B		SCH4B_4119 45% 2e-140		SCH4B_0372 79% 0.0	SCH4B_2076 87% 0.0	

Organism	DmdB1	DmdB2	DmdC1	DmdC2	DmdC3	DmdD
Ruegeria sp. TW15	N/A 79% 0.0	N/A 85% 0.0		N/A 81% 0.0	N/A 91% 0.0	
Ruegeria sp. R11	RR11_1578 62% 0.0	RR11_964 55% 0.0		RR11_2806 77% 0.0	RR11_2398 88% 0.0	
Sagittula stellata E-37	SSE37_11144 49% 8e-173	SSE37_21745 56% 0.0		SSE37_19952 71% 0.0	SSE37_10113 77% 0.0	
Sulfitobacter sp. EE36	EE36_15767 68% 0.0 EE36_00485 63% 0.0	EE36_03673 55% 0.0		EE36_04173 75% 0.0	EE36_03638 85% 0.0	EE36_13798 80% 3e-156
Sulfitobacter NAS-14.1	NAS141_05893 68% 0.0	NAS141_08601 54% 0.0		NAS141_09101 75% 0.0	NAS141_08566 85% 0.0	NAS141_18839 80% 8e-157
Thalassiobium R2A62		TR2A62_3433 51% 0.0		TR2A62_2200 71% 0.0	TR2A62_3444 77% 0.0	
Wenxinia marina DSM24838		N/A 54% 0.0		N/A 69% 0.0	N/A 75% 0.0	

Locus tags (if available), identity to *R. pomeroyi* protein sequence and E-value are shown for each strain.

5.2.7.2 The cleavage route

As described in **Chapter 4**, the DMSP-derived acrylate in *Ruegeria pomeroyi* is thought to be metabolised via the propionate catabolism pathway. To re-cap, acrylate is converted to acryloyl-CoA by the CoA-ligase SPO2934, and then reduced to propionyl-CoA by AcuI (SPO1914). Propionyl-CoA is catabolised to methylmalonyl-CoA by the Pcc complex, encoded by SPO1094 and SPO1101, and then methylmalonyl-CoA is converted to succinyl-CoA by SPO1105 (see **Figure 4.23)**. The presence of all five of these enzymes in the Roseobacter clade was investigated by using a BLASTp search against individual genome sequences with the *R. pomeroyi* peptide sequences as queries.

These searches revealed that every strain had good homologues to all five enzymes (**Table 5.9**). This is unsurprising since propionate metabolism genes are found from bacteria to humans, with highly conserved amino acid sequences. For example, *pccA* and *pccB* of humans and *R. pomeroyi* share 54% and 65% sequence identity, respectively (Huang *et al*., 2010). The AcuI enzyme is also highly conserved amongst both marine bacteria that catabolise DMSP, and non-marine species which do not. Therefore, even though *R. pomeroyi* metabolises acrylate via the propionate pathway, the different growth phenotypes of Roseobacters on DMSP or acrylate cannot be explained purely by the presence of this pathway.

Table 5.9 Acrylate catabolism enzymes in Roseobacter strains (continued overleaf)

Organism	PrpE	AcuI	SPO1101	SPO1094	SPO1105
Citreicella sp. SE45	CSE45_0942 78% 0.0	CSE45_4841 53% 5e-112	CSE45_2834 82% 0.0	CSE45_2840 90% 0.0	CSE45_2829 86% 0.0
Citreicella sp. 357	C357_13632 77% 0.0	C357_04487 80% 7e-180	C357_22650 81% 0.0	C357_22685 90% 0.0	C357_22630 87% 0.0
Dinoroseobacter shibae DFL12	Dshi_0825 82% 0.0	Dshi_2319 83% 0.0	Dshi_0723 85% 0.0	Dshi_0718 86% 0.0	Dshi_0726 87% 0.0
Jannaschia sp. CCS1	Jann_2298 63% 0.0	Jann_2378 82% 1e-176	Jann_3370 82% 0.0	Jann_3374 86% 0.0	Jann_3367 87% 0.0
Loktanella hongkongensis DSM17492	Lokhon_01077 70% 0.0	Lokhon_01293 80% 0.0	Lokhon_02544 71% 0.0	Lokhon_02543 85% 0.0	Lokhon_02545 80% 0.0
Loktanella vestfoldensis SKA53	SKA53_09399 79% 0.0	SKA53_03559 84% 0.0	SKA53_01671 81% 0.0	SKA53_01656 86% 0.0	SKA53_01681 86% 0.0
Maritimibacter alkaliphilus HTCC2654	RB2654_20788 77% 0.0	RB2654_18026 43% 4e-83	RB2654_08782 87% 0.0	RB2654_08767 90% 0.0	RB2654_08797 88% 0.0
Oceanibulbus indolifex HEL45	OIHEL45_09205 75% 0.0	OIHEL45_13350 85% 0.0	OIHEL45_10268 87% 0.0	OIHEL45_10238 91% 0.0	OIHEL45_10278 85% 0.0
Oceanicola batsensis HTCC2597	OB2597_06365 79% 0.0	OB2597_08944 50% 1e-107	OB2597_15305 87% 0.0	OB2597_15275 92% 0.0	OB2597_15315 91% 0.0
Oceanicola granulosus HTCC2516	OG2516_13444 73% 0.0	OG2516_15269 75% 1e-147	OG2516_00349 79% 0.0	OG2516_00319 87% 0.0	OG2516_00364 84% 0.0
Oceanicola sp. S124	N/A 77% 0.0	N/A 78% 0.0	N/A 87% 0.0	N/A 92% 0.0	N/A 91% 0.0
Oceaniovalibus guishaninsula JLT2003	OCGS_1663 79% 0.0	OCGS_2034 79% 0.0	OCGS_0523 79% 0.0	OCGS_0517 87% 0.0	OCGS_0525 79% 0.0
Octadecabacter arcticus 238	OA238_c33230 81% 0.0	OA238_c20420 85% 0.0	OA238_c05970 80% 0.0	OA238_c05940 85% 0.0	OA238_c06000 83% 0.0
Octadecabacter antarcticus 307	OAN307_c37100 79% 0.0	OAN307_c32580 83% 0.0	OAN307_c11470 80% 0.0	OAN307_c11430 84% 0.0	OAN307_c11500 83% 0.0

Organism	PrpE	AcuI	SPO1101	SPO1094	SPO1105
Pelagibaca bermudensis HTCC2601	R2601_10072 79% 0.0	R2601_25396 84% 0.0	R2601_19539 82% 0.0	R2601_19502 90% 0.0	R2601_19564 88% 0.0
Phaeobacter gallaeciensis DSM17395	PGA1_c11900 86% 0.0	PGA1_c13870 90% 0.0	PGA1_c21540 88% 0.0	PGA1_c21600 96% 0.0	PGA1_c21510 93% 0.0
Phaeobacter gallaeciensis 2.10	PGA2_c11900 86% 0.0	PGA2_c13770 90% 0.0	PGA2_c20490 88% 0.0	PGA2_c20530 96% 0.0	PGA2_c20460 93% 0.0
Rhodobacterales bacterium KLH11	RKLH11_2805 89% 0.0	RKLH11_2520 88% 0.0	RKLH11_2359 90% 0.0	RKLH11_2817 95% 0.0	RKLH11_163 93% 0.0
Rhodobacterales bacterium HTCC2083	RB2083_2488 82% 0.0	RB2083_240 86% 0.0	RB2083_2983 86% 0.0	RB2083_3197 91% 0.0	RB2083_3816 87% 0.0
Rhodobacterales bacterium HTCC2150	RB2150_14531 75% 0.0	RB2150_02904 87% 0.0	RB2150_16232 78% 0.0	RB2150_16197 88% 0.0	RB2150_16247 82% 0.0
Roseobacter denitrificans Och114	RD1_3986 79% 0.0	RD1_2290 88% 0.0	RD1_2032 87% 0.0	RD1_2028 88% 0.0	RD1_2035 87% 0.0
Roseobacter litoralis Och149	RLO149_c004510 79% 0.0	RLO149_c022340 88% 0.0	RLO149_c011790 87% 0.0	RLO149_c011750 88% 0.0	RLO149_c011820 86% 0.0
Roseobacter sp. AzwK-3b	RAZWK3B_07859 81% 0.0	RAZWK3B_13014 82%	RAZWK3B_20346 83% 0.0	RAZWK3B_20371 87% 0.0	RAZWK3B_20326 87% 0.0
Roseobacter sp. CCS2	RCCS2_03859 82% 0.0	RCCS2_18181 82% 0.0	RCCS2_14664 81% 0.0	RCCS2_14644 86% 0.0	RCCS2_14674 86% 0.0
Roseobacter sp. GAI101	RGAI101_1556 77% 0.0	RGAI101_1919 86% 0.0	RGAI101_2932 86% 0.0	RGAI101_529 90% 0.0	RGAI101_2082 87% 0.0
Roseobacter sp. MED193	MED193_21871 85% 0.0	MED193_02805 84% 0.0	MED193_21751 85% 0.0	MED193_21711 94% 0.0	MED193_21771 92% 0.0
Roseobacter sp. SK209-2-6	RSK20926_19992 85% 0.0	RSK20926_18027 84% 0.0	RSK20926_22419 88% 0.0	RSK20926_22464 95% 0.0	RSK20926_22399 93% 0.0
Roseovarius nubinhibens ISM	ISM_09576 81% 0.0	ISM_00165 80% 0.0	ISM_16820 88% 0.0	ISM_16845 89% 0.0	ISM_16800 89% 0.0
Roseovarius sp. TM1035	RTM1035_06488 81% 0.0	RTM1035_03435 88% 0.0	RTM1035_12858 84% 0.0	RTM1035_12888 92% 0.0	RTM1035_12838 89% 0.0

Organism	PrpE	AcuI	SPO1101	SPO1094	SPO1105
Roseovarius sp. 217	ROS217_16780 81% 0.0	ROS217_14476 89% 0.0	ROS217_23577 83% 0.0	ROS217_23612 92% 0.0	ROS217_23552 89% 0.0
Ruegeria lacuscaerulensis ITI-1157	SL1157_2163 91% 0.0	SL1157_2966 87% 3e-172	SL1157_0076 91% 0.0	SL1157_0081 95% 0.0	SL1157_0073 93% 0.0
Ruegeria pomeroyi DSS-3	SPO2934 100% 0.0	SPO1914 100% 0.0	SPO1101 100% 0.0	SPO1094 100% 0.0	SPO1105 100% 0.0
Ruegeria sp. TM1040	TM1040_1574 86% 0.0	TM1040_1443 84% 0.0	TM1040_1869 86% 0.0	TM1040_1877 91% 0.0	TM1040_1865 92% 0.0
Ruegeria sp. Trich CH4B	SCH4B_2051 85% 0.0	SCH4B_2197 85% 0.0	SCH4B_1883 87% 0.0	SCH4B_1874 91% 0.0	SCH4B_1887 92% 0.0
Ruegeria sp. TW15	N/A 91% 0.0	N/A 88% 0.0	N/A 90% 0.0	N/A 95% 0.0	N/A 92% 0.0
Ruegeria sp. R11	RR11_746 86% 0.0	RR11_3555 92% 0.0	RR11_3052 89% 0.0	RR11_2064 96% 0.0	RR11_3206 92% 0.0
Sagittula stellata E-37	SSE37_21680 85% 0.0	SSE37_07013 81% 0.0	SSE37_08883 84% 0.0	SSE37_08918 87% 0.0	SSE37_08858 88% 0.0
Sulfitobacter sp. EE36	EE36_03723 76% 0.0	EE36_15567 86% 0.0	EE36_11798 87% 0.0	EE36_11773 90% 0.0	EE36_11813 86% 0.0
Sulfitobacter NAS-14.1	NAS141_08651 75% 0.0	NAS141_05688 86% 0.0	NAS141_17024 87% 0.0	NAS141_16999 90% 0.0	NAS141_17039 86% 0.0
Thalassiobium R2A62	TR2A62_3423 79% 0.0	TR2A62_0838 85% 0.0	TR2A62_1401 80% 0.0	TR2A62_1406 91% 0.0	TR2A62_1398 83% 0.0
Wenxinia marina DSM24838	N/A 73% 0.0	N/A 49% 2e-101	N/A 82% 0.0	N/A 89% 0.0	N/A 84% 0.0

Locus tags (if available), identity to *R. pomeroyi* protein sequence and E-value are shown for each strain.

5.2.7.3 Arrangement of propionate metabolism genes in the Roseobacter clade

In *R. pomeroyi*, the *pcc* genes SPO1101 and SPO1094 and the methylmalonyl-CoA mutase gene SPO1105 are closely linked on the chromosome separated by an unusual arrangement of very small genes encoding hypothetical proteins and lipoproteins, plus some large intergenic spaces. Inspection of the corresponding region in all the other Roseobacter genomes, revealed that, in all cases, *pccA* and *pccB*, and the methylmalonyl-CoA mutase gene *mcm* were also closely linked, and in the same order as in *R. pomeroyi*. And, with one exception, the other strains resembled *R. pomeroyi* in that the *pcc* and *mcm* genes were always separated by a series of small genes encoding hypothetical proteins. However, the numbers of these intervening genes and their sequences varied considerably, almost providing a strain-specific fingerprint. **Figure 5.14** shows gene maps for a selected few strains to demonstrate the variance in this region. All of the intervening genes were predicted to encode either hypothetical proteins or lipoproteins, with the exception of one gene (Dshi_0720) for a putative endoribonuclease in *D. shibae*. Some homologous hypothetical protein and lipoprotein genes occur in more than one genome. Indeed, two hypothetical genes, labelled hypothetical 2 and 4 in **Figure 5.14** are present in all strains. However, this is the only consistency between the different strains. Otherwise, there is a rather eclectic mix of different small genes. Most are not limited to only one strain, but there are a few genes that are not found in any other Roseobacter, or, indeed in any organism on the NCBI database. These genes are coloured in black in **Figure 5.14**.

5.2.7.4 Lipoproteins in the *pcc* region

There are three different genes in the *pcc* region that occur in several different species and which are predicted to encode lipoproteins (**Figure 5.14**). Bacterial lipoproteins are cell surface components, characterized by a conserved N-terminal lipid-modified cysteine residue that allows the hydrophilic protein to anchor onto bacterial cell membranes. There is no common function for lipoproteins; it is a general term for a diverse group of proteins which are anchored to the cell surface and either have a structural or catalytic function. The *pcc* region of *R. pomeroyi* and *Roseobacter* SK209-2-6 has all three different lipoprotein genes, termed lipoprotein A, B and C (**Figure 5.14**). Two of these, lipoprotein A and C are also present in *Citreicella* sp. SE45. The remaining strains shown in **Figure 5.14** do not have homologues to these lipoproteins anywhere in their genomes.

5.2.7.5 Hypothetical proteins in the *pcc* region

A similar situation is seen for genes encoding hypothetical proteins in the *pcc* region, whereby some genes are present in several different strains, whereas others are restricted to just one or two (**Figure 5.14**). As mentioned, hypothetical protein genes 2 and 4 are present in all strains. The remaining six hypothetical proteins appear in some strains but not others. Since the

functions of these encoded proteins are not known, it is not possible to draw any conclusions about why they are present in some strains and not others, and/or whether they are connected in any way to propionate catabolism in these organisms.

5.2.7.6 *Loktanella hongkongensis*

Interestingly, there is at least one known example of a Roseobacter strain, in which the *pcc* genes are uninterrupted by lipoprotein or hypothetical genes. Thus, in *Loktanella hongkongensis*, the *pccA, pccB* and *mcm* genes are completely contiguous and, given the absence of any intergenic spaces between them, it is very likely that they form a single transcriptional unit (**Figure 5.14**). Interestingly, *L. hongkongensis* does have a cluster of the small genes encoding hypothetical proteins and lipoproteins, namely hypothetical 2 and 4, and lipoprotein B, but these are elsewhere in the genome (**Figure 5.15**). Given that hypothetical protein genes 2 and 4 are also conserved in all other Roseobacter strains, it may be that these proteins do have an important, unknown role in cellular processes.

5.2.7.7 The *pcc* region in other bacteria

The *pcc* and *mcm* genes are usually contiguous in other bacterial taxa, as in the case of *L. hongkongensis.* For example, other α-proteobacteria like *Brucella* spp. and *Sinorhizobium* spp. have close homologues to the Roseobacter *pcc* and *mcm* genes but lack the interrupting small genes in between. Some exceptions to this were *Rhizobium* spp. and *Agrobacterium* spp., whose *pccA* and *pccB* genes were separated by one or two genes, respectively (**Figure 5.16**). Interestingly, in the case of *Agrobacterium* spp., the product of the smaller of the two intervening genes (Agau_L100147 in **Figure 5.16**) is a close homologue of hypothetical protein 2 in the Roseobacters, with ~60% sequence identity. Based on this information, it may be possible that most bacteria at some point had the intervening genes but have now lost all, or most of them. Even when *Rhizobium* and *Agrobacterium* spp. are taken into account, the *pcc* region of the Roseobacter clade is strikingly different to any other group of bacteria. It would be interesting to know whether this region confers any kind of advantage to these marine organisms, and what the functions of the tiny encoded proteins may be.

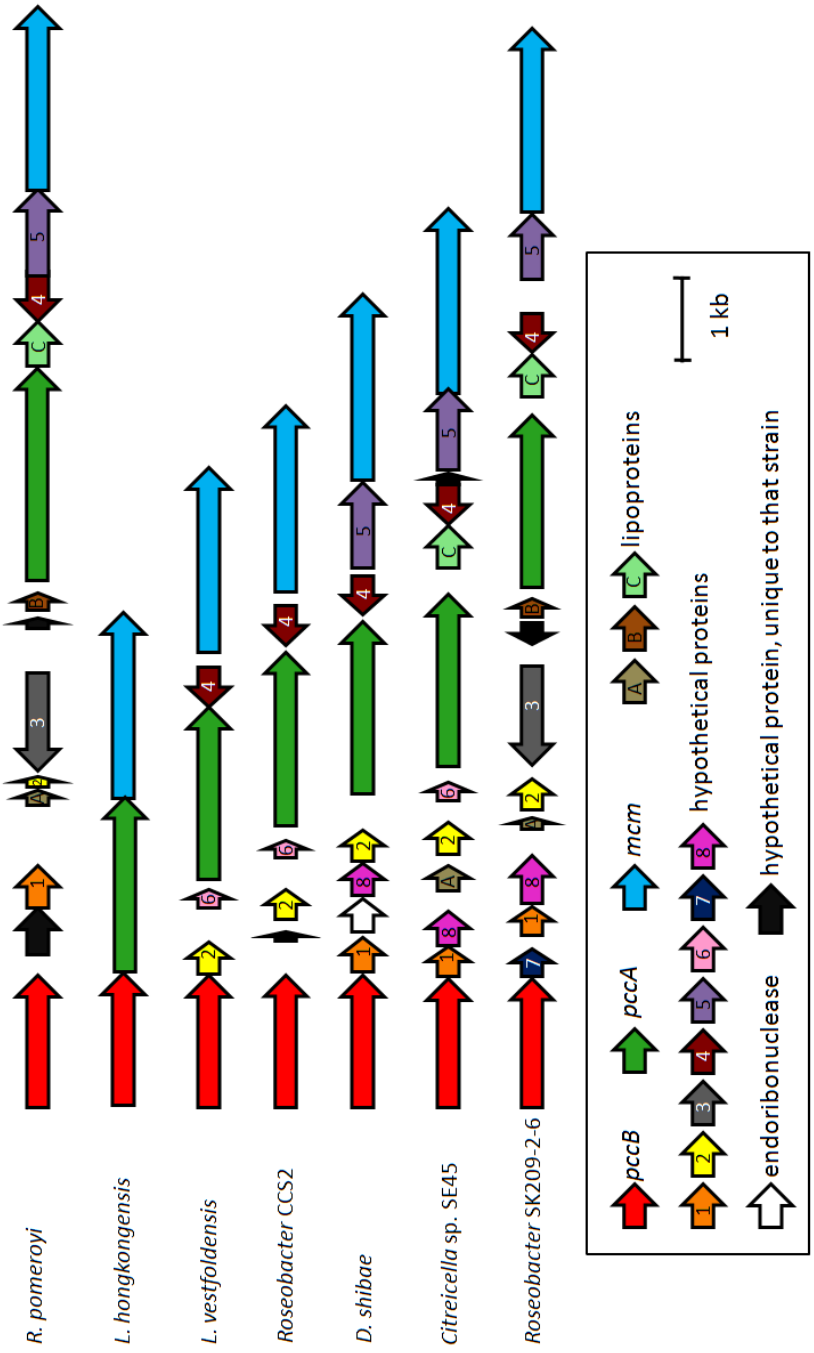

R. pomeroyi
L. hongkongensis
L. vestfoldensis
Roseobacter CCS2
D. shibae
Citreicella sp. SE45
Roseobacter SK209-2-6
pccB
pccA
mcm
lipoproteins
hypothetical proteins
1 kb
endoribonuclease
hypothetical protein, unique to that strain

Figure 5.16 Arrangement of propionate catabolism genes in Roseobacter strains. Genomic location of genes encoding propionyl-CoA carboxylase, subunits A (*pccA*) and B (*pccB*), and the methylmalonyl-CoA mutase (*mcm*) in selected Roseobacter strains. In most cases these genes are separated by a number of small genes encoding hypothetical proteins (represented by arrows with a number. Arrows with the same numbers are genes encoding homologous proteins), or lipoproteins (represented by arrows with a letter)

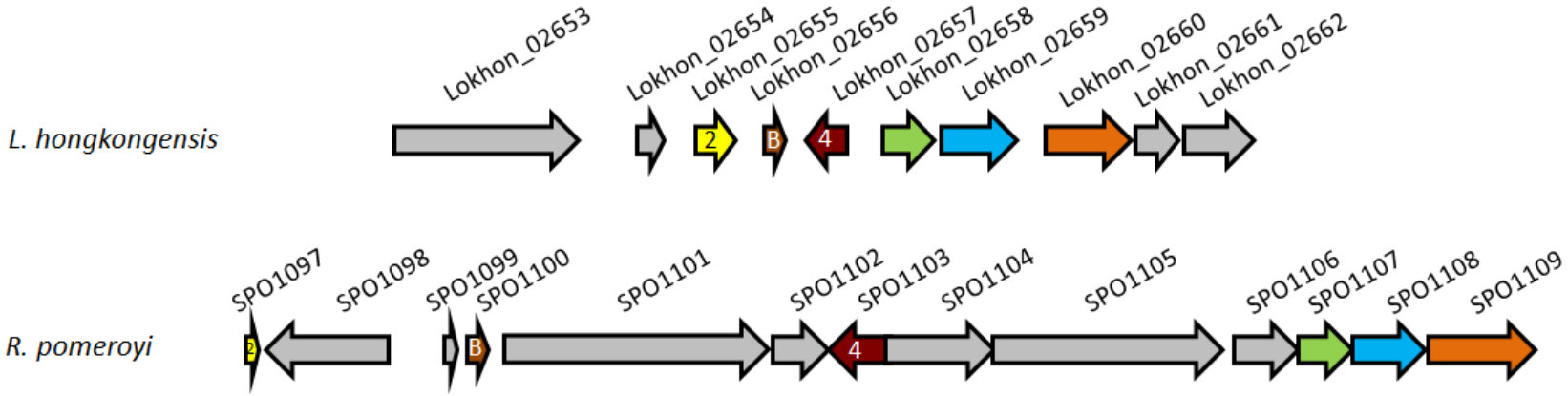


Figure 5.17 Genomic location of *L. hongkongensis* hypothetical genes 2 and 4. Hypothetical genes 2 and 4, and lipoprotein gene B (represented by yellow arrows, dark brown arrows and light brown arrows, respectively) are positioned contiguously in *L. hongkongensis*, next to genes encoding homologues of the *R. pomeroyi* SPO1107, SPO1108 and SPO1109 genes. Thus, the products of SPO1107 and Lokhon_02658 (green arrows) are 43% identical and are putative acetyltransferase-superfamily proteins. Lokhon_02659 and SPO1108 (blue arrows) encoded proteins are 67% identical and are annotated as DnaJ-like. The

product of Lokhon_02660 is 54% identical to that of SPO1109 (orange arrows), predicted to be a member of the endonuclease/exonuclease/phosphatase protein family. In *R. pomeroyi*, hypothetical genes 2 and 4, and lipoprotein B, and SPO1107, SPO1108 and SPO1109 are interrupted by *pccA* and *mcm* genes, and four other genes, predicted to encode lipoproteins (SPO1099 and SPO1102), and hypothetical proteins (SPO1104, SPO1106). In *L. hongkongensis* the remaining genes shown are predicted to encode a choline dehydrogenase (Lokhon_02653), a succinoglycan biosynthesis protein (Lokhon_02654), an ATPase (Lokhon_02661) and a phosphoglycolate phosphatase (Lokhon_02662).

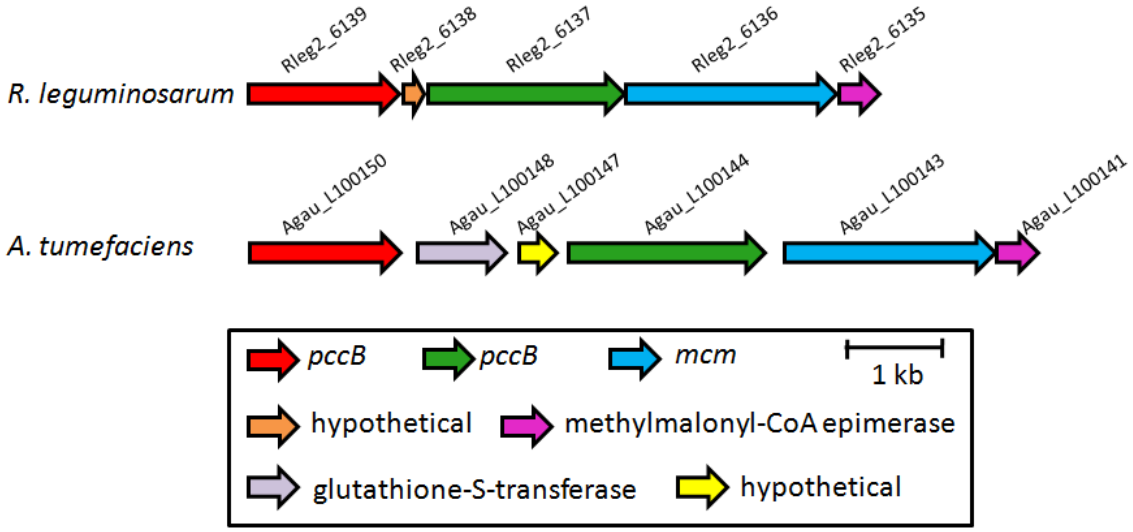


Figure 5.18 Arrangement of propionate catabolism genes in *Rhizobium leguminosarum* and *Agrobacterium tumefaciens*. Genomic location of *pccA*, *pccB* and *mcm* in other α-proteobacteria. In contrast to the Roseobacters, the genes in these species are also linked to a gene encoding a putative methylmalonyl-CoA epimerase. However, *pccB* and *pccA* are also interrupted by hypothetical proteins, in a similar situation to Roseobacter species.

5.3 Summary

It is clear from this work that DMSP catabolism genes are abundant amongst sequenced strains of Roseobacters. It is also not unusual for a single strain to have multiple copies of different *ddd* genes, as well as the DMSP demethylase *dmdA* – a situation not seen in many other sequenced bacteria outside of this clade. However, other Roseobacter strains, like *P. bermudensis*, do not have any known DMSP genes, and also do not make MeSH or DMS from DMSP. Indeed the number of DMSP-related genes in a single species ranges from zero to five, with *R. pomeroyi* having a copy of *dmdA*, *dddP*, *dddQ*, *dddW* and *dddD*. The number and type of DMSP genes in a strain is apparently not linked to its phylogeny, so more closely related strains do not necessarily have the same set of genes. It might be that the acquisition of DMSP genes by Roseobacter strains was a relatively recent event, and occurred independently in each strain, or that their common ancestor had multiple lyases, some of which have since been lost in some strains.
One pattern which emerges from the plethora of different DMSP genes, is that those strains with *dmdA* usually also have a copy of *dddP*. It is not known why that should be, since they are not closely linked in any genome. Another interesting pattern is that strains with *dddL* usually lack any other known DMSP-related genes.

The second part of this work examined the synteny of the different DMSP-related genes in each genome, which also lead to some interesting observations. For example, homologues of DddD fall into two distinct groups amongst the Roseobacters, of which only one group has been shown to have a functional copy of DddD. Interestingly, from the synteny of the *dddD* genes, the 'functional' *dddD*s are all linked to a *dddA* homologue, whereas the 'non-functional' *dddD* genes are not near any gene known to be involved in DMSP catabolism. This observation is intriguing, and requires further empirical work to confirm whether all the DddD enzymes of group A are indeed functional, and vice versa for group B.

Excitingly, this work also revealed that *dmdA* of *P. gallaeciensis* was linked to all three genes encoding the rest of the DMSP demethylation pathway, i.e *dmdB2*, *dmdC* and *dmdD*. A similar situation is seen in the SAR11 strain *P. ubique* HTCC1062, whereby *dmdB* and *dmdC* are adjacent, and separated from *dmdA* by only one gene which encodes a hydratase, but has no homology to DmdD and does not function as a MTA-CoA hydratase. In *P. ubique*, the closest DmdD homologue (25% identical to SPO3805) is found elsewhere in the genome, but also does not possess DmdD activity (Reisch *et al*., 2011). In contrast, DmdA, DmdB2, DmdC and DmdD of *P. gallaeciensis* all have very high homology to the *R. pomeroyi* demethylation enzymes, and

it would be of interest to confirm whether they are functional in the DMSP demethylation pathway in this strain.

Finally, the comparison of the *pcc* region of different Roseobacters has led to some interesting observations. In many bacteria, *pccA*, *pccB*, and *mcm* are contiguous in the genome, and likely co-transcribed, but this was only the case for one Roseobacter strain, namely *Loktanella hongkongensis*. In all other sequenced Roseobacters, these three genes were interrupted by an unusual arrangement of varying numbers and types of small genes encoding hypothetical proteins or lipoproteins. None of these genes encode proteins of known function, and it would be very interesting to investigate what they do and why their genes seemingly interrupt the co-transcription of the propionate catabolism genes.

The work in this chapter has revealed several interesting observations, and really demonstrates the power of genome comparison as a tool for explaining empirical findings, raising new, interesting questions, and building hypotheses. In this case, the comparison of regions involved in DMSP catabolism in different Roseobacter species has led to some exciting findings, which will hopefully be explored further by future researchers.

Chapter 6

General Discussion

6.1 Preamble

The purpose of the work described in this thesis was to deepen our understanding of DMSP catabolism in marine bacteria. This has been achieved in several different ways. A thorough genomic comparison of the important Roseobacter clade yielded some interesting patterns in the synteny and presence of DMSP-related genes. Secondly, I identified an enzyme with DMSP-cleaving activity in one of the most abundant bacteria on earth; SAR11 strain HTCC1062. An investigation into the γ-proteobacterium *O. doudoroffii*, revealed that the presence of multiple DMSP lyases is not restricted to the α-proteobacteria. Finally, I identified a potential pathway of DMSP carbon assimilation in the model organism, *Ruegeria pomeroyi*. This chapter explores the results of this work in more detail.

6.2 DMSP metabolism in *Ruegeria pomeroyi*

Before this project started, nothing was known of how the model marine organism *R. pomeroyi* was able to metabolise the acrylate that was generated *via* DMSP catabolism. The work described in **Chapter 4** provides substantial evidence for the metabolism of acrylate through acryloyl-CoA and thence propionyl-CoA in this bacterium. However, given the prominence of *R. pomeroyi* as a model Roseobacter in DMSP research, and the importance in understanding how this abundant sulphur molecule is used, it is perhaps unsurprising that research groups at the University of Georgia were simultaneously unravelling the mechanism of acrylate catabolism in this organism.

Thus, work by Reisch *et al*. (2013) on *R. pomeroyi* metabolism of acrylate was published in July 2013, just as my research on the same project was coming to a close. In what follows, I will present their findings and describe how my own work both complements and provides additional support for their conclusions.

