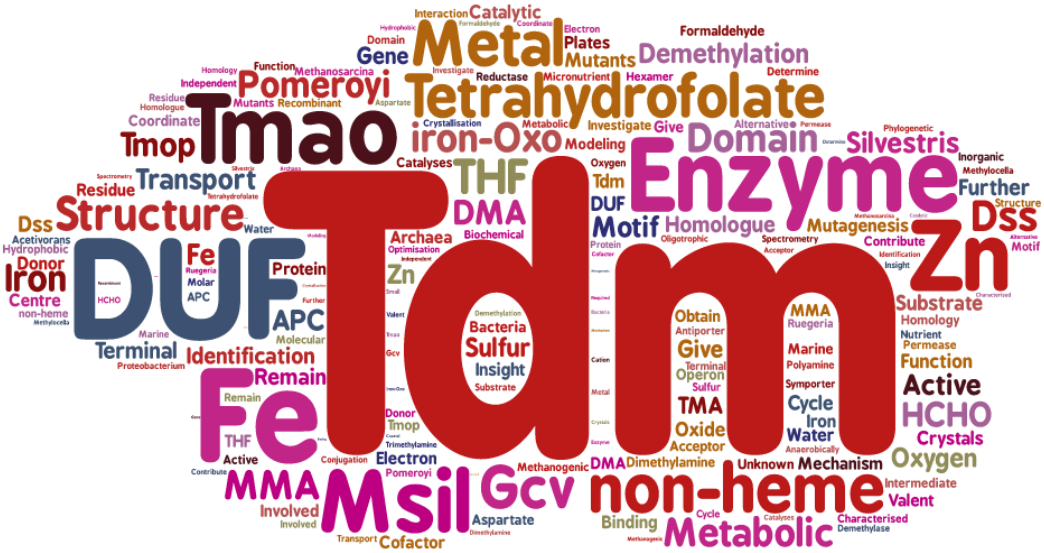
GENETICS AND MOLECULAR MECHANISM OF TRIMETHYLAMINE *N*-OXIDE DEMETHYLASE OF *METHYLOCELLA SILVESTRIS* BL2

Yijun Zhu

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

September 2016

University of Warwick Coventry, UK



Contents

Acknowledgments

I would like to thank my supervisors Dr. Yin Chen and Dr. Hendrik Schäfer for all of their guidance, unrelenting support and encouragement throughout this process. I would like to pay tribute to my PhD committee Prof. Eric Holub and Dr. Kevin Purdy for their advice and helpful discussions.

A special mention for Dr. Claudia Blindauer and Prof. Tim Bugg, for our insightful and detailed discussions. Many Thanks to Prof. Alexander Cameron, for his expert guidance in crystallisation and structure analysis.

I would also like to thank all past and current members of D14, especially Julie and Ellie for all their technical support and encouragement. Thanks go to the Ian, Alastair, Tiantian, Eileen and Michi, whose friendships have made this PhD an enjoyable experience.

My family has been encouraging, supportive and shown belief in me and my work. Without their support, this would not be possible. Finally I owe much to my other half Xueming, without whose love and understanding I would not have completed this work.

Declaration

This thesis is submitted to the University of Warwick in support of my application for the degree of Doctor of Philosophy. It has been composed by myself and has not been submitted in any previous application for any degree.

The work presented (including data generated and data analysis) was carried out by the author except in the cases outlined below:

1) Generation of the various *Methylocella silvestris* BL2 knockout mutants presented in Chapter 3 was carried out by Dr. Yin Chen, University of Warwick.
2) The plasmid pET28a-Tdm presented in Chapter 3 and 4 was generated by Dr Yin Chen, University of Warwick.
3) The plasmid pET28a_Tdm_DSS3 and pET28a_Tdm_HIMB59 presented in Chapter 5 were generated by Dr Ian Lidbury, University of Warwick.
4) Metal quantification of Tdm presented in Chapter 4 and 5 using ICP-OES/ MS was performed by Ms Amira Ksibe and Dr Lijiang Song, University of Warwick.
5) Electron paramagnetic resonance analysis presented in Chapter 4 was performed by Dr Christopher Wedge.

Parts of this thesis have been published by the author:

Zhu Y, Jameson E, Parslow RA, Lidbury I, Fu T, Dafforn TR *et al* (2014). Identification and characterization of trimethylamine *N*-oxide (TMAO) demethylase and TMAO permease in *Methylocella silvestris* BL2. *Environ Microbiol* **16:** 3318-3330.

Zhu Y, Ksibe A, Schäfer H, Blindauer C, Bugg T, Chen. Y (2016) O_2-independent demethylation of trimethylamine *N*-oxide by Tdm of *Methylocella silvestris*. *FEBS J*. DOI: 10.1111/febs.13902.

Abstract

Methylocella silvestris, an alphaproteobacterium isolated from a forest soil, can grow on trimethylamine *N*-oxide (TMAO) as a sole nitrogen source, however, the molecular and biochemical mechanisms underpinning its growth remain unknown. Marker-exchange mutagenesis enabled the identification of several genes involved in TMAO metabolism, including *Msil_3606*, a permease of the amino acids-polyamine (APC) superfamily, and *Msil_3603*, consisting a N-terminal domain of unknown function (DUF1989) and a C-terminal tetrahydrofolate-binding domain. Null mutants of *Msil_3603* and *Msil_3606* can no longer grow on TMAO. Purified Msil_3603 from recombinant *Escherichia coli* can convert TMAO to dimethylamine and formaldehyde (1 TMAO → 1 dimethylamine + 1 formaldehyde), confirming that it encodes a bona fide TMAO demethylase (Tdm).

Site-directed mutagenesis, homology modelling and metal analyses by inorganic mass spectrometry have been applied to gain insight into metal stoichiometry and underlying catalytic mechanism of Tdm of *M. silvestris*. Herein, it is demonstrated that active Tdm has 1 molar equivalent of Zn^{2+} and 1 molar equivalent of non-haem Fe^{2+}. Further investigation of Zn^{2+} and Fe^{2+}-binding sites through homology modelling and site-directed mutagenesis revealed that Zn^{2+} is coordinated by a 3-sulfur-1-O motif. An aspartate residue (D198) likely bridges Fe^{2+} and Zn^{2+} centres. H276, and maybe H256, contribute to Fe^{2+} binding. Site-directed mutagenesis of Tdm also led to the identification of three hydrophobic aromatic residues likely involved in substrate coordination (F259, Y305, W321), potentially through a cation-π interaction. Furthermore, a cross-over experiment using a substrate intermediate analogue gave direct evidence that a trimethylamine-alike intermediate was produced during the Tdm catalytic cycle, suggesting TMAO has a dual role of being both a substrate and an oxygen donor for formaldehyde formation.

In this thesis, I attempted to resolve 3D-structure of Tdm to investigate structure-function relationship. Various Tdm homologues and mutants have been screened for crystallisation. Tdm from *Ruegeria pomeroyi* DSS-3 forms small 2D plates, hence warrants further refinements. Although Tdm of *R.pomeroyi* has comparable activity to that of *M. silvestris*, it is different from the one of *M.silvestris* in that a trimer and more susceptible to EDTA chelator.

Together, this study has contributed to the understanding of the genetic and biochemical mechanisms for TMAO degradation in *M. silvestris*, and provides novel insight into the role of Zn^{2+} and Fe^{2+} in the catalysis of TMAO demethylation by this unique oxygen-independent enzyme.

Abbreviations

ADH	Alcohol dehydrogenase	**DMS**	Dimethylsulphide
Asc	Ascorbic acid	**DMSO**	Dimethylsulfoxide
EDTA	Ethylenediaminetetraacetic acid	**DMSP**	Dimethylsulfoniopropionate
HCHO	Formaldehyde	**dNTP**	Deoxynucleotide triphosphate
ABC	ATP-binding cassette	**DUF**	Domain of unknown function
ANOVA	Analysis of variance	**EPR**	Electron paramagnetic resonance
ATP	Adenosine triphosphate	**ESI**	Electrospray ionization
BLAST	Basic Local Alignment Search Tool	**FAD**	Flavin adenine dinucleotide
BS-DMA	Benzenesulfonyl dimethylamine adduct	**GBT**	Glycine betaine
BS-MEA	Benzenesulfonyl methylethylamine adduct	**GC**	Gas chromatography
BSA	Bovine serum albumin	**GCV_T**	Glycine cleavage T protein
BSC	Benzenesulfonyl chloride	**GMA**	γ-glutamylmethylamide
CDA	Cytidine deaminase	**GMAS**	γ-glutamylmethylamide synthetase
CntA/B	Carnitine oxygense	**GOS**	Global Ocean Survey
CutC	Choline Trimethylamine-Lyase	**GrdH**	Betaine reductase
Da	Dalton	**HEPES**	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
DMA	Dimethylamine	**HMS^-**	Hydroxymethanesulfonate
DMADH	Dimethylamine dehydrogenase	**HNLC**	High Nutrient–Low Chlorophyll
DMEA	dimethyethylamine	**ICP**	Inductively coupled plasma-mass spectrometry
DMPO	5,5-dimethyl- pyrroline-N-oxide	**IPTG**	Isopropyl β-D-1-thiogalactopyranoside