6.2.1 SPO2934 is an acryloyl-CoA ligase

In **Chapter 4** I hypothesised that the *R. pomeroyi* gene SPO2934 could be an acryloyl-CoA ligase, based on the sequence of its predicted gene product, annotated as a propionyl-CoA ligase (PrpE), and its up-regulation in the presence of DMSP and acrylate in the microarray data (M. Kirkwood). I obtained further evidence in support of this model, by cloning and expressing the gene in *E. coli*, where it conferred extreme sensitivity to acrylate, probably due to a build-up of toxic acryloyl-CoA. In an attempt to identify a pathway of DMSP catabolism, Reisch *et al*. (2013) assayed acryloyl-CoA activity in crude cell extracts of *R. pomeroyi* grown with DMSP as

a sole carbon source. When cell-free extracts were supplied with acrylate, HS-CoA plus ATP, they produced acryloyl-CoA at a rate of 24 nmol min^{-1} mg $protein^{-1}$. Based on their own regulatory studies, they came to the same hypothesis that the enzyme responsible was the so-called "propionyl-CoA ligase". Microarray data from Reisch *et al*. (2013) showed that *prpE* was up-regulated between 3.5- and 7-fold in the presence of DMSP compared to glucose, in agreement with our microarray data (M. Kirkwood) which showed a 3.26-fold up-regulation in the presence of DMSP compared to succinate. To confirm that *prpE* did have acryloyl-CoA ligase activity, Reisch *et al*. (2013) also cloned the gene and expressed it in *E. coli*; cell-free extracts of the recombinant *E. coli* had CoA-ligase activity with both acrylate and propionate as substrates, compared to wild type *E. coli* which did not.

Reisch *et al*. (2013) made an insertional mutation in *prpE* and found that the mutant could still grow on DMSP or on and acrylate, just as I had found in a similar examination of its phenotype (**Chapter 4**). However, in their case, they found that the growth was somewhat delayed compared to the wild type, confirming the importance of PrpE in DMSP and acrylate catabolism. For reasons that are not clear, I did not see such an inhibitory effect.

The ability of the *prpE*$^{-}$ mutant to grow on DMSP was expected, since DMSP carbon is also thought to be assimilated via the demethylation route (Reisch *et al*., 2011). A more surprising result was that the mutant still grew on acrylate. Reisch *et al.* hypothesised that this could be down to a redundancy in acryloyl-CoA ligase activity, although no further work was carried out to confirm this. I came to the same hypothesis and to investigate the possibility of functional redundancy directly, I cloned two additional putative CoA-ligase genes from *R. pomeroyi* and expressed them in *E. coli*. These CoA-ligases also conferred an acrylate hypersensitivity phenotype, but not to the extent of PrpE. Thus I showed that there is indeed a redundancy in acryloyl-CoA ligase activity, but that PrpE is still a good candidate for the most important enzyme.

6.2.2 SPO0147 is an acryloyl-CoA hydratase

Reisch *et al*. (2013) found that acryloyl-CoA was rapidly converted to 3-hydroxypropionyl-CoA (3HP-CoA) in cell-free extracts. They purified this acryloyl-CoA hydratase activity and identified the enzyme responsible as the product of SPO0147. Interestingly, the sequence of this enzyme, annotated as an enoyl-CoA hydratase, is 55% identical to AcuK from *Halomonas* sp. HTNK1. In *Halomonas*, *acuK* is co-transcribed with *acuN*, which encodes a putative CoA transferase. Work in the UEA laboratory showed that AcuK and AcuN from *Halomonas* work in tandem to convert acrylate to 3HP. Although a 3HP-CoA intermediate was not seen in this case,

it was hypothesised that AcuK acts as an acryloyl-CoA hydratase (Todd *et al*., 2010a). The findings of Reisch *et al*. (2013) are consistent with this hypothesis.

Despite this finding, the importance of the SPO0147 product to acrylate metabolism could not be confirmed, since several attempts to make a mutation in the gene were unsuccessful, both in Georgia and our laboratory (unpublished observations; Reisch *et al*., 2013). There is no evidence that the product of SPO0147 is essential for survival, but it is possible that an insertion in this gene also disrupts the downstream SPO0148, that is predicted to encode 30S ribosomal protein S20, and this could be detrimental to the cell. Indeed, deletion of the S20 gene in a *Salmonella enterica* strain, while not lethal, did confer a significant reduction in the rate of mRNA binding to ribosomes (Tobin *et al*., 2010). Nonetheless, Reisch *et al*. (2013) set out to identify the fate of 3HP-CoA, and found that it was converted to propionyl-CoA in an NADH/NADPH-dependent manner. They hypothesised that this could be due to the reverse activity of the acryloyl-CoA hydratase plus an acryloyl-CoA reductase.

The production of 3HP, or its CoA intermediate, from acryloyl-CoA is supported by some preliminary metabolomics work carried out by Mark Kirkwood. This work was carried out originally to try and identify the products made in wild type *R. pomeroyi* in comparison with an AcuI$^-$ mutant, following the addition of acrylate. Cultures of the wildtype and mutant strains were grown in the presence or absence of acrylate, and samples of each were analysed for metabolites by nuclear magnetic resonance (Dr. Gwen Legal, Institute of Food Research, Norwich). The NMR showed a greater concentration of 3HP in the AcuI$^-$ mutant compared to the wild type strain, when each of these were grown with acrylate. Therefore, more acryloyl-CoA may be routed via a 3HP pathway if a propionate route is unavailable due to mutation (see **Figure 6.1**). Additionally, although the SPO0147 gene is not upregulated in our microarray data, a gene encoding a putative methylmalonate semialdehyde dehydrogenase (SPO2203) is induced ~4- and ~3-fold in the presence of DMSP and acrylate, respectively (M. Kirkwood). The product of SPO2203 could convert 3HP-derived malonate semialdehyde to acetyl-CoA, and thus play an important part in a 3HP-route of acrylate catabolism. Interestingly, this gene is even further induced by DMSP (~12-fold) in the AcuI$^-$ mutant (M. Kirkwood), consistent with the prediction that such a mutant would accumulate more 3HP.

6.2.3 SPO1914 (AcuI) is an acryloyl-CoA reductase

SPO1914 had been shown to be connected to acrylate catabolism by the UEA lab, and annotated as *acuI* (Sullivan *et al*., 2011, Todd *et al*., 2012b), and further work in *Rhodobacter sphaeroides* showed that AcuI in that organism is an acryloyl-CoA reductase (Schneider *et al*., 2012). Reisch *et al*. (2013) confirmed for the first time that AcuI from *R. pomeroyi* also has acryloyl-CoA reductase activity by cloning and expressing it in *E. coli* and assaying for activity. It had been shown by Todd *et al*. (2012b), and is also shown in **Chapter 4,** that a SPO1914$^-$ mutant is hypersensitive to the presence of DMSP and acrylate. Reisch *et al*. (2013) also noted that a SPO1914$^-$ mutant was unable to use DMSP or acrylate as sole carbon sources, and although they did not carry out toxicity tests, they suggested that the growth phenotype of the SPO1914$^-$ may be caused by the subsequential build-up of acryloyl-CoA sequestering essential coenzyme A. I attempted to test this theory (**Chapter 4)**, by trying to relieve the sensitivity phenotype of SPO1914$^-$ by adding the coenzyme A precursor pantothenate. However, this had no effect on the ability of any mutant strain to grow in the presence of acrylate, suggesting a lack of coenzyme A is not a cause of hypersensitivity. Instead, I showed that the sensitivity phenotype was slightly relieved by adding glutathione. It therefore seems more likely that the strong electrophilic nature of acryloyl-CoA has a detrimental effect on essential nucleophilic molecules, and that this might be relieved by the presence this powerful reducing agent.

6.2.4 DMSP-grown cells have enhanced propionyl-CoA carboxylase activity

In **Chapter 4,** I showed that mutations in the *R. pomeroyi* propionyl-CoA carboxylase genes *pccA* and *pccB*, and in the methylmalonyl-CoA mutase gene SPO1105 conferred hyper-sensitivity to the presence of DMSP, acrylate and propionate, due, most likely, to the build-up of toxic coenzyme A intermediates. Significantly, though, this would rely on the fact that at least some of the carbon from DMSP is routed via the propionyl-CoA pathway. Findings by Reisch *et al*. (2013) support this. They showed that *R. pomeroyi* cells grown with DMSP as a carbon source possessed propionyl-CoA carboxylase activity of 38 nmol min^{-1} mg protein^{-1}, almost 10-fold greater than glucose-grown cells.

6.2.5 Future work on DMSP metabolism in *R. pomeroyi*

The work described in this thesis and the work of Reisch *et al*. (2013) has increased our understanding of DMSP catabolism in *R. pomeroyi*, but several unanswered questions remain. For example, we now know that DMSP can be metabolised via its cleavage to acrylate through a propionate (and possibly 3HP) route, and also via its demethylation to MMPA and the DmdB, DmdC and DmdD pathway (see **Chapter 1**; Reisch *et al*., 2011). However, there may be alternative routes of DMSP catabolism which remain to be discovered. One possibility is a direct cleavage of DMSP-derived MMPA to methanethiol (MeSH) plus either acrylate or propionate. However this pathway has not been confirmed in any organism, and no genes for MMPA-cleavage have yet been identified. Nevertheless, I obtained some preliminary evidence for this pathway in *R. pomeroyi* (which was not presented in the results of this Thesis) by testing the propionyl-CoA carboxylase$^-$ and methylmalonyl-CoA mutase$^-$ mutant strains of *R. pomeroyi* for growth on MMPA as a sole carbon source. None of these mutants could use MMPA as a sole carbon source, whereas the wild type did. This hints at the possibility that carbon might be assimilated from MMPA via propionate (or acrylate) in *R. pomeroyi*. However, it should be stressed that, although there were biological repeats, the growth tests were very preliminary and should be repeated. Regardless, based on the initial MMPA growth tests, I also attempted to isolate an MMPA-cleaving enzyme that would produce MeSH in a single enzymatic step. To do this, I screened an *R. pomeroyi* library in *Rhizobium leguminosarum* for MMPA-dependent MeSH production. Of ~500 cosmids screened, several conferred the MeSH-producing phenotype to *R. leguminosarum* but, when sequenced, all of these cosmids were found to contain *dmdB2* (SPO0677), and no good candidates for an MMPA 'demethiolase'. More work needs to be done on these cosmids, to verify the gene responsible, but it seems likely that, while wild type *R. leguminosarum* is not able to produce MeSH from MMPA, it may have part of the Dmd demethylation pathway, and the addition of SPO0677 complements this by providing MMPA-CoA ligase activity. Therefore, although I was not able to isolate an MMPA-cleaving gene, the possibility of an MMPA cleavage pathway in *R. pomeroyi* has not been ruled out, and this very interesting and important question requires more work.

It is also interesting that *R. pomeroyi* can use both DMSP and acrylate as sole carbon sources at all, since the studies conducted on Roseobacters so far suggest this is a rare trait amongst this group of α-proteobacteria. Even several strains that are equipped with several DMSP lyases, plus the DmdA demethylase and the necessary downstream pathways do not grow on DMSP, as will be described next.

6.2.5.1 Growth on DMSP and acrylate unusual amongst the α-proteobacteria.

All previous studies on the use of DMSP as a sole carbon source had focussed on species of γ- or β-proteobacteria with either *dddD* or *dddY*, most significantly *Halomonas* HTNK1 and *Alcaligenes faecalis* M3A (see **Chapter 1**). Both of these strains have a cluster of genes surrounding their respective DMSP lyase genes which were shown to be involved in the assimilation of carbon from DMSP and acrylate, namely *dddA*, *dddC*, *acuN* and *acuK*. Clusters of these *ddd* and *acu* genes are also found in other γ-proteobacteria with *dddD*, such as *Marinomonas* MWYL1, *Pseudomonas* J465, *Psychrobacter* J466 and *Oceanimonas doudoroffii*, and these bacteria too, can use DMSP as a sole source of carbon (Curson *et al*., 2010; Curson *et al*., 2012; see **Chapters 1** and **3**). Additionally, strains of bacteria isolated from the environment on the basis of growth on DMSP as a sole carbon source tend to be γ- or β-proteobacteria (J. Todd, personal communication).

In contrast, many α-proteobacteria containing the DMSP lyases DddP, DddQ, DddW and DddL are not able to grow on DMSP as a sole carbon source, at least under laboratory conditions. Indeed, out of 13 Roseobacter strains tested (in this work, and in other studies) only three grew on both DMSP and acrylate, namely *Ruegeria pomeroyi*, *Roseovarius* sp. 217 and *Sagittula stellata* E-37. In addition, *Sulfitobacter* EE-36 grew on DMSP, but not acrylate, and *Ruegeria* TrichCH4B grew only on acrylate (**Table 6.1**; González *et al*., 1999). The other eight did not grow on either carbon source under the conditions used. This is interesting, since the genomic comparison analysis presented in **Chapter 5** showed that all of the sequenced Roseobacter strains contain the acrylate pathway genes identified in this work, and by Reisch *et al*. (2013). Therefore, each of the strains that were tested has a homologue of SPO2934 (the acryloyl-CoA ligase), as well as the *acuI*, *pccA* and *pccB* and the methylmalonyl-CoA mutase genes. Each strain does have a different arrangement of primary DMSP enzymes, but the growth phenotypes do not form a pattern that correlates with any particular individual or combination of DMSP lyases, or with the presence or absence of DmdA. Indeed, since only four strains grew on DMSP, it is difficult to draw any real conclusions about the presence of certain genes and the growth phenotype of strains. However, it is interesting to note that all of the "DMSP growers" have the MMPA-CoA ligase, DmdB1, whereas the majority of the non-growers lack this enzyme. The exceptions to this are *Ruegeria* sp. R11, *Phaeobacter gallaeciensis* DSM17395 and *Roseobacter* sp. GAI101, all of which have DmdB1 but do not grow on DMSP. This is worthy of note, since it was suggested that DmdB1 might be more important than DmdB2 during growth on DMSP, following the observation that *dmdB1* transcripts were always higher than *dmdB2* during growth on DMSP as a sole carbon source (Bullock *et al*., 2014).

It would be of interest to further explore the different growth phenotypes seen in different Roseobacter strains by testing more strains for their use of DMSP and acrylate. It could also be revealing to transform some of the strains with an *R. pomeroyi* genomic library to identify any *R. pomeroyi* genes that confer a DMSP growth phenotype to non-growing Roseobacters.

Table 6.1 Growth of Roseobacter strains on 5 mM DMSP or acrylate as sole carbon sources.

Strain	DMSP	Acrylate	Source
Ruegeria pomeroyi	✓	✓	This work
Roseovarius sp. 217	✓	✓	This work; Schäfer *et al*., 2005
Sagitulla stellata E-37	✓	✓	González *et al*., 1999
Sulfitobacter EE-36	✓	×	González *et al*., 1999
Ruegeria sp. TrichCh4B	×	✓	This work
Dinoroseobacter shibae	×	×	This work
Pelagibaca bermudensis	×	×	This work
Phaeobacter gallaeciensis DSM17395	×	×	This work
Roseobacter denitrificans	×	×	This work
Ruegeria sp. R11	×	×	This work
Roseobacter litoralis	×	×	This work
Roseobacter sp. GAI101	×	×	González *et al*., 1999
Roseovarius nubinhibens ISM	×	×	González *et al*., 1999

All of the growth tests were carried out in liquid minimal media with 5 mM acrylate or DMSP as the sole carbon source (González *et al*., 1999; Schäfer *et al*., 2005).

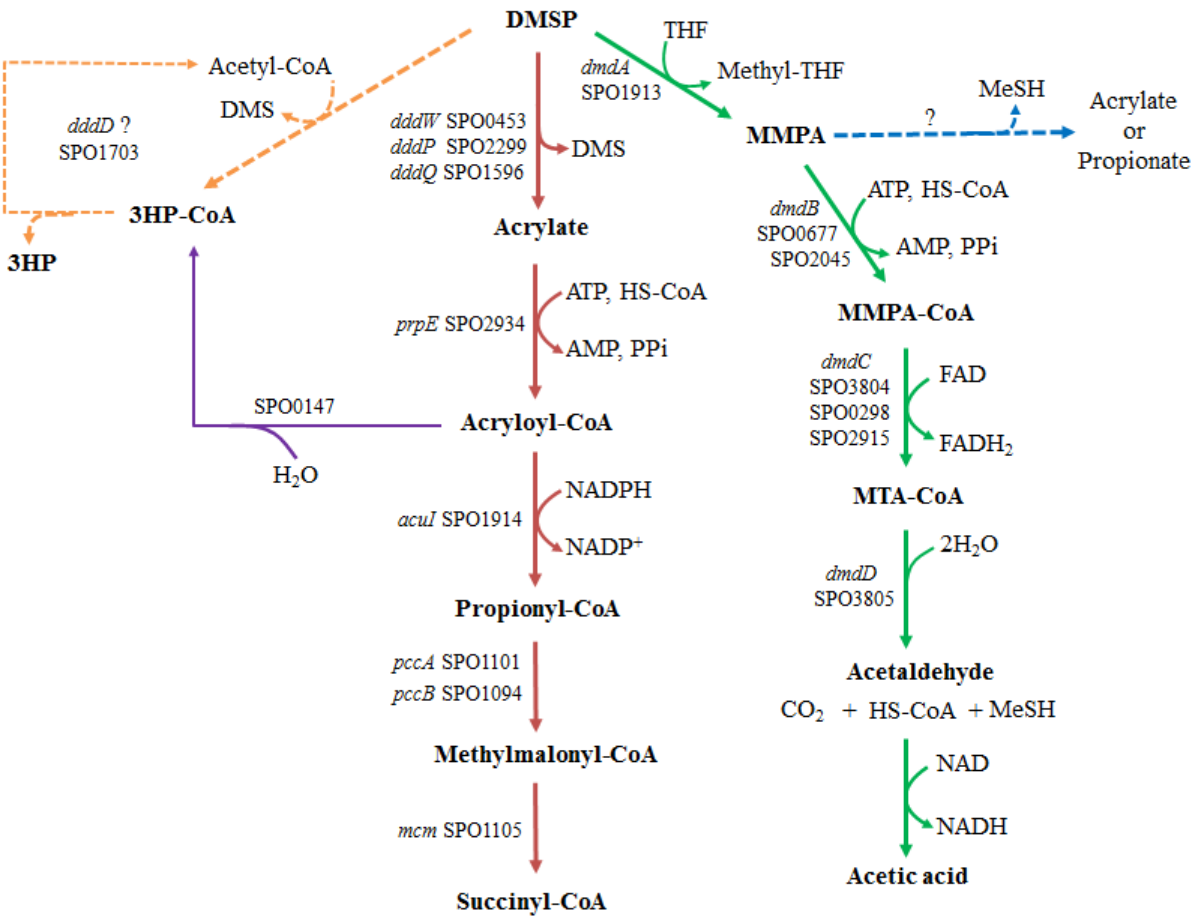


Figure 6.1 Pathways of DMSP catabolism in *Ruegeria pomeroyi* DSS-3. Confirmed (solid arrows) and proposed (dashed arrows) pathways of DMSP metabolism in *R. pomeroyi* DSS-3. The gene name and locus tag (SPO) for each enzyme is given where this is known. Question marks indicate a putative pathway that has not yet been confirmed experimentally. Green arrows show the DMSP demethylation pathway catalysed by DmdA, DmdB, DmdC and DmdD enzymes (Reisch *et al*., 2011). In this pathway, tetrahydrofolate (THF) serves as a methyl acceptor in the demethylation of DMSP to methylmercaptopropionate (MMPA). MMPA is then converted to the coenzyme A intermediate MMPA-CoA, in an ATP-dependent reaction that produces AMP. MMPA-CoA is then dehydrogenated to methylthioacryloyl-CoA (MTA-CoA), which is subsequently hydrated, forming acetaldehyde, free coenzyme A, carbon dioxide and methanethiol (MeSH). Acetaldehyde is then oxidised to acetic acid. A second possible fate for MMPA is indicated by the blue dashed arrow, whereby it is simply cleaved to release MeSH and either acrylate or propionate. However, this is a purely hypothetical pathway. The red arrows show the DMSP cleavage pathway. DMSP can be acted on by one of three DMSP lyases – DddW, DddP or DddQ (Todd *et al*., 2012a; 2009; 2010b), each producing dimethylsulphide (DMS) and acrylate. The acrylate is converted to acryloyl-CoA by PrpE, in an ATP dependent reaction, and the acryloyl-CoA is subsequently reduced to propionyl-CoA by AcuI. This is converted to methylmalonyl-CoA in a carboxylation reaction catalysed by PccA and

PccB, before conversion to succinyl-CoA by a methylmalonyl-CoA mutase (Reisch *et al*., 2013; this work). Alternatively, acryloyl-CoA may be converted to 3-hydroxypropionyl-CoA (3HP-CoA) by SPO0147 (purple arrow, Reisch *et al*., 2013). It is also possible that DMSP is acted on by DddD, since *R. pomeroyi* also has a homologue of this enzyme. This pathway, shown in orange, would convert DMSP to 3HP-CoA in an acetyl-CoA dependent reaction which releases DMS. The 3HP-CoA could then be converted back to acetyl-CoA, with the release of 3HP (Alcolombri *et al*., 2014). However, studies *in vitro* could not confirm the functionality of *R. pomeroyi* DddD (Todd *et al*., 2010b).

6.3 The diversity of DMSP genes in the Roseobacter clade

Through the genomic comparison of Roseobacter species (**Chapter 5**), I showed that there is a remarkable diversity amongst this clade in terms of DMSP catabolism genes. Although the Roseobacter members are closely related, there is a huge variation in the numbers and types of DMSP lyases and in the presence or absence of DmdA, with some species (e.g. *R. pomeroyi*) possessing five primary DMSP catabolism genes, and others (*Oceanicola* sp. S124) which has no known *ddd* or *dmdA* genes. Even different strains of the same species can vary considerably; thus *Phaeobacter gallaeciensis* DSM17395 has a copy of *dmdA* plus *dddP*, but *P. gallaeciensis* 2.10 lacks *dmdA*. Another example is *Loktanella vestfoldensis* DSM16212 which has no known DMSP genes, but *L. vestfoldensis* SKA53 has DddL. Of course, while bioinformatics is an extremely powerful tool for making predictions about the prevalence of genes, it can be difficult to devise cut-off points for homology, and much more direct experimental work remains to be done to confirm the functions of the DMSP lyase homologues amongst most members of the Roseobacter clade. However, the variation in genotype regarding DMSP-degrading enzymes certainly seems to be a trait that is largely restricted to α-proteobacteria, and the Roseobacter clade in particular (the exception to this being *O. doudoroffii* and one other γ-proteobacterium - see below). In fact, the Roseobacters are one of only three bacterial lineages with both the DMSP cleavage and DMSP demethylation pathways (the others being SAR116 and SAR11), highlighting the importance of DMSP as a nutrient source to these bacteria. Although the significance of this variation in DMSP pathways is unknown, the ecology of the Roseobacter clade can provide clues as to why this trait exists. For example, Roseobacter members are known to form associations with a number of phytoplankton groups, such as dinoflagellates and coccolithophorids (Moran *et al*., 2007' Luo and Moran, 2014). These organisms are prolific producers of DMSP, and so the abundance of DMSP-related genes amongst the Roseobacters may be an adaptation to living in an environment likely to contain higher concentrations of DMSP (around 1.6 mM during some phytoplankton blooms {van Duyl *et al*., 1997}). It would be interesting to see if Roseobacter species that are not found in association with eukaryotic DMSP producers still show the same propensity to DMSP genes.

An analysis of the synteny of DMSP-related genes also revealed differences between strains, although there were some intriguing patterns. For example, *dmdA* was almost without exception found immediately upstream, and likely co-transcribed with, *acuI*. In *Ruegeria pomeroyi* and in *Roseobacter* sp. MED193, *dmdA* is divergently transcribed from a gene, termed *dmdR*, which encodes a GntR-type regulator. A *dmdR*-like gene is also found nearby to *dmdA* in almost every other strain (the exception being *P. gallaeciensis* DSM17395), but in each case it is separated from *dmdA* by a few intervening genes with no known function in DMSP catabolism. There are other examples where *acuI* is located next to a DMSP lyase gene, for example in *R. sphaeroides* where it is co-transcribed with *dddL*

(Sullivan *et al*., 2011) and in *Alcaligenes faecalis* where it forms part of the DMSP catabolism gene cluster along with *dddY* (Curson *et al*., 2011). However, the association with *dmdA* in the Roseobacter clade is particularly intriguing given that the product of the DmdA mediated reaction is not acrylate, but MMPA. This raises some interesting questions about the regulation of the demethylation and cleavage pathways in the Roseobacters. According to the *R. pomeroyi* microarray data, its *dmdA-acuI* operon is considerably up-regulated in the presence of DMSP and of acrylate (M. Kirkwood, personal communication). It makes sense that an increase in acrylate would require more AcuI enzyme to catabolise the toxic acryloyl-CoA product. This up-regulation also has the effect of increasing transcription of *dmdA*, which may lead to an increase in the amount of DMSP that is demethylated to MMPA. An additional factor is that the regulator, *dmdR*, is also up-regulated by DMSP and acrylate, and has been shown to repress the *dmdA* promoter in *R. pomeroyi*. However, this repression could not be relieved by DMSP, acrylate, MMPA, DMS or methanethiol (M Kirkwood, personal communication), and so there may be an unidentified player in the regulation of *dmdA* and *acuI*, and consequently the demethylation and cleavage pathways.

Related to this, one of the most striking findings from the synteny analysis was that in *Phaeobacter gallaeciensis* DSM 17395 *dmdA* is not next to *acuI* and *dmdR*, but is in an apparent operon with *dmdB2*, *dmdC1*, and *dmdD* (**Figure 6.2**). This is the only known example where *dmdA* and *dmdB* are adjacent to each other, or to *dmdC* and *dmdD* (although in SAR11 strain HTCC1062, *dmdA* and *dmdB* are separated by only one gene {Reisch *et al*., 2011}). Apart from this *dmdABCD* operon, DMS 17395 is similar to another *P. gallaeciensis* strain; 2.10 (**Figure 6.2, see legend for description**). The absence of *dmdABCD* in 2.10 indicates that strain DSM17395 acquired the demethylation operon recently, following the divergence of *P. gallaeciensis* from its ancestor. It would be very interesting to study DMSP demethylation in strain DSM17395 to see how it is affected by the different arrangement of its *dmd* genes. Note that the *acuI* and *dmdR* genes of DSM 17395 are still divergently transcribed from each other, so it would be of interest to study the regulation of both *dmdA* and *acuI* in DSM 17395 to see if this differs from that in *R. pomeroyi*. In particular, it would be interesting to determine if the DSM 17395 *dmdR* can repress both *acuI* and *dmdA* expression, despite no longer being linked to *dmdA* in the genome.

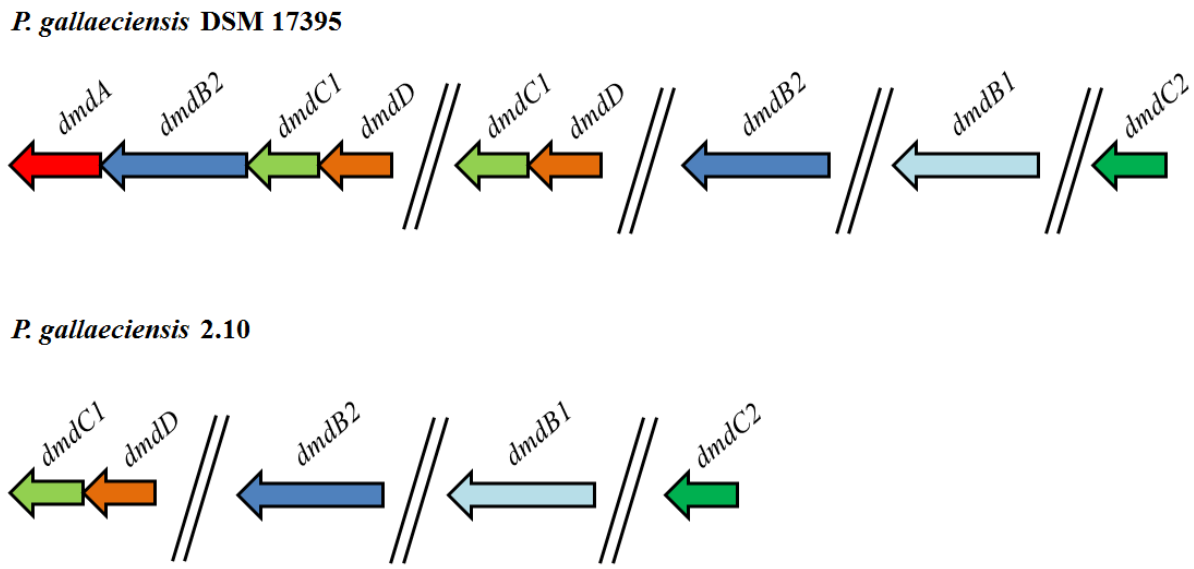


Figure 6.2 Comparison of *dmd* genes in two strains of *Phaeobacter gallaeciensis*. The two strains of *Phaeobacter gallaeciensis*, DSM 17395 and 2.10, have different arrangements of *dmd* genes. Strain 2.10 lacks a copy of *dmdA*, and its *dmdB2*, *dmdB1* and *dmdC2* homologues are all located separately. However, it does have a copy of *dmdC1* and *dmdD*, which are located together, likely as a single transcriptional unit. Similarly, DSM 17395 has a *dmdC1*:*dmdD* pair, and separate *dmdB2*, *dmdB1*, *dmdC2* genes. However, this strain also has a copy of *dmdA* and additional *dmdB2*, *dmdC1* and *dmdD* genes which all likely form a single operon. The black double lines in the figure indicate a separate locus in the genome.

Indeed, one of the most important unanswered questions is how any bacteria with both the DMSP cleavage and demethylation pathways control the flux through each route, since only the cleavage pathway leads to the production of the climatically relevant DMS. The "bacterial switch" hypothesis proposes that bacteria shift between producing more or less DMS and MeSH (Simo, 2001), through the expression of genes encoding either the cleavage or the demethylation pathways. It has been suggested that the expression of each pathway is dependent on the concentration of DMSPd and the bacterial sulphur demand (Kiene *et al*., 1999). In theory, a greater sulphur demand would lead to more DMSP being routed through the demethylation pathway, since sulphur can be assimilated from MeSH, but not DMS. Thus, a low concentration of DMSP and a high sulphur demand would increase expression of the demethylation pathway, whereas a high concentration of DMSP and a low sulphur demand would result in more DMSP being routed through the cleavage pathway. However, the empirical evidence to support this hypothesis is slim. A recent study sampled coastal waters for DMSP gene transcript abundance, while simultaneously measuring the presence of phytoplankton communities to try and identify the ecological and physiological factors affecting DMSP flux. The

abundance of *dmdA* and *dddP* transcripts from *Roseobacter* sp. HTCC2255 were measured at different time points for the month of October, in Monterey Bay, CA, USA. This work showed that *dddP* transcripts were more abundant when a mixed diatom and dinoflagellate community was present, whereas *dmdA* transcripts were greater when only dinoflagellates were recorded (Varaljay *et al*., 2015). Although a weak (non-significant) correlation, this finding supports the bacterial sulphur demand hypothesis if it is assumed that a greater variety of reduced sulphur compounds are present in the mixed community, reducing the dependency on DMSP demethylation for sulphur assimilation, and thus down-regulating this pathway. However, this study certainly has limitations, such as being conducted in one place, over a period of only one month, and as yet, no study has identified at a molecular level in a single organism the mechanism of regulation of each of the DMSP pathways. The fact that *dmdA* is usually linked to *acuI* in the Roseobacters is contradictory to the "bacterial switch", since these two genes, despite encoding steps in the separate pathways, are both regulated by *dmdR*. It may be possible that further work comparing regulation of the *dmdA* operon in *P. gallaeciensis* which is now distant from *dmdR*, to other Roseobacters, could shed some light on this problem.

6.4 *Oceanimonas doudoroffii* has multiple DMSP lyases

The first biochemical studies on DMSP catabolism by bacteria were carried out by Yoch's group in the 1990s using *Alcaligenes faecalis* and *Oceanimonas* (then *Pseudomonas*) *doudoroffii* (de Souza and Yoch 1995a,b). The work described in **Chapter 3** expanded this earlier work, and presented some interesting and surprising findings.