Abbreviations

K_i	Inhibition constant	**OES**	Optical emission spectrometer
K_m	Michaelis-Menten constant	**PAGE**	Polyacrylamide gel electrophoresis
LC	Liquid chromatography	**PCR**	Polymerase chain reaction
m/z	Mass to charge ratio	**PDB**	Protein data bank
MADH	Methylamine dehydrogenase	**PEG**	Polyethylene glycol
MaoXII	Methylamine oxidase	**PIPES**	1,4-piperazinediethanesulfonic acid
MAs	Methylated amines	**SBP**	Substrate binding protein
mau	Methylamine utilization genes	**SDS**	sodium dodecyl sulphate
MEA	Methyethylamine	**Tdm**	TMAO demethylase
MES	2-(N-morpholino)ethanesulfonic acid	**THF**	Tetrahydrofolate
MMA	monomethylamine	**TMA**	Trimethylamine
MRC	Marine roseobacter clade	**TMADH**	Trimethylamine dehydrogenase
MS	Mass spectrometer	**TMAO**	Trimethylamine *N*-oxide
NAD	Nicotinamide adenine dinucleotide (oxidised form)	**Tmm**	TMA monooxygenase
NADH	Nicotinamide adenine dinucleotide (reduced form)	**TmoP**	TMAO permease
NADPH	Nicotinamide adenine dinucleotide phosphate	**TmoV**	Transmembrane permease protein of TMAO ABC transporter
NMG	*N*-methylglutamate	**TmoW**	ATP-binding domain protein of TMAO ABC transporter
NMGDH	*N*-methylglutamate dehydrogenase	**TmoX**	Substrate binding protein of TMAO ABC transporter
NMGS	*N*-methylglutamate synthase	**Tris**	Tris (hydroxymethyl) aminomethane
		V_{max}	Maximal velocity

List of Figures

List of Tables

List of appendices

Chapter 1 Introduction

1.1 Methylated amines

1.1.1 Environmental and health significance

Methylated amines (MAs) namely, trimethylamine (TMA), dimethylamine (DMA) and monomethylamine (MMA), together with other aliphatic amines, have been identified as some of the important components of trace gases in the atmosphere. Because MAs are basic, they can neutralise acidic clusters and may therefore play an important role in the formation of new atmospheric particles which can act as cloud condensation nuclei (**Figure 1-1**) (Almeida et al 2013, Cape et al 2011, Erupe et al 2011, Ge et al 2011, Murphy et al 2007, Youn et al 2015). Such aerosol particles have great potential to cause a net cooling effect of the climate by scattering sunlight and by leading to the formation of cloud droplets, which can counteract global warming induced by greenhouse gases (Ramanathan et al 2001).

Animal husbandry, biomass burning and oceans are three major sources of MAs of the atmosphere (**Figure 1-1**)(Ge et al 2011). The annual flux of MAs into the atmosphere is estimated to be in the order of 285 ± 78 Gg globally, of which TMA makes up approximately 170 Gg that are mainly derived from animal husbandry (Ge et al 2011). The composition of MAs varies spatially. In semi-arid (Tucson, Arizona) and marine (central coast of California, north and tropical Atlantic) areas, DMA has been reported as the most abundant methylated amine (Facchini et al 2008, Muller et al 2009, Youn et al 2015), suggesting that the oceans are a DMA source.

Being an important component of organic nitrogen in the atmosphere, the cycling of MAs between land and oceans can also affect global biogeochemical cycles of nitrogen through atmospheric deposition (**Figure 1-1**) (Cape et al 2011).

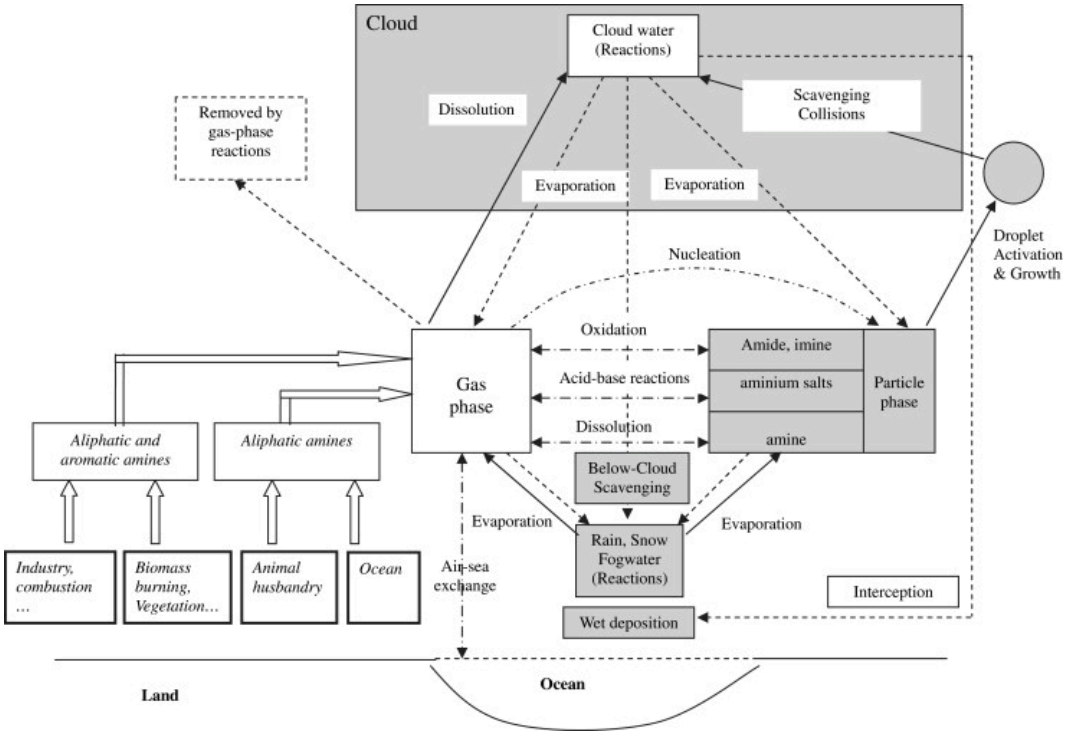


Figure 1-1 Schematic diagram of the amine cycle in the atmosphere. Adapted from Ge et al (2011).

The presence of such large quantities of MAs in the atmosphere can significantly affect human well-being. For example, MAs can be precursors for carcinogens such as *N*-nitrosodimethylamine, causing concerns for public health (Mitch et al 2003).

1.1.2 Methylated amine metabolism by microorganisms

MAs have been recognised as an important carbon and energy source for microbes, primarily due to research interest in methylotrophy (Anthony 1982, Chistoserdova et al 2009). Each methyl group of MAs is oxidised to one molecule of formaldehyde, which can be further oxidised to produce energy and reductants in the form of NADH or NADPH or assimilated into cell biomass (Anthony 1982, Chistoserdova et al 2009). But recently, it has been demonstrated that non-methylotrophic bacteria can utilise MAs as supplementary energy source and/or N source by sequestering N from MAs without assimilating C (Chen et al 2010b, Chen 2012, Lidbury et al 2015, Sun et al 2011, Wischer et al 2015).

Bacterial MA metabolism pathways have been proposed decades ago and some of the key enzymes have been purified and characterised. TMA can be oxidised directly to DMA by a trimethylamine dehydrogenase (TMADH) or indirectly through the formation of TMAO as the key intermediate (Anthony 1982). DMA can be further oxidised to MMA by a DMA monooxygenase (Dmm) or a DMA dehydrogenase (DMADH) (Anthony 1982, Yang et al 1995) (**Figure 1-2**). MMA oxidation is catalysed either directly by a periplasmic methylamine dehydrogenase (MADH) or a methylamine oxidase to form formaldehyde and ammonium, or indirectly through the formation of *N*-methylated amino acid intermediates, *i.e. N*-methylglutamate (NMG) and γ-glutamylmethylamide (GMA) (**Figure 1-2**) (Anthony 1982, Chen et al 2010c, Chistoserdov et al 1994, Husain and Davidson 1987, Zhang et al 1993). The *N*-methylglutamate (NMG) pathway, originally described in *Aminobacter aminovorans*

(previously known as *Pseudomonas* strain MA and strain MS), is a complex, multistep pathway in which the methyl group from MA is transferred to glutamate by either *N*-methylglutamate synthase (NMGS) or γ-glutamylmethylamide synthetase (GMAS) to form two novel amino acid derivatives, NMG and γ-glutamylmethylamide (GMA), which are then oxidized by *N*-methylglutamate dehydrogenase (NMGDH) to release formaldehyde (**Figure 1-2**) (Chen et al 2010c, Levitch 1976, Shaw et al 1966).

Although some of the enzymes involved in microbial MMA metabolism were purified and characterised over 3 decades ago, the encoding genes for several enzymes in various methylotrophs and non-methylotrophic heterotrophs have only been identified very recently (**Table 1-1**). The genes encoding Tdm and Dmm were unidentified when this project started.

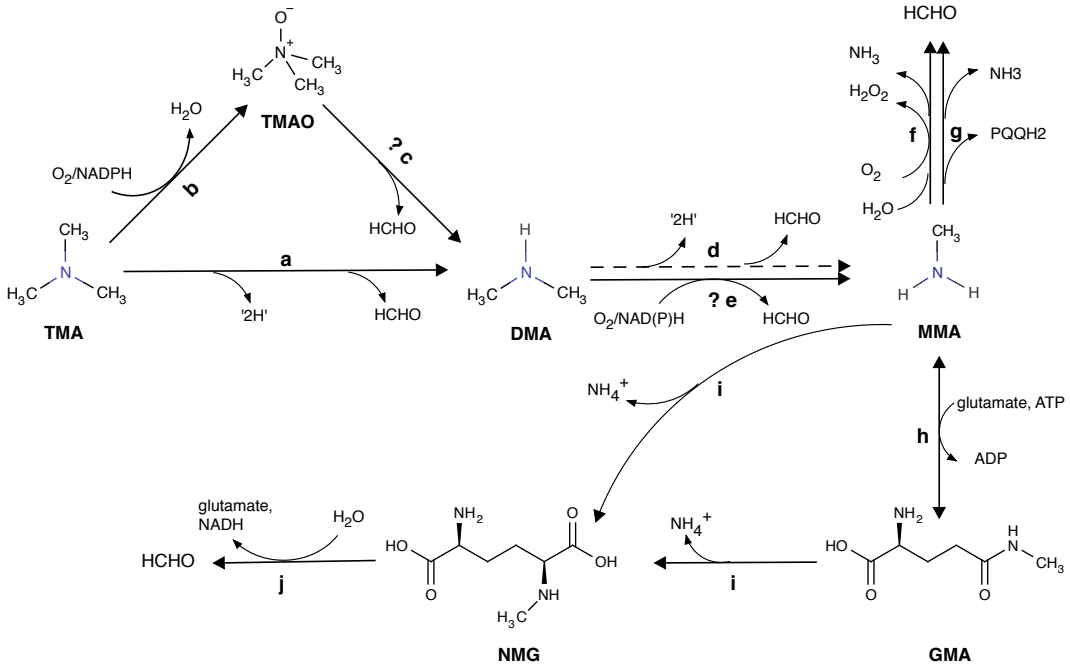


Figure 1-2 The oxidation of methylated amines to formaldehyde. The enzymes catalysing these reactions are as follows: (a) trimethylamine dehydrogenase (TMADH); (b) trimethylamine monooxygenase (Tmm); (c) trimethylamine *N*-oxide demethylase (aldolase) (Tdm); (d) dimethylamine dehydrogenase (DMADH) (anaerobic *Hyphomicrobia*); (e) dimethylamine monooxygenase (Dmm); (f) amine oxidase (Mao) (in *Arthrobacter* and methazotrophic yeasts); (g) methylamine dehydrogenase (Mau); (h) γ-glutamylmethylamide synthase (GMAS); (i) *N*-methylglutamate synthase (NMGS); (j) *N*-methylglutamate dehydrogenase (NMGDH). ?: encoding gene unidentified. Modified from (Anthony 1982, Chen et al 2010c, Chistoserdova et al 2009, Latypova et al 2010, Nayak and Marx 2014).

Table 1-1 Selected identified genes involved in MAs metabolism

Gene	Locus tag	Organism	Reference
tmadh	DHTM_METME	*Methylophilus methylotrophus*	(Boyd et al 1992)
tmm	Msil_3604	*Methylocella silvestris* BL2	(Chen et al 2011)
	Spo_1551	*Rugeria pomeroyi*	(Chen et al 2011)
dmadh	DHDM_HYPSX	*Hyphomicrobium* sp.	(Yang et al 1995)
maoXII	AMO2_ARTS1	*Arthrobacter* strain P1	(Zhang et al 1993)
mau	MexAM1_META1p2769-2779	*Methylobacterium extorquens* AM1	(Chistoserdov et al 1994)
	Pden_4730-4738	*Paracoccus denitrificans* PD1222	(Delorme et al 1997)
gmaS	Msil_2635	*Methylocella silvestris* BL2	(Chen et al 2010c)
	METUNv1_03599	*Methyloversatilis universalis* FAM5	(Latypova et al 2010)
	METDI2327	*Methylobacterium extorquens* DM4	(Gruffaz et al 2014)
	Atu4230	*Agrobacterium tumefaciens*	(Chen et al 2010b)
nmgS	METUNv1_03600–03602	*Methyloversatilis universalis* FAM5	(Latypova et al 2010)
	METDI2324–2326	*Methylobacterium extorquens* DM4	(Gruffaz et al 2014)
	Msil_2632–2634	*Methylocella silvestris* BL2	(Chen et al 2010c)
	Atu4227-4229	*Agrobacterium tumefaciens*	(Chen et al 2010b)
nmgdh	METUNv1_03595–03598	*Methyloversatilis universalis* FAM5	(Latypova et al 2010)
	METDI2319–2322	*Methylobacterium extorquens* DM4	(Gruffaz et al 2014)
	Msil_2636–2639	*Methylocella silvestris* BL2	(Chen et al 2010c)
	Atu4221_4224	*Agrobacterium tumefaciens*	(Chen et al 2010b)

1.1.3 The indirect MA metabolism pathway may play an important role in global C and N cycling

TMAO is a key intermediate in the indirect pathways of methylated amines metabolism. TMA is oxidised by TMA monooxygenase (Tmm), a NADPH-dependent flavin monooxygenase (Anthony 1982, Chen et al 2011). TMAO is further degraded by TMAO demethylase (adolase) (Tdm) to form a molecule of formaldehyde and DMA, which is further metabolised to form a molecule of formaldehyde and MMA (Anthony 1982). Recently, the genes encoding key enzymes involved in the indirect MA-utilization pathway, including Tmm (Msil_3604), GmaS, NMG synthase and NMG dehydrogenase have been identified in terrestrial (*i.e. M. silvestris* (Chen et al 2011), *M. universalis* FAM5 (Latypova et al 2010), *M. extorquens* DM4 (Gruffaz et al 2014), *A. tumefaciens* (Chen et al 2010b)) and marine microbes (*Roseobacter* clade, *i.e. R. pomeroyi* (Chen et al 2011)), but the genes encoding Tdm and Dmm remained unknown when the project started.

Recent studies have shown that non-methylotrophic heterotrophs can oxidise MAs to sequester N and generate ATP without assimilating C (Chen et al 2010b, Chen 2012, McClelland et al 2001). Indeed, Tmm homologues occur in many non-methylotrophic *Alphaproteobacteria*, in particular the *Rhodobacteraceae* (marine *Roseobacter* clade, MRC) and the SAR11 clade (Chen et al 2011, Chen 2012), both of which are key players in the oceanic carbon cycle (Rusch et al 2007). Blast analysis with the *tmm* gene against the Global Ocean Sampling (GOS) dataset revealed that ~20% of the bacteria in the surface ocean contain *tmm*, implying that the indirect MAs pathway play an important role in global C and N cycles (Chen et al 2011). Phylogenetic analysis

indicated that the Tmm homologues from the GOS database could be placed into three groups (**Figure 1-3**), the majority of which were closely related to Tmm homologues in the SAR11 clade bacteria (94.1%) and the MRC bacteria (4.7%), respectively (Chen et al 2011). MRC accounts for around 10% of bacterioplankton in the open ocean (Buchan et al 2005, Labbe and Rettmer 1989, Rappé and Giovannoni 2003) and up to 25% in coastal waters (DeLong 2005, Giebel et al 2011, Suzuki et al 2001). SAR11 clade of the *Alphaproteobacteria* comprises about 25% cells of coastal, estuary and open sea habitats (Malmstrom et al 2004).

The significance of the indirect MA oxidation pathway utilised by marine bacterioplankton is also supported by the presence of comparable abundance of the *gmaS* and the absence of the *tmadh* gene in the MRC isolates (Chen 2012). Therefore, one of the key aims of this thesis is to identify the *tdm* gene encoding TMAO demethylase to gain more insight into the indirect pathway of MA metabolism.

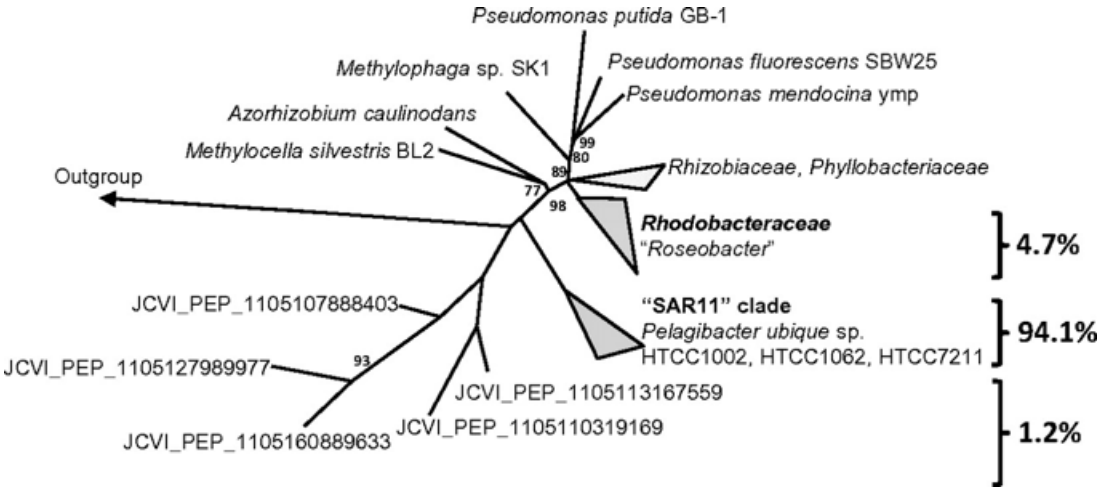


Figure 1-3 An unrooted tree showing Tmm homologues retrieved from sequenced bacterial genomes and the Global Ocean Sampling expedition data set. The neighbour-joining tree was constructed using sequences retrieved from sequenced bacterial genomes (~450 amino acids). Environmental sequences were added by parsimony. Bootstrap values were calculated based on 100 replicates. *Homo sapiens* FMO3 was used as the outgroup. Adapted from Chen et al (2011).