Although the original expectation had been that *O. doudoroffii* would have a copy of DddY, based on Yoch's purification and N-terminal sequencing of this enzyme (de Souza and Yoch 1996b), this enzyme could not be found, despite good coverage of the genome sequence (Curson *et al*., 2012). However, and unexpectedly, *O. doudoroffii* did have a functional copy of DddD, plus two functional DddP enzymes. As shown in this thesis, the occurrence of multiple DMSP lyases in a single strain is common to the Roseobacter clade, but it is rarely seen outside of this group. Indeed, this was the first example of a γ-proteobacterium possessing more than one gene for a DMSP lyase. Another marine γ-proteobacterium, *Leucothrix mucor* also has a copy of DddD and DddP, although no work has been done to confirm the functionality of either lyase (unpublished, genome available on NCBI database). As is the case for several other γ-proteobacteria, the *L. mucor dddD* is found in a cluster with homologues of *dddA, dddC, dddT* and *dddZ,* and so it would be interesting to see if this bacterium, like *O. doudoroffii,* can grow on DMSP as a sole carbon source.

Recently, two further strains of *Oceanimonas* have had their genomes sequenced - *Oceanimonas* sp. GK1 (Yeganeh *et al*., 2012) and *O. smirnovii* ATCC BAA-899 (Kyrpides *et al*., 2014). However,

neither of these strains have homologues to any known DMSP lyase. In addition, no other *Oceanimonas* species have been assayed for Ddd^+ activity, so currently it is not known how common DMSP catabolism is amongst this genus.

The benefit (if any) to organisms conferred by the presence of multiple DMSP lyases remains to be seen. One possibility is that each of the DMSP lyases is regulated differently, and may be expressed under particular environmental conditions. In *O. doudoroffii* I showed that only the DddD promoter, and not *dddP1* or *dddP2*, was responsive to the presence of DMSP, suggesting that this enzyme was responsible for the overall increase in Ddd^+ activity seen when cells were pre-grown in DMSP. Other studies have also shown this variance in regulation of different lyases. In some cases, the products of the DMSP cleavage reactions were found to induce lyase expression, for example *dddD* of *Halomonas* HTNK1 and *dddY* of *Alcaligenes faecalis*, which are both induced by 3HP and acrylate (Todd *et al*., 2010a; Curson *et al*., 2011). This form of regulation is unusual, but not unprecedented - a similar scenario is seen for the regulation of *myo*-inositol catabolism in *Rhizobium leguminosarum* (Fry *et al*., 2001). The earlier studies by Yoch's group showed that DMSP lyases from *A. faecalis* and *O. doudoroffii* required different pH and salt conditions for optimum activity (de Souza and Yoch, 1995b). Interestingly, *A. faecalis* lyase activity had a higher K_m for DMSP (2 mM) than *O. doudoroffii* (20 µM), perhaps reflecting normal DMSP concentrations in each organism's environment. *A. faecalis* was isolated from a salt marsh containing the grass *Spartina alterniflora* (de Souza and Yoch, 1995a), which produces high concentrations of DMSP (up to 250 µmol g^{-1} dry weight, Otte *et al*., 2004). Thus, *A. faecalis* likely encounters higher concentrations than *O. doudoroffii* which was isolated from the open ocean (Baumann *et al*., 1972). These studies show that different species may contain different DMSP lyases, optimal to their environments, which may explain the existence of multiple DMSP lyases. The presence of multiple DMSP lyases in a single organism is suggestive of a fluctuating environment, in which the organism may benefit from having multiple isofunctional enzymes that work optimally under different conditions. Fluctuating conditions are indeed present in marine environments. For example, the normal concentration of DMSP in open seawater is 1.2 nM, but this can exceed 1 mM in times of phytoplankton blooms, such is in the case of *Phaeocystis* (van Duyl *et al*., 1997) since some phytoplankton groups produce DMSP in high concentrations. For example, the coccolithophore *Emiliania huxleyi* is known to produce high intracellular DMSP concentrations of 50-250 mM (Steinke *et al*., 1998), and also forms massive seasonal blooms, sometimes in excess of 100,000 square kilometres (Brown and Yoder, 1994). During *E. huxleyi* blooms the concentrations of DMSPd fluctuate, and have been measured as increasing from 25 nM to 70 nM, then declining again as the bloom collapses (Levasseur *et al*., 1996). The decline in DMSPd was accompanied by an increase in bacterial activity and cell number, so it is likely that DMSP is being consumed by microorganisms. Since some bacteria, including *O. doudoroffii*, have substrate-

inducible DMSP lyases, it is likely that an increase in DMSP concentration primes the lyase pathway during algal blooms.

In addition to having multiple DMSP lyases, *O. doudoroffii* also has several *dddT*-like transporter genes, and these are located near to the lyase genes. Strikingly, there are no less than four of these BCCT-type transporters in the vicinity of dddD. Possessing so many betaine-type transporters is unusual, but this trait is also found in other DMSP-degrading bacteria, like *R. pomeroyi* which has five such systems, although none have yet been confirmed as *bona fide* DMSP transporters (Moran *et al*., 2004). It would be interesting to determine if the DddT homologues in *O. doudoroffii* are capable of DMSP transport, and, as with the multiple DMSP lyases, determine their relative importance to DMSP catabolism in *O. doudoroffii* in natural environments.

Perhaps one of the most puzzling findings from the work on *O. doudoroffii* was the absence of any *dddY* gene in the genome, which seemingly contradicts the findings of de Souza and Yoch (1996b), in which they purified a DMSP lyase from *O. doudoroffii* with an N-terminal sequence homologous to DddY from *A. faecalis.* Although the genome of *O. doudoroffii* was sequenced with approximately 98.5% coverage, there remains a slim chance that *dddY* was not represented amongst this. De Souza and Yoch reported similar properties of the purified lyases from *A. faecalis* and *O. doudoroffii* and showed that they were immunologically cross-reactive (de Souza and Yoch, 1996b). Therefore, a future project would be to use antibody raised against *A. faecalis* DddY to probe for the corresponding DMSP lyase in *O. doudoroffii*.

6.5 The discovery of novel DMSP lyase DddK

The work in **Chapter 2** was carried out in collaboration with Stephen Giovannoni's group from the Oregon State University, Corvallis. This group showed that the abundant SAR11 strain *Candidatus* Pelagibacter ubique HTCC1062 was able to produce methanethiol from DMSP, consistent with the presence of DmdA, but the first time that this had been demonstrated experimentally. Intriguingly, and excitingly, they also showed that this strain produced DMS from DMSP, yet it had no convincing homologues of known DMSP lyases in its deduced proteome. I showed that the cupin-domain containing protein designated DddK from HTCC1062 is able to cleave DMSP to DMS and acrylate, *in vitro*, albeit with a very high K_m for DMSP of ~50 mM. Although relatively high K_m values for DMSP-acting enzymes are not uncommon – indeed DmdA from strain HTCC1062 has a K_m of 13.2 mM (Reisch *et al*., 2008), more work needs to be done to fully characterise DddK. For example, it might be that the high K_ms of DddK and DddQ are a result of the expression and purification process. It is generally assumed that small Histidine-tags do not affect the properties of the enzyme, and,

indeed, this has repeatedly shown to be the case (Carson *et al*., 2006). However, there are some reports of native and His-tagged proteins behaving differently (Freydank *et al*., 2008), and so each enzyme purification should be considered on an individual basis. Another factor arising from the purification process that should be considered is the possibility of nickel ions from the affinity column binding to the native enzyme. This has been observed for the *E. coli* protein YodA, which bound nickel following purification using a nickel-nitrilotriacetic acid affinity column (David *et al*., 2003).

This is particularly relevant for DddK, since preliminary results showed that the metal-chelator EDTA reduced activity of the enzyme. This was expected, since metal-binding domains are conserved in the cupin-type enzymes (Dunwell *et al*., 2004) and studies on the structure of DddQ predicted that Zn^{2+} ions are required for its activity (Li *et al*., 2014). Therefore it is likely DddK has metal binding sites, which could become saturated with nickel, inhibiting binding of any other metal cofactors. The cofactor required by DddK is not yet known, but it would be of interest to identify, especially to see if additions of these cofactors to the assay buffer affect the K_m of the enzyme for DMSP.

Another reasonable explanation for the high K_m values is that DMSP is not major or "natural" substrate of DddK or DddQ. These enzymes may instead act on an analogue of DMSP such as dimethylsulfonioacetate or dimethylsulfoxide, which could still lead to the production of DMS, so these types of potential substrates should be tested in future studies on DMSP lyases. However, DddK and other lyases may also act on a substrate which did not yield DMS, in which case it would be very difficult to know which assay to use.

If DMSP is indeed not the natural substrate of DddK, it could mean that, while DddK is certainly able to cleave DMSP into DMS and acrylate, it is not the sole DMSP lyase in this strain, and another unknown lyase is present with a greater affinity to DMSP. In other DMSP-degrading organisms, mutations have been made to show that their DMSP lyase is solely responsible for DMSP-dependent DMS production (Todd *et al*., 2007; Curson *et al*., 2008; Curson *et al*., 2011). Ideally, a mutant strain of HTCC1062 lacking DddK would also be made, to understand its individual contribution to DMS production. However, microbes with streamlined genomes are often troublesome to grow, or completely escape cultivation, likely because they are so well adapted to the specific environment in which they reside (Giovannoni *et al*., 2014). This makes it very difficult to carry out thorough phenotyping and genetics on these strains. Since the SAR11 strains can only grow in liquid media (in sea water), it is not possible to make any mutations in genes of interest, and so we cannot yet verify DddK as the sole (or even major) DMSP lyase in this strain. However, a thorough screening of an HTCC1062 genomic library for Ddd^+ cosmids could be used to search for more candidate lyases in this strain.

Although streamlined organisms are difficult to work with, one of the most exciting revelations from the work on SAR11 is that such a tiny genome should still contain the genes required to make both DMS and MeSH from DMSP. This gives an indication of how important DMSP breakdown must be to these abundant ocean microbes, and, excitingly, SAR11 strains are not the only streamlined bacteria to contain homologues to DMSP-acting enzymes. Recent work in Georgia used single cell sequencing to obtain partial genome sequences of four uncultivated bacteria of the Roseobacter clade, and showed that, as with SAR11 strains, these cells also have small genomes, ranging between 2.6 and 3.5 Mb (Luo *et al.*, 2014). Significantly, all four cells had a good homologue to DmdA (36%-64% identical to SPO1913), and one, strain AAA076-CO3, also had a homologue of DddP (68% identical to ISM_05385, with an *E*-value of 0.0). These enzymes are yet to be verified for DMSP demethylase/lyase function, and the work on "DddP" in SAR11 strain HTCC7211 showed how important direct experimental testing of enzyme function is, since while PB7211_1082 had significant homology to ISM_05385 (35% identical), but it did not have DMSP lyase activity (see **Chapter 2**). Nevertheless, it is exciting that potential DMSP enzymes are also found in these streamlined genomes. Furthermore, the total coverage of each genome ranged between 23% and 76%, so there may be yet more homologues to known DMSP lyases which were not covered by the sequencing. The uncultivable organisms are of great importance, not least because they make up the vast majority of ocean bacteria. Furthermore, they often have reduced, streamlined genomes, like the SAR11 and Roseobacter strains mentioned above, and so the genes that become fixed in these cells are likely to be essential or extremely beneficial adaptations to their environment. Thus, DddK and DddQ may be very important enzymes to the SAR11 clade of bacteria, and possibly reflect the significance of DMSP as a nutrient source to ocean bacteria.

6.6 Abundance of DMSP genes in the environment

Throughout this thesis there has been a theme – the abundance and diversity of DMSP-related genes in bacteria. It is therefore apt to finish this discussion with a broader consideration of the global presence of these genes in the environment. Given that DMSP is found predominantly in marine environments, it is reassuring that homologues of *ddd* genes and *dmdA* are largely found in bacterial species that have been isolated from such environments. This is also reflected in metagenomic data. Consider the data presented in **Figure 6.3** – eight different sets of metagenomic data from different environments were interrogated for hits to DddW, DddL, DddQ, DddK, DddD, DddP and DmdA peptide sequences (Carrión *et al*., 2015). These environments ranged from terrestrial such as forest and grassland soils, and the rhizosphere and phyllosphere of rice, to marine open waters. It is immediately obvious that the DMSP genes differ significantly in their abundances in these different

locations. The most prolific peptides in the datasets examined here were DmdA and DddP, which were each most abundant in marine environments. In the three marine datasets (the Global Ocean Survey {GOS}, the North Atlantic spring bloom and the Monterey Bay coastal environment) DddP and DmdA were present in similar abundances, ranging from 10-15% of cells for the latter two datasets, and 25% for the GOS. This is interesting, since in the Roseobacter clade, DddP usually co-occurs with DmdA (see **Chapter 5**), and the similar abundance of these two genes in the marine metagenomic data possibly reflects this co-occurrence. The profusion of DddP and DmdA sequences over other DMSP lyases is likely due to their presence in some of the most prolific bacteria in the oceans. The majority of sequenced Roseobacter species have DddP and DmdA peptides, and this clade makes up 20% of coastal bacteria. Homologues of both peptides are also found in SAR11 strains (**Chapter 2**), and the SAR116 clade. Indeed, very recent work investigating the diversity of *dddP* in the North Western Pacific Ocean showed that the major *dddP*-containing bacteria in coastal waters were Roseobacters, but SAR116 strains dominated *dddP* bacteria in surface waters of the oligotrophic ocean. One strain, *Candidatus* Puniceispirillum marinum IMCC1322 had a Ddd^+ phenotype and a homologue of DddP was upregulated in the presence of DMSP (Choi *et al*., 2015).

Intriguingly, DddP is also fairly abundant in the soil metagenomes (2% of cells in the forest soil dataset, and 12% of cells in grassland soil). Indeed, homologues of DddP are found in some terrestrial species of bacteria and fungi (Todd *et al*., 2009). For example, within the Ascomycete, *Fusarium graminearum* PH1 and *Fusarium oxysporum* f. sp. *lycopersici* 4286 are pathogens of cereals and tomatoes, respectively, and both have a good homologue of DddP. The *dddP* genes from closely related *Fusarium* spp. have been cloned and shown to confer a Ddd^+ phenotype to *E. coli*. Sequences from organisms belonging to the Ascomycetes Order are present in both of the soil metagenomic databases, so it is possible that at least some of the hits to DddP in these environments are from fungal species. Despite the presence of some lyases in terrestrial environments, non-marine DMSP concentrations are minute and the production of DMS, while still significant, is much less than in marine environments (Yoch, 2002). Sources of terrestrial DMS include methoxylated aromatics and sulphide, and the methylation of methanethiol and degradation of DMSO (Bak *et al*., 1992; Kiene and Hines, 1995). However, despite the low concentrations, there is a precedent for DMSP-dependent DMS production in freshwater environments. Thus, Yoch *et al*. (2001) showed that river sediments, when incubated in minimal media with added DMSP, produced DMS. Additionally they isolated a Gram-positive bacterium of the family Nocardiaceae from the sediment that could grow on DMSP and had substrate inducible DMSP lyase activity (Yoch *et al*., 2001). Ironically, the first Ddd^+ bacterium to be isolated was also a freshwater species - *Clostridium* sp. was isolated from DMSP-enriched river mud samples. Incubation of the strain with DMSP resulted in the disappearance of DMSP, and the appearance of acrylate and DMS (Wagner and Stadtman, 1962). This Ddd^+ potential of species which do not encounter DMSP in significant concentrations is puzzling. It may be that

DMSP is not the natural substrate for the "lyase" enzyme in these organisms - a theory that could be explored further by testing the effect of DMSP inhibitors and DMSP analogues on DMS production in freshwater environments.

The other Ddd enzymes occur less frequently than DddP and DmdA. Indeed, DddW is only present in an estimated 0.1% of cells in the GOS, and returns no hits in any other dataset. This reflects the fact that only two sequenced species have a copy of DddW – *Ruegeria pomeroyi* DSS-3 and *Roseobacter* sp. MED193. Another gene which is not very abundant in the ocean datasets is DddD, with a frequency of only 0.4%. This is also not surprising, since amongst the sequenced strains of the Roseobacter clade, only 6 have a homologue to DddD, and this lyase is also not found in the abundant SAR11 clade. In fact the majority of DddD homologues occur in γ-proteobacteria isolated from salt marshes, or associated with marine eukaryotic organisms, rather than strains isolated from ocean waters. Intriguingly, DddD is more abundant in the Rice Phyllosphere and Rhizophere datasets (2.7% and 0.5%, respectively) than the ocean metagenomes. It has previously been shown that the N_2-fixing, symbiotic *Rhizobium* NGR234, which induces nodules on legumes, and the β-proteobacterium *Burkholderia cepacia* AMMD which is found on the roots of angiosperms each have a functional copy of DddD (Todd *et al*., 2007). Therefore it is not unprecedented that DddD should occur in strains which reside in terrestrial environments. However, both of these strains have an extremely wide host range, and so it could be those species with DddD enzymes which occur in the rice metagenomic datasets also have plant hosts which are capable of producing DMSP.

The cupin-type DMSP lyase enzymes DddL, DddQ and the newly found DddK, are all present in the GOS, at frequencies of 1.4%, 5.9% and 5.3%, respectively. Homologues of DddK are also present in the North Atlantic spring bloom and Monterey Bay coastal environment datasets. Neither DddK nor DddQ are found in the non-marine datasets, in keeping with the absence of these enzymes from any sequenced non-marine strains. Amongst sequenced strains, DddK is restricted to a few members of the SAR11 clade (see **Chapter 2**), but since members of this clade are among the most abundant organisms on the planet it is reasonable that DddK, and the other DMSP enzymes found in this clade, (DddP, DddQ and DmdA) are the most frequent DMSP peptides in ocean metagenomes. The GOS data for DddL shows that this gene occurs at a slightly lower frequency than DddQ and DddK, with an abundance of 1.4%. This lower frequency could be explained by the fact that while DddL is present in some Roseobacters, it has not been found in any sequenced SAR11 strain.

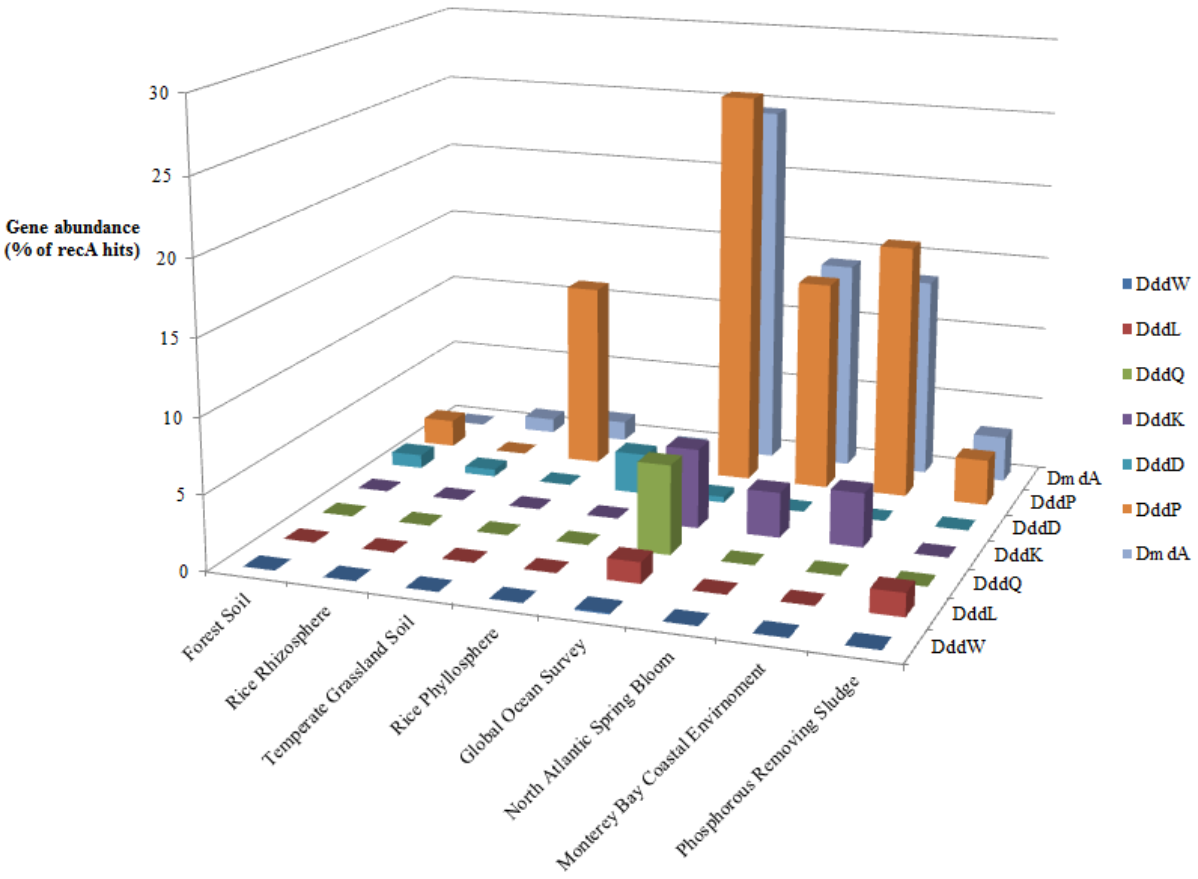


Figure 6.3 Abundance of DMSP genes in different environments. Metagenomic datasets were probed using peptide sequences of functionally ratified enzymes, normalized to the number of unique RecA sequences. Figure adapted from Carrión *et al*, 2015.

6.7 Final remarks

It is clear that significant progress has been made over the last decade in identifying the molecular basis of bacterial DMSP catabolism, and hopefully the work described in this thesis further contributes to our overall knowledge of this field. I have shown that the presence of multiple DMSP lyases is not restricted to the α-proteobacteria through the discovery of DddD, and two DddPs in *O. doudoroffii*, and that, in keeping with many other DddD-containing γ-proteobacteria, *O. doudoroffii* can use DMSP as a sole carbon source. In contrast, this growth phenotype is unusual amongst the α-proteobacteria. However, I, and others, showed that a few Roseobacters including *R. pomeroyi* can use DMSP and acrylate as carbon sources, and independently identified a route of DMSP metabolism in this organism. Additionally, I presented a detailed analysis of the presence and synteny of DMSP-related genes in the Roseobacter clade, identifying some interesting patterns, which warrant further experimental studies. Finally, I have shown through the identification of DddK in one of the most abundant organisms on the planet, that there are still likely to be novel DMSP lyases yet to be discovered (although the fact DddK is another cupin-type lyase suggests that we may be approaching the limit of completely new DMSP enzymes). In addition to expanding our knowledge of DMSP catabolism in microbes, the work in this thesis has certainly highlighted the sheer importance of DMSP to marine bacteria, from the diversity of enzymes and pathways involved, to the retention of multiple DMSP genes in very streamlined, tiny genomes. Given the importance of DMSP, and the diversity of enzymes involved, the next step is surely to gain a much better understanding of the regulation of these different pathways, and their individual contributions to the release of important sulphurous gases.

Chapter 7

Materials and Methods

Table 7.1 Bacterial strains used in this study

Strain	Characteristics	Source
Escherichia coli 803	Met^-; used for transformation of large plasmids.	Wood, 1966
E. coli BL21	Used for over-expression of proteins.	Yanisch-Perron *et al*., 1985
E. coli JM101	$LacZ^-$; used for transformation of small plasmids.	Messing, 1979
E. coli A118	Contains chromosomally located copy of Tn*5lacZ*; used for transposon mutagenesis; Kan^R	Simon *et al*., 1989
Rhizobium leguminosarum J391	Wild type strain; $Strep^R$ mutant	Young *et al*., 2006
Oceanimonas doudoroffii DSM 7028	Wild type strain	DSMZ, Braunschweig
J495	*O doudoroffii* DSM 7028; Rif^R mutant	Curson *et al*., 2012
Ruegeria pomeroyi DSS-3	Wild type strain	González *et al*., 2003
J470	*Ruegeria pomeroyi* DSS-3; Rif^R mutant	Todd *et al*., 2010b
Roseobacter denitrificans OCh 114	Wild type strain	Shiba, 1991
Dinoroseobacter shibae DFL-12	Wild type strain	Biebl *et al*., 2005
Pelagibaca bermudensis HTCC2601	Wild type strain	Cho and Giovannoni, 2006
Phaeobacter inhibens (previously *Roseobacter gallaeciensis*) DSM17395	Wild type strain	Ruiz-Ponte *et al*., 1998; Martens *et al*., 2006
Roseobacter litoralis OCh 149	Wild type strain	Shiba, 1991

Roseovarius nubinhibens ISM	Wild type strain	González *et al*., 2003
Roseovarius sp. 217	Wild type strain	Schäfer *et al*., 2005
Ruegeria sp. Trich-CH4B	Wild type strain	Roe *et al*., 2012
J527	J470 with pK19 insertional mutation in *acuI*	Todd *et al*., 2012b
J471	J470 with pK19 insertional mutation in *dmdA*	Todd *et al*., 2012b
J559	J470 with pK19 insertional mutation in SPO1094	This work
J560	J470 with pK19 insertional mutation in SPO1101	This work
J561	J470 with pK19 insertional mutation in SPO1105	This work
J562	J470 with pK19 insertional mutation in SPO2934	This work
J563	J470 with pK19 insertional mutation in SPO1912	This work

Table 7.2 Plasmids used in this study

Plasmid	Characteristics	Source
pRK2013	Used for mobilising plasmids in bacterial conjugation: Kan^R	Figurski and Helinski, 1979
pBIO1879	Spc^R derivative of suicide plasmid pK19*mob*, used for insertional mutagenesis: Spc^R; Kan^R	Todd *et al*., 2010a
pET16b	Overexpression plasmid, T7 promoter, N-terminal His-tag; Amp^R.	*Novagen*
pET21a	Overexpression plasmid, T7 promoter, Optional C-terminal His-tag; Amp^R	*Novagen*
pRK415	Wide host range plasmid cloning vector; Tet^R	Keen *et al*., 1988
pMP220	Wide host-range *lacZ* reporter plasmid; Tet^R	Spaink *et al*., 1987
pBIO1878	Wide host-range *lacZ* reporter plasmid, based on pMP220; Spc^R, Tet^R	Todd *et al*., 2012a
pBluescript M13	M13 phagemid, T7 T3, *lacZ lacI* Amp^R	Short *et al*., 1988
pLAFR3	Wide host-range cosmid	Staskawicz *et al*., 1987
pBIO1932	pLAFR3 containing J495 genomic DNA, including *dddD*	Curson *et al*., 2012
pBIO1930	pLAFR3 containing J495 genomic DNA, including *dddP1*	Curson *et al*., 2012
pBIO1931	pLAFR3 containing J495 genomic DNA, including *dddP2*	Curson *et al*., 2012
pBIO1933	pET21a containing intact *O. doudoroffii dddP1*	Curson *et al*., 2012
pBIO1934	pET21a containing intact *O. doudoroffii dddP2*	Curson *et al*., 2012
pBIO1951	pMP220-based *O. doudoroffii dddP1-lacZ* transcriptional fusion plasmid; Tet^R	This work
pBIO1952	pMP220-based *O. doudoroffii dddD-lacZ* transcriptional fusion plasmid; Tet^R	This work
pBIO1953	pMP220-based *O. doudoroffii* $dddT^{D1}$-*lacZ* transcriptional fusion plasmid; Tet^R	This work

pBIO1954	pMP220-based *O. doudoroffii* $dddT^{P2}$-*lacZ* transcriptional fusion plasmid; Tet^R	This work
pBIO1955	pMP220-based *O. doudoroffii* $hcaE^{P2}$-*lacZ* transcriptional fusion plasmid; Tet^R	This work
pBIO1958	pMP220-based *O. doudoroffii dddP2-lacZ* transcriptional fusion plasmid; Tet^R	This work
pBIO2007	pK19 containing internal fragment of SPO1094	This work
pBIO2037	pLAFR3 containing 24 kb contiguous fragment of J470 genome, spanning the chromosomal SPO1087-SPO1110 genes*	This work
pBIO2044	pK19 containing internal fragment of SPO1105	This work
pBIO2049	pK19 containing internal fragment of SPO1101	This work
pBIO2093	pBluescript containing intact SPO2528	This work
pBIO2094	pBluescript containing intact SPO2934	This work
pBIO2095	pBluescript containing intact SPO1014	This work
pBIO2096	pK19 containing internal fragment of SPO2934	This work
pBIO2204	pET16b-derivative containing intact HIMB5_00000220 from α-proteobacteria HIMB5.	This work
pBIO2207	pET16b-derivative containing intact PB7211_1082 from *Candidatus* Pelagibacter ubique HTCC7211	This work
pBIO2206	pET16b-derivative containing intact SAR11_0394 from *Candidatus* Pelagibacter ubique HTCC1062	This work

*; SPO refers to gene tags in the *Ruegeria pomeroyi* DSS-3 genome

7.1 Media and Growth Conditions

Media for bacterial growth were prepared as shown below. All media and glassware were sterilized in by autoclaving at 121°C. *Escherichia coli* strains were grown at 37°C and all Roseobacter strains and *Rhizobium leguminosarum* were grown at 28°C. Solid plate cultures were grown on 1.5% agar (*Formedium*). Liquid cultures were incubated with shaking at 200 rpm. All 5 ml cultures were grown in 20 ml glass universals with plastic screw caps, and 100 ml cultures were grown in 250 ml conical flasks, sealed by a foam bung and topped with aluminium foil.

Note: * as shown below, indicates where a constituent was added after autoclaving, from a sterile stock solution.

7.1.1 Lysogeny Broth

Lysogeny broth (LB) (Maniatus *et al*., 1982) was used for the routine growth of *E. coli* strains and contained, per litre of dH_2O:

5 g NaCl
10 g Tryptone
5 g Yeast Extract
1.5 g Glucose

The pH was adjusted to 7.2 using HCl or NaOH

7.1.2 M9 Minimal Medium

M9 minimal medium was used for growth of *E. coli* under defined conditions (Maniatus *et al*., 1982) and contained per litre of dH_2O:

12.8 g $Na_2HPO_4.7H_2O$
3 g KH_2PO_4
0.5 g NaCl
1 g NH_4Cl
1 ml 0.1 M $CaCl_2$*
2 ml 1 M $MgSO_4$*
1 ml 30 mg/ml thiamine-HCl*
1 ml 30 mg/ml methionine*

7.1.3 ½YTSS Medium

Yeast extract, Tryptone and Sea Salts (½YTSS) medium was used for the routine growth of *R. pomeroyi* and all other Roseobacter strains (González *et al*., 2003) and contained, per litre dH_2O:

1.25 g Tryptone
2 g yeast extract
20 g Sea Salts (*Sigma*)

The pH was adjusted to 7.0 using HCl.

7.1.4 Marine Basal Medium

Marine basal media was used for growing *R. pomeroyi* and other Roseobacter strains (González *et al*., 1997) under defined conditions and contained per litre dH_2O:

20 g Sea Salts (*Sigma*)
71.43 ml 1 M Tris (pH 7.5)
41.4 mg K_2HPO_4
710 mg NH_4Cl
25 mg Ethylenediamine tetra-acetic acid; ferric-sodium salt (Fe EDTA)*
1 ml MBM vitamin stock solution*

MBM vitamin stock

Per 100 ml:
2 mg biotin
2 mg folic acid
10 mg pyridoxine
5 mg riboflavin
5 mg thiamine
5 mg nicotinic acid
5 mg pantothenic acid
0.1 mg cyanocobalamin
5 mg p-aminobenzoic acid

The pH was adjusted to 7.0 using HCl.