1.2 TMAO's biological and physiological role

1.2.1 Osmolyte and osmoprotectant

TMAO is a ubiquitous organic osmolyte that occurs in a wide variety of marine biota, including algae, zooplankton and fish, to maintain cell turgor in osmotic environment (Yancey 2005). Organic osmolytes are small, highly soluble, organic molecules, which include sugars and polyols (*e.g.* trehalose, glycerol) and α- and β-amino acids and their derivatives, including methylamines (*e.g.* TMAO, glycine betaine) and DMSP (Roesser and Müller 2001, Yancey et al 1982, Yancey 2005). These organic osmolytes do not interfere with the central metabolism, even if they are accumulated to high concentrations (Brown 1976).

In addition to TMAO's role in protecting cells from osmotic pressure, it is also recognised as a potent protein stabiliser to counteract the effect of destabilizers (*i.e.* temperature, urea, pressure), and can enhance protein folding (Singh et al

2005, Street et al 2006, Yancey 2005). In fact, TMAO is the most commonly used protein stabilizing agent in biotechnology (Mello and Barrick 2003).

Although the molecular mechanism of how TMAO stabilises protein is unclear, experimental and simulation studies suggest that TMAO is directly excluded preferentially from protein surface (Canchi et al 2012, Rosgen and Jackson-Atogi 2012), meanwhile, it forms H-bond with water, thus weakening water's H-bonding to protein (Ma et al 2014, Rosgen and Jackson-Atogi 2012) (**Figure 1-4 A**). On the contrary, urea tends to bind to protein backbone and some amino acid side groups to unfold and destabilise protein (**Figure 1-4 C**) (Canchi and García 2013, Hua et al 2008). TMAO is able to counteract the destabilising effect of urea, the major organic osmolyte in elasmobranch fishes and a highly concentrated waste product in mammalian kidneys and urine. For example, intracellular urea concentrations were estimated to reach up to 400 mmol l^{-1} in shark and mammal renal tissue (Yancey 2005). Urea destabilizes many macromolecular structures and inhibits their functions at such physiological concentrations. Therefore, to offset the perturbing effect of urea, shark and mammal renal tissue are known to accumulate as high as ~200 mmol l^{-1} TMAO, giving an urea:TMAO ratio of 2:1 (Yancey 2005).

Such counteracting phenomenon of TMAO has inspired many studies to gain insight into the underlying mechanism. Although the molecular mechanism of TMAO-urea counteraction remains unclear, it is generally believed that TMAO interacts with urea with the mediation of water to prevent urea binding to protein (**Figure 1-4 B**).

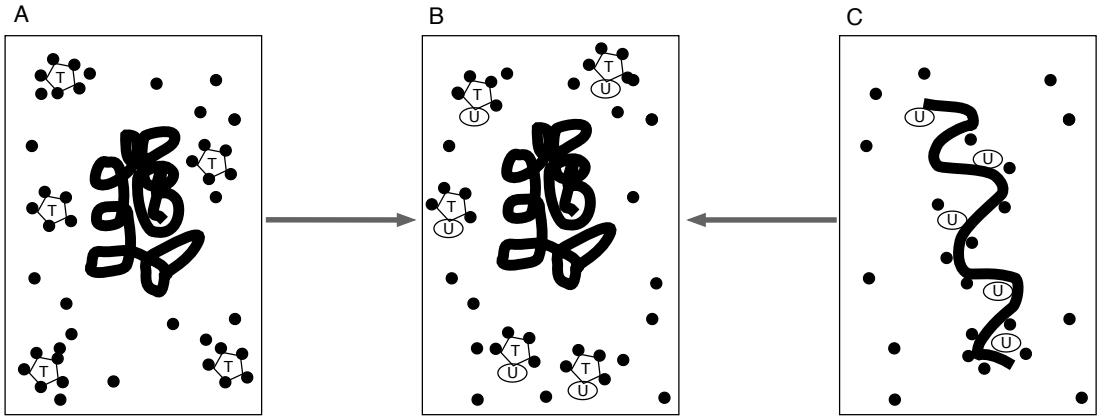


Figure 1-4 TMAO counteracts the protein denaturing effects of urea. Small spheres represent water molecules. TMAO stabilises protein by exclusion from protein surface and weakened water's hydrogen-bonding to protein (A). TMAO interacts with urea to counteract its denaturing effects (B). Urea binds to protein backbone and side chain, thus cause protein unfolding (C). Modified from (Yancey 2005).

TMAO is also known to counteract pressure effects (Martin et al 2002). It has been reported that TMAO concentrations in tissues of marine fish and other vertebrates increase with depth (**Figure 1-5**) (Kelly and Yancey 1999, Samerotte et al 2007), with a decrease of urea to maintain osmotic homeostasis. The muscle urea:TMAO ratio in shallow water environment (from sea shore to the beginning of the reef wall) is near a 2:1 ratio, while a species from 2850 m depth yields a "reversed ratio" of nearly 1:2 (Laxson et al 2011). In fact, levels of TMAO in the Kermadec snail fish captured at 7000 m were the highest recorded of ~400 mmol kg^{-1} (wet mass), supporting the notion that TMAO may biochemically restrain marine fish from inhabiting deepest ocean depth (**Figure 1-5**) (Yancey et al 2014). Although it is not clear why TMAO increases with depth, several hypothesises have been proposed, including reducing osmotic costs, aiding buoyancy and counteracting pressure, and higher lipid production in deep-sea animals as TMAO may accumulate passively during lipid production (Samerotte et al 2007, Seibel and Walsh 2002, Yancey et al 2001, Yancey and Siebenaller

2015, Bockus and Seibel 2016). As a result of ubiquitous distribution of TMAO in marine fish and other vertebrates, TMAO has been frequently detected in marine waters ((Gibb and Hatton 2004), also see section 1.3), which in turn represents a key source of nutrients for marine microbes as a supplementary energy source as well as a key nitrogen source (Lidbury et al 2014, Lidbury et al 2015, Sun et al 2011).

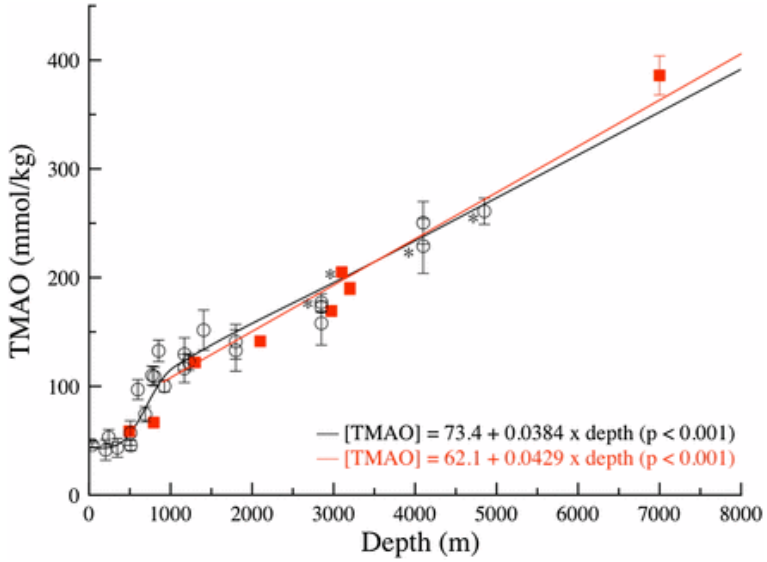


Figure 1-5 Muscle osmolyte contents as a function of depth. Circles, with standard deviation bars, are published data (Gillett et al 1997, Kelly and Yancey 1999, Samerotte et al 2007, Yancey et al 2004), with a linear fit (black line) from 900–4,850 m. Solid red squares without standard deviation bars are data (n = 1 each) for a snailfish *C. melanurus* (793 m), four eelpout, and two grenadier species from Yancey et al (2014). The solid red square with standard deviation bars is the hadal snailfish *N. kermadecensis* from 7,000 m (n = 5), with a new linear fit (the red line) for all new and old data for 900–7,000 m. **C. armatus* (abyssal grenadier) at four depths (note that the specimen at 4,850 m was from the northeast Atlantic, whereas the others were from the northeastern Pacific). Adapted from Yancey et al (2014).

1.2.2 TMAO is associated with cardiovascular disease in humans

In addition to its role in the marine ecosystem, TMAO also plays a contributory role in the development and progression of coronary atherosclerotic plaque burden and cardiovascular risk through its interaction with macrophages and

lipid metabolism (**Figure 1-6**) (Arumugam et al 2011, Tang et al 2014, Tang et al 2015, Wang et al 2011, Zhu et al 2016). TMAO is derived from microbial metabolism of dietary quaternary amines, *e.g.* choline, L-carnitine, glycine betaine (GBT) and phosphatidylcholine, to TMA, which is subsequently oxidized to TMAO by the host hepatic flavin monooxygenases. Two TMA-forming enzymes have been identified recently in gut microorganisms, namely CutC/CutD, CntA/CntB, encoding a choline-TMA lyase and a carnitine monooxygenase respectively (Craciun and Balskus 2012, Zhu et al 2014).