7.1.5 TY Medium

Tryptone Yeast (TY) medium was used for the routine growth of *R. leguminosarum* (Beringer, 1974), and contained per litre of dH_2O:

0.9 g $CaCl_2.2H_2O$
3 g yeast extract
5 g Tryptone

The pH was adjusted to 6.8 using HCl or NaOH

7.1.6 *Rhizobium* Minimal Media

Rhizobium minimal (RM) media (Beringer, 1974) was used for growth of *R. leguminosarum* under defined conditions and contained per litre of dH_2O:

3 g Tris(hydroxymethyl)aminomethane HCl
0.1 g $MgSO_4.7H_2O$
0.22 g $CaCl_2.6H_2O$
0.22 g K_2HPO_4
0.02 g $FeCl_3$
0.75 mg biotin
0.75 mg thiamine
0.75 mg DL-pantothenic acid Ca salt
10 ml 1 M NH_4Cl*

The pH was adjusted to 6.8 using HCl or NaOH

Table 7.3 Supplements to media

Antibiotic	**Solvent**	**Final Concentration ($\mu g\ ml^{-1}$)**
Ampicillin	70% ethanol	100
Tetracycline	70% ethanol	5
Rifampicin	Methanol	20
Spectinomycin	dH_2O	200
Streptomycin	dH_2O	400
Kanamycin	dH_2O	20
Gentamicin	dH_2O	5
Supplement	**Solvent**	**Final Concentration ($\mu g\ ml^{-1}$)**
Isopropyl β-D-1-thiogalactopyranoside (IPTG)	dH_2O	200
5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-gal)	2,2-dimethylformamide	40
Ortho-nitrophenyl-β-galactoside (ONPG)	dH_2O	800
Carbon source	**Solvent**	**Final concentration (mM)**
Succinate	dH_2O	10
Glycerol	dH_2O	10
Dimethylsulphoniopropionate (DMSP)	dH_2O	5
Acrylate	dH_2O	2
Methylmercaptopropionate (MMPA)	dH_2O	2
Propionate	dH_2O	5
3-hydroxypropionate (3-HP)	dH_2O	5
Dimethylsulphide (DMS)	-	1
Methanethiol (MeSH)	dH_2O	1

7.2 Strain Storage

Strains were grown to stationary phase in appropriate complete media (LB, TY or ½YTSS as appropriate), and aliquots of culture were mixed with 25% (v/v) glycerol. Cells were flash frozen and stored at -80°C. All Roseobacter strains were stored in the presence of 15% (v/v) Dimethylsulphoxide (DMSO) as a cryoprotectant (González *et al*., 2003).

7.3 Nucleic Acid Preparations

7.3.1 Plasmid preparation by alkaline lysis and phenol chloroform extraction

For low yield plasmid preparations, an alkaline lysis-phenol chloroform extraction method was performed, using buffers provided with the *Qiagen* Plasmid Midi Kit, and according to the following protocol:

1) A single colony of *E. coli* or *R. leguminosarum* containing plasmid DNA was inoculated into 5 ml rich media and grown overnight in the presence of appropriate antibiotics.
2) Approximately 1.5 ml of culture was harvested and pelleted by centrifugation at 13,000 rpm for 2 minutes.
3) The supernatant was discarded and the cell pellet was resuspended in 250 µl ice-cold P1 Resuspension Buffer (50 mM Tris-Cl, pH 8.0; 10 mM EDTA; 100 µg/ml RNase A).
4) Next, 250 µl Lysis Buffer (200 mM NaOH; 1% SDS (w/v)) was added, mixed by inversion and samples were incubated at room temperature for 5 minutes, after which 350 µl ice-cold neutralization buffer (3 M potassium acetate, pH 5.0) was added.
5) Samples were mixed by inversion, left on ice for 15 minutes, and centrifuged at 13,000 rpm for 15 minutes to pellet precipitated proteins and cell debris.
6) The resultant supernatant containing nucleic acids (*ca*. 700 µl) was removed to a clean microfuge tube.
7) Next, 400 µl phenol chloroform was added (phenol:chloroform:isoamyl alcohol 25:24:1 {v/v}, *Sigma*) and samples were vortexed for 10 seconds.
8) Samples were centrifuged at 13,000 rpm for 2 minutes, and then the upper aqueous layer (about 700 µl) was removed to a clean microfuge tube containing an equal volume of 100% ethanol.

9) Samples were mixed by inversion, and centrifuged at 13,000 rpm for 15 minutes to pellet precipitated plasmid DNA.
10) The supernatant was removed and discarded and pellets were washed in 500 µl 70% ethanol, then pelleted by centrifugation at 13,000 rpm for 2 minutes.
11) The ethanol was completely removed and pellets were rehydrated in 50 µl dH_2O.
12) Plasmid preparations were stored at -20°C.

7.3.2 Plasmid preparation using *Qiagen* midi-prep columns

High yield plasmid preparations were performed using a *Qiagen* Plasmid Midi Kit, using the buffers provided and according to the following protocol:

1) A single colony of *E. coli* containing plasmid DNA was inoculated into 100 ml LB and grown overnight in the presence of appropriate antibiotics.
2) Cells were transferred to 50 ml Falcon tubes and pelleted by centrifugation at 6,000 rpm for 10 minutes.
3) Cells were resuspended in 4 ml ice cold P1 Resuspension Buffer, and then mixed with 4 ml P2 lysis buffer by gentle inversion.
4) Samples were incubated at room temperature for 5 minutes and then 4 ml ice cold P3 Neutralization Buffer (3 M potassium acetate, pH5.0) was added.
5) Samples were incubated on ice for 15 minutes, and then centrifuged at 10,000 rpm for 45 minutes to pellet precipitated proteins and cell debris.
6) A QIAGEN-tip 100 column was equilibrated using 4 ml QCF Equilibration Buffer and the supernatant from step 5 was loaded onto the equilibrated column and allowed to drip through.
7) Once the supernatant had passed through the column, 10 ml of QC Wash Buffer was loaded and allowed to drip through.
8) This step was repeated, and then the plasmid DNA was eluted from the column into a sterile plastic universal container using 5 ml QF Elution Buffer.
9) To precipitate plasmid DNA, 3.5 ml isopropanol was added to the eluate and mixed by inversion. The sample was then aliquoted into 6 microfuge tubes, which were centrifuged at 13,000 rpm for 30 minutes to pellet plasmid DNA.
10) The supernatant from each tube was discarded, and pellets were washed in 500 µl 70% ethanol and combined into one tube.
11) The combined pellets were centrifuged for 2 minutes at 13,000 rpm, and the ethanol supernatant was removed.

12) The pellet was left to air-dry for 15 minutes, and then rehydrated using 100 µl dH_2O.
13) Plasmid preparations were stored at -20°C.

7.3.3 Genomic DNA preparations

Genomic DNA preparations were carried out using the *Promega* Wizard® Genomic DNA Purification Kit according to the following protocol:

1) A single colony of the desired strain was used to produce an overnight culture.
2) Of that culture, 1 ml was added to a 1.5 ml microfuge tube and centrifuged at 13,000 rpm for 2 minutes to pellet cells.
3) Cells were resuspended in 600 µl Nuclei Lysis Solution, and incubated at 80°C for 5 minutes.
4) Lysed cells were cooled to room temperature, then 3 µl RNase Solution was added, mixed by inversion and incubated at 37°C for 1 hour.
5) Next, 200 µl of Protein Precipitation Solution was added to the RNase-treated cell lysate, and mixed vigorously by vortexing.
6) Samples were incubated on ice for 5 minutes, then centrifuged at 13,000 rpm for 3 minutes to pellet protein precipitate and cell debris.
7) The supernatant was transferred to a 1.5 ml tube containing 600 µl isopropanol, and mixed gently by inversion until threads of precipitated DNA became visible.
8) The samples were centrifuged at 13,000 rpm for 2 minutes to pellet DNA.
9) The supernatant was discarded and the DNA pellet washed with 600 µl 70% ethanol
10) The samples were then centrifuged at 13,000 rpm for 2 minutes, and the 70% ethanol was completely removed.
11) The DNA pellet was rehydrated in 100 µl dH_2O and incubated overnight at 4°C.
12) Rehydrated DNA preparations were stored at -20°C.

7.3.4 Quantification of nucleic acid preparations

All DNA preparations were quantified using absorbance readings at 260 nm as measured by a NanoDrop 2000 (*Thermo Scientific*). The purity of DNA preparations was determined using ratios of absorbances at 260/280 nm for protein contamination, and 260/230 nm for solvent contamination. A 260/280 ratio of *ca*. 1.8 and a 260/230 ratio in the range of 2.0-2.2 were considered acceptable.

7.4 Transfer of genetic material

7.4.1 Transformations

7.4.1.1. Preparation of competent E. coli *cells for transformation:*

1) An overnight starter culture of the desired *E. coli* strain was used to inoculate 100 ml LB (1:100).
2) The culture was incubated at 37°C, shaking at 200 rpm until the cells reached early exponential growth phase.
3) Cells were harvested by centrifugation at 6,000 rpm for 5 minutes, and the supernatant was removed.
4) Cell pellets were resuspended in 15 ml ice cold 0.1 M $CaCl_2$, and incubated on ice for 30 minutes.
5) Cells were then centrifuged as above, and the supernatant removed.
6) Cells were finally resuspended in 2 ml ice-cold 0.1 M $CaCl_2$ and stored at 4°C overnight before use.

7.4.1.2 Transformation using heat shock:

1) Plasmid DNA was mixed with 100 µl competent cells in a 1.5 ml microfuge tube and incubated on ice for 1 hour.
2) The cells were heat-shocked at 42°C for 3 minutes, and returned to ice for 1 minute, before the adding 500 µl LB medium.
3) The cells were then left to express antibiotic resistance genes at 37°C for 30-45 minutes before spreading onto appropriate selective media.

7.4.2 Bacterial Conjugations

The conjugational transfer of plasmids from *E. coli* to other species of bacteria was carried out via tri-parental mating using a patch cross (Johnston *et al*., 1978) or a filter cross (Beringer and Hopwood, 1976). *E. coli* 803 containing the conjugation helper plasmid pRK2013 was used to mobilize non-self-transmissible plasmid DNA from the host *E. coli* strain to the desired recipient.

7.4.2.1 Patch Cross

The donor, helper and recipient strains were cultured on agar plates containing appropriate antibiotics prior to patch crossing. A loopful of each strain was mixed together onto the surface of an agar plate with appropriate media for the recipient strain. Negative controls using just two

of the parent strains were also performed. The bacteria were incubated at 28°C for 2 days, before streaking out onto fresh, selective media. Selective plates were incubated at 28°C until single colonies appeared.

7.4.2.2 Filter Cross

For rare conjugation events, such as insertional mutagenesis, crosses were carried out nitrocellulose filters (*Whatman*). Overnight liquid cultures of the *E. coli* donor strain, the helper *E. coli* strain and the recipient were prepared. Then 500 µl of each of the *E. coli* strains was mixed with 1 ml of the recipient strain. The mixture was then centrifuged at 13,000 rpm for 2 minutes to pellet cells and all but 100 µl of the supernatant was removed. The cells were resuspended in the remaining 100 µl and then spread onto a 47 mm sterile nitrocellulose filter (*Whatman*) that was placed on a plate with medium that supported the growth of all three strains. After 2 day's incubation at 28°C, the filters were then removed to a sterile plastic universal, 1 ml of minimal media was added to the tube and the bacteria washed from the filter by vortexing. Serial dilutions were performed on the washed cells and spread onto selective media. Selective plates were incubated at 28°C until single colonies appeared.

7.5 Polymerase Chain Reaction (PCR) amplification of DNA

Amplification of DNA by PCR was carried out in a Techne TC-512 PCR machine. The reactions were set up as follows:

1 µl DNA template (*ca*. 50 ng/µl)
10 µl *Qiagen* Mastermix, containing Taq DNA polymerase, buffer and dNTPs
1 µl forward primer (20 pmol/µl)
1 µl reverse primer (20 pmol/µl)
7 µl dH_2O

Table 7.4 PCR cycle used for DNA amplification

Step	No. of cycles	Function	Temperature	Time (s)
1	1	Initial denaturation	95°C	300
2	30	Denaturation	95°C	50
		Annealing	53°C	30
		Extension	72°C	60
3	1	Final extension	72°C	300

7.6 Purification of PCR Products

PCR products were purified using a High-Pure PCR product purification kit (*Roche*), according to manufacturer's instructions. All buffers were supplied with the kit, and the protocol was as follows:

1) The PCR product was added to 500 µl Binding Buffer and mixed by inversion.
2) This mixture was loaded onto a High-Pure spin filter tube, and this was centrifuged at 13,000 rpm for 1 minute.
3) Flow-through from the filter was discarded, 500 µl Wash Buffer was applied to the filter and centrifuged for 1 minute.
4) Flow-through was discarded, and the filter centrifuged for 1 minute to remove any residual buffer.
5) The column was transferred to a clean 1.5 ml tube and 30 µl dH_2O was applied to the filter and allowed to soak for 2 minutes, before centrifugation for 1 minute to elute the clean PCR product.

Purified PCR products were stored at -20°C until needed.

7.7 Restriction enzyme digestion of DNA

Restriction enzymes were purchased from *Roche* or *Promega* and used according to manufacturer's instructions, with their supplied buffers. Where two different enzymes were used in the same reaction, buffers were checked for compatibility with both enzymes. The buffer leading to the most efficient digestion by both enzymes was chosen. Typically, 200 ng PCR product or plasmid DNA was digested in a 20 µl reaction, using 2 µl 10x buffer and 1 µl enzyme. Digestions were incubated at 37°C for 2 hours, unless otherwise specified. Enzyme reactions were stopped by heating at 80°C for 10 minutes.

7.8 Alkaline Dephosphorylation

To reduce the likelihood of plasmid DNA re-ligating, digested plasmids were treated with rAPid alkaline phosphatase (*Roche*) according to manufacturer's instructions. For every 20 µl for plasmid digestion, 2.5 µl alkaline phosphatase buffer, 1 µl alkaline phosphatase and 1.5 µl dH_2O was added. The reaction was incubated at 37°C for 1 hour and the enzyme inactivated by heating at 80°C for 10 minutes.

7.9 Ligation Reactions

Digested insert and plasmid DNA were ligated to form recombinant plasmids using T4 DNA ligase (*Roche*), according to manufacturer's instructions. In a 20 µl reaction, 2 µl 10x T4 ligase buffer was used, with 1 µl T4 ligase. Typically, the remaining 17 µl comprised a 6:1 ratio of insert:plasmid DNA, although in some cases this ratio varied. Ligation reactions were incubated at 4°C overnight.

7.10 DNA Gel Electrophoresis

DNA fragments were separated by electrophoresis using 1% (w/v) agarose gels containing TAE buffer [40 mM Tris, 20 mM glacial acetic acid, 1 mM EDTA (pH 8.0)] and 1 µg/ml ethidium bromide. To analyse very small DNA fragments, 2.5% agarose was used. DNA samples were mixed with 0.2 volumes 5 x loading dye [0.25% bromophenol blue (v/v), 30% glycerol (v/v)]

prior to loading. Electrophoresis was carried out in SCIE-PLUS HU13/HU6 horizontal tanks, at 80V for 1-2 hours with TAE as running buffer, and 1 kb Plus ladder (*Invitrogen*).

7.11 Gel extractions

DNA fragments were extracted from agarose gels using a QIAquick Gel Extraction Kit (*Qiagen*), according to the following protocol:

1) The desired DNA fragment was cut from the gel using a clean scalpel, and incubated at 50°C in 300 µl QG buffer for approximately 10 minutes, to dissolve the agarose gel.
2) The sample was then applied to a QIAquick column and centrifuged at 13,000 rpm for 1 minute.
3) The flow-through was discarded and 500 µl QG buffer was added to the column, and centrifuged for 1 minute.
4) The flow-through was discarded once more, and 750 µl PE buffer containing was applied to wash the column.
5) The column was centrifuged for 1 minute, the flow-through was discarded, and then the column was centrifuged for 1 minute to remove any residual buffer.
6) To elute the DNA, the column was placed in a clean 1.5 ml tube, and 30 µl dH_2O was added to the column. This was allowed to soak for 2 minutes before centrifuging for 1 minute.

DNA samples were stored at -20°C until needed.

7.12 Plasmid integration mutants

Primers were designed to PCR-amplify a 500-1000 bp internal fragment of the gene to be mutated, prior to its ligation into pBIO1879, which can be mobilised from *E. coli* into many Gram negative bacteria but which cannot replicate in non-Enteric hosts (Schäfer *et al*., 1994). Recombinant pK19-based plasmids were transferred by conjugation from *E. coli* 803 were mobilised into Rif^R *R. pomeroyi* J470 via a tri-parental filter cross (see **Section 7.4.2.2**), with *E. coli* 2013 as a helper strain. Since pBIO1879 is unable to replicate in *R. pomeroyi*, colonies growing on rifampicin, spectinomycin and kanamycin are candidates for those transconjugants in which the recombinant pBIO1879 had integrated into the genome via a single-crossover

homologous recombination event. Potential mutants were verified using Southern blot analysis, as described in **Section 7.15**.

7.13 Protein Over-expression and Purification

7.13.1 Over-expression of proteins in a heterologous host

Genes encoding proteins of interest were cloned into the plasmid expression vectors pET16a or pET21b, and then transformed into *E. coli* BL21. Fresh transformants were used to inoculate 5 ml LB containing ampicillin, and incubated at 37°C, shaking at 200 rpm, for 3-4 hours. Once cultures reached early exponential phase, a suitable concentration of Isopropyl β-D-1-thiogalactopyranoside (IPTG) was added to induce expression of genes under the control of the T7 promoter. The cultures were then incubated at 28°C, shaking at 200 rpm, overnight.

The cells were then pelleted and resuspended in an appropriate buffer. They were put on ice, and then lysed by sonication for 4 x 10 second bursts. Sonicated samples were centrifuged for 2 minutes at 13,000 rpm to pellet cell debris and the resultant supernatant, was removed to a clean tube. Soluble fractions were stored at 4°C until needed.

7.13.2 Separation of proteins by SDS-PAGE

Proteins were separated for analysis using polyacrylamide gel electrophoresis (PAGE), with sodium dodecyl sulphate (SDS) used to provide denaturing conditions. Gels were prepared according to Sambrook *et al*. (1989). A 15% resolving gel was prepared (see below) and immediately poured between two glass plates, sealed with a rubber gasket, to a level 2 cm from the top. The gel was levelled using H_2O, allowed to polymerise, and then topped up with a 6% stacking gel (see below). A 12-well comb was placed into the stacking gel before it polymerised.

Samples were mixed with SDS-PAGE loading buffer (see below) and were run alongside Precision Plus Protein™ Dual Color Standard (*Bioline*).

Gels were run in vertical tanks (ATTO AE-6450) at 150 V for 2 hours in PAGE running buffer [25 mM Tris, 200 mM glycine, 0.1% SDS (w/v)]. Gels were then removed from glass plates and stained with InstantBlue™ (*Expedeon*) for 15 minutes.

Resolving gel

5 ml 4 x resolving buffer [1.5 M Tris pH 8.8, 0.4% SDS (w/v)].

5 ml dH_2O

10 ml Protogel 30% acrylamide solution

250 µl 10% APS (w/v)

25 µl Tetramethylethylenediamine (Temed)

Stacking gel

2.5 ml 4 x stacking buffer [0.5 M Tris pH 6.8, 0.4% SDS (w/v)].

5.5 ml dH_2O

2 ml Protogel 30% acrylamide solution (*National Diagnostics*)

100 µl 10% ammonium persulphate (APS) (w/v)

20 µl Temed

SDS-PAGE loading buffer (4 x stock)

2 ml Tris HCl (1 M, pH 6.8)

0.8 g SDS

4 ml glycerol

0.4 ml β-mercaptoethanol

1 ml EDTA (0.5 M)

8 mg Bromophenol blue

3 ml dH_2O

7.13.3 Purification of His-tagged proteins

Histidine-tagged proteins were purified from *E. coli* BL21 using a Ni-NTA spin column (*Qiagen*). Buffers were prepared as follows:

Lysis Buffer (NPI-10)

50 mM sodium phosphate

300 mM sodium chloride

10 mM imidazole

pH 8.0

Wash Buffer (NPI-30)
50 mM sodium phosphate
300 mM sodium chloride
30 mM imidazole
pH 8.0

Wash Buffer (NPI-300)
50 mM sodium phosphate
300 mM sodium chloride
300 mM imidazole
pH 8.0

The purification protocol was as follows:

1) A 50 ml culture of *E. coli* containing over-expressed protein was prepared in 1.4 ml Buffer NPI-10.
2) An Ni-NTA spin column was equilibrated using 600 µl Buffer NPI-10, then centrifuged at 2,900 rpm for 2 minutes
3) The soluble fraction was loaded onto the equilibrated column and centrifuged at 1,600 rpm for 5 minutes. This step was carried out twice in order to load the entire sample. Each time, the flow-through was kept for SDS-PAGE analysis.
4) The Ni-NTA column was washed three times using 600 µl Buffer NPI-30 with centrifugation at 2,900 rpm for 2 minutes after each step. Each flow-through was kept for SDS page analysis.
5) The protein was eluted from the column using 300 µl Buffer NPI-300 with centrifugation at 2,900 rpm for 2 minutes. This step was carried out twice to remove any residual protein.
6) Eluted protein was stored at 4°C.

7.14 Bradford's Assay

Protein concentrations were estimated using Bradford's assays (Bradford, 1976). An appropriate volume of sample was added to dH_2O to a total volume of 800 µl. Then, 200 µl Bradford's reagent (*Bio-Rad*) was added and mixed by inversion. A standard curve was created using bovine serum albumin at concentrations of 0, 10, 20 or 40 mg/ml, plotting their absorbances at 595 nm (*Jenway Genova* spectrophotometer). This curve was used to calculate the concentration of unknown samples based on their absorbance at 595 nm.

7.15 Southern Blot

7.15.1 Probe design

In order to verify pK19 insertional mutations, PCR products of the intact gene of interest were used as probes against genomic DNA of the putative mutant strains.

7.15.2 Digestion and gel electrophoresis of samples

Genomic DNA preparations of wild type and mutant strains were digested using suitable restriction enzymes that were predicted to produce diagnostic fragments. Digested samples were then loaded onto a 1% agarose gel and run at 80V for 2-3 hours, alongside a 1 kb Plus ladder (*Invitrogen*). Gel images under UV light were captured alongside a ruler for reference.

7.15.3 Preparation of the gel

The agarose gel containing digested samples was washed in depurination buffer (0.2 M HCl), following by denaturation buffer (0.5 M NaOH, 1.5 M NaCl) and finally neutralistion buffer [1 M Tris-HCl (pH 8.0), 1.5 M NaCl]. Each wash step lasted 20 minutes, and the gel was rinsed in dH_2O after each wash.

7.15.4 Blot Assembly

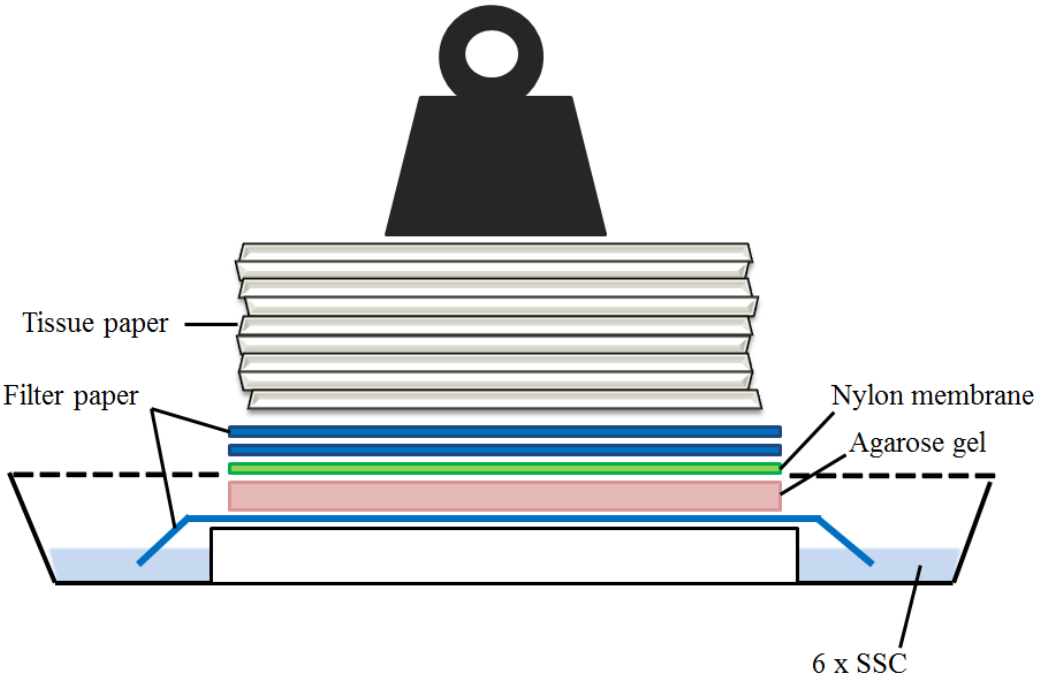


Figure 7.1 Southern blot assembly. Representation of the Southern blot technique used in this study. A tray was partially filled with 6 x SSC buffer and a filter paper "bridge" created over a plastic base. Once the filter paper became saturated with 6 x SSC, the agarose gel was placed on top, followed by a Hybond-N^+ Nylon membrane (*Amersham Biosciences*) cut to the exact dimensions of the gel. Two similarly sized sheets of filter paper were placed on that, followed by a block of tissue paper and a weight. To ensure efficient blotting, a sheet of cling film (dotted line) was placed around the edges of the assembly, preventing the filter bridge coming into contact with any other blotting components.

The blot was assembled according to **Figure 7.1**. The 6 x SSC buffer was drawn up through the gel and Hybond-N+ nylon membrane (*Amersham*) by capillary action (Sambrook *et al*., 1989). The blot was left overnight at room temperature, allowing the DNA to transfer onto the positively charged membrane. The membrane was then removed and dried, and the DNA was fixed using UV crosslinking. The membrane was stored, wrapped in filter paper at room temperature until ready to probe.

7.15.5 Probe labelling, hybridisation and detection

The probes were labelled, hybridised and detected using the digoxigenin (DIG) method, using DIG High-Prime DNA Labelling and Detection Starter Kit I (*Roche*), according to the following protocol:

1) The PCR product to be used as a probe was extracted and purified from a 1% agarose gel. The DNA was eluted in a final volume of 30 µl.
2) Next, 16 µl of gel extracted DNA was boiled for 10 minutes to denature, and then cooled rapidly on salty ice.
3) Once cool, 4 µl of DIG High-Prime (containing random primers, nucleotides, DIG-dUTP, Klenow enzyme and buffer components) was added to the denatured DNA and incubated at 37°C overnight.
4) After 16 hours, the labelled probe mixture was incubated at 65°C for 10 minutes and then frozen until required.
5) For the hybridisation of the probe and membrane, the DIG Easy-Hyb Granules were dissolved in 64 ml dH_2O, and warmed to 42°C.
6) The membrane was pre-hybridised in 20 ml of warmed Easy-Hyb solution, at 42°C for 30 minutes, using a suitable container and gentle agitation.
7) The probe was denatured at 100°C for 10 minutes and then cooled rapidly on a mixture of ice and salt.
8) Next, 4 µl of denatured probe was mixed with 4 ml of Easy-Hyb solution, and added to the membrane in a sealed bag.
9) The membrane was left to hybridise overnight at 42°C with gentle agitation.
10) Post-hybridisation membrane washes were carried out as follows:

    - 2 X 5 minutes [2 X SSC, 0.1% SDS] at room temperature
    - 2 X 15 minutes [pre-warmed 0.5% SSC, 0,1% SDS] at 68°C

11) The membrane was then rinsed in Washing Buffer (0.1 M Maleic acid, 0.15 M NaCl, pH 7.5, 0.3% Tween 20).
12) Next, the membrane was incubated in Blocking solution (0.1 volume 10 x Blocking solution, diluted in 0.1 M Maleic acid, 0.15 M NaCl, pH 7.5) for 30 minutes at room temperature.
13) The Anti-digoxigenin-AP antibody was diluted to 150 mU/ml in 20 ml Blocking Solution. The mixture was added to the membrane and incubated for 30 minutes at room temperature.

14) The membrane was washed in Washing Buffer for 15 minutes. This step was repeated once more.
15) The membrane was then washed in 20 ml Detection Buffer (0.1 M Tris-HCl, 0.1 M NaCl, pH 9.5) for 5 minutes at room temperature.
16) For detection, the membrane was sealed in a bag with 5 ml Colour Substrate Solution (100 µl nitro-blue tetrazolium chloride (NBT)/ 5-bromo-4-chloro-3'-indolyphosphate (BCIP) stock, 5 ml Detection Buffer) and protected from light
17) Once bands became visible, the membrane was thoroughly washed in dH_2O to stop the reaction.

7.16 Colony Blot

Colony hybridisation to ^{32}P labelled probes was used to search for genes of interest in genomic libraries of *R. pomeroyi* genomic DNA, cloned in the cosmid pLAFR3, as follows:

7.16.1 Preparation of filters

Single colonies of *E. coli* 803 containing the pLAFR3-based: *R. pomeroyi* genomic library were picked to gridded Hybond-N+ filter discs (*Amersham*) on LB agar containing tetracycline (pLAFR3 confers resistance to this antibiotic). Plates were incubated overnight at 37°C, and filters were replica plated to fresh media. The replica plates were incubated overnight at 37°C and then stored at 4°C.

The filters were transferred to suitable containers and washed as follows (the denaturation and neutralisation buffers were prepared at stated in **Section 7.15.3**):

Wash 1 - denaturation buffer for 7 mins
Wash 2 - neutralisation buffer for 3 mins
Wash 3 - neutralisation buffer for 3 mins
Wash 4 - 6X SSC, where the cell debris was removed from the filters
Wash 5 - Rinsed in clean 6X SSC to wash away any remaining cell debris

The filters were then transferred to filter paper and left to dry before UV cross-linking to fix the DNA.

7.16.2 Probe labelling, hybridisation and detection

The probes used in colony blotting were gel-extracted PCR products of the gene of interest.

These were labelled, hybridised to the filters and detected as follows:

1) Probes were labelled by a random-primed method (Feinberg and Vogelstein, 1983) with α-^{32}P dCTP (10 mCi/ml *Perkin Elmer*) using hexanucleotide primers.
2) A 10 µl aliquot of gel-extracted DNA was added to 2 µl 20 mM dithiothreitol, 2 µl 10 x labelling buffer [0.9 M 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (pH 6.6), 0.1 M $MgCl_2$], 2 µl dNTP solution (5 mM of dGTP, dATP, and dTTP), 1 µl random hexonuleotide primers, and 1 µl dH_2O.
3) The probe mixture was then heat-denatured at 95°C for 5 minutes, then cooled on ice for 1 minute.
4) Once cool, 1 µl Klenow DNA polymerase *(Invitrogen*) was added to the mixture, followed by 1 µl α-^{32}P dCTP.
5) The mixture was then incubated at room temperature for 3 hours.
6) Unincorporated radionucleotides were removed from the labelling mixture using a NICK column (*Amersham*) as follows:

    - The NICK column was rinsed once with equilibration (TE) buffer (10 mM Tris HCl, pH 7.5, 1 mM EDTA).
    - The labelling mixture (20 µl) was added to 80 µl TE buffer.
    - Then, 3 ml TE buffer was added to the column and allowed to enter the gel bed, followed by 100 µl labelling mixture, and 400 µl TE buffer.
    - Waste was collected and discarded.
    - The clean, labelled probe was eluted from the column in 400 µl TE buffer.

7) To prepare the filters for hybridisation, they were soaked in 6 x SSC, rolled in muslin and placed in a hybridisation tube with 50 ml pre-heated (65°C) pre-hybridisation solution (15 ml 20 x SSC, 5 ml 100 x Denhardt's soltution, 2.5 ml 10% SDS, 27.5 ml dH_2O).
8) A 1 ml aliquot of Herring Sperm DNA was heated to 95°C for 5 minutes and added to the filters to block any further binding of DNA to the filter paper.
9) The filters were incubated at 65°C for 6 hours in a rolling incubator.