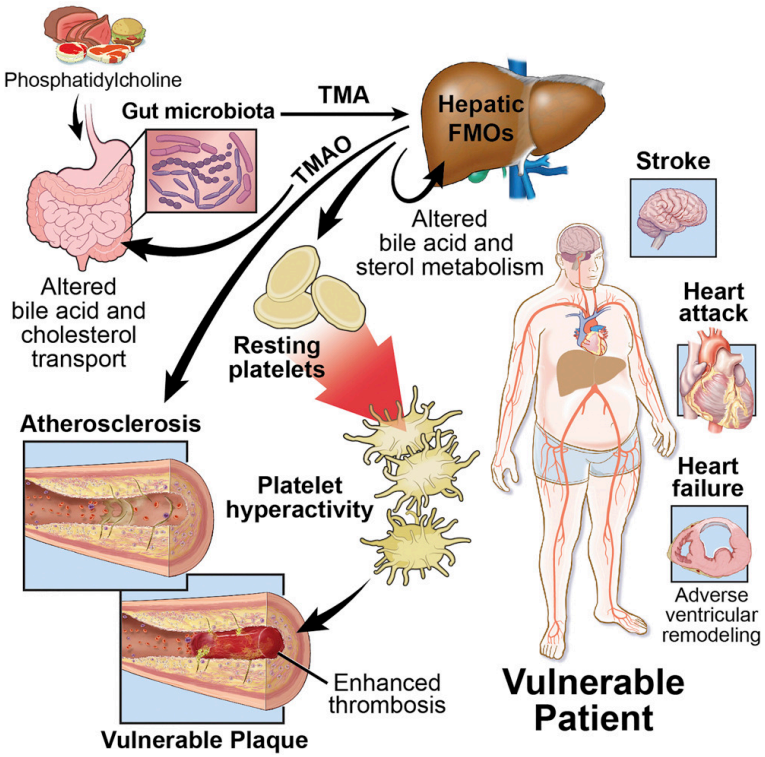


Figure 1-6 A schematic diagram illustrating metaorganismal pathways linking dietary sources of choline (abundant in a western diet), gut microbiota and host hepatic FMOs, resulting in TMAO production, and subsequent development of hyperresponsive platelet phenotype and enhanced thrombotic event risk. The pro-atherosclerotic effects of TMAO and the potential involvement of TMAO in the development of vulnerable plaque are also shown. EC, endothelial cell; FMOs, flavin monooxygenases; MΦ, macrophage; TMA, trimethylamine; TMAO trimethylamine *N*-oxide. Adapted from Zhu et al (2016).

1.3 TMAO in the environment

1.3.1 TMAO in marine surface water

Although the marine ecosystem has been recognised as one of the three major sources of MAs, the study of MAs concentration of marine surface water has been rare due to technical difficulty in measuring ambient MAs at nanomolar concentrations (Gibb and Hatton 2004). Gibb & Hatton (Gibb and Hatton 2004) used a coupled flow injection-ion chromatographic technique to determine TMAO in natural seawater and found that TMAO ranged from below the analytical detection limit (1.65 nmol•l^{-3}) to 76.9 nmol•l^{-3} in the coastal waters off the Antarctic Peninsula.

1.3.2 TMAO in soil

Studies of TMAO in soils are limited. TMAO in soil solution (soil water extraction) has been detected from a sub-alpine grassland (Warren 2013a, Warren 2013b), but was not quantified. Soil water contained a large pool of several quaternary amines (*e.g.* carnitine, acetyl carnitine, betaine, choline, ergothioneine), which are potential precursors of TMAO. The combined pool size of these quaternary amines is approximately 25% of the size of the common amino acids pool (Warren 2013a, Warren 2013b). Given the large amount of and diverse types of quaternary amines in soil, it is proposed that quaternary amines have a dual role in central metabolism and osmoprotection in plants and associated microbes against several environmental stresses, such as heat and desiccation (Warren 2013a, Warren 2013b).

1.4 Microbiology of TMAO production from quaternary amines via TMA

TMAO is oxidised from TMA derived from degradation of quaternary amines, *e.g.* choline, glycine betaine, carnitine (**Figure 1-7**). Genes encoding the respective choline TMA-lyase (CutC), betaine reductase (GrdH), and carnitine oxygenase (CntA/B) have been identified in bacteria (Craciun and Balskus 2012, Meyer et al 1995, Zhu et al 2014).

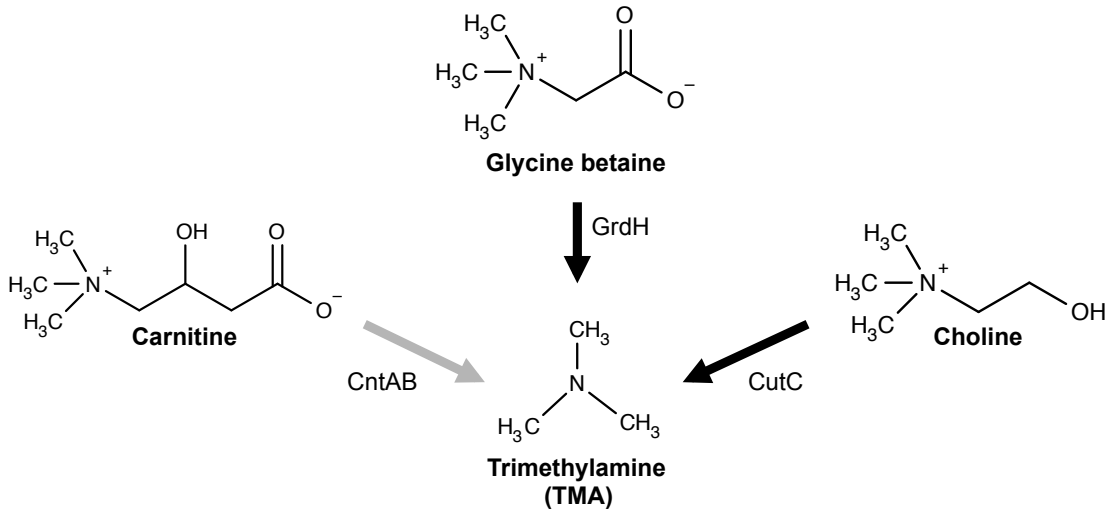


Figure 1-7 TMA production from choline, glycine betaine and carnitine. CutC, choline TMA-lyase, GrdH, betaine reductase, CntA, carnitine oxygenase. The grey arrows indicate aerobic pathway and the black arrows indicate anaerobic pathway.

1.5 Microbial TMAO degradation

TMAO can be converted to TMA by a periplasmic TMAO reductase (EC 1.6.6.9) (Barrett and Kwan 1985). In the absence of O_2, it is known that many bacteria can use TMAO as a terminal electron acceptor anaerobically (Barrett and Kwan 1985). TMAO can also be converted to DMA by a TMAO demethylase (EC 4.1.2.32) aerobically. TMAO demethylation to DMA via a TMAO demethylase is the focus in this thesis.

1.5.1 Bacterial TMAO transporter

When the project started, the only known microbial TMAO transporter in the literature was an ATP-dependent active transporter of the ABC superfamily found in *Aminobacter aminovorans* (Raymond and Plopper 2002), however, the gene encoding this transporter remained unknown. ABC transporters, which consist of three components, a transmembrane domain that is bound to an inner membrane-bound ATP-binding domain and a periplasmic substrate-binding protein (SBP), which binds a given ligand, are essential for bacteria in that they are responsible for the high affinity uptake of a wide range of nutrients, such as sugars, amino acids, metals, and vitamins, at the expense of ATP (Davidson and Chen 2004). This ABC-type transporter for TMAO has recently been identified in a variety of marine bacteria, and the genes encode the periplasmic SBP (TmoX), the ATP-binding domain protein (TmoW), and the transmembrane permease protein (TmoV) have been uncovered (**Figure 1-8**) (Lidbury et al 2014).

The TMAO transporter was initially annotated as a putative GBT/proline betaine-type ABC transporter, which are involved in the uptake of structurally-related compounds, such as choline, glycine betaine, proline betaine and

carnitine (Berntsson et al 2010, Thomas 2010). However, phylogenetic analysis of the SBP of GBT/proline betaine-type ABC transporter superfamily members showed that TmoX forms a distinctive group from the other previously characterised SBPs (**Figure 1-9**) (Lidbury et al 2014).

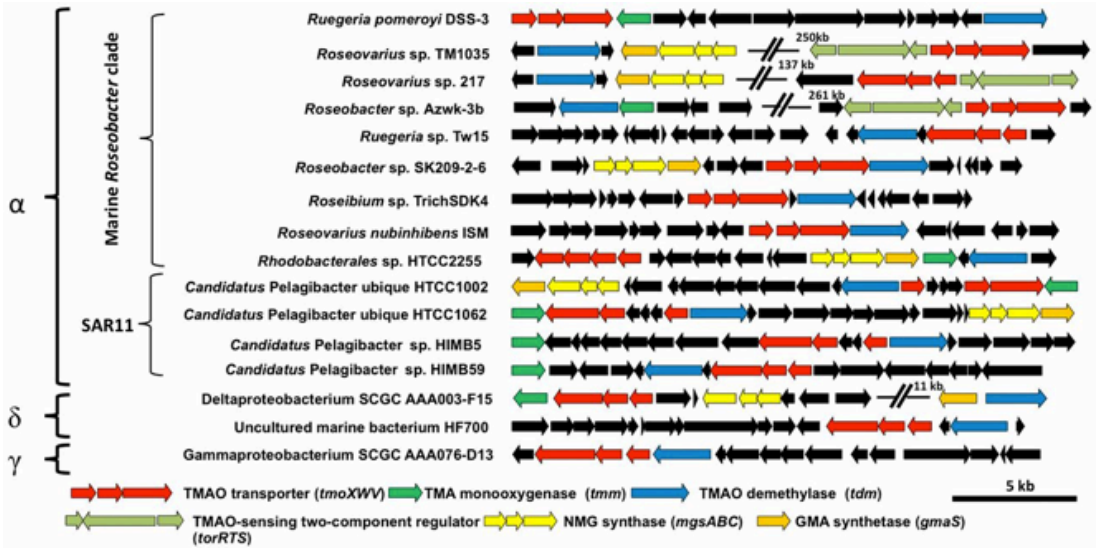


Figure 1-8 Genetic neighbourhoods of the genes (*tmoXWV*) that encode the TMAO transporter (red) among representative genome-sequenced marine bacteria. All genes coloured black have no confirmed functional relationship with TMAO metabolism. α, *Alphaproteobacteria*; δ, *Deltaproteobacteria*; γ, *Gammaproteobacteria*; GMA, γ-glutamylmethylamide; NMG, *N*-methylglutamate. Adapted from Lidbury et al (2014).