10) Then the clean, labelled probe was denatured by heating to 95°C for 5 minutes, and added to the filter set up, which was then incubated at 65°C overnight.
11) The hybridisation mixture was decanted and the filters were washed twice in [2 X SSC, 0.1% SDS], twice in [1 X SSC, 0,1% SDS] and once in [0.1 X SSC, 0.1% SDS]. All washes lasted 10 minutes and were incubated at 65°C, and wash solutions were added in 50 ml volumes and decanted after each incubation:
12) The filters were removed from the muslin, wrapped in Clingfilm and placed in a storm cassette with a Kodak phosphor screen to expose overnight.
13) The phosphor screen was scanned using a "Typhoon FLA 9500" laser scanner (*GE Healthcare Life Sciences*).
14) Areas of radioactivity were identified, and the corresponding colony on the replica plate was used for cosmid preparation and further analyses.

7.17 Genomic library preparation

Genomic libraries used in this study were prepared by ARJ Curson, using the wide host range cosmid pLAFR3. Briefly, 25 µg pLAFR3 was digested with *Eco*RI and dephosphorylated before phenol-chloroform extraction and ethanol precipitation followed by resuspension in an appropriate volume of H_2O. Genomic DNA was prepared using the Wizard® Genomic DNA Purification Kit (*Promega*), and 10 µg of the prepared DNA was digested with *Eco*RI for an appropriate digestion time, as predetermined through test digestions. The digestion reaction was stopped abruptly by flash freezing the reaction in a tube containing 100% ethanol and 3 M sodium acetate. The genomic DNA mix was thawed, ethanol precipitated and resuspended in 30 µl H_2O. Next, 2.5 µg digested genomic DNA was ligated to 1 µg digested pLAFR3 using T4 DNA ligase (*Promega*). Approximately 0.7 µg ligation reaction in a volume of 1-4 µl was then packaged into λ phage using the Gigapack III XL packaging mix (*Stratagene*) and used to transfect *E. coli* 803.

7.18 Assays for DMS production

Gas chromatography was used to measure DMS production. Most gas chromatography assays were performed using a flame photometric detector (Agilent 7890A gas chromatograph fitted with a 7693 autosampler) and an HP-INNOWax 30 m x 0.320 mm capillary column (Agilent Technologies J&W Scientific). However, assays for work on *Oceanimonas doudoroffii* were

carried out using a flame photometric detector (Focus GC; Thermo Scientific) and a 30 m x 0.53 mm ID-BP1 5.0 µm capillary column (SGE Europe, Milton Keynes, UK). The column flow was 2 ml min^{-1}, with a split ratio of 2:1 and a pressure of 13.095 psi. The hydrogen gas flow was 50 ml min^{-1}, and the air and make up (nitrogen) flow were 60 ml min^{-1}. The carrier gas was helium. The injector and detector temperatures were 250°C, with an oven temperature of 40°C. All assays were carried out in 2 ml glass crimp-top vials, sealed with an 11 mm PTFE/rubber/aluminium crimp cap (*Thermo Scientific*), with a total reaction volume of 300 µl, and an injection volume of 50 µl. The retention time of the peak representing DMS was 2.29 minutes.

To quantify DMS concentrations in the headspace of vials, the assay was calibrated using the peak areas produced by seven standard concentrations DMSP (1 µM, 2 µM, 10 µM, 20 µM, 100 µM, 200 µM and 300 µM) mixed with sodium hydroxide for complete alkaline lysis of DMSP into equimolar concentrations of DMS and acrylate. To do this, 500 mM NaOH was added to each vial, and the appropriate amount of DMSP, dissolved in dH_2O was pipetted onto the septum of each vial lid. Thus the NaOH and DMSP was mixed only when lids were inverted and sealed onto the vials, ensuring minimal escape of released DMS. Vials were incubated overnight at 28°C in the dark, and then assayed at room temperature. Assuming complete lysis of all the DMSP to DMS, 1 µM of DMS would be produced from 1 µM of DMSP. This is equivalent to 0.3 nmol total DMS in the 300 µl liquid reaction volume, and of this, 0.15 nmol DMS is transferred to the 1.7 ml vial headspace. Thus the peak area produced in the calibration was converted to pmols DMS per headspace of each vial.

7.18.1 Assays *in vivo*

Cell cultures to be assayed for DMS production were grown overnight in appropriate media, in the presence or absence of potential co-inducer molecules. Cells were then pelleted and washed in minimal or rich media. Washed cells were added to vials with an appropriate concentration of DMSP, and vials were sealed immediately. Samples were incubated at 30°C, and then assayed at room temperature.

7.18.2 Assays *in vitro*

DMSP lyase proteins were over-expressed in *E. coli* BL21 as described above. Assays were performed on lysed cells, or partially pure protein. For lysed cells, 1 ml of induced culture was centrifuged to pellet cells, which were then resuspended in NPI-10 buffer. The resuspended cells were lysed by sonication for 4 x 10 second bursts at full power. Then, 297 µl lysed cell material was added to 3 µl 100 mM DMSP in a vial and sealed immediately. Samples were incubated at 28°C and then assayed at room temperature.

For enzyme kinetics using partially pure protein, samples were prepared as above using varying concentrations of DMSP. For each concentration of substrate, gas chromatograph readings were taken every 4.5 minutes for the first 30 minutes. For these measurements, samples were always at room temperature, due to short assay times.

DMS production for all samples was calculated as nmol DMS $minute^{-1}$ µg $protein^{-1}$. Total protein estimations of samples were carried out using Bradford's assay, as described above.

7.19 β-galactosidase Assay

Potential promoter regions of interest were cloned 5' of a promoterless *lacZ* gene in the wide host-range *lacZ* reporter plasmids pMP220 or pBIO1878. Expression of *lacZ* was measured as units of β-galactosidase activity, with *ortho*-nitrophenyl-β-galactoside (ONPG) as the substrate. Cleavage of this colourless substrate β-galactosidase yields galactose and *ortho*-nitrophenyl, the latter's yellow colour being measured spectrophotometrically at OD_{420}.

The protocol was as follows:

1) Starter cultures of wild type strains containing pBIO1878-/pMP220-based plasmids were inoculated into complete or minimal media containing 2-5 mM potential inducer, and incubated overnight at 28°C, shaking at 200 rpm.
2) Optical density of the induced cultures was measured at 600 nm.
3) A 500 µl aliquot of the culture was mixed with 500 µl Z buffer (per 50 ml dH_2O: 1 ml 3 M $Na_2PO_4.7H_2O$; 0.5 ml 4 M $NaH_2PO_4.2H_2O$; 0.5 ml 1 M KCl; 0.5 ml 0.1 M $MgSO_4.7H_2O$; 175 µl β-mercaptoethanol) in clean 2 ml microfuge tubes.
4) Using a Pasteur pipette, 2 drops of chloroform and 1 drop of 0.1% SDS were added to the samples, and vortexed for 10 seconds, before incubating at 28°C for 5 minutes.
5) Samples were removed and 200 µl ONPG (4 mg ml^{-1}) was added.
6) Samples were incubated at 28°C until a sufficient yellow colour developed, and the reaction was stopped using 500 µl 1 M Na_2CO_3. The total reaction time was recorded.
7) Samples were centrifuged at 13,000 rpm for 2 minutes to pellet cell debris and OD_{420} readings of the supernatant were taken.
8) Miller units of β-galactosidase activity were calculated using the following formula:

$$\text{Miller Units} = (1000 \times OD_{420})/(t \times v \times OD_{600})$$

t = time of assay (min)

v = volume of cell culture used (ml)

7.20 Growth Curves

Overnight cultures were adjusted to equivalent OD_{600} values, and diluted 1:100 into 100 ml media in 250 ml conical flasks, at 28°C with shaking at 200 rpm. Absorbance readings at 600 nm were measured periodically until cultures reached stationary phase.

7.21 Sensitivity studies

Sensitivity of bacteria to different compounds was tested on complete or minimal media agar plates with increasing concentrations of DMSP, propionate, MMPA or acrylate. Where minimal media was used, 10 mM succinate was provided as the additional carbon source. Cultures were grown to stationary phase and adjusted to equivalent OD_{600} values. Cells were pelleted, washed and resuspended using minimal media with no added carbon source. Spots (10 µl) of various dilutions of the culture were applied to the different agar plates and incubated at 28°C until growth appeared.

7.22 Detection of DMSP catabolites by NMR

From a 5 ml culture of *E. coli* BL21 containing over-expressed DddQ or DddK, a 2 ml aliquot was re-suspended in a total volume of 1 ml 20 mM deuterium oxide (pH 6.75). Cells were lysed by sonication, cell debris removed by centrifugation and the soluble fraction was incubated in the presence of 3 mM 3^{13}C-DMSP. Perchloric acid was added to a final concentration of 15 µl ml^{-1} and incubated on ice for 15 minutes. Samples were centrifuged and the supernatant was added to NMR tubes. Proton-decoupled ^{13}C NMR spectra were measured at 75 MHz with a Varian Gemini 2000.

7.23 Chemical syntheses

Chemical syntheses were performed by Dr Yohan Chan (School of Chemistry, UEA).

7.23.1 DMSP

To synthesise DMSP, DMS (15.3 ml, equivalent to 0.21 mol) was mixed with HCl (100 ml of a 2 M solution) and then acrylic acid (10 g, equivalent to 0.14 mol). To increase the reaction rate, the mixture was heated under reflux (95°C) for 2 hours, then cooled to room temperature and concentrated under reduced pressure. The residue was triturated using ethanol and diethyl ether, and the resulting solid was filtered and washed first in ethanol, then twice in diethyl ether, yielding 21.8 g DMSP (92%). The identity and purity of DMSP was determined by melting point (134-135°C) and infra-red and NMR spectroscopy [ν_{max} (solid/cm^{-1}) 3013, 2621, 2549, 2478, 2426, 1787, 1691, 1414, 1396, 1247, 1183, 906; δ_H (400 MHz; D_2O) 2.94 (6H, s, 2 x CH_3), 2.98 (2H, t, J = 6.8 Hz), 3.53 (2H, t, J = 6.8 Hz); δ_C (75 MHz; D_2O) 24.51, 27.84, 38.05, 173.17].

To make ^{13}C-DMSP, ^{13}C-acrylic acid (100 mg, 1.39 mmol) was added to HCl (5 ml, 2 M), then DMS (2.78 mmol) was added and the mixture was treated as above, yielding 210 mg ^{13}C-DMSP (89%).

7.23.2 MMPA

To synthesise methylmercaptopropionate (MMPA), methyl 3-(methylthio)propionate (20.0 g, 18.64 mL, 0.15 mol, 1 equiv.) was dissolved in aqueous sodium hydroxide solution (2M, 150 mL) and the reaction mixture was stirred overnight. An aqueous hydrochloric acid solution (1 M) was then added till the ph reached 1-2. The aqueous layer was extracted with diethyl ether (3x 100mL) to yield MMPA (12.2g, 68% purity).

7.24 Oligonucleotide design

DNA oligonucleotides were manufactured by Eurofins MWG Operon, Ebersberg, Germany, and designed using Primer3Plus software (Untergasser *et al*., 2012) and ARTEMIS (*Sanger*; Rutherford *et al*., 2000). Primers were chosen to have a length of 20-30 bp, a melting temperature of 60°C and a GC content of 40-60%, if possible. Where necessary, restriction sites were incorporated into the 5' end of the primer, and capped with appropriate bases, according to instructions from *New England Biolabs*. Sequences of primers used in this study are shown in **Table 7.5**.

7.25 DNA sequencing

Sequencing services were carried out by Genome Enterprise Ltd. (John Innes Centre, UK), or Eurofins MWG Operon using an AbiPrism 3730 capillary sequencer, or an ABI 3730XL sequencer, respectively. Plasmids were prepared for sequencing using a *Qiagen* midi-prep kit – see above. Universal sequencing primers M13F and M13R were provided by each company for sequencing pK19-, pLAFR3- and pBluescript-based recombinant plasmids. For other plasmids, appropriate primers were supplied with the sample at a concentration of 1.5 pmol μl^{-1}.

7.26 *In silico* analyses

7.26.1 Sequence alignment

Multiple alignments of protein sequences were done using **MU**ltiple **S**equence **C**omparison by **L**og-**E**xpectation (MUSCLE; Edgar, 2004) through the MEGA 6.0 software (Tamura *et al*., 2013).

7.26.2 Phylogenetic trees

Unrooted phylogenetic trees were estimated from multiple sequence alignments using the optimum model as determined by MEGA 6.0 software. The length of each branch represents the distance between sequence pairs, and numbers at the base of each branch show bootstrap values.

7.26.3 Database searches

Searches for homologous protein sequences in the NCBI non-redundant protein database were done using BLAST.

7.26.4 Statistical analysis

Statistical analyses of data were conducted in R v. 2.3 (R core team, 2015). Where appropriate, a t-test was used to conduct single or pairwise comparisons. Data that were not normally distributed were analysed using the non-parametric Wilcoxon rank sum for single pair comparisons. One way ANOVAs were used for analysing differences between groups.

Table 7.5 Primers used in this study

Primer Name	Sequence
SPO1094PK19FOR	TGCAGAATTCTTTCGGGATGGAAAAACAAC
SPO1094PK19REV	ACTGTCTAGATTTCGAAACAGTCGCAGAAG
SPO1101PK19FOR	TGCAGAATTCGGCTTCCTTTCGGAAAACTC
SPO1101PK19REV	ATGCTCTAGACAGCGTGACCACCCATTC
SPO1105PK19FOR	ACGTGAATTCAAGAGTTCATGGTCCGCAAC
SPO1105PK19REV	CAGTTCTAGACATCGCCATGCTGATCTCT
SPO1094pS4FOR2	CAGTTCTAGACCGATCACCACCTGCATC
SPO1094pS4REV2	CTGACTGCAGGAGCGAAGCGAAAAGAGTTT
SPO1101pS4FOR	GTCAGAATTCTAAGACGGACCAATCCGAAG
SPO1101pS4REV	TGCATCTAGATTCCAGCTTGGACTCGAAAT
SPO1105pRKFOR2	GCATAAGCTTTATGCCGATGACCTGATGTG
SPO1105pRKREV2	CAGTTCTAGAGCGAAATGATCCAGTTCCAT
SPO2528FOR	GGGATCTAGAGGCGCTTGCATCAGAAG
SPO2528REV	TCGTCTGCAGAGACCTCTTTGCGGG
SPO1014FOR	CTGGTCTAGAACATGGGGGGATCGG
SPO1014REV	CGCGCTGCAGCATGGCGATCTGGCCA
SPO2934pK19FOR	CATGTCTAGACGTGCATTCGGTGGTCTT
SPO2934pK19REV	CGTACTGCAGTAGAGGTACCCGTCCTCGTC
SPO2934FOR	CGGCTCTAGAAGCGGCTTTGCAAAAG
SPO2934REV	CAAGCTGCAGGGCGGATGCGATGCG
OdP2ProXbaFOR	GGCCCTGGGGCTCTAGAGCCAGGGCC
OdP2ProPstREV	CCCTGACCCGCTGCAGGCGATAGGTCCGG
OdDddTproEcoFor	GCGAATTCTCGAAGAAAACATGACCATCTG
OdDddTproXbaRev1	GCTCTAGAAGGAGTTGAACATGGCGAGT
OdReiskeproEcoFor	GCGAATTCCCAGGTTTTCGCTGAAAAA
OdReiskeproXbaRev	GCTCTAGACCAGGTTTTCTTGAATATTTGC
dddDproEcoFor1	GCGAATTCATACCAGGCCAGCTCTTCG
dddDproXbaRev1	GCTCTAGACGATTGAGAATGGCATTGG
dddDTproEcoFor1	GCGAATTCCGATTGAGAATGGCATTGG
dddDTproXbaRev1	GCTCTAGAATACCAGGCCAGCTCTTCG

List of References

Alber, B.E., and Fuchs, G. (2002) Propionyl-coenzyme A synthase from *Chloroflexus aurantiacus*, a key enzyme of the 3-hydroxypropionate cycle for autotrophic CO_2 fixation. *J Biol Chem* **277**: 12137-12143.

Albertsen, M., Hansen, L.B.S., Saunders, A.M., Nielsen, P.H., and Nielsen, K.L. (2012) A metagenome of a full-scale microbial community carrying out enhanced biological phosphorus removal. *ISME J* **6**: 1094-1106.

Alcolombri, U., Laurino, P., Lara-Astiaso, P., Vardi, A., and Tawfik, D.S. (2014) DddD is a CoA-transferase/lyase producing dimethyl sulphide in the marine environment. *Biochemistry* **53**: 5473-5475.

Amo, L., Rodríguez-Gironés, M., and Barbosa, A. (2013) Olfactory detection of dimethyl sulphide in a krill-eating Antarctic penguin. *Mar Ecol Prog Ser* **474**: 277-285.

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

Andrews, S.C., Robinson, A.K., and Rodríguez-Quiñones (2006) Bacterial iron homeostasis. *FEMS Microbiol Rev* **27**: 215-237.

Ansede, J.H., Pellechia, P.J., and Yoch, D.C. (1999) Metabolism of acrylate to β-hydroxypropionate and its role in dimethylsulfoniopropionate lyase induction by a salt marsh sediment bacterium, *Alcaligenes faecalis* M3A. *Appl Environ Microb* **65**: 5075-5081.

Attwood, P.V., and Wallace, J.C. (2002) Chemical and catalytic mechanisms of carboxyl transfer reactions in biotin-dependent enzymes. *Acc Chem Res* **35**: 113-120.

Ayra, A.S., Lee, S.A., and Eiteman, M.A. (2013) Differential sensitivities of the growth of *Escherichia coli* to acrylate under aerobic and anaerobic conditions and its effect on product formation. *Biotechnol Lett* **35**: 1839-1843.

Bachhawat, A.K., Thakur, A., Kaur, J. and Zulkifli, M. (2013) Glutathione transporters. *Biochim Biophys Acta* **1830**: 3154-3164.

Bacic, M.K., and Yoch, D.C. (1998). *In vivo* characterization of dimethylsulfoniopropionate lyase in the fungus *Fusarium lateritium*. *Appl Environ Microbiol* **64**: 106.

Bahr, M., Hobbie, J.E., and Sogin, M.L. (1996) Bacterial diversity in an arctic lake: a freshwater SAR11 cluster. *Aquat Microb Ecol* **11**: 271-277.

Bak, F., Finster, K. and Rothfuß, F. (1992) Formation of dimethylsulfide and methanethiol from methoxylated aromatic compounds and inorganic sulfide by newly isolated anaerobic bacteria. *Arch Microbiol* **157**: 529-534.

Bates, T.S., Lamb, B.K., Guenther, A., Dignon, J., and Stoiber, R.E. (1992) Sulfur emissions to the atmosphere from natural sources. *J Atmos Chem* **14**: 315-337.

Baumann, L., Baumann, P., Mandel, M., and Allen, R.D. (1972) Taxonomy of aerobic marine eubacteria. *J Bacteriol* **110**: 402-429.

Bazan, J.F., Weaver, L.H., Roderick, S.L., Huber, R. and Matthews, B.W. (1994) Sequence and structure comparison suggest that methionine aminopeptidase, prolidase, aminopeptidase P, and creatinase share a common fold. *Proc Natl Acad Sci USA* **91**: 2473-2477.

Begley, T.P., Kinsland, C. and Strauss, E. (2001) The biosynthesis of coenzyme A in bacteria. *Vitam Horm* **61**: 157-171.

Bentley, R., and Chasteen, T.G. (2004) Environmental VOSCs - formation and degradation of dimethyl sulfide, methanethiol and related materials. *Chemosphere* **55**: 291-317.

Beringer, J.E. (1974) R factor transfer in *Rhizobium leguminosarum*. *J Gen Microbiol* **84**: 188-198.

Beringer, J.E., and Hopwood, D.A. (1976) Chromosomal recombination and mapping in *Rhizobium leguminosarum*. *Nature* **264**: 261-293.

Biebl, H., Allgaier, M., Tindall, B.J., Koblizek, M., Lünsdorf, H., Pukall, R., and Wagner-Döbler, I. (2005) *Dinoroseobacter shibae* gen. nov., sp. nov., a new aerobic phototrophic bacterium isolated from dinoflagellates. *Int J Syst Evol Micr* **55**: 1089-1096.

Boden, R., Kelly, D.P., Murrell, J.C., and Schäfer, H. (2010) Oxidation of dimethylsulfide to tetrathionate by *Methylophaga thiooxidans* sp. nov.: a new link in the sulfur cycle. *Environ Microbiol* **12**: 2688-2699.

Boden, R., Borodina, E., Wood, A.P., Kelly, D.P., Murrell, J.C., and Schäfer, H. (2011) Purification and characterization of dimethylsulfide monooxygenase from *Hyphomicrobium sulfonivorans*. *J Bacteriol* **193**: 1250-1258.

Bordoli, L., Kiefer, F., Arnold, K., Benkert, P., Battey, J., and Schwede, T. (2008) Protein structure homology modeling using SWISS-MODEL workspace. *Nat Protoc* **4**: 1-13.

Bradford, M.M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* **72**: 248-254.

Bratbak, G., Lavasseur, M., Michaud, S., Cantin, G., Fernandez, E., Heimdal, B.R., and Heldal, M. (1996) Viral activity in relation to *Emiliania huxleyi* blooms: a mechanism of DMSP release? *Mar Ecol Prog Ser* **128**: 133-142.

Brinkhoff, T., Giebel, H.A., and Simon, M. (2008) Diversity, ecology, and genomics of the Roseobacter clade: a short overview. *Arch Microbiol* **189**: 531-539.

Broadbent, A.D., and Jones, G.B. (2004) DMS and DMSP in mucus ropes, coral mucus, surface films and sediment pore waters from coral reefs in the Great Barrier Reef. *Mar Fresh Res* **55**: 849-845.

Broadbent, A.D., Jones, G.B., and Jones, R.J. (2002) DMSP in corals and benthic algae from the Great Barrier Reef. *Estuar Coast Shelf S* **55**: 547-555.

Brown, G.R., Sutcliffe, I.C., and Cummings, S.P. (2001) Reclassification of [*Pseudomonas*] *doudoroffii* (Baumann *et al*. 1983) into the genus *Oceanimonas* gen. nov. as *Oceanomonas doudoroffii* comb. nov., and adescription of a phenol-degrading bacterium from estuarine water as *Oceanimonas baumannii* sp. nov. *Int J Syst Evol Micr* **51**: 67-72.

Brown, C. W. and Yoder, J.A. (1994) Coccolithophorid blooms in the global ocean. *J. Geophys Res* **99**: 7467–7482.

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

Buchan, A., González, J.M., and Moran, M.A. (2005) Overview of the marine roseobacter lineage. *Appl Environ Microbiol* **71**: 5665-5677.

Bullock, H.A., Reisch, C.R., Burns, A.S., Moran, M.A. and Whitman, W.B. (2014) Regulatory and functional diversity of methylmercaptopropionate coenzyme A ligases from the dimethylsulphoniopropionate demethylation pathway in *Ruegeria pomeroyi* DSS-3 and other proteobacteria. *J Bacteriol* **196**: 1275-1285.

Campbell, S., Harada, R.M., and Li, Q.X. (2007) *Ferrimonas senticii* sp. nov., a novel gammaproetobacterium isolated from the mucus of a puffer fish caught in Kaneohe Bay, Hawai'i. *Int J Syst Evol Microbiol* **57**: 2670-3.

Cantoni, G.L., and Anderson, D.G. (1956) Enzymatic cleavage of dimethylpropiothetin by *Polysiphonia lanosa*. *J Biol Chem* **222**: 171-177.

Carr, E.L., Kämpfer, P., Patel, B.K.C., Gürtler, V., and Seviour, R.J. (2003) Seven novel species of *Acinetobacter* isolated from activated sludge. *Int J Syst Evol Microbiol* **53**: 953-963.

Carrión, O., Curson, A.R.J., Kumaresan, D., Fu, Y., Lang, A.S. , Mercadé, E. and Todd, J.D. (2015) A novel pathway producing dimethylsulphide in bacteria is widespread in soil envirnoments. *Nat Commun* **6**: no. 6579.

Carson, M., Johnson, D.H., McDonald, H., Brouillette, C., Delucas, L.J. (2007) His-tag impact on structure. *Acta Crystallogr D Biol Crystallogr* **63**: 295-301.

Caruana, A.M.N., and Malin, G. (2014) The variability in DMSP content and DMSP lyase activity in marine dinoflagellates. *Prog Oceanogr* **120**: 410-424.

Challenger, F., and Simpson, M.I. (1948) 320. Studies on biological methylation. Part XII. A precursor of the dimethyl sulphide evolved by *Polysiphonia fastigiata* - dimethyl-2-carboxyethylsulphonium hydroxide and its salts. *J Chem Soc*: 1591-1597.

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

Chen, Y. (2012) Comparative genomics of methylated amine utilization by marine *Roseobacter* clade bacteria and development of functional gene markers *(tmm, gmaS*). *Environ Microbiol* **14**: 2308-2322.

Cheo, D.L., Bayles, K.W., and Yasbin, R.E. (1991) Cloning and characterization of DNA damage-inducible promoter regions from *Bacillus subtilis*. *J Bacteriol* **173**: 1696-1703.

Chichester, D.F. and Tanner, F.W. (1972) Antimicrobial food additives. In: CRC Handbook of Food Additives, Second Edition, Volume 1. Furia, T.E. (ed): *CRC Press*, pp. 115-184.

Cho, J.C., and Giovannoni, S.J. (2006) *Pelagibaca bermudensis* gen. nov., sp. nov., a novel marine bacterium within the Roseobacter clade in the order Rhodobacterales. *Int J Syst Evol Micr* **56**: 855-859.

Choi, D.H., Park, K-T., An, S.M., Lee, K., Cho, J-C., Lee, J-H., Kim, D., Jeon, D. and Noh, J.H. (2015) Pyrosequencing revealed SAR116 clade as dominant *dddP*-containing bacteria in oligotrophic NW Pacific Ocean. *PLoS One* **10**: e0116271.

Colmer, T.D., Teresa W-M, F., Läuchli, A., and Higashi, R.M. (1996) Interactive effects of salinity, nitrogen and sulphur on the organic solutes in *Spartina alterniflora* leaf blades. *J Exp Bot* **47**: 369-375.

Cooper, D.R., Grelewska, K., Kim, C.Y., Joachimiak, A., Derewenda, Z.S. (2010) The structure of DinB from *Geobacillus stearothermophilus*: a representative of a unique four-helix-bundle superfamily. *Acta Crystallogr Sect F Struct Biol Cryst Commun* **66**: 219-224.

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

Coude, F.X., Sweetman, L. and Nyhan, W.L. (1979) Inhibition by propionyl-coenzyme A of N-acetylglutamate synthetase in rat liver mitochondria. A possible explanation for hyperammonemia in propionic and methylmalonic acidemia. *J Clin Invest* **64**: 1544-1551.

Courties, A., Riedel, T., Jarek, M., Intertaglia, L., Lebaron, P., and Suzuki, M.T. (2013) Genome sequence of strain MOLA814, a proteorhodopsin-containing representative of the beta-proteobacteria common in the ocean. *Genome announc* **1**: e01062-01013.

Cunningham, G.B., Strauss, V., and Ryan, P.G. (2008) African penguins (*Spheniscus demersus*) can detect dimethyl sulphide, a prey-related odour. *J Exp Biol* **211**: 3123-3127.

Curson, A.R.J., Burns, O., Voget, S., Daniel, R., Todd, J.D., McInnis, K., Wexler, M., and Johnston, A.W.B. (2014) Screening of metagenomic and genomic libraries reveals three classes of bacterial enzymes that overcome the toxicity of acrylate. *PLoS ONE* **9**: e97660.

Curson, A.R.J., Sullivan, M.J., Todd, J.D., and Johnston, A.W.B. (2011) DddY, a periplasmic dimethylsulfoniopropionate lyase found in taxonomically diverse species of Proteobacteria. *ISME J* **5**: 1191-1200.

Curson, A.R.J., Sullivan, M.J., Todd, J.D., and Johnston, A.W.B. (2010) Identification of genes for dimethyl sulfide production in bacteria in the gut of Atlantic Herring (*Clupea harengus*). *ISME J* **4**: 144-146.

Curson, A.R.J., Rogers, R., Todd, J.D., Brearley, C.A., and Johnston, A.W.B. (2008) Molecular genetic analysis of a dimethylsulfoniopropionate lyase that liberates the climate-changing gas dimethylsulfide in several marine alpha-proteobacteria and *Rhodobacter sphaeroides*. *Environ Microbiol* **10**: 757-767.

Curson, A.R.J., Fowler, E.K., Dickens, S., Johnston, A.W.B., and Todd, J.D. (2012) Multiple DMSP lyases in the γ-proteobacterium *Oceanimonas doudoroffii*. *Biogeochemistry* **110**: 109-119.

Dacey, J.W.H., King, G.M., and Wakeham, S.G. (1987) Factors controlling emission of dimethylsulphide from salt marshes. *Nature* **330**: 643-645.

David, G., Blondeaus, K., Schiltz, M., Penel, S. and Lewit-Bentley, A. (2003) YodA from *Esherichia coli* is a metal-binding lipocalin-like protein. *J Biol Chem* **278**: 43728-43795.

De Bont, J.A.M., Van Dijken, J.P., and Harder, W. (1981) Dimethyl sulphoxide and dimethyl sulphide as a carbon, sulphur and energy source for growth of *Hyphomicrobium S*. *J Gen Microbiol* **127**: 315-323.

de Souza, M.P., and Yoch, D.C. (1995a) Purification and characterization of dimethylsulfoniopropionate lyase from an *Alcaligenes*-like dimethyl sulfide-producing marine isolate. *Appl Environ Microbiol* **61**: 21-26.

de Souza, M.P., and Yoch, D.C. (1995b) Comparative physiology of dimethyl sulfide production by dimethylsulfoniopropionate lyase in *Pseudomonas doudoroffii* and *Alcaligenes* sp. Strain M3A. *Appl Environ Microbiol* **61**: 3986-3991.

de Souza, M.P., and Yoch, D.C. (1996a) Dimethylsulfoniopropionate lyase from the marine macroalga *Ulva curvata*: purification and characterization of the enzyme. *Planta* **199**: 433-438.

de Souza, M.P., and Yoch, D.C. (1996b) N-terminal amino acid sequences and comparison of DMSP lyases from *Pseudomonas doudoroffii* and *Alcaligenes* strain M3A. In *Biological and environmental chemistry of DMSP and related sulfonium compounds*. Kiene, R.P., Visscher, P.T., Keller, M.D., and Kirst, G.O. (eds). New York: Plenum Press, pp. 131-142.

DeBose, J.L., Lema, S.C., and Nevitt, G.A. (2008) Dimethylsulfoniopropionate as a foraging cue for reef fishes. *Science* **319**: 1356.

Dickschat, J.S., Zell, C. and Brock, N.L. (2010) Pathways and substrate specificity of DMSP catabolism in marine bacteria of the Roseobacter clade. *Chembiochem* **11**: 417-425.

Dlakić, M. (2000) Functionally unrelated signalling proteins contain a fold similar to Mg^{2+}-dependent endonucleases. *Trends Biochem Sci* **25**: 272-273.

Dunwell, J.M., Purvis, A., and Khuri, S. (2004) Cupins: the most functionally diverse protein superfamily? *Phytochemistry* **65**: 7-17.

Edgar, R.C. (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* **32**: 1792-97.

Eitinger, T., Rodionov, D.A., Grote, M., and Scheider, E. (2011) Canonical and ECF-type ATP-binding cassette importers in prokaryotes: diversity in modular organization and cellular functions. *FEMS Microbiol Rev* **35**: 3-67.

Elssner, T., Engemann, C., Baumgart, K., and Kleber, H.P. (2001) Involvement of coenzyme A esters and two new enzymes, an enoyl-CoA hydratase and a CoA-transferase, in the hydration of crotonobetaine to L-carnitine by *Escherichia coli*. *Biochemistry* **40**: 11140-11148.

Endres, C.S., and Lohmann, K.J. (2012) Perception of dimethyl sulfide (DMS) by loggerhead sea turtles: a possible mechanism for locating high-productivity oceanic regions for foraging. *J Exp Biol* **215**: 3535-3538.

Feinberg, A.P., and Vogelstein, B. (1983) A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. *Anal Biochem* **132**: 6-13.

Figurski, D.H., and Helinski, D.R. (1979) Replication of an origin-containing derivative of plasmid RK2 dependent on a plasmid function provided *in trans*. *P Natl Acad Sci USA* **76**: 1648-1652.