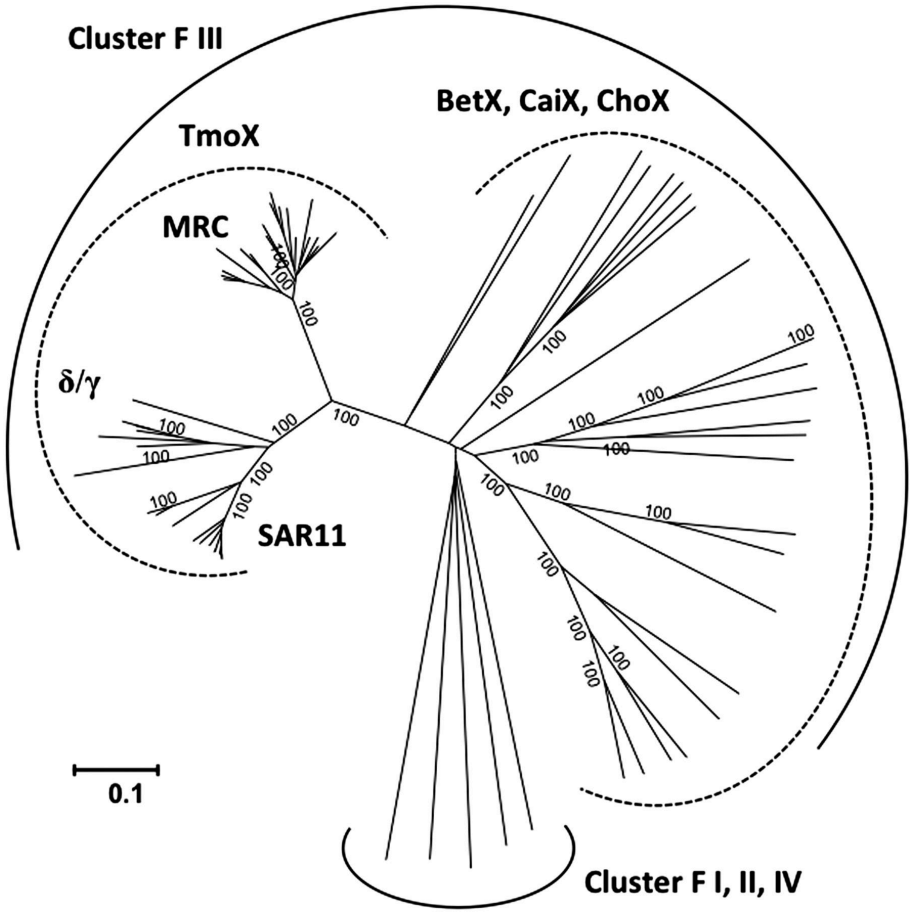


Figure 1-9 Phylogenetic analysis of the SBP, TmoX, of the TMAO-specific transporter in relation to other characterized SBPs. Currently known SBPs specific for osmolytes, such as choline, glycine betaine, and carnitine, fall into the cluster F of the ABC superfamily (Berntsson et al 2010). The evolutionary history was inferred using the neighbour-joining method (Saitou and Nei 1987). Bootstrap values (500 replicates) greater than 99% are shown. The scale bar represents the number of amino acid differences per site. The analysis involved 69 SBP sequences. There were a total of 296 amino acids positions in the alignment. Evolutionary analyses were conducted in MEGA5.1 (Tamura et al 2011) δ, *Deltaproteobacteria*; γ, *Gammaproteobacteria*; BetX, glycine betaine/proline betaine SBP; CaiX, carnitine SBP; ChoX, choline SBP. Adapted from Lidbury et al (2014).

1.5.2 TMAO demethylase (Tdm)

Tdm was first discovered in *Bacillus* sp. PM6, and later in *Aminobacter aminovorans* (previously known as *Pseudomonas aminovorans*). It catalyses non-oxidative and non-hydrolytic cleavage of TMAO to DMA and HCHO (Large 1971, Myers and Zatman 1971). Tdm has been partially purified from *Bacillus* sp. PM6 and *Aminobacter aminovorans*, characterisation of which revealed contrasting characteristics, such as molecular weight and cofactor (**Table 1-2**) (Anthony 1982, Large 1971, Myers and Zatman 1971). Tdm from *Bacillus* sp. PM6 displayed a molecular weight of 37-50 kDa whilst that of *Aminobacter aminovorans* showed 240-280 kDa. Ferrous iron (Fe^{2+}) was found to stimulate Tdm activity of *Bacillus* sp. PM6, whilst it had no impact on Tdm of *Aminobacter aminovorans*.

Table 1-2 Comparison of previously characterised Tdm

	***Bacillus* sp. PM6**	***Aminobacter aminovorans***
K_m (mM)	2.85	2
Optimum pH	7.5	6
Optimum temperature (℃)	-	25
Molecular Weight (kDa)	37-50	240-280
Cofactor	Fe^{2+}	No impact of Fe^{2+} addition

-: not determined

Tdm activity has also been found in a range of eukaryotic organisms, such as fish and squid (Fu et al 2006, Kimura et al 2000, Parkin and Hultin 1986). Eukaryotic Tdm encoded by aspolin1 and aspolin2 genes, has been identified as an extremely aspartic acid-rich proteins with a molecular weight of 25.8 kDa and 39 kDa, respectively (Takeuchi et al 2003). Searching bacterial genomes did not reveal any homologue of asplin1 nor aspolin2, suggesting that bacteria employ a different Tdm.

As the project proceeded, Ian Lidbury, another PhD student in the group, has identified Tdm in marine bacteria, the pelagic *α*-proteobacterium SAR11 strain HIMB59 and the coastal MRC strain *Ruegeria pomeroyi* DSS-3 (**Figure 1-10**) (Lidbury et al 2014). Similar to the distribution of the ABC-type TMAO transporter (TmoX), Tdm is widely distributed in MRC and SAR11 clade marine bacteria (Lidbury et al 2014). Approximately half of the genome-sequenced MRC isolates contain Tdm. Blast Tdm against the GOS dataset reference revealed that about 21% of bacterial cells in the oceanic surface waters contain Tdm. The abundance of Tdm (21%) in marine bacterioplankton is comparable to that of Tmm (20%) and GmaS (22%), two enzymes involved in the indirect TMA metabolic pathway (Chen et al 2011, Lidbury et al 2014). Similar to the distribution of Tmm in marine bacteria, Tdm is also predominantly found in the SAR11 clade (92%) and the MRC clade (5%) (Lidbury et al 2014).

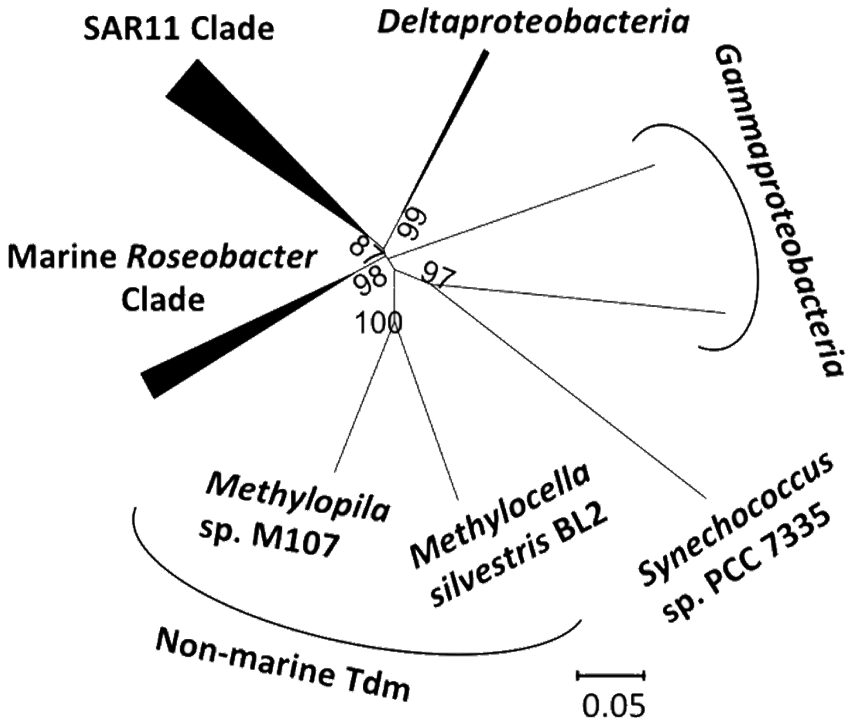


Figure 1-10 Neighbour-joining phylogenetic analysis of Tdm retrieved from the genomes of sequenced marine bacteria. Bootstrap values (500 replicates) greater

than 60% are shown. The scale bar denotes the number of amino acid differences per site. The analysis involved 49 Tdm sequences. There were a total of 468 amino acid residues in the alignment. Evolutionary analyses were conducted in MEGA5.1 (Tamura et al 2011). Adatped from Lidbury et al (2014).