Francalanci, F., Davis, N.K., Fuller, J.Q., Murfitt, D. and Leadlay, P.F. (1986) The subunit structure of methylmalonyl-CoA mutase from *Propionibacterium shermanii*. *Biochem J* **236**: 489-494.

Franklin, D.J., Steinke, M., Young, J., Probert, I., and Malin, G. (2010) Dimethylsulphoniopropionate (DMSP), DMSP-lyase activity (DLA) and dimethylsulphide (DMS) in 10 species of coccolithophore. *Mar Ecol Prog Ser* **410**: 13-23.

Friedberg, E.C., Feaver, W.J., and Gerlach, V.L. (2000) The many faces of DNA polymerases: strategies for mutagenesis and for mutational avoidance. *Proc Natl Acad Sci USA* **97**: 5681-5683.

Fry, J., Wood, M. and Poole, P.S. (2001) Investigation of *myo*-inositol catabolism in *Rhizobium leguminosarum* bv. *viciae* and its effect on nodulation competitiveness. *Mol Plant Microbe Interact* **14**: 1016-25.

Fuse, H., Takimura, O., Murakami, K., Yamaoka, Y., and Omori, T. (2000) Utilization of dimethyl sulfide as a sulfur source with the aid of light by *Marinobacterium* sp. strain DMS-S1. *Appl Environ Microb* **66**: 5527-5532.

Fuse, H., Ohta, M., Takimura, O., Murakami, K., Inoue, H., Yamaoka, Y. Oclarit, J.M., and Omori, T. (1998) Oxidation of trichloroethylene and dimethyl sulfide by a marine

Methylomicrobium strain containing soluble methane monooxygenase. *Biosci Biotech Bioch* **62**: 1925-1931.

Freydank, A-C., Brandt, W. and Dräger, B. (2008) Protein structure modeling indicates hexahistidine-tag interference with enzyme activity. *Proteins* **72**: 173-183.

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

Gibson, J.A.E., Garrick, R.C., Burton, H.R., and McTaggart, A.R. (1990) Dimethylsulfide and the alga *Phaeocystis pouchetii* in Antarctic coastal waters. *Mar Biol* **104**: 339-346.

Giovannoni, S.J., Britschgi, T.B., Moyer, C.L., and Field, K.G. (1990) Genetic diversity in Sargasso Sea bacterioplankton. *Nature* **345**: 60-3.

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

Giovannoni, S.J., Tripp, H.J., Givan, S., Podar, M., Vergin, K.L., Baptista, D., Bibbs, L., Eads, J., Richardson, T.H., Noordewier, M., Rappé, M.S., Short, J.M., Carrington, J.C., and Mathur, E.J. (2005) Genome streamlining in a cosmopolitan oceanic bacterium. *Science* **309**: 1242-5.

González, J.M., Covert, J.S., Whitman, W.B., Henriksen, J.R., Mayer, F., Scharf, B., Schmitt, R., Buchan, A., Fuhrman, J.A., Kiene, R.P. and Moran, M.A. (2003) *Silicibacter pomeroyi* sp. nov. and *Roseovarius nubinhibens* sp. nov., dimethylsulfoniopropionate-demethylating bacteria from marine environments. *Int J Syst Evol Micr* **53**: 1261-1269

González, J.M., Mayer, F., Moran, M.A., Hodson, R.E., and Whitman, W.B. (1997) *Sagittula stellata* gen. nov., sp. nov., a lignin-transforming bacterium from a coastal environment. *Int J Syst Bacteriol* **47**: 773-780.

González, J.M., Kiene, R.P., and Moran, M.A. (1999) Transformation of sulfur compounds by an abundant lineage of marine bacteria in the α-subclass of the class Proteobacteria. *Appl Environ Microbiol* **65**: 3810-3819.

Gould, W.D., and Kanagawa, T. (1992) Purification and properties of methyl mercaptan oxidase from *Thiobacillus thioparus* TK-m. *J Gen Microbiol* **138**: 217-221.

Greene, R.C. (1962) Biosynthesis of dimethyl-beta-propiothetin. *J Biol Chem* **237**: 2251-2254.

Grote, J., Trash, J.C., Huggett, M.J., Landry, Z.C., Carini, P., Giovannoni, S.J., and Rappé, M.S. (2012) Streamlining and core genome conservation amongst highly divergent members of the SAR11 clade. *mBio* **3**: e00252-12.

Guo, Y., and Oliver, D.J. (2012) *E. coli* propionyl-CoA synthetase is regulated *in vitro* by an intramolecular disulfide bond. *Prikl Biokhim Mikrobiol* **48**: 289-93.

Haas, P. (1935) The liberation of methyl sulphide by seaweed. *Biochem J* **29**: 1297-1299.

Haller, T., Buckel, T., Rétey, J. and Gerlt, J.A. (2000) Discovering new enzymes and metabolic pathways: conversion of succinate to propionate by *Escherichia coli*. *Biochemistry* **39**: 4622-4629.

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

Hehemann, J-H., Law, A., Redecke, L., and Boraston, A.B. (2014) The structure of *Rp*DddP from *Roseobacter denitrificans* reveals that DMSP lyases in the DddP-family are metalloenzymes. *PLoS ONE* **9**: e103128.

Herrmann, G., Selmer, T., Jessen, H.J., Gokarn, R.R., Selifonova, O., Gort, S.J., and Buckel, W. (2005) Two beta-alanyl-CoA:ammonia lyases in *Clostridium propionicum*. *FEBS J.* **272**: 813-21.

Hill, C.H. (1952) Studies on the inhibition of growth of *Streptococcus faecalis* by sodium propionate. *J Biol Chem* **199**: 329-332.

Hoffman, C., Schweitzer, T.R. and Dalby, G. (1939) Fungistatic properties of the fatty acids and possible biochemical significance. *J Food Sci* **4**: 539-545.

Horinouchi, M., Kasuga, K., Nojiri, H., Yamane, H., and Omori, T. (1997) Cloning and characterization of genes encoding an enzyme which oxidizes dimethyl sulfide in *Acinetobacter* sp. strain 20B. *FEMS Microbiol Lett* **155**: 99-105.

Horswill, A.R. and Escalante-Semerena, J.C. (1999) The *prpE* gene of *Salmonella typhimurium* LT2 encodes propionyl-CoA synthetase, *Microbiology* **145**: 1381-1388.

Howard, E.C., Sun, S., Biers, E.J., and Moran, M.A. (2008) Abundant and diverse bacteria involvec in DMSP degradation in marine surface waters. *Environ Microbiol* **10**: 2397-2410.

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., Kiene, R.P., Whitman, W.B. and Moran, M.A. (2006) Bacterial taxa that limit sulfur flux from the ocean. *Science* **314**: 649-652.

Huang, C.S., Sadre-Bazzaz, K., Shen, Y., Deng, B., Hong Zhou, Z., and Tong, L. (2010) Crystal structure of the $\alpha_6\beta_6$ holoenzyme of propionyl-coenzyme A carboxylase. *Nature* **466**: 1001-1005.

Husband, J.D., Kiene, R.P., and Sherman, T.D. (2010) Oxidation of dimethylsulfoniopropionate (DMSP) in response to oxidative stress in *Spartina alterniflora* and protection of a non-DMSP producing grass by exogenous DMSP and acrylate. *Environ Exp Bot* **79**: 44-48.

Ivanova, E.P., Nedashkovskaya, O.I., Zhukova, N.V., Nicolau, D.V., Christen, R., and Mikhailov, V.V. (2003) *Shewanella waksmanii* sp. nov., isolated from a sipuncula (*Phascolosoma japonicum*). *Int J Syst Evol Microbiol* **53**: 1471-7.

Ivanova, E.P., Sawabe, T., Hayashi, K., Gorshkova, N.M., Zhukova, N.V., Nedashkovskaya, O.I., Mikhailov, V.V., Nicolau, D.V., and Christen, R. (2003) *Shewanella fidelis* sp.nov., isolated from sediments and sea water. *Int J Syst Evol Microbiol* **53**: 577-82.

Jackowski, S. (1996) Biosynthesis of pantothenic acid and coenzyme A. In: *Escherichia coli* and *Salmonella typhimurium*: cellular and molecular biology, second edition. Neidhardt, F.C. *et al.* (eds). *American Society for Microbiolgy, Washington, DC* pp. 687-694.

Jackowski, S. and Rock, C.O. (1981) Regulation of coenzyme A biosynthesis. *J Bacteriol* **148**: 926-932.

Jacob, F., and Monod, J. (1961) Genetic regulatory mechanisms in the synthesis of proteins. *J Mol Biol* **3**: 318-356.

James, F., Nolte, K.D., and Hanson, A.D. (1995a) Purification and properties of S-adenosyl-L-methionine:L-methionine S-methyltransferase from *Wollastonia biflora* leaves. *J Biol Chem* **270**: 22344-22350.

James, F., Paquet, L., Sparace, S.A., Gage, D.A., and Hanson, A.D. (1995b) Evidence implicating dimethylsulfoniopropionaldehyde as an intermediate in dimethylsulfoniopropionate biosynthesis. *Plant Physiol* **108**: 1439-1448.

Johnston, A.W.B., Setchell, S.M., and Beringer, J.E. (1978) Interspecific crosses between *Rhizobium leguminosarum* and *R. meliloti*: Formation of haploid recombinants and of R primes. *J Gen Microbiol* **104**: 209-218.

Juliette, L.Y., Hyman, M.R., and Arp, D.J. (1993) Inhibition of ammonia oxidation in *Nitrosomonas europaea* by sulfur compounds: thioethers are oxidized to sulfoxides by ammonia monooxygenase. *Appl Environ Microb* **59**: 3718-3727.

Kanehisa, M., and Goto, S. (2000) KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res* **28**: 27-30.

Kanehisa, M., Goto, S., Sato, Y., Kawashima, M., Furumichi, M., and Tanabe, M. (2014) Data, information, knowledge and principle: back to metabolism in KEGG. *Nucleic Acids Res* **42**: D199-D205.

Karsten, U., Kück, K., Vogt, C., and Kirst, G.O. (1996) Dimethylsulfoniopropionate Production in Phototrophic Organisms and its Physiological Functions as a Cryoprotectant. In *Biological and Environmental Chemistry of DMSP and Related Sulfonium Compounds*. Kiene, R., Visscher, P., Keller, M., and Kirst, G. (eds): Springer US, pp. 143-153.

Karsten, U., Wiencke, C., and Kirst, G.O. (1990) The β-dimethylsulfoniopropionate (DMSP) content of macroalgae from Antarctica and southern Chile. *Bot Mar* **33**: 143-146.

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

Kettle, A.J., and Andreae, M.O. (2000) Flux of dimethylsulfide from the oceans: A comparison of updated data sets and flux models. *J Geophys Res - Atmos* **105**: 26793-26808.

Kiene, R.P. (1992) Dynamics of dimethyl sulfide and dimethylsulfoniopropionate in oceanic water samples. *Mar Chem* **37**: 29-52.

Kiene, R.P. and Hines, M.E. (1995) Microbial formation of dimethyl sulfide in anoxic Sphagnum peat. *Appl Environ Microbiol* **61**: 2720-2726.

Kiene, R.P., and Taylor, B.F. (1988) Demethylation of dimethylsulfoniopropionate and production of thiols in anoxic marine sediments. *Appl Environ Microbiol* **54**: 2208-2212.

Kiene, R.P., Linn, L.J., and Bruton, J.A. (2000) New and important roles for DMSP in marine microbial communities. *J Sea Res* **43**: 209-224.

Kiene, R.P., Linn, L.J., González, J., Moran, M.A., and Bruton, J.A. (1999) Dimethylsulfoniopropionate and methanethiol are important precursors of methionine and protein-sulfur in marine bacterioplankton. *Appl Environ Microb* **65**: 4549-4558.

Kiesel, A. (1913) The action of different acids and acid salts upon the development of *Aspergillus niger*. *Ann Inst Pasteur* **27**: 391-420.

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

Kirkwood, M., Todd, J.D., Rypien, K.L., and Johnston, A.W.B. (2010b) The opportunistic coral pathogen *Aspergillus sydowii* contains *dddP* and makes dimethyl sulfide from dimethylsulfoniopropionate. *ISME J* **4**: 147-150.

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

Kocsis, M.G., and Hanson, A.D. (2000) Biochemical evidence for two novel enzymes in the biosynthesis of 3-dimethylsulfoniopropionate in *Spartina alterniflora*. *Plant Physiol* **123**: 1153-1161.

Kocsis, M.G., Nolte, K.D., Rhodes, D., Shen, T.L., Gage, D.A., and Hanson, A.D. (1998) Dimethylsulfoniopropionate biosynthesis in *Spartina alterniflora*1. Evidence that S-methylmethionine and dimethylsulfoniopropylamine are intermediates. *Plant Physiol* **117**: 273-281.

Kowalewsky, S., Dambach, M., Mauck, B., and Dehnhardt, G. (2006) High olfactory sensitivity for dimethyl sulphide in harbour seals. *Biol Lett* **2**: 106-109.

Kyrpides, N.C., Woyke, T., Eisen, J.A., Garrity, G., Lilburn, T.G., Beck, B.J., Whitman, W.B., Hugenholtz, P. and Klenk, H.P. (2013) Genomic encyclopedia of type strains, phase I: the one thousand microbial genomes (KMG-I) project. *Stand Genomic Sci* **9**:1278-1284.

Lana, A., Bell, T.G., Simó, R., Vallina, S.M., Ballabrera-Poy, J., Kettle, A.J., Dachs, J., Bopp, L., Saltzman, E.S., Stefels, J., Johnson, J.E. and Liss, P.S. (2011) An updated climatology of surface dimethlysulfide concentrations and emission fluxes in the global ocean. *Global Biogeochem Cy* **25**: GB1004

Larher, F., Hamelin, J., and Stewart, G.R. (1977) L'acide dimethylsulfonium-3 propanoique que de *Spartina anglica*. *Phytochemistry* **16**: 2019-2020.

Lau, S.C.K., Tsoi, M.M.Y., Li, X., Plakhotnikova, I., Dobretsov, S., Wong, P-K., and Qian, P-Y. (2005) *Gramella portivictoriae* sp. nov., a novel member of the family *Flavobacteriaceae* isolated from marine sediment. *Int J Syst Evol Micr* **55**: 2497-2500.

Lavesseur, M., Michaud, S., Egge, J., Cantin, G., Nejstgaard, J.C., Sanders, R., Fernandez, E., Solberg, P.T., Heimdal, B. and Gosselin, M. (1996) Production of DMSP and DMS during a mesocosm study of an *Emiliania huxleyi* bloom: influence of bacteria and *Calanus finmarchicus* grazing. *Mar Biol* **126**: 609-618.

Ledyard, K.M., DeLong, E.F., Dacey, J.W.H. (1993) Characterization of a DMSP-degrading bacterial isolate from the Sargasso Sea. *Arch Microbiol* **160**:312-318.

Ledyard, K.M., and Dacey, J.W.H. (1994) Dimethylsulfide production from dimethylsulfoniopropionate by a marine bacterium. *Mar Ecol-Prog Ser* **110**: 95-103.

Lesniak, J., Barton, W.A., and Nikolov, D.B. (2003) Structural and functional features of the *Escherichia coli* hydroperoxide resistance protein OsmC. *Protein Sci* **12**: 2838-2843.

Levasseur, M., Michaud, S., Egge, J., Cantin, G., Nejstgaard, J.C., Sanders, R., Fernandez, E., Solberg, P.T., Heimdal, B. and Gosselin, M. (1996) Production of DMSP and DMS during a mesocosm study of an *Emiliania huxleyi* bloom: influence of bacteria and *Calanus finmarchicus* grazing. *Mar Biol* **126**: 609-618.

Li, C.Y., Wei, T.D., Zhang, S.H., Chen, X.L., Gao, X., Wang, P., Xie, B.B., Su, H.N., Qin, Q.L., Zhang, X.Y., Yu, J., Zhang, H.H., Zhou, B.C., Yang, G.P., Zhang, Y.Z. (2014) Molecular insight into bacterial cleavage of oceanic dimethylsulfoniopropionate in dimethyl sulfide. *Proc Natl Acad Sci USA* **111**: 1036-31.

Lovelock, J.E., Maggs, R.J., and Rasmussen, R.A. (1972) Atmospheric dimethyl sulphide and the natural sulphur cycle. *Nature* **237**: 452-453.

Luo, H., Swan, B.K., Stepanauskas, R., Hughes, A.L. and Moran, M.A. (2014) Evolutionary analysis of a streamlined lineage of surface ocean Roseobacters. *ISME J* **8**: 1428-1439.

Luo, H., and Moran, M.A. (2014) Evolutionary ecology of the marine Roseobacter clade. *Micro Mol Biol Rev* **78**: 573-587.

Lyon, B.R., Lee, P.A., Bennett, J.M., DiTullio, G.R., and Janech, M.G. (2011) Proteomic Analysis of a Sea-Ice Diatom: Salinity Acclimation Provides New Insight into the Dimethylsulfoniopropionate Production Pathway. *Plant Physiol* **157**: 1926-1941.

Maddocks, S.E. and Oyston, P.C.F. (2008) Structure and function of the LysR-type transcriptional regulator (LTTR) family proteins. *Microbiology* **154**: 3609-3623.

Magnuson, K., Jackowski, S., Rock, C.O., and Cronan, J.E. (1993) Regulation of fatty acid biosynthesis in *Escherichia coli*. *Microbiol Rev* **57**: 522-542.

Maniatis, T., Sambrook, J., and Fritsch, E.F. (1982) *Molecular cloning: a laboratory manual*. Cold Spring Harbor, N.Y. Cold Spring Harbor Laboratory.

Martens, T., Heidorn, T., Pukall, R., Simon, M., Tindall, B.J., and Brinkhoff, T. (2006) Reclassification of *Roseobacter gallaeciensis* Ruiz-Ponte *et al*. 1998 as *Phaeobacter*

gallaeciensis gen. nov., comb. nov., description of *Phaeobacter inhibens* sp. nov., reclassification of *Ruegeria algicola* (Lafay *et al*. 1995) Uchino *et al*. 1999 as *Marinovum algicola* gen. nov., comb. nov., and emended descriptions of the genera *Roseobacter*, *Ruegeria* and *Leisingera*. *Int J Sys Evol Micr* **56**: 1293-1304.

Maruyama, K. and Kitamura, H. (1985) Mechanisms of growth inhibition by propionate and restoration of the growth by sodium bicarbonate or acetate in *Rhodopseudomonas sphaeroides*. *Biochem J* **98**: 819-824.

Masip, L., Veeravalli, K., and Georgiou, G. (2006) The many faces of glutathione. *Antioxid Redox Sign* **8**: 753-762.

McCarthy, A.A., Baker, H.M., Shewry, S.C., Patchett, M.L., and Baker, E.N. (2001) Crystal structure of methylmalonyl-coenzyme A epimerase from *P. shermanii*: a novel enzymatic function on an ancient metal binding scaffold. *Structure* **9**: 637-646.

McClung, C.R., and Patriquin, D.G. (1980) Isolation of a nitrogen-fixing *Campylobacter* species from the roots of *Spartina alterniflora* Loisel. *Can J Microbiol* **26**: 881-886.

McLennan, A.G. (2006) The NUDIX hydrolase superfamily. *Cel Mol Life Sci* **63**: 123-143.

Messing, J. (1979) A multipurpose cloning system based on the single-stranded DNA bacteriophage M13. *Recomb DNA Tech Bul* **3**: 43-48.

Miller, T.R., and Belas, R. (2004) Dimethylsulfoniopropionate metabolism by *Pfiesteria*-associated *Roseobacter* spp. *Appl Environ Microbiol* **70**: 3383-3391.

Miyamoto, E., Watanabe, F., Charles, T.C., Yamaji, R., Inui, H. and Nakano, Y. (2003) Purification and characterization of homodimeric methylmalonyl-CoA mutase from *Sinorhizobium meliloti*. *Arch Microbiol* **180**: 151-154.

Mohapatra, B.R., Rellinger, A.N., Kieber, D.J., and Kiene, R.P. (2013) Comparative functional characteristics of DMSP lyases extracted from polar and temperate *Phaeocystis* species. *Aquat Biol* **18**: 185-195.

Moran, M.A., Belas, R., Schell, M.A., González, J.M., Sun, F., Sun, S., Binder, B.J., Edmonds, J., Ye, W., Orcutt, B., Howard, E.C., Meile, C., Palefsky, W., Goesmann, A.,

Ren, Q., Paulson, I., Ulrich, L.E., Thompson, L.S., Saunders, E. and Buchan, A. (2007) Ecological genomics of Marine Roseobacters. *Appl Environ Microbiol* **73**: 4559-4569.

Moran, M.A., Buchan, A., González, J.M., Heidelberg, J.F., Whitman, W.B., Kiene, R.P., Henriksen, J.R., King, G.M., Belas, R., Fuqua, C., Brinkac, L., Lewis, M., Johri, S., Weaver, B., Pai, G., Eisen, JA., Rahe, E., Sheldon, W.M., Ye, W., Miller, T.R., Carlton, J., Rasko, D.A., Paulsen, I.T., Ren, Q., Daugherty, S.C., Deboy, R.T., Dodson, R.J., Durkin, A.S., Madupu, R., Nelson, W.C., Sullivan, S.A., Rosovitz, M.J., Haft, D.H., Selengut, J., Ward, N. (2004) Genome sequence of *Silicibacter pomeroyi* reveals adaptations to the marine environment. *Nature* **432**: 910-913.

Moran, M.A., Reisch, C.R., Kiene, R.P., and Whitman, W. (2012) Genomic insights into bacterial DMSP transformations. *Ann Rev Mar Sci* **4**: 523-542.

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

Nakagawa, T., Iino, T., Suzuki, K., and Harayama, S. (2006) *Ferrimonas futtsuensis* sp. nov. and *Ferrimonas kyonanensis* sp. nov., selenate-reducing bacteria belonging to the Gammaproteobacteria isolated from Tokyo Bay. *Int J Syst Evol Microbiol* **56**: 2639-45.

Neidhart, D.J., Kenyon, G.L., Gerlt, J.A., and Petsko, G.A. (1990) Madelate racemase and muconate lactonizing enzyme are mechanistically distinct and structurally homologus. *Nature* **347**: 692-694.

Nemec, A., Musílek, M., Sedo, O., De Baere, T., Maixnerová, M., van der Reijden, T.J., Zdráhal, Z., Vaneechoutte, M., Dijkshoorn, L. (2010) *Acinetobacter bereziniae* sp. nov. and *Acinetobacter guillouiae* sp. nov., to accommodate *Acinetobacter* genomic species 10 and 11, respectively. *Int J Syst Evol Microbiol* **60**: 896-903.

Nevitt, G.A., and Bonadonna, F. (2005) Sensitivity to dimethyl sulphide suggests a mechanism for olfactory navigation by seabirds. *Biol Letters* **1**: 303-305.

Newton, G.L., Leung, S.S., Wakabayashi, J.I., Rawat M., and Fahey, R.C. (2011) The DinB superfamily includes novel mycothiol, bacillithiol and glutathione *S*-transferases. *Biochemistry* **50**: 10751-10760.

Niki, T., Kunugi, M., and Otsuki, A. (2000) DMSP-lyase activity in five marine phytoplankton species: its potential importance in DMS production. *Mar Biol* **136**: 759-764.

Nishiguchi, M.K., and Somero, G.N. (1992) Temperature- and concentration-dependence of compatibility of the organic osmolyte beta-dimethylsulfoniopropionate. *Cryobiology* **29**: 118-124.

Omori, T., Saiki, Y., Kasuga, K., and Kodama, T. (1995) Desulfurization of Alkyl and Aromatic Sulfides and Sulfonates by Dibenzothiophene-desulfurizing *Rhodococcus* sp. Strain SY1. *Biosci Biotech Bioch* **59**: 1195-1198.

Otte, M.L., and Morris, J.T. (1994) Dimethylsulphoniopropionate (DMSP) in *Spartina alterniflora* Loisel. *Aquat Bot* **48**: 239-259.

Otte, M.L., Wilson, G., Morris, J.T., and Moran, B.M. (2004) Dimethylsulphoniopropionate (DMSP) and related compounds in higher plants. *J Exp Bot* **55**: 1919-1925.

Patel, S.S. and Walt, D.R. (1987) Substrate specificity of acetyl coenzyme A synthetase. *J Biol Chem* **262**: 7132-1134.

Puigbò, P., Guzmán, E., Romeu, A., and Garcia-Vallvé, S. (2007) OPTIMIZER: A web server for optimizing the codon usage of DNA sequences. *Nucleic Acids Res*. **35**: W126-31.

Qiao, X., Sun, Y., Qiao, J., and Mindich, L. (2008) The role of host protein YajQ in the temporal control of transcription in bacteriophage Φ6. *P Natl Acad Sci* **105**: 15956-15960.

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

R Core Team (2015). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

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

Rajashekhara, E. and Watanabe, K. (2004) Propionyl-coenzyme A synthetases of *Ralstonia solanacearum* and *Salmonella choleraesuis* display atypical kinetics. *FEBS lett* **556**: 143-147.

Rappé, M.S., Connon, S.A., Vergin, K.L., and Giovannoni, S.J. (2002) Cultivation of the ubiquitous SAR11 marine bacterioplankton clade. *Nature* **418**: 630-3

Regnier, P., Grunberg-Manago, M., and Portier, C. (1987) Nucleotide sequence of the *pnp* gene of *Escherichia coli* encoding polynucleotide phosphorylase. Homology of the primary structure of the protein with the RNA-binding domain of ribosomal protein S1. *J Biol Chem* **262**: 63-68.

Reisch, C.R., Crabb, W.M., Gifford, S.M., Teng, Q., Stoudemayer, M.J., Moran, M.A. and Whitman, W.B. (2013) Metabolism of dimethylsulphoniopropionate by *Ruegeria pomeroyi* DSS-3. *Mol Microbiol* **89**: 774-791.

Reisch, C.R., Moran, M.A., and Whitman, W.B. (2008) Dimethylsulfoniopropionate-dependent demethylase (DmdA) from *Pelagibacter ubique* (sic) and *Silicibacter pomeroyi*. *J Bacteriol* **190**: 8018-8024.

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

Rhodes, D., Gage, D.A., Cooper, A., and Hanson, A.D. (1997) S-Methylmethionine conversion to dimethylsulfoniopropionate: Evidence for an unusual transamination reaction. *Plant Physiol* **115**: 1541-1548.

Roe K.L., Barbeau, K., Mann, E.L., and Haygood, M.G. (2012) Acquisition of iron by *Trichodesmium* and associated bacteria in culture. *Environ Microbiol* **14**: 1681-1695.

Rosselló-Mora, R.A., Ludwig, W., Kämpfer, P., Amann, R., Schleifer, K-H. (1995) *Ferrimonas balearica* gen. nov., spec. nov., a new marine facultative Fe(III)-reducing Bacterium. *Syst Appl Microbiol* **18**: 196-202.

Ruiz-Ponte, C., Cilia, V., Lambert, C., and Nicolas, J.L. (1998) *Roseobacter gallaeciensis* sp. nov., a new marine bacterium isolated from rearings and collectors of the scallop *Pecten maximus. Int J Syst Bacteriol* **48 Pt 2**: 537-542.

Rusch, D.B., Halpern, A.L., Sutton, G., Heidelberg, K.B., Williamson, S., Yooseph, S., Wu, D., Eisen, J.A., Hoffman, J.M., Remington, K., Beeson, K., Tran, B., Smith, H., Baden-Tillson, H., Stewart, C., Thorpe, J., Freeman, J., Andrews-Pfannkoch, C., Venter, J.E., Li, K., Kravitz, S., Heidelberg, J.F., Utterback, T., Rogers, Y.H., Falcón, L.I., Souza, V., Bonilla-Rosso, G., Eguiarte, L.E., Karl, D.M., Sathyendranath, S., Platt, T., Bermingham, E., Gallardo, V., Tamayo-Castillo, G., Ferrari, M.R., Strausberg, R.L., Nealson, K., Friedman, R., Frazier, M., Venter, J.C. (2007) The Sorcerer II global ocean sampling expedition: northwest Atlantic through eastern tropical Pacific. *PLoS Biol* **5**: e77.

Rutherford, K., Parkhill, J., Crook, J., Horsnell, T., Rice, P., Rajandream, M.A., and Barrell, B. (2000) Artemis: sequence visualization and annotation. *Bioinformatics* **16**: 944-5.

Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989) Molecular cloning: a laboratory manual. *Cold Spring Harbor, N.Y.; Cold Spring Harbor Laboratory*.

Schäfer, H., McDonald, I.R., Nightingale, P.D., and Murrell, J.C. (2005) Evidence for the presence of a CmuA methyltransferase pathway in novel marine methyl halide-oxidizing bacteria. *Environ Microbiol* **7**: 839-852.

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

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

Schiffmann, R., Neugebauer, A., and Klein, C.D. (2006) Metal-mediated inhibition of *Escherichia coli* methionine aminopeptidase: structure-activity relationships and development of a novel scoring function for metal-ligand interactions. *J Med Chem* **49**: 511-522.

Schneider, K., Asao, M., Carter, M.S., and Alber, B.E. (2012) *Rhodobacter sphaeroides* uses a reductive route via propionyl coenzyme A to assimilate 3-hydroxypropionate. *J Bacteriol* **194**: 225-232.

Shiba, T. (1991) *Roseobacter litoralis* gen. nov., sp. nov., and *Roseobacter denitrificans* sp. nov., Aerobic Pink-Pigmented Bacteria which Contain Bacteriochlorophyll a. *Syst Appl Microbiol* **14**: 140-145.

Short, J.M., Fernandez, J.M., Sorge, J.A., and Huse, W.D. (1988) λ ZAP: a bacteriophage λ expression vector with *in vivo* excision properties. *Nucleic Acids Res* **16**: 7583-7600.

Sieburth, J.M. (1961) Antibiotic properties of acrylic acid, a factor in the gastrointestinal antibiosis of polar marine animals. *J Bacteriol* **82**: 72-79.

Simo, R. (2001) Production of atmospheric sulfur by oceanic plankton: biogeochemical, ecological and evolutionary links. *Trends Ecol Evol* **16**: 287-294.

Simon, R., Quandt, J.R., and Klipp, W. (1989) New derivatives of transposon Tn*5* suitable for mobilization of replicons, generation of operon fusions and induction of genes in Gram-negative bacteria. *Gene* **80**: 161-169.

Slezak, D.M., Puskaric, S. and Herndl, G.J. (1994) Potential role of acrylic acid in bacterioplankton communities in the sea. *Mar Ecol Prog Ser* **105**: 191-197.

Sorokin, D.Y., Jones, B.E., and Kuenen, J.G. (2000) An obligate methylotrophic, methane-oxidizing *Methylomicrobium* species from a highly alkaline environment. *Extremophiles* **4**: 145-155.

Spaink, H.P., Okker, R.J.H., Wijffelman, C.A., Pees, E. and Lugtenberg, B.J.J. (1987) Promoters in the nodulation region of the *Rhizobium leguminosarum* Sym plasmid pRL1JI. *Plant Mol Biol* **9**: 27-39.

Staskawicz, B., Dahlbeck, D., Keen, N., and Napoli, C. (1987) Molecular characterization of cloned avirulence genes from race 0 and race 1 of *Pseudomonas syringae* pv. *glycinea*. *J Bacteriol* **169**: 5789-5794.

Stefels, J., Steinke, M., Turner, S., Malin, G., and Belviso, S. (2007) Environmental constraints on the production and removal of the climatically active gas dimethylsulphide (DMS) and implications for ecosystem modelling. *Biogeochemistry* **83**: 245-275.

Stefels, J., and van Boekel, W.H.M. (1993) Production of DMS from dissolved DMSP in axenic cultures of the marine phytoplankton species *Phaeocystis* sp. *Mar Ecol Prog Ser* **97**: 11-18.

Steinke, M., Daniel, C., and Kirst, G.O. (1996) DMSP Lyase in Marine Macro- and Microalgae. In: *Biological and Environmental Chemistry of DMSP and Related Sulfonium Compounds*. Springer US 317-324.

Steinke, M., and G.O. Kirst (1996) Enzymatic cleavage of dimethylsulfoniopropionate (DMSP) in cell-free extracts of the marine macroalga *Enteromorpha clathrata* (Roth) Grev. (Ulvales, Chlorophyta). *J Exp Mar Biol Ecol* **201**: 73-85.