1.6 Metal dependency of Tdm

It has been suggested that metal ions may play a role in Tdm catalysis (Myers and Zatman 1971). For example, the partially purified Tdm of *Bacillus* sp. PM6 was strongly activated by ferrous iron (Myers and Zatman 1971). In agreement with the putative role of metals in catalysis, purified Tdm of *Methylocella silvestris* does not contain either a flavin adenine dinucleotide (FAD) or a nicotinamide adenine dinucleotide (NAD) (detailed information in **Chapter 3**). Further characterization revealed that Tdm is a Zn^{2+}, Fe^{2+}-dependent protein (see **Chapter 4**).

The dependency of Zn^{2+} and Fe^{2+} raises the question whether Zn^{2+} or Fe^{2+} in Tdm could be replaced by other ions, for instance when Zn^{2+} or Fe^{2+} of the ocean is at scarce, *e.g.* in high nutrient–low chlorophyll (HNLC) regions that take over ~20% of open sea waters including the Southern Ocean, the equatorial Pacific, and parts of the Antarctic Ocean (Carol and Timothy 1997, Moore et al 2013). HNLC regions are replete with light and macronutrients (nitrate, phosphate and silicate), yet the standing stock of phytoplankton biomass remains low (**Figure 1-11**) (Martin et al 1994). Iron deficiency and grazing by microzooplankton are the two main explanations that have been put forward to explain the low primary production in HNLC regions (Landry et al 1997). HNLC oceans typically have subnanomolar concentrations of dissolved iron (Martin et al 1994, Mawji et al 2015). Such low concentrations of dissolved iron have been identified as limiting nutrient for both primary productivity and bacterial growth (**Figure 1-11**) (Behrenfeld and Kolber 1999, Church et al 2000, Moore et al 2013, Tagliabue et

al 2014). Thus iron fertilisation triggered massive primary production in HNLC region (Coale et al 1996).

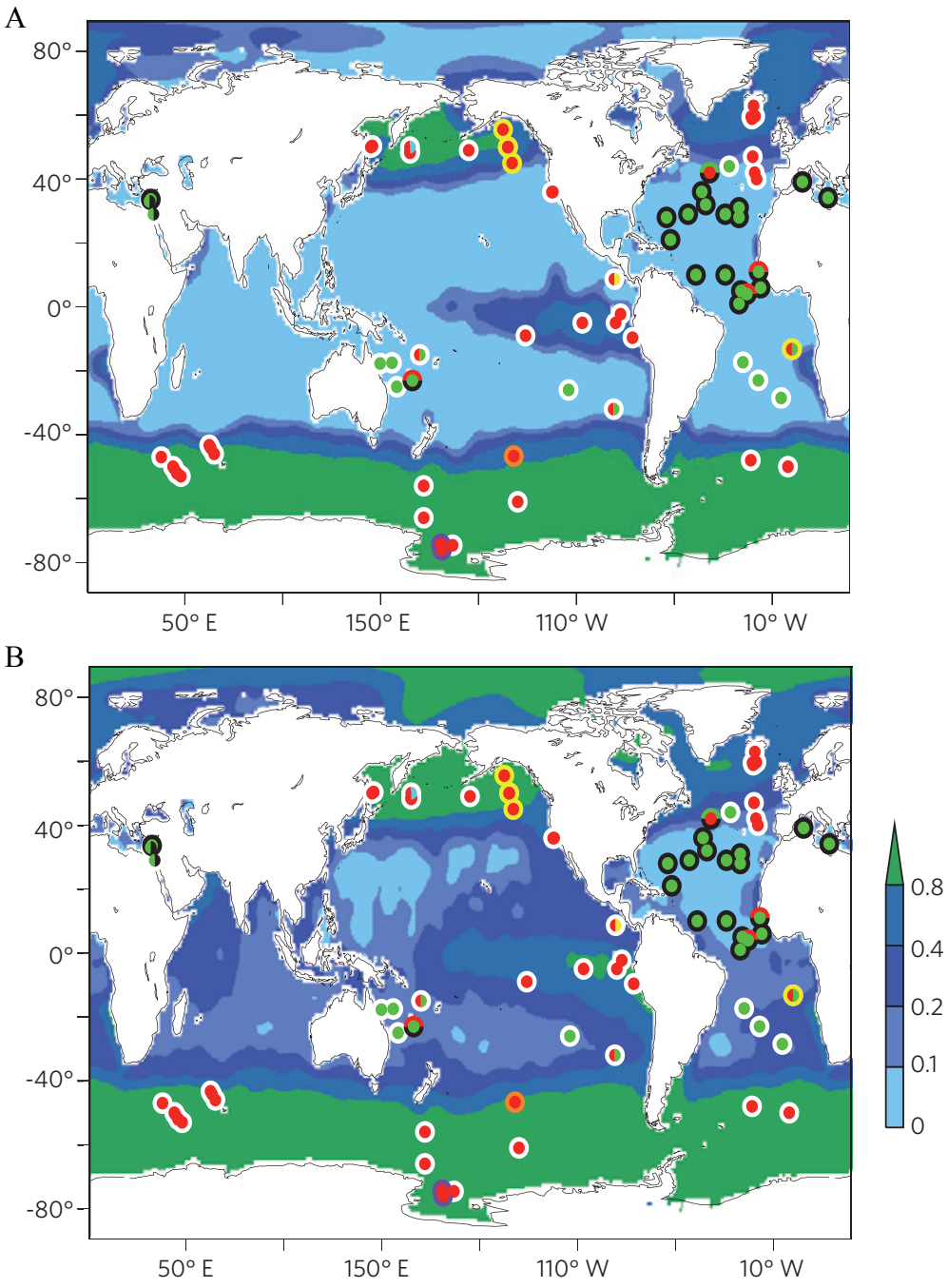


Figure 1-11 Patterns of nutrient limitation. Backgrounds indicate annual average concentrations of nitrate (A) and phosphate (B) in μmol kg^{-1}. Symbols indicate the primary (central circles) and secondary (outer circles) limiting nutrients as inferred from chlorophyll and/or primary productivity increases following artificial amendment of: N (green), P (black), Fe (red), Si (orange), Co (yellow), Zn (cyan) and vitamin B12 (purple). Divided circles indicate potentially co-limiting elements. White outer circles indicate that no secondary limiting nutrient was identified, which in many cases will be because of the lack of a test. Adapted from Moore et al (2013).

Zn, after Fe, is the second most abundant element ion in most organisms, and the only metal ion known to be present in enzymes from six classes established by the International Union of Biochemistry (Sousa et al 2007). Zn plays a key role in the productivity of the ocean in that it is the essential cofactor of the enzymes involved in primary production, such as carbonic anhydrase for carbon fixation and alkaline phosphatase for organic phosphorous acquisition (Auld 2001, Morel and Price 2003). The vertical profile of Zn in the world's oceans resembles those nutrient-like depth profiles of elements that are taken up from surface waters and regenerated at depth as dead organisms sink to the deep ocean and decompose. Dissolved Zn concentrations in the surface waters of the Pacific Ocean, the Atlantic Ocean and the Southern Ocean are in the subnanomolar range (Lohan et al 2002, Zhao et al 2014). ~98% of Zn is bound strongly to uncharacterized organic ligands such that the concentration of bioavailable Zn is in the picomolar range (Andersen et al 2011, Lohan et al 2002). At such concentration, phytoplankton growth can be limited, which was observed in the laboratory, probably due to reduced activities of Zn-containing metalloenzymes (Anderson et al 1978, Brand et al 1983, Morel et al 1994, Shaked et al 2006, Tortell et al 2000). Therefore, Morel (et al 1994) proposed a "Zn-hypothesis", suggesting that Zn as well as Fe may be limiting phytoplankton growth. However, the "Zn-hypothesis" was proposed based on laboratory experiments, and unlike Fe^{2+}, few experiments have been conducted with natural assemblages. In most cases, such field experiments showed minimal stimulation effects of Zn additions on the bulk phytoplankton community (Coale et al 2003, Crawford et al 2003, Ellwood and Hunter 2000, Jakuba et al 2012).

Two possible explanations were proposed for the lack of observed Zn-limitation during field studies. Firstly, Zn limitation in the oceans may be obscured by co-limitation by the more widespread occurrence of Fe limitation. Secondly, Zn^{2+} at active sites of zinc enzymes can be substituted by other metal ions, *e.g.* cobalt (Co) and cadmium (Cd). Such metal replacement has been observed in carbonic anhydrase of *Thalassiosira weissflogii* (Lane and Morel 2000, Price and Morel 1990) when Zn^{2+} was limited.

1.7 DUF domains

Sequence analysis of Tdm showed the presence of a domain of unknown function (DUF1989) at its N-terminus (more information in **Chapter 3**). Proteins having DUF domains currently represent more than a quarter of sequence entries in public databases such as Pfam (**Figure 1-12**) (Punta et al 2012). It has been estimated that the number of DUFs will soon outnumber the families with known function being added to Pfam (Bateman et al 2010). Functional annotation of DUFs remains a great challenge for the scientific community since they not only represent a major knowledge gap between protein structure and functional relationship but also prevent the complete understanding of cellular functions from completed genomes (Galperin and Koonin 2010). By investigating structure and function of Tdm, I aimed to shed some light on the function of DUF1989.