Steinke, M., Stefels, J., and Stamhuis, E. (2006) Dimethyl sulfide (DMS) triggers search behavior in copepods. *Limnol Oceanogr* **51**:1925-1930.

Steinke, M., Wolfe, G.V., and Kirst, G.O. (1998) Partial characterisation of dimethylsulfoniopropionate (DMSP) lyase isozymes in 6 strains of *Emiliania huxleyi*. *Mar Ecol-Prog Ser* **175**: 215-225.

Steinke, M., Evans, C., Lee, G.A., and Malin, G. (2007) Substrate kinetics of DMSP lyases in axenic culture and mesocosm populations of *Emiliana huxleyi*. *Aquat Sci* **69**: 352-359.

Stöcker, W., Wolz, R.L., Zwilling, R., Strydom, D.J., and Auld, D.S. (1988) *Astacus* protease, a Zinc metalloenzyme. *Biochemistry* **27**: 5026-5032.

Strom, S., Wolfe, G., Slajer, A., Lambert, S., and Clough, J. (2003) Chemical defense in the microplankton II: Inhibition of protist feeding by dimethylsulfoniopropionate (DMSP). *Limnol Oceanogr* **48**: 230-237.

Struck, A-W., Thompson, M.L., Wong, L.S., and Micklefield, J. (2012) S-Adenosyl-methionine-dependent methyltransferases: highly versatile enzymes in biocatalysis, biosynthesis and other biotechnological applications. *ChemBioChem* **13**: 2642-2655.

Sullivan, M.J., Curson, A.R.J., Shearer, N., Todd, J.D., Green, R.T., and Johnston, A.W.B. (2011) Unusual regulation of a leaderless operon involved in the catabolism of dimethylsulfoniopropionate in *Rhodobacter sphaeroides*. *PLoS One* **6**: e15972.

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

Sun, L., Curson, A.R.J., Todd, J.D., and Johnston, A.W.B. (2012) Diversity of DMSP transport in marine bacteria revealed by genetic analyses. *Biogeochemistry* **110**: 121-130.

Sunda, W., Kieber, D.J., Kiene, R.P., and Huntsman, S. (2002) An antioxidant function for DMSP and DMS in marine algae. *Nature* **418**: 317-320.

Suylen, G.M.H., Large, P.J., van Dijken, J.P. and Kuenen, J.G. (1987) Methyl mercaptan oxidase, a key enzyme in the metabolism of methylated sulphur compounds by *Hyphomicrobium* EG. *J Gen Microbiol* **133**: 2989-2997.

Suzuki, H., Koyanagi, T., Izuka, S., Onishi, A., and Kumagai, H. (2005) The *yliA, -B, -C* and *-D* genes of *Escherichia coli* K-12 encode a novel glutathione importer with an ATP-binding cassette. *J Bacteriol* **187**: 5861-5867.

Tamura, K., Stecher, G., Peterson, D., Filipski, A., and Kumar, S. (2013) MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. *Mol Biol Evol* **30**: 2725-2729.

Taylor, B.F., and Gilchrist, D.C. (1991) New routes for aerobic biodegradation of dimethylsulfoniopropionate. *Appl Environ Microb* **57**: 3581-3584.

Teufel, R., Kung, J.W., Kockelkorn, D., Alber, B.E., and Fuchs, G. (2009) 3-hydroxypropionyl-coenzyme A dehydratase and acryloyl-coenzyme A reductase, enzymes of the autotrophic 3-hydroxypropionate/4-hydroxybutyrate cycle in Sulfolobales. *J Bacteriol* **191**: 4572-81.

Thompson, F.L., Hoste, B., Thompson, C.C., Goris, J., Gomez-Gil, B., Huys, L., De Vos, P. and Swings, J. (2002) Enterovibrio norvegicus gen. nov., sp. nov., isolated from the gut of turbot (*Scophthalmus maximus*) larvae: a new member of the family Vibrionaceae. *Int J Syst Evol Micr* **52**: 2015-2022.

Tobin, C., Mandava, C.S., Ehrenberg, M., Andersson, D.I. and Sanyal, S. (2010) Ribosomes lacking protein S20 are defective in mRNA binding and subunit association. *J Mol Biol* **397**: 767-776.

Todd, J.D., Kirkwood, M., Newton-Payne, S., and Johnston, A.W.B. (2012a) DddW, a third DMSP lyase in a model Roseobacter marine bacterium, *Ruegeria pomeroyi* DSS-3. *ISME J* **6**: 223-226.

Todd, J.D., Curson, A.R.J., Dupont, C.L., Nicholson, P., and Johnston, A.W.B. (2009) The *dddP* gene, encoding a novel enzyme that converts dimethylsulfoniopropionate into dimethyl sulfide, is widespread in ocean metagenomes and marine bacteria and also occurs in some Ascomycete fungi. *Environ Microbiol* **11**: 1376-1385.

Todd, J.D., Curson, A.R., Kirkwood, M., Sullivan, M.J., Green, R.T., and Johnston, A.W.B (2010b) DddQ, a novel, cupin-containing, dimethylsulfoniopropionate lyase in marine roseobacters and in uncultured marine bacteria. *Environ Microbiol* **13**: 427-438.

Todd, J.D., Curson, A.R.J., Sullivan, M.J., Kirkwood, M., and Johnston, A.W.B. (2012b) The *Ruegeria pomeroyi acuI* gene has a role in DMSP catabolism and resembles *yhdH* of *E. coli* and other bacteria in conferring resistance to acrylate. *PLoS One* **7**: e35947.

Todd, J.D., Curson, A.R.J., Nikolaidou-Katsaraidou, N., Brearley, C.A., Watmough, N.J., Chan, Y., Page, P.C., Sun, L., Johnston, A.W.B. (2010a) Molecular dissection of bacterial acrylate catabolism-unexpected links with dimethylsulfoniopropionate catabolism and dimethyl sulfide production. *Environ Microbiol* **12**: 327-343.

Todd, J.D., Rogers, R., Li, Y.G., Wexler, M., Bond, P.L., Sun, L., Curson, A.R.J., Malin, G., Steinke, M. and Johnston, A.W.B. (2007) Structural and regulatory genes required to make the gas dimethyl sulfide in bacteria. *Science* **315**: 666-669.

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

Trossat, C., Nolte, K.D., and Hanson, A.D. (1996a) Evidence That the Pathway of Dimethylsulfoniopropionate Biosynthesis Begins in the Cytosol and Ends in the Chloroplast. *Plant Physiol* **111**: 965-973.

Trossat, C., Nolte, K.D., Rathinasabapathi, B., and Hanson, A.D. (1996b) Subcellular localization of the DMSP synthesis pathway in *Wollastonia biflora* (Asteraceae). *Plant Physiol* **111**: 405-405.

Trossat, C., Rathinasabapathi, B., Weretilnyk, E.A., Shen, T.L., Huang, Z.H., Gage, D.A., and Hanson, A.D. (1998) Salinity promotes accumulation of 3-dimethylsulfoniopropionate and its precursor S-methylmethionine in chloroplasts. *Plant physiol* **116**: 165-171.

Untergrasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faricloth, B.C., Remm, M., and Rozen, S.G. (2012) Primer3 – new capabilities and interfaces. *Nucleic Acids Res* **40**: e115.

Vairavamurthy, A., Andreae, M.O., and Iverson, R.L. (1985) Biosynthesis of dimethylsulfide and dimethylpropiothetin by *Hymenomonas carterae* in relation to sulfur source and salinity variations. *Limnol Oceanogr* **30**: 59-70.

Van Alstyne, K., Koellermeier, L., and Nelson, T. (2007) Spatial variation in dimethylsulfoniopropionate (DMSP) production in *Ulva lactuca* (Chlorophyta) from the Northeast Pacific. *Mar Biol* **150**: 1127-1135.

Van Alstyne, K., Wolfe, G.V., Freidenburg, T.L., Neill, A., and Hicken, C. (2001) Activated defense systems in marine macroalgae: evidence for an ecological role for DMSP cleavage. *Mar Ecol-Prog Ser* **213**: 53-65.

Van Bergeijk, S.A., Van der Zee, C., and Stal, L.J. (2003) Uptake and excretion of dimethylsulphoniopropionate is driven by salinity changes in the marine benthic diatom *Cylindrotheca closterium*. *Eur J Phycol* **38**: 341-349.

van der Maarel, M.J.E.C., Aukema, W., and Hansen, T.A. (1996a) Purification and characterization of a dimethylsulfoniopropionate cleaving enzyme from *Desulfovibrio acrylicus*. *FEMS Microbiol Lett* **143**: 241-245.

van der Maarel, M.J.E.C., van Bergeijk, S., van Werkhoven, A.F., Laverman, A.M., Meijer, W.G., Stam, W.T., and Hansen, T.A. (1996b) Cleavage of dimethylsulfoniopropionate and reduction of acrylate by *Desulfovibrio acrylicus* sp. nov. *Arch Microbiol* **166**: 109-115.

van Duyl, F.C., Gieskes, W.W.C., Kop, A.J., and Lewis, W.E. (1998) Biological control of short-term variations in the concentration of DMSP and DMS during a *Phaeocystis* spring bloom. *J Sea Res* **40**: 221-231.

Varaljay, V.A., Robidart, J., Preston, C.M., Gifford, S.M., Durham, B.P., Burns, A.S., Ryan, J.P., Marin, R., Kiene, R.P., Zehr, J.P., Scholin, C.A. and Moran, M.A. (2015) Single-taxon field measurements of bacterial gene regulation controlling DMSP fate. *ISME J* **doi:10.1038/ismej.2015.23** (ahead of print).

Vergauwen, B., Elegneert, J., Dansercoer, A., Devreese, B. and Savvides, S.N. (2010) Glutathione import in *Haemophilus influenzae* Rd is primed by the periplasmic heme-binding protein HbpA. *P Natl Acad Sci USA.* **107**: 13270-13275.

Villaseñor, T., Brom, S., Dávalos, A., Lozano, L., Romero, D. and García-de los Santos, A. (2011) Housekeeping genes essential for pantothenate biosynthesis are plasmid-encoded in *Rhizobium etli* and *Rhizobium leguminosarum*. *BMC Microbiol* **11**: 66.

Visscher, P.T., and Taylor, B.F. (1993) A new mechanism for the aerobic catabolism of dimethyl sulfide. *Appl Environ Microb* **59**: 3784-3789.

Visscher, P.T., and Taylor, B.F. (1994) Demethylation of dimethylsulfoniopropionate to 3-mercaptopropionate by an aerobic marine bacterium. *Appl Environ Microb* **60**: 4617-19.

Vogt, C., Rabenstein, A., Rethmeier, J.r., and Fischer, U. (1998) Alkali-labile precursors of dimethyl sulfide in marine benthic cyanobacteria. *Arch Microbiol* **169**: 263-266.

Wagner, C., and Stadtman, E.R. (1962) Bacterial fermentation of dimethyl-β-propiothetin. *Arch Biochem Biophys* **98**: 331-336.

Wagner, J., Gruz, P., Kim, S.R., Yamada, M., Matsui, K., Fuchs, R.P., Nohmi, T. (1999) The *dinB* gene encodes a novel *E. coli* DNA polymerase, DNA pol IV, involved in mutagenesis. *Mol Cell* **4**: 281-286.

Wagner-Döbler, I., and Biebl, H. (2006) Environmental biology of the marine Roseobacter lineage. *Annu Rev Microbiol* **60**: 255-280.

Wolfe, G.V., Sherr, E., and Sherr, B.F. (1994) Release and consumption of DMSP from *Emiliania huxleyi* during grazing by *Oxyrrhis marina*. *Mar Ecol Prog Ser* **111**: 111-119.

Wolfe, G.V., Steinke, M., and Kirst, G.O. (1997) Grazing-activated chemical defence in a unicellular marine alga. *Nature* **387**: 894-897.

Wood, W.B. (1966) Host specificity of DNA produced by *Escherichia coli*: Bacterial mutations affecting the restriction and modification of DNA**.** *J Mol Biol* **16**: 118-113.

Xiao, X., Wang, P., Zeng, X., Bartlett, D.H., and Wang, F. (2007) *Shewanella psychrophila* sp. nov. and *Shewanella piezotolerans* sp. nov., isolated from west Pacific deep-sea sediment. *Int J Syst Evol Microbiol* **57**: 60-5.

Yanisch-Perron, C., Vieira, J., and Messing, J. (1985) Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene* **33**: 103-119.

Yeganeh, L.P., Azarbaijani, R., Sarikhan, S., Mousavi, H., Ramezani, M., Amoozegar, M.A., Fazeli, A.S. and Salekdeh, G.H. (2012) Complete genome sequence of *Oceanimonas* sp. GK1, a halotolerant bacterium from Gavkhouni Wetland in Iran. *J Bacteriol* **194**: 2123-2124.

Yoch, D.C. (2002) Dimethylsulfoniopropionate: Its sources, role in the marine food web , and biological degradation to dimethylsulfide. *Appl Environ Microbiol* **68**: 5804-5815.

Yoch, D.C., Caraway, R.H., Friedman, R. and Kulkarni, N. (2001) Dimethylsulfide (DMS) production from dimethylsulfoniopropionate by freshwater river sediments: phylogeny of Gram-positive DMS-producing isolates. *FEMS Microbiol Ecol* **37**: 31-37.

Young, J.P.W., Crossman, L.C., Johnston, A.W.B., Thomson, N.R., Ghazoui, Z.F., Hull, K.H., Wexler, M., Curson, A.R., Todd, J.D., Poole, P.S., Mauchline, T.H., East, A.K., Quail, M.A., Churcher, C., Arrowsmith, C., Cherevach, I., Chillingworth, T., Clarke, K., Cronin, A., Davis, P., Fraser, A., Hance, Z., Hauser, H., Jagels, K., Moule, S., Mungall, K., Norbertczak, H., Rabbinowitsch, E., Sanders, M., Simmonds, M., Whitehead, S., Parkhill, J. (2006) The genome of *Rhizobium leguminosarum* has recognizable core and accessory components. *Genome Biol* **7**: R34.

Zeyer, J., Eicher, P., Wakeham, S.G., and Schwarzenbach, R.P. (1987) Oxidation of dimethyl sulfide to dimethyl sulfoxide by phototrophic purple bacteria. *Appl Environ Microb* **53**: 2026-2032.

Zhang, L., Kuniyoshi, I., Hirai, M., and Shoda, M. (1991) Oxidation of dimethyl sulfide by *Pseudomonas acidovorans* DMR-11 isolated from peat biofilter. *Biotechnol Lett* **13**: 223-228.

Ziegler, C., Bremer, E., and Krämer, R. (2010) The BCCT-family of carriers: from physiology to crystal structure. *Mol Microbiol* **78**: 13-34.

Appendix

Abbreviations

ABC	ATP-binding cassette
Acu	Acrylate-utilising
AMP	Adenosine monophosphate
ATP	Adenosine triphosphate
BCCT	Betaine-carnitine-choline transporter
BCIP	5-bromo-4-chloro-3-indolyphosphate
bp	Base pair(s)
CCN	Cloud condensation nuclei
CO_2	Carbon dioxide
CoA	Coenzyme A
Da	Dalton
Ddd	DMSP-dependent DMS
dH_2O	Distilled water
DIG	Digoxigenin
DMS	Dimethylsulphide
DMSO	Dimethylsulphoxide
DMSP	Dimethylsulphonio-propionate
DMSHB	4-dimethylsulphonio-2-hydroxybutyrate
DNA	Deoxyribonucleic acid
dNTP	deoxynucleotide triphosphate
DNVS	Dissolved non-volatile degradation products
EDTA	Ethylenediaminetetra-acetic acid
EtBr	Ethidium bromide
FAD	Flavin adenine dinucleotide
g	Gram
GC	Gas chromatography
GB	Glycine betaine
GOS	Global Ocean Survey
HGT	Horizontal gene transfer
3HP	3-hydroxypropionate
HPLC	High performance liquid chromatography
IPTG	Isopropyl β-D-1-thiogalactopyranoside
k	Kilo
kb	Kilo base pair(s)
kDa	Kilo Daltons
KEGG	Kyoto encyclopedia of genes and genomes
LB	Lysogeny broth
l	Litre
m	Milli
µ	Micro
M	Molar
MBM	Marine basal medium
MCM	Methylmalonyl-CoA mutase
MeSH	Methanethiol
mol	Moles
MMPA	Methylmercaptopropionate
MPA	3-mercaptopropionate
MTA	Methylthioacryloyl
MTHB	4-methyl-thio-2-hydroxybutyrate
MTOB	4-methylthio-2-oxobutyrate
MTPA	4-dimethylsulphonio-2-hydroxybutyrate
n	Nano
NAD	Nicotinamide adenine dinucleotide
NADP	Nicotinamide adenine dinucleotide phosphate

NBT	nitro-blue tetrazolium
NMR	Nuclear magnetic resonance
NPI	Sodium phosphate imidazole buffer
OD	Optical density
ONPG	*ortho*-nitrophenyl-β-D-galactopyranoside
PAGE	Polyacrylamide gel electrophoresis
PCC	Propionyl-CoA carboxylase
PCR	Polymerase chain reaction
p	Pico
PPi	Inorganic pyrophosphate
RBS	Ribosome binding site
RM	*Rhizobium* minimal media
mRNA	Messenger ribonucleic acid
rRNA	Ribosomal ribonucleic acid
ROS	Reactive Oxygen Species
rpm	Revolutions per minute
SDS	Sodium dodecyl sulphate
SMM	S-methyl-methionine
sp.	Species
T	Tera
TEMED	N,N,N,N,-tetramethylethyleneamide
THF	Tetrahydrofolate
Tris	Tris (hydroxymethyl) aminomethane
TY	Tryptone yeast
UV	Ultra-violet
X-gal	5-bromo-4-chloro-3-indolyl-β-D-galactoside
YTSS	Yeast tryptone sea salt media

Publication

Biogeochemistry (2012) 110:109–119
DOI 10.1007/s10533-011-9663-2

Multiple DMSP lyases in the γ-proteobacterium *Oceanimonas doudoroffii*

Andrew R. J. Curson · Emily K. Fowler · Shilo Dickens · Andrew W. B. Johnston · Jonathan D. Todd

Received: 15 March 2011 / Accepted: 27 September 2011 / Published online: 13 October 2011

Abstract The marine γ-proteobacterium *Oceanimonas doudoroffii* was shown to have at least three different enzymes, each of which can cleave dimethylsulfoniopropionate (DMSP), an abundant compatible solute made by different classes of marine phytoplankton. These various DMSP lyases have similarities, but also some differences to those that had been identified in other bacteria. This was demonstrated by cloning each of the corresponding genes and transferring them into other species of bacteria in which backgrounds they conferred the ability to catabolise DMSP, releasing dimethyl sulfide (DMS) as one of the products (Ddd^+ phenotype; DMSP-dependent DMS). One of these genes resembled *dddD,* which was in a cluster with other *ddd* genes variously involved in subsequent steps of DMSP catabolism, in DMSP import and in DMSP-dependent transcriptional regulation. The other two gene products both had sequence similarity to the previously identified DddP lyase. However, these two *Oceanimonas* DddP polypeptides were not particularly similar to each other and were in two different subbranches compared to those that had been studied in strains of the Roseobacter clade of bacteria. One of these *O. doudoroffii* enzymes, DddP1, most closely resembled gene products in a disparate group of microbes that included two bacteria, *Vibrio orientalis* and *Puniceispirillum marinum* and, more strikingly, some Ascomycete fungi that can catabolise DMSP. Previously, the only bacteria known to have multiple ways to catabolise DMSP were in the Roseobacter clade, which were also the only bacteria that had been shown to have functional DddP DMSP lyases. Thus *Oceanimonas doudoroffii* is unusual on more than one count and likely acquired its *dddD, dddP1* and *dddP2* genes by independent horizontal gene transfer events.

Keywords DddD · DddP · DMSP lyases · Gene regulation · Horizontal gene transfer · *Oceanimonas*

Introduction

In the 1990s, a number of laboratories described several individual species of microbes that catabolised dimethylsulfoniopropionate (DMSP) and conducted a series of biochemical and physiological experiments that revealed some key features of this important process (reviewed in Yoch 2002). Two very different general pathways were recognised; one of these

Electronic supplementary material The online version of this article (doi:10.1007/s10533-011-9663-2) contains supplementary material, which is available to authorized users.

A. R. J. Curson · E. K. Fowler · A. W. B. Johnston · J. D. Todd (✉)
School of Biological Sciences, University of East Anglia, Norwich Research Park, Norwich NR4 7TJ, UK
e-mail: Jonathan.Todd@uea.ac.uk

S. Dickens
Department of Biochemistry, University of Cambridge, Tennis Court Road, Cambridge CB2 1GA, UK

involved an initial demethylation step, to form methylmercaptopropionate (MMPA), the other required cleavage of DMSP, releasing dimethyl sulfide (DMS) as one of the products. Although the catch-all term 'DMSP lyase' was (and still is) used to describe the enzymatic activity that cleaves DMSP in such a way as to liberate dimethyl sulfide, it was apparent early on that there must be several rather different types of polypeptide with this catalytic activity. This was because the lyases in different bacteria varied in their calculated sizes, their K_m and V_{max} values, their responses to potential inhibitors, and even such fundamental features as their proposed sub-cellular locations and the identities of the initially formed C3 catabolites (Yoch 2002).

The recent application of genetic methodology has begun to explain the basis of this heterogeneity at a molecular level. For example, we have described six different genes, namely *dddD* (Todd et al. 2007, 2010), *dddL* (Curson et al. 2008), *dddP* (Todd et al. 2009), *dddQ* (Todd et al. 2011a), *dddY* (Curson et al. 2011) and *dddW* (Todd et al. 2011b), which occur in a range of different bacteria and, in the case of *dddP*, in some fungi. All these genes encode enzymes that cleave DMSP, releasing DMS, yet their gene products are in different polypeptide families. In most cases (DddL, DddP, DddQ, DddY and DddW), the "other" C3 catabolite of the cleavage is acrylate, but the DddD enzyme generates 3-OH-propionate (3HP) as the first detectable C3 product. Furthermore, the different *ddd* genes are subject to different sorts of transcriptional regulation, mediated by a variety of transcriptional regulators, in response not only to the substrate DMSP, but in some cases to its catabolites acrylate or 3HP (Todd et al. 2010; Sullivan et al. 2011).

One of the earliest strains to be studied biochemically was a marine γ-Proteobacterium isolated off Hawaii, which was originally named *Pseudomonas doudoroffii* by Baumann et al. (1972), before its genus was reclassified as *Oceanomonas* (Brown et al. 2001) prior to a final correction to *Oceanimonas* (Anon 2001). This strain generated DMS from DMSP (Ledyard et al. 1993) and further studies in Yoch's laboratory indicated that it had a cytoplasmic DMSP lyase, some of whose features (e.g. pH optimum, response to inhibitors, and inducibility) were described (de Souza and Yoch 1995; Yoch et al. 1997). Indeed, a DMSP lyase from *O. doudoroffii* was purified and its N-terminal sequence was presented. Interestingly, this sequence was very similar to that of another DMSP lyase, made by the β-proteobacterium *Alcaligenes faecalis,* but biochemical evidence indicated that the *A. faecalis* enzyme was located at the bacterial cell surface, whereas it was cytoplasmic in *O. doudoroffii* (de Souza and Yoch 1995, 1996). We recently showed that the *Alcaligenes* DMSP lyase was encoded by the *dddY* gene, whose deduced gene product contained an N-terminal leader which, if cleaved by a signal peptidase, would yield a mature protein that would be targeted to the periplasm. Significantly, the sequence of the predicted N-terminus of this processed form of DddY corresponded to that which had been determined experimentally in *A. faecalis* (de Souza and Yoch 1996; Curson et al. 2011). Taken together, these observations raised the interesting possibility that *A. faecalis* and *O. doudoroffii* both contained versions of the DddY DMSP lyase but that these differed in their sub-cellular locations.

We set out to identify the gene(s) in *O. doudoroffii* that encoded its DMSP lyase(s) and found that this bacterium had a more complex assembly of such enzymes than had been anticipated.

Materials and methods

Strains, plasmids and growth conditions

Oceanimonas doudoroffii strain DSM 7028 was obtained from the DSMZ culture collection, Braunschweig, Germany. *O. doudoroffii* and *E. coli* strains were routinely grown in LB complete medium (Sambrook et al. 1989) at 28°C and 37°C respectively and *Rhizobium leguminosarum* strain 3841 (Young et al. 2006) was grown in TY complete medium at 28°C (Beringer 1974). Antibiotics were used at the following concentrations: ampicillin (100 μg ml^{-1}), streptomycin (200 μg ml^{-1}), tetracycline (5 μg ml^{-1}). *O. doudoroffii* was also grown in M9 minimal medium (Sambrook et al. 1989) for DMS assays with 10 mM succinate as carbon source. For growth tests, 1 ml of an overnight LB culture of *Oceanimonas* was spun down and the pellet washed three times with M9 buffer. Washed *Oceanimonas* cells were then inoculated (1:100) into M9 without added carbon source or supplemented with DMSP (1 mM, 5 mM), acrylate (1 mM, 5 mM), 3HP (1 mM, 5 mM) or DMS (1 mM, 5 mM).

In vivo and in vitro genetic manipulations

Plasmids were transferred by triparental conjugation into a newly made rifampicin resistant derivative of *O. doudoroffii* (strain J495) using the helper plasmid pRK2013 (Figurski and Helinski 1979) and into *E. coli* strain 803 (Wood 1966) by transformation as described in Wexler et al. (2001). Recombinant plasmids based on the expression vector pET21a (Merck4Biosciences, Darmstadt, Germany) were made by PCR amplification using primers shown in Supplementary Table S1, and transformed into *E. coli* BL21 (Studier and Moffat 1986). A library of *Oceanimonas* genomic DNA was made in the wide-host range cosmid pLAFR3 (Staskawicz et al. 1987) essentially as in Curson et al. (2008), as follows. A culture of *O. doudoroffii* was grown to late exponential phase in LB medium. Genomic DNA was isolated using a Qiagen genomic DNA kit and partially digested for various times with *Eco*R1 and aliquots were examined following electrophoresis in agarose gels to determine the approximate sizes of the genomic fragments. A treatment of 10 min was found to generate fragments that were 20–30 kb in size, suitable for cloning into pLAFR3, which accepts inserts of ca. 25 kb. This sample of DNA was ligated to *Eco*R1-digested pLAFR3, prior to in vitro packaging into heads of bacteriophage lambda and transfection into *E. coli* strain 803. The transfected cells were plated and counted for colony-forming units on LB agar containing 5 μg ml^{-1} tetracycline, to which pLAFR3 confers resistance; this showed that the library comprised ca 5,000 primary transfectants. Cosmid DNA, isolated from a random sample of five such colonies, was examined by gel electrophoresis, following digestion with *Eco*R1. In all cases, the recombinant cosmids contained different regions of genomic DNA and the total sizes of the cloned DNA in each cosmid was in the range 20–30 kb.

In order to make transcriptional fusions to *dddD*, *dddP1* and *dddP2*, primers were designed to amplify the relevant regions (see Fig. 1 and Supplementary Table 1) from the *O. doudoroffii* genome, with primers containing restriction sites (*Eco*R1 and *Xba*1 for *dddD* and *dddP1*; *Xba*1 and *Pst*1 for *dddP2*) to facilitate cloning into the wide host range *lacZ* reporter plasmid pMP220 (Spaink et al. 1987).

DMS assays

To assay DMS production in *Oceanimonas*, an overnight culture was diluted (1:100) into M9 minimal medium containing 10 mM succinate, with or without 2 mM DMSP, acrylate or 3HP as potential inducers and incubated for 16 h at 28°C. Then, 1 ml culture was spun down and washed three times in M9 minimal medium lacking any carbon source or inducer and 285 μl of washed cells were added to 2 ml GC vials (Alltech Associates) in a final volume of 300 μl with the substrate DMSP (5 mM). Assay vials were incubated at 22°C before DMS production was quantified by gas chromatography with a flame photometric detector (Focus GC; Thermo Scientific) and a 30 m × 0.53 mm ID-BP1 5.0 μm capillary column (SGE Europe, Milton Keynes, UK). DMS concentrations were calculated by regression analysis based on an eight-point calibration with standard DMS solutions (1–300 μM DMS). To assay DMS production in *E. coli* strain BL21 expressing cloned *ddd* genes, an overnight culture was inoculated (1:100) into LB and grown to an OD_{600} of ~0.7 before induction with 0.2 mM isopropyl-*β*-D-1-thiogalactopyranoside (IPTG) for 16 h at 30°C then the cells were spun down, resuspended in 0.5 ml M9 minimal medium and sonicated (4 × 10 s). Following centrifugation to remove debris, lysates were assayed for DMS production as above. Rates of DMS production are expressed as pmol DMS μg $protein^{-1}$ min^{-1}, with protein concentrations estimated using a BIO-Rad protein assay. Screening the *Oceanimonas* library for DMS production in *Rhizobium leguminosarum* was done as above using individual *Rhizobium* transconjugants grown overnight in TY medium supplemented with 5 mM DMSP. DMS degradation by *Oceanimonas* was tested by diluting cells (1:100) from an overnight culture in complete medium to vials containing minimal medium plus 10 mM succinate and 0.1 mM DMS. The concentrations of DMS that remained in the headspace were measured after incubation at 22°C for 20 h.

Bioinformatics and *in silico* analysis

Searches for homologues and sequence analyses were done using NCBI BLAST and the DNAstar-Lasergene v6 package. Sequences were aligned using ClustalV. Sequencing of the cosmids that contained the *dddD,*

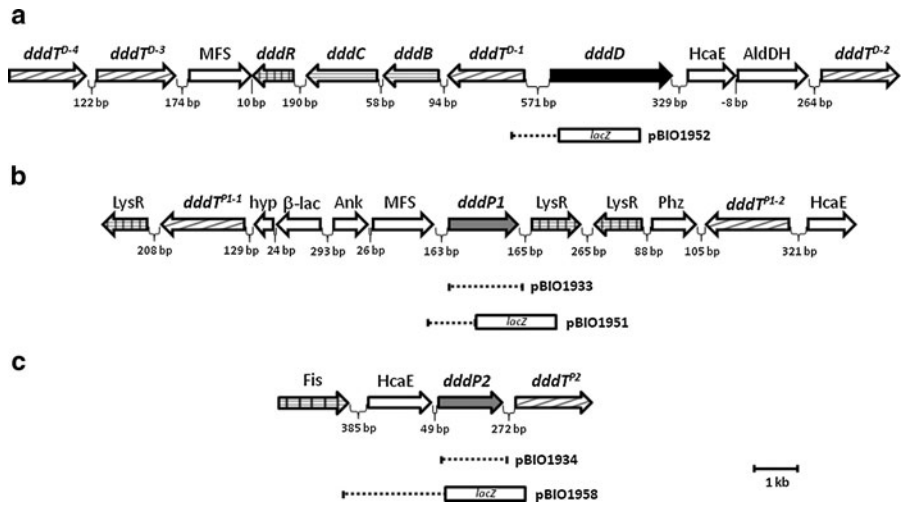


Fig. 1 Arrangement of genes near *dddD* (**a**), *dddP1* (**b**), and *dddP2* (**c**) in *Oceanimonas doudoroffii*. Genes marked with diagonal lines encode BCCT-type transporters, those with horizontal lines are predicted to be involved in downstream steps in DMSP catabolism and those that are hatched are regulatory. The numbers of base pairs in the intergenic spaces are marked and cloned regions in the recombinant plasmids pBIO1933, pBIO1934, pBIO1951, pBIO1952 and pBIO1958 are shown below the genes. *MFS* major facilitator superfamily transporter, *HcaE* oxygenase family polypeptide with C-terminal Rieske domain, *AldDH* aldehyde dehydrogenase, *hyp* hypothetical protein, *β-lac* β-lactamase, *Ank* ankyrin repeat protein, *Phz* phenazine biosynthesis protein. *Fis* and *LysR* are transcriptional regulators in the *Fis* and *LysR* super-families respectively

dddP1 and *dddP2* genes and the draft genome of *Oceanimonas* DSM 7028 were done at the Department of Biochemistry, University of Cambridge, UK. Sequencing of the genome was done using a Roche 454 Genome Sequencer FLX and produced 118,560 reads and 46.8 Mb of raw data, representing an average 10× coverage of the genome. The reads assembled to 193 contigs, with an average size of 27.6 kb, the largest contig being 155.9 kb and totalling 3,829,948 bp. Cosmid sequences were deposited at NCBI Genbank with accession numbers (JN541238, JN541239, JN541240) and the *Oceanimonas* partial genome sequence was deposited at IMG (http://img.jgi.doe.gov/).

Results

We showed that *O. doudoroffii* DSM 7028 grew well on M9 minimal medium in which DMSP (5 mM) was the sole carbon source, with single colonies appearing on plates after 72 h. However, it did not grow at the expense of either 5 mM or 1 mM acrylate or 3-OH-propionate (3HP), the C3 catabolites that are known to be formed following cleavage of DMSP by DddD (for 3HP) or by DddL, DddP, DddQ, DddW or DddY, all of which generate acrylate (see above), and which support the growth of some other bacteria that catabolise DMSP (Todd et al. 2010; Curson et al. 2011). Given that *Oceanimonas* can grow on DMSP, but not its C3 catabolites, we tested whether it could catabolise DMS, the other product of DMSP breakdown. *Oceanimonas* did not grow on 5 mM or 1 mM DMS as sole carbon source and did not degrade DMS when grown in minimal medium containing 10 mM succinate and 0.1 mM DMS.

When grown with DMSP in minimal media that also contained 10 mM succinate, the cells of *O. doudoroffii* emitted DMS (Ddd^+ phenotype). In order to test for induction of DMS production, *O. doudoroffii* was pregrown with or without DMSP, acrylate or 3HP as potential inducers. DMS production was increased, ~4-fold (0.215 pmol ug^{-1} min^{-1} compared to 0.046 pmol ug^{-1} min^{-1} without inducer), when the cells had been pregrown in the presence of DMSP. No induction of activity was seen with either acrylate or 3HP, two DMSP catabolites that are co-inducers in other bacteria (Yoch 2002; Todd et al. 2010; Curson et al. 2011; Sullivan et al. 2011).

These preliminary observations tally with earlier ones on this strain made by de Souza and Yoch (1995) and by Yoch et al. (1997). The levels of DMS

production described here for *O. doudoroffii* are similar to those reported for *Pseudomonas* J465 and *Psychrobacter* J466, but were lower than those for *Halomonas* sp. HTNK1, all of which also contain the *dddD* gene (Todd et al. 2010; Curson et al. 2010).

Cloning the *ddd* genes from *O. doudoroffii*

To identify the gene(s) responsible for the Ddd^+ phenotype of *O. doudoroffii,* we used a functional genetic approach, as has been done previously to clone the *ddd* genes from other bacteria (e.g. Todd et al. 2007). We made a genomic library of *O. doudoroffii* DNA, cloned in pLAFR3, a cosmid vector that accepts inserts of ~25 kb and which can be transferred by conjugation to a wide range of proteobacteria. Following transfection of *E. coli* strain 803 with the ligation mix, ~5,000 primary transfectants were obtained. These were pooled and used *en masse* as a donor culture in a conjugational, triparental cross in which *Rhizobium leguminosarum* strain 3841 (Young et al. 2006) was used as the recipient. The reason for choosing *R. leguminosarum* is that it more effectively expresses heterologous genes, probably because of its relatively large number of RNA polymerase sigma factors (Young et al. 2006). Individual transconjugant colonies were assayed for their ability to make DMS when grown with DMSP; of ~400 that were tested, three had a Ddd^+ phenotype. The cosmids were isolated from these three Ddd^+ transconjugants, transformed into *E. coli*, then mobilised back to *Rhizobium* by conjugation. In all cases, the newly constructed *Rhizobium* transconjugants had a Ddd^+ phenotype, confirming that the cosmids contained functional *ddd* genes.

Restriction digests of these three cosmids showed that they contained different regions of cloned *O. doudoroffii* genomic DNA, since they did not share any fragments of the same size. To identify the relevant *ddd* genes, we sequenced the cloned DNA in each cosmid and found that each of them contained a gene that encoded a product that resembled a previously identified Ddd enzyme. Thus, one cosmid (pBIO1932) contained a gene whose product is ~70% identical to the DddD enzymes that cleave DMSP into DMS plus 3HP. Functional *dddD* genes occur in a range of other γ-Proteobacteria, including strains of *Marinomonas, Halomonas* and *Pseudomonas,* and in a few strains of *Rhizobium* and *Burkholderia* (α- and β-Proteobacteria respectively), which may have acquired *dddD* by horizontal gene transfer (HGT) (Todd et al. 2007; Raina et al. 2010).

The other two cosmids, pBIO1930 and pBIO1931, each contained genes, termed *dddP1* and *dddP2* respectively, whose products were 50% identical to each other and had sequence similarity to DddP, a DMSP lyase in the M24 metallopeptidase family (Todd et al. 2009; Kirkwood et al. 2010a). Previously, *dddP* had only been found in some strains of Roseobacters, an abundant group of marine α-Proteobacteria, and, more surprisingly, in some Ascomycete fungi (Kirkwood et al. 2010b; Todd et al. 2009).

Bioinformatics-based description of the *ddd* and nearby genes in *Oceanimonas*

As shown in Fig. 1, the *dddD*-like gene of *O. doudoroffii* is transcribed divergently from a predicted four-gene operon whose promoter-proximal gene, $dddT^{D\text{-}1}$, encodes a BCCT-type transporter that is likely involved in the uptake of DMSP (see Sun et al. 2011). Downstream of $dddT^{D\text{-}1}$ are *dddB* and *dddC,* which respectively encode an Fe-containing dehydrogenase and an enzyme related to methylmalonate-semialdehyde dehydrogenase, with the promoter-distal gene, *dddR,* encoding a polypeptide in the LysR super-family of transcriptional regulators (Fig. 1a). This arrangement is the same as in the γ-Proteobacteria *Marinomonas* sp. MWYL1 and *Pseudomonas* sp. J465, which also grow well on DMSP (Todd et al. 2007; Curson et al. 2010) and the products of all the genes in the *O. doudoroffii dddD* and *dddTBCR* operons are very similar (at least 70% identical) to their corresponding homologues in both these strains. However, one novel feature of this region in *O. doudoroffii* was the presence of three further genes ($dddT^{D\text{-}2}$, $dddT^{D\text{-}3}$ and $dddT^{D\text{-}4}$) that are predicted to encode BCCT-type transporters (Fig. 1a) that likely import DMSP, though this has not yet been ratified experimentally (Sun et al. 2011). None of the other genes in the *dddD* cluster (Fig. 1a) has any known link with DMSP catabolism in other bacteria.

Turning to the two *O. doudoroffii dddP*-like genes, the outcomes of BLAST-based comparisons of their products with those in the databases were more unusual. As mentioned above, the *dddP* gene had only been described in Roseobacter strains among the

Bacteria, but it also occurs in the eukaryotic fungal pathogens *Aspergillus oryzae*, *A. sydowii* and *Fusarium culmorum* (Todd et al. 2009; Kirkwood et al. 2010b). Interestingly, the *O. doudoroffii* DddP1 polypeptide was more similar (>67% identical, *E* value $<1e^{-176}$) to these fungal sequences than to the previously identified Roseobacter types (~55% identical). Furthermore, two recently available bacterial genome sequences, of the marine γ-Proteobacterium *Vibrio orientalis* CIP 102891 (Yang et al. 1983; http://www.ebi.ac.uk/ena/data/view/Project:40487) and *Puniceispirillum marinum* IMCC1322, in the SAR116 clade of α-proteobacterium (Oh et al. 2010) have deduced polypeptides (gene tags VIA_000771 and SAR116_1427 for *V. orientalis* and *P. marinum* respectively) that cluster with these fungal DddPs and with DddP1 of *O. doudoroffii.* Indeed, DddP of *V. orientalis* is more similar (~80% identical) to *O. doudoroffii* DddP1 than to any other polypeptide in current databases.

The *O. doudoroffii dddP1* gene likely forms a single gene transcriptional unit, lying 5′ of a predicted regulatory gene in the LysR family (see Fig. 1b), whose product has no particular similarity to any known transcriptional regulator of any other *ddd* genes. The only other nearby genes of note encode two transporters of the BCCT type, one of which is 5′ of *dddP1* and separated by four intervening genes and the other 3′ of *dddP1* and separated by three intervening genes (see Sun et al. 2011).

The *dddP2* gene and its deduced product differ from that of *dddP1* in a number of ways. Although a member of the M24 family of peptidases (*E* value 7.8 e^{-42}), DddP2 is rather distantly related to those ratified versions of the enzyme that have DMSP lyase activity in other bacteria, in the Roseobacter clade (43% identity to *Roseovarius nubinhibens* DddP). Further, DddP2 is not particularly similar to the polypeptides represented by DddP1, to which it is ~50% identical. Instead, DddP2 is closest in sequence to polypeptides in the deduced proteomes of strains of some unrelated bacteria that have no known links to DMSP catabolism, including those in the genera *Burkholderia* and *Rhizobium/Sinorhizobium* (β- and α-Proteobacteria, respectively).

Based on the size of its upstream intergenic region, it is likely that *dddP2* of *O. doudoroffii* is co-transcribed with a gene whose product is in the HcaE family (COG4638). Members of this family are dioxygenases, with a catalytic C-terminal domain linked to an N-terminal region that resembles the family of Rieske proteins, which are involved in transferring the electrons necessary for the oxygenase reactions. There is no known role for such a protein in DMSP catabolism but we noted that Rieske family polypeptides are also encoded by a gene found downstream of *dddD* in *Oceanimonas* and by a gene divergently transcribed from the BCCT transporter 3′ of *Oceanimonas dddP1* (see Fig. 1). The *dddP2* of *O. doudoroffii* lies upstream of a gene that encodes a BCCT-type transporter, which is therefore a candidate for being involved in DMSP import, although, in this case, the intergenic spacing suggests that this gene would not be co-transcribed with *dddP2* (Fig. 1c; see Sun et al. 2011).

Ratification of the function of the *Oceanimonas doudoroffii dddP* genes

Given that both DddP1 and DddP2 of *Oceanimonas* were rather different from those DddP polypeptides that we had studied previously, we set out to confirm that they had functional DMSP lyase activity as follows. The *dddP1* and *dddP2* genes were individually cloned into the expression vector pET21a (in pBIO1933 and pBIO1934 respectively), following their amplification from genomic DNA, using primers that corresponded to sequences immediately 5′ and 3′ of these genes (see Fig. 1 and Supplementary Table 1). The resultant recombinant plasmids were each transformed into *E. coli* strain BL21 and the cell-free extracts were assayed for their Ddd phenotypes. In both cases, the transformants generated DMS when they were grown with DMSP, those with the cloned *dddP2* being considerably higher than those with *dddP1* (values of 151 and 12.4 pmol μg protein^{-1} min^{-1}, respectively).

Regulation of *Oceanimonas dddD*, *dddP1* and *dddP2* genes

It was shown previously (de Souza and Yoch 1995), and confirmed here (see above), that DMS production in *Oceanimonas* is induced by DMSP. We therefore examined if this was associated with enhanced expression of one or more of the *ddd* genes that we had identified. To do this, we made a series of transcriptional fusions in the wide host-range promoter-probe

plasmid pMP220 (see "Materials and methods"), which contains a reporter *lacZ* gene that lacks its own promoter. Fragments that spanned the promoter regions of *dddD, dddP1* and *dddP2* (see Fig. 1) were amplified from genomic DNA and cloned into pMP220 and the resultant plasmids were individually mobilised into *O. doudoroffi* strain J495 (rif^R derivative) by conjugation. The transconjugants were then assayed for β-galactosidase activity (encoded by *lacZ*) after pre-growth in the presence or absence of 2 mM DMSP, acrylate or 3HP. It was found that expression of the *dddD-lacZ* transcriptional fusion (pBIO1952) was dramatically increased, ca. 50-fold by pre-growth in DMSP, with values of 119 and 5779 Miller units following growth in the absence and presence of DMSP respectively. No induction of this *dddD-lacZ* fusion was obtained following growth in acrylate or 3HP, consistent with the failure of these compounds to induce DMS production. In contrast to the *dddD* fusion, the *dddP1-lacZ* fusion (pBIO1951) was expressed constitutively, with low-level expression irrespective of the presence or absence of any of the potential inducers, and the *dddP2-lacZ* fusion (pBIO1958) was expressed at even lower levels, which were barely detectable under all conditions tested.

Thus the enhanced Ddd^+ phenotype seen when cells of *Oceanimonas* were grown in DMSP is most likely due to the increase in expression of its *dddD* gene.

Failure to find a *dddY*-like gene in the genome of *Oceanimonas doudoroffii*

As mentioned above, de Souza and Yoch (1996) purified a DMSP lyase from *O. doudoroffii*. This enzyme had a similar size (48 kDa) to those of the deduced DddP1 and DddP2 polypeptides (50 and 48 kDa respectively). However, the experimentally determined N-terminal sequence (AQFQSQDDV KPASIDAWSGK), which resembles that of the processed DddY polypeptide of *Alcaligenes* (de Souza and Yoch 1996; Curson et al. 2011) does not match that of either the *dddP1* or *dddP2* gene products. The DddD polypeptide is much larger, with a deduced Mr of 92 kDa; not surprisingly, the deduced N-terminal sequence of the *O. doudoroffii* DddD does not correspond to that found by de Souza and Yoch (1996).

In an attempt to identify the *Oceanimonas* gene whose product includes this proposed N-terminal sequence, we obtained a near-complete genomic sequence of this strain, comprising a total of 3,829,948 bp, in 193 contigs. All of the predicted single-copy bacterial genes that we searched for, namely *recA,* 16S rRNA*, rpoB, rpoD, rpoS, rpoZ, gyrA, ftsZ, dnaK, infB, atpD, groEL, sodA,* as well as the regions represented by the cosmids described above, were present in the available sequence, indicating that it contained the great majority of the *O. doudoroffii* genome. Based on analysis of the read numbers and sequencing coverage of the genome, it was predicted to have covered 98.5% of the genome.

However, a search of this newly acquired genome yielded no sequences that matched the N-terminal sequence (see above) of the DMSP lyase described by de Souza and Yoch (1996). It is not clear if this was because it was encoded by a gene in a region that had not been sequenced or if, perhaps, the gene had been lost by spontaneous deletion in the intervening years since they analysed this strain.

We also noted that the partial *O. doudoroffii* genome did not contain any genes whose products resembled the other known DMSP lyases namely DddL, DddW or DddQ, nor was there any sign of a polypeptide sequence that corresponded to the DmdA DMSP demethylase (Howard et al. 2006).

Discussion

The work presented here complements and extends earlier physiological and biochemical studies on the ability of the marine bacterium *Oceanimonas doudoroffii* to catabolise DMSP (de Souza and Yoch 1995; Yoch et al. 1997). This strain contains at least three enzymes with DMSP cleavage activity, representing two very different families, the DddD CoA-transferases and the DddP M24 peptidases. The presence of multiple mechanisms of DMSP catabolism has been found in other bacteria, but, to our knowledge, these have been confined to the Roseobacter clade. Thus, it has been known for some time that some individual Roseobacter strains can both demethylate DMSP and can cleave it via "DMSP lyase" activities that liberate DMS (González et al. 1999, 2003). Recent genetic and genomic analyses have shown that most strains of Roseobacter whose genomes have been sequenced

contain *dmdA,* which encodes DMSP demethylase (Howard et al. 2006; Newton et al. 2010), plus at least one DMSP lyase. Indeed, some strains contain multiple DMSP lyases, with their own particular portfolios, encoded by different *ddd* genes. For example, *Ruegeria pomeroyi* DSS-3 has *dddQ*, *dddP* and *dddW,* whereas *Roseovarius nubinhibens* has two versions of *dddQ* plus *dddP* (Todd et al. 2009, 2011a, b).

This multiplicity of DMSP catabolic systems in the Roseobacters is in keeping with the importance of DMSP in the lifestyles of these bacteria (Newton et al. 2010). By the same token, our finding of multiple *ddd* genes in a γ-Proteobacterium implies that DMSP may be a key substrate for *O. doudoroffii,* as further illustrated by the *dddT*-like genes near *dddD, dddP1* and *dddP2,* with no less than four such predicted BCCT transporter genes being seen in the vicinity of *dddD.* It will be of interest to gauge the relative contributions of these different transporters in the natural environment (see Sun et al. 2011). To date, there are no reports of DMSP catabolism in other *Oceanimonas* strains and no other genome sequences of this genus or the closely related *Oceanisphaera* (Ivanova et al. 2004) are available, so there are no bioinformatic data on the prevalence of *ddd* and *dmdA* genes in this genus.

Given that *O. doudoroffii* contains both DddP and DddD, it was surprising that neither acrylate nor 3HP, respectively the initial catabolites of these enzymes, acted as carbon sources, unlike the situation in some other DMSP-catabolising bacteria (e.g. Todd et al. 2010). However, a strain of *Pseudomonas* that contains *dddD* and grows well on DMSP also fails to use 3HP as a carbon source (Curson et al. 2010). One possible explanation for this is that such bacteria lack effective acrylate and/or 3HP transporters. Some bacteria can grow at the expense of DMS as sole carbon source (Schäfer et al. 2010), so one other possibility is that *O. doudoroffii* might catabolise some of the DMS formed by DMSP cleavage. However, we found no evidence that this strain could grow when provided with DMS as sole carbon source. Indeed, when O. *doudoroffii* was grown in medium that contained both DMS plus a conventional carbon source (succinate), there was no detectable removal of the exogenous DMS.

DMS production in *O. doudoroffii* had been shown previously to be induced by DMSP (de Souza and Yoch 1995) and we obtained strong evidence that this is due to the markedly increased levels of transcription of its *dddD* gene when the cells were pregrown in DMSP. It has been noted that the expression of *dddD* in several other bacteria is also highly induced by the DMSP substrate (Todd et al. 2007, 2010). In contrast, transcription of the *Oceanimonas dddP1* gene was at low levels and was unaffected by any of the potential co-inducers tested here. More strikingly, *dddP2* was not expressed at detectable levels in any of our conditions. Although the *dddP*-like genes in some bacteria of the Roseobacter clade are induced by DMSP (Todd et al. 2009), the factor of increase is modest (2–4-fold). It remains to be seen if either or both of these *dddP*-like genes in *Oceanimonas* are expressed in response to some unknown environmental signal that has so far eluded us in laboratory conditions.

The sequence and the local genomic geography of the *O. doudoroffii dddD* gene was conventional in the sense that *dddD* was clustered with other genes (*dddT, dddB, dddC* and *dddR*) that were in the same, or very similar, relative positions in other γ-proteobacteria such as *Marinomonas* and *Pseudomonas* (Todd et al. 2007; Curson et al. 2010), and the sequences of the corresponding products of all these genes were very similar in these different strains.

A different situation pertains to the two *Oceanimonas dddP* genes. The DddP1 polypeptide was very similar to the gene products in a miscellany of other microbes, ranging from two unrelated bacteria, *Vibrio orientalis* and *Puniceispirillum marinum* through to some eukaryotic fungi. The finding of members of this out-branch of the DddP polypeptides in such a disparate group of organisms is strong evidence that these were acquired by repeated instances of HGT. This conclusion is bolstered by the fact that of >20 strains of *Vibrio* that have been sequenced, *V. orientalis* CIP 102891 is the only one that contains *dddP*. It is not clear why this subset of the *dddP* gene should, apparently, be more prone to HGT to other bacteria than the forms that are represented by the DddP enzymes in the Roseobacters.

In its primary sequence, DddP2 differs not only from DddP1, but also from the DddP polypeptides in the Roseobacter clade (Todd et al. 2009). Nevertheless it can be a functional enzyme, as shown by the high level of DMSP-dependent production seen in *E. coli* containing *dddP2*. Indeed, its activity when expressed in the heterologous host *E. coli* was significantly

higher than that with the cloned *dddP1*. The behaviour of *dddP2* has some similarities to that of another gene, *dddL,* which encodes a wholly different type of DMSP lyase, in the Roseobacter species *Dinoroseobacter shibae* DFL 12. This gene, when cloned and expressed in *E. coli,* confers the ability to make DMS from DMSP even though *D. shibae* DFL 12 itself does not make DMS (Dickschat et al. 2010) and the *dddL* gene is not expressed at detectable levels (JD Todd, unpublished observations).

There have been several reports on the frequencies of the different *ddd* and *dmd* genes in the massive metagenomic databases, most importantly those in the Global Ocean Sampling (GOS) data set (Rusch et al. 2007). These have shown that the *dmdA* demethylase is the most abundant, and that *dddP* and *dddQ* are widespread, compared to other *ddd* genes such as *dddL* and *dddD* (Todd et al. 2009; Newton et al. 2010). The census-taking of the DddP-type of DMSP lyase in the GOS was based on the numbers of close homologues of the originally described version of this enzyme, identified in the Roseobacters. Performing the same sort of survey with the newly identified *Oceanimonas* DddP1 and DddP2 polypeptides shows that these, too, have close homologues (*E* values $<e^{-86}$) in the GOS, and that they are both relatively abundant, with census numbers that are each approximately a fifth of the number of metagenomic homologues that closely resembled the original DddP enzyme, as exemplified by the versions found in the Roseobacters and described by Todd et al. (2009).

De Souza and Yoch (1996) predicted that *O. doudoroffii* has a DMSP lyase that resembles the recently characterized periplasmic DddY of *Alcaligenes faecalis* (Curson et al. 2011), but that it has an important difference, in that it was proposed to be cytoplasmic in *Oceanimonas.* However, despite sequencing an estimated 98.5% of the *O. doudoroffii* genome, we did not find any match for the sequence of the published N-terminal region of this enzyme. Since the DMSP lyases of *A. faecalis* and of *O. doudoroffii* are immunologically cross-reactive (de Souza and Yoch, 1996), we plan to use antibody against over-expressed DddY of the former species to examine the production of the corresponding protein in *Oceanimonas* and will attempt to identify the corresponding gene, in case it is in a region which, by chance, was not represented in our partial genome sequence of this strain.

In strains of different Roseobacters that have multiple ways of catabolising DMSP, the relative importance of the different pathways was shown to depend on environmental factors such as the concentration of the DMSP substrate (González et al. 2003). Having now found a multiplicity of DMSP catabolic pathways in a different type of bacterium, it will be of interest to determine which of the DddP- and DddD-mediated routes are most important for *Oceanimonas* and to establish if and how the relative contributions of these enzymes are affected by factors in the natural environments of this bacterium.

Acknowledgments The work was funded by grants from the United Kingdom Biotechnology and Biological Sciences Research Council and the Natural Environment Research Council and a Tyndall Centre Studentship to EKF. We thank Jennifer Sampson and Pamela Wells for technical support.

References

Anon (2001) Notification that new names and new combinations have appeared in volume 51 part 1, of the IJSEM. Int J Syst Evol Microbiol 51:269

Baumann L, Baumann P, Mandel M, Allen RD (1972) Taxonomy of aerobic marine eubacteria. J Bacteriol 110:402–429

Beringer JE (1974) R factor transfer in *Rhizobium*. J Gen Microbiol 84:188–198

Brown GR, Sutcliffe IC, Cummings SP et al (2001) Reclassification of [*Pseudomonas*] *doudoroffii* (Baumann et al. 1983) into the genus *Oceanomonas* gen. nov. as *Oceanomonas doudoroffii* comb. nov., and description of a phenol-degrading bacterium from estuarine water as *Oceanomonas baumannii* sp. nov. I. J Syst Evol Microbiol 51:67–72

Curson ARJ, Rogers R, Todd JD, Brearley CA, Johnston AWB (2008) Molecular genetic analysis of a dimethylsulfoniopropionate lyase that liberates the climate-changing gas dimethylsulfide in several marine alpha-proteobacteria and *Rhodobacter sphaeroides*. Env Microbiol 10:757–767

Curson ARJ, Sullivan MJ, Todd JD, Johnston AWB (2010) Identification of genes for dimethyl sulfide production in bacteria in the gut of Atlantic Herring (*Clupea harengus*). ISME J 4:144–146

Curson ARJ, Sullivan MJ, Todd JD, Johnston AWB (2011) DddY, a periplasmic dimethylsulfoniopropionate lyase found in taxonomically diverse species of proteobacteria. ISME J 5:1191–1200

de Souza MP, Yoch DC (1995) Comparative physiology of dimethyl sulfide production by dimethylsulfoniopropionate lyase in *Pseudomonas doudoroffi*i and *Alcaligenes* sp. Strain M3A. Appl Env Microbiol 61:3986–3991

de Souza MP, Yoch DC (1996) N-terminal amino acid sequences and comparison of DMSP lyases from *Pseudomonas doudoroffii* and *Alcaligenes* strain M3A. In: Kiene RP, Visscher PT, Keller MD, Kirst GO (eds) Environmental and

biological chemistry on dimethylsulfoniopropionate and related sulfonium compounds. Plenum Press, New York, pp 293–304

Dickschat JS, Zell C, Brock NL (2010) Pathways and substrate specificity of DMSP catabolism in marine bacteria of the Roseobacter clade. Chem Biochem 15:417–425

Figurski DH, Helinski DR (1979) Replication of an origin-containing derivative of plasmid RK2 dependant on a plasmid function provided *in trans*. Proc Nat'l Acad Sci USA 76:1648–1652

González JM, Kiene RP, Moran MA (1999) Transformation of sulfur compounds by an abundant lineage of marine bacteria in the α-subclass of the class *Proteobacteria*. Appl Env Microbiol 65:3810–3819

González JM, Covert JS, Whitman WB, Henriksen JR, Mayer F, Scharf B, Schmitt R, Buchan A, Fuhrman JA, Kiene RP, Moran MA (2003) *Silicibacter pomeroyi* sp. nov. and *Roseovarius nubinhibens* sp. nov., dimethylsulfoniopropionate-demethylating bacteria from marine environments. Int J Syst Evol Microbiol 53:1261–1269

Howard EC, Henriksen JR, Buchan A, Reisch CR, Bürgmann H et al (2006) Bacterial taxa that limit sulfur flux from the ocean. Science 314:649–652

Ivanova EP, Flavier S, Christen R (2004) Phylogenetic relationships among marine *Alteromonas*-like proteobacteria: emended description of the family Alteromonadaceae and proposal of Pseudoalteromonadaceae fam. nov., Colwelliaceae fam. nov., Shewanellaceae fam. nov., Moritellaceae fam. nov., Ferri. Int J Syst Evol Microbiol 54:1773–1788

Kirkwood M, Le Brun NE, Todd JD, Johnston AWB (2010a) The *dddP* gene of *Roseovarius nubinhibens* encodes a novel lyase that cleaves dimethylsulfoniopropionate into acrylate plus dimethyl sulfide. Microbiol 156:1900–1906

Kirkwood M, Todd JD, Rypien KL, Johnston AWB (2010b) The opportunistic coral pathogen *Aspergillus sydowii* contains *dddP* and makes dimethyl sulfide from dimethylsulfoniopropionate. ISME J 4:147–150

Ledyard KM, DeLong EF, Dacey JWH (1993) Characterization of a DMSP-degrading bacterial isolate from the Sargasso Sea. Arch Microbiol 160:312–318

Newton RJ, Griffin LE, Bowles KM, Meile C, Gifford S, Givens CE et al (2010) Genome characteristics of a generalist marine bacterial lineage. ISME J 4:784–798

Oh HM, Kwon KK, Kang I, Kang SG, Lee JH, Kim SJ, Cho JC (2010) Complete genome sequence of "Candidatus *Puniceispirillum marinum*" IMCC1322, a representative of the SAR116 clade in the alphaproteobacteria. J Bacteriol 192:3240–3241

Raina JB, Dinsdale EA, Willis BL, Bourne DG (2010) Do the organic sulfur compounds DMSP and DMS drive coral microbial associations? Trends Microbiol 18:101–108

Rusch DB, Halpern AL, Sutton G, Heidelberg KB, Williamson S et al (2007) The *Sorcerer II* Global Ocean Sampling Expedition: Northwest Atlantic through Eastern Tropical Pacific. PLoS Biol 5(3):e77. doi:10.1371/journal.pbio.0050077

Sambrook S, Fritsch EF, Maniatis T (1989) "Molecular cloning: a laboratory manual". Cold Spring Harbor Laboratory, Cold Spring Harbor

Schäfer H, Myronova N, Boden R (2010) Microbial degradation of dimethylsulfide and related C1-sulfur compounds: organisms and pathways controlling fluxes of sulfur in the biosphere. J Exp Bot 61:315–334

Spaink HP, Okker RJH, Wijffelman CA, Pees E, Lugtenberg BJJ (1987) Promoters in the nodulation region of the *Rhizobium leguminosarum* Sym plasmid pRL1JI. Plant Mol Biol 9:27–39

Staskawicz B, Dahlbeck D, Keen N, Napoli C (1987) Molecular characterization of cloned avirulence genes from race 0 and race 1 of *Pseudomonas syringae* pv. *glycinea*. J Bacteriol 169:5789–5794

Studier FW, Moffat BA (1986) Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes. J Mol Biol 189:113–130

Sullivan MJ, Curson ARJ, Shearer N, Todd JD, Green RT, Johnston AWB (2011) Unusual regulation of a leaderless operon involved in the catabolism of dimethylsulfoniopropionate in *Rhodobacter sphaeroides*. PLoS One 6(1):e15972

Sun L, Curson ARJ, Todd JD, Johnston AWB (2011) Diversity of DMSP transport in marine bacteria, revealed by genetic analyses. Biogeochemistry. doi:10.1007/s10533-011-9666-z

Todd JD, Rogers R, Li YG, Wexler M, Bond PL et al (2007) Structural and regulatory genes required to make the gas dimethyl sulfide in bacteria. Science 315:666–669

Todd JD, Curson ARJ, Dupont CL, Nicholson P, Johnston AWB (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. Env Microbiol 11:1376–1385

Todd JD, Curson ARJ, Nikolaidou-Katsaridou N, Brearley CA, Watmough NJ, Chan Y, Page PC, Sun L, Johnston AWB (2010) Molecular dissection of bacterial acrylate catabolism–unexpected links with dimethylsulfoniopropionate catabolism and dimethyl sulfide production. Env Microbiol 12:327–343

Todd JD, Curson ARJ, Kirkwood M, Sullivan MJ, Green RT, Johnston AWB (2011a) DddQ, a novel, cupin-containing, dimethylsulfoniopropionate lyase in marine roseobacters and in uncultured marine bacteria. Env Microbiol 13:427–438

Todd JD, Kirkwood M, Newton-Payne S, Johnston AWB (2011b). DddW, a third DMSP lyase in a model Roseobacter marine bacterium, *Ruegeria pomeroyi DSS*-3. ISME J. doi: 10.1038/ismej.2011.79. [Epub ahead of print]

Wexler M, Yeoman KH, Stevens JB, de Luca NG, Sawers G, Johnston AWB (2001) The *Rhizobium leguminosarum tonB* gene is required for the uptake of siderophore and haem as sources of iron. Mol Microbiol 41:801–816

Wood WB (1966) Host specificity of DNA produced by *Escherichia coli*: bacterial mutations affecting the restriction and modification of DNA. J Mol Biol 16:118–133

Yang Y, Yeh L, Cao Y, Baumann L, Baumann P, Tang JS, Beaman B (1983) Characterization of marine luminous bacteria isolated off the coast of China and description of *Vibrio orientalis* sp. nov. Curr Microbiol 8:95–100

Yoch DC (2002) Dimethylsulfoniopropionate: its sources, role in the marine food web, and biological degradation to dimethylsulfide. Appl Env Microbiol 68:5804–5815

Yoch DC, Ansede JH, Rabinowitz KS (1997) Evidence for intracellular and extracellular dimethylsulfoniopropionate (DMSP) lyases and DMSP uptake sites in two species of marine bacteria. Appl Env Microbiol 63:3182–3188

Young JPW, Crossman LC, Johnston AWB, Thomson NR, Ghazoui ZF, Hull KH et al (2006) The genome of *Rhizobium leguminosarum* has recognizable core and accessory components. Genome Biol 7:R34