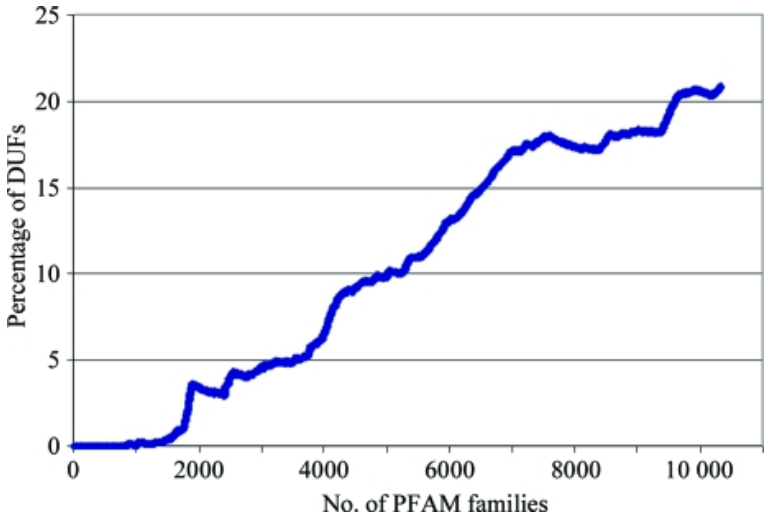


Figure 1-12 A graph showing the growth of DUFs as a percentage of all families added to Pfam at the time of release 23.0. Adapted from Bateman (2010).

1.8 *Methylocella silvestris* BL2

Methylocella species, including *Methylocella palustris* (Dedysh et al 2000), *Methylocella silvestris* (Dunfield et al 2003) and *Methylocella tundra* (Dedysh et al 2004) are authenticated facultative methanotrophs that are able to use methane, methanol and MAs, as well as multi-carbon compounds, such as acetate, pyruvate, succinate, malate and ethanol, as sole carbon and energy sources (Dedysh et al 2005). *Methylocella* species form a distinct taxonomic cluster of acidophilic, methanotrophic bacteria. Although they belonging to the *Alphaproteobacteria* (type II methanotrophs), they are more closely related to the nonmethanotrophic heterotroph *Beijerinckia indica* (Dedysh et al 2005).

Of all the isolated *Methylocella* species, *Methylocella silvestris* BL2 grows most robustly, hence it was used to investigate indirect MAs metabolic pathway. *Methylocella silvestris* BL2 was isolated from a forest soil in Germany and can utilise MAs, including TMA, DMA and MMA as carbon, nitrogen and energy sources (Chen et al 2010a, Chen et al 2010c, Chen et al 2011, Dunfield et al 2003). *M. silvestris* BL2 employs an indirect pathway involving TMAO as the

key intermediate for the degradation of TMA to ammonium and formaldehyde as illustrated in **Figure 1-2** (Chen et al 2011).

1.9 *N*-demethylation catalysed by iron-containing enzymes

It has been suggested that metal ions may play a role in Tdm catalysis (Myers and Zatman 1971). For example, the partially purified Tdm of *Bacillus* sp. PM6 was strongly activated by ferrous iron (Myers and Zatman 1971). Although a crystal structure for Tdm has yet to be solved, structures of three DUF1989-domain containing proteins (3ORU, 3SIY, 3DI4) available in the PDB database all contain Zn^{2+}.

As the project went on, it has been revealed that Tdm is a Zn^{2+}, Fe^{2+}-dependent enzyme and Fe^{2+} plays a catalytic role (more detailed information in Chapter 4). Iron-containing enzymes have been found to perform a wide variety of reactions, including *N*-demethylation, such as cytochrome P450 mediated *N*,*N*-dimethylaniline demethylation (Roberts and Jones 2010), DNA dealkylase (AlkB) (Mishina and He 2006), histone demethylase (JHDM1, JMJD6) (Chang et al 2007, Tsukada et al 2006) and Rieske-type demethylase (Daughtry et al 2012, Summers et al 2012). O_2 is required as oxygen donor for all the aforementioned reactions to form the high Fe(IV)-oxo complex, the common active oxidant for attacking the C-H bond of saturated carbon centres. It is hypothesised that a high-valent oxidant (*e.g.* Fe(IV)-oxo) is required for TMAO demethylation and a tertiary amine intermediate (*i.e.* TMA) is formed.

1.10 Project aims

The aims of the project were as follows:

1. To identify *tdm* in *Methylocella silvestris* BL2 (Chapter 3)
2. To establish the genetic, biochemical and biophysical characteristics of Tdm (Chapters 3, 4)
3. To elucidate the structure and functional mechanism of Tdm (Chapter 4)
4. To investigate the structure and function of DUF1989 domain of Tdm (Chapters 4, 5)
5. To investigate the metal dependency of Tdm homologues in Zn^{2+} and/or Fe^{2+}-depleted environment (Chapter 5)

Chapter 2 General methods

The methods presented in this chapter are referred to routinely elsewhere in this thesis. More specific methods are referred to within the relevant chapters to maintain coherence and clarity.

2.1 Materials

Analytical-grade chemicals were obtained from Sigma-Aldrich Corporation (St Louis, MO, USA), or Fisher Scientific UK (Loughborough, UK). General purpose buffers and chemicals were prepared according to Sambrook and Russell (2001). Custom oligonucleotide primers were obtained from Sigma-Aldrich (St Louis, MO, USA).

2.2 Growth media and conditions

2.2.1 DNMS

M. silvestris wild type and mutant strains were grown at 25 °C under natural light conditions in 125-ml serum vials containing 20 ml diluted mineral salt medium (DNMS) (Crombie 2011, Dunfield et al 2003) (**Table 2-1**) with an inoculum size of 10% (v/v). Methylated amines (final concentration 1.5 mM), *i.e.* TMA, TMAO, DMA and MMA, were used as the nitrogen source. Either methanol (10 mM) or succinate (5 mM) was used as the carbon source.

Table 2-1 Composition of Diluted Nitrate Mineral Salts medium

Diluted Nitrate Mineral Salts (DNMS)	
Components	**Amount (per litre)**
$MgSO_4•7H_2O$	0.1 g
KNO_3	0.1 g
$CaCl_2•H_2O$	0.02 g
3.8% (w/v) solution Fe-EDTA	0.01 ml
0.1% (w/v) $NaMo•4H_2O$	0.05 ml
Trace element solution (recipe below)	1.0 ml

Trace element solution (1X)	
Ingredient	**Amount (per litre)**
$FeSO_4•7H_2O$	500 mg
$ZnSO_4•7H_2O$	400 mg
$MnCl_2•7H_2O$	20 mg
$CoCl_2•6H_2O$	50 mg
$NiCl_2•6H_2O$	10 mg
H_3BO_3 (boric acid)	15 mg
EDTA	250 mg

Phosphate stock solution (1X) pH 5.8	
Ingredient	**Amount (per litre)**
KH_2PO_4	26 g
$Na_2HPO_4•7(H_2O)$	62 g

DNMS and phosphate stock solution was autoclaved separately. 10 ml $litre^{-1}$ of autoclaved phosphate stock was added to DNMS before use.

2.2.2 Luria Broth and Super Optimal Broth with Catabolite repression (SOC) mediums

Escherichia coli strains were routinely grown in Luria broth (LB) containing 10 g l^{-1} bacto tryptone, 5 g l^{-1} yeast extract and 10 g l^{-1} NaCl. Solid LB media contained 1.5 % (w/v) Bacto agar in addition. Transformed *E. coli* strains were grown in SOC medium containing 20 g l^{-1} Bacto tryptone, 5 g l^{-1} yeast extract, 0.5 g l^{-1} NaCl, 0.186 g l^{-1} KCl, 0.952 g l^{-1} $MgCl_2$.7H2O, 3.603 g l^{-1} glucose.

2.2.3 M9 minimal media

The defined minimal medium contained NH_4Cl (1 g l^{-1}), NaCl (0.5 g l^{-1}), KH_2PO_4 (3 g l^{-1}), $Na_2HPO_4 \cdot 7H_2O$ (12.8 g l^{-1}), $MgSO_4 \cdot 7H_2O$ (0.5 g l^{-1}), $CaCl_2 \cdot 2H_2O$ (0.15 g l^{-1}), $Na_2MoO_4 \cdot 2H_2O$ (0.5 mg l^{-1}), and a mix of the following vitamins or supplements, including biotin (0.4 mg l^{-1}), folic acid (0.4 mg l^{-1}), pyridoxine hydrochloride (2 mg l^{-1}), thiamine hydrochloride (1 mg l^{-1}), riboflavin (1 mg l^{-1}), nicotinic acid (1 mg l^{-1}), pantothenic acid (1 mg l^{-1}), vitamin B12 (20 mg l^{-1}), 4-aminobenzoic acid (1 mg l^{-1}) and lipoic acid (1 mg l^{-1}).

2.3 Nucleic acid manipulation techniques

2.3.1 Quantification of DNA

DNA purity and size were estimated by agarose gel electrophoresis and comparison with a known quantity of 1 kb DNA ladder (Fermentas). DNA concentration was determined using ND-1000 spectrophotometer (NanoDrop Technologies Inc., Wilmington, DE, USA).

2.3.2 Polymerase chain reaction (PCR)

Polymerase chain reactions (PCR) were performed using a T3000 (Biometra, Goettingen, Germany) or a Tetrad thermal cycler (Bio-Rad, USA). Kapa Taq DNA polymerase (Kapa) was used for rutine PCR check purpose. Q5 high fidelity DNA polymerase (NEB)

and Kapa high fidelity DNA polymerase (Kapa) were used for high fidelity applications. 50 µl reaction volumes were used containing 1 × buffer, $MgCl_2$ (1.5 mM), dNTPs (0.2 mM of each), forward and reverse primer (0.4 µM) and Taq DNA polymerase (2.5 units). For direct amplification from colonies or cultures, DMSO (5%, v/v) and BSA (0.07%, w/v) were included in addition. Cycling conditions were typically: initial denaturation at 95 °C, 3 min; 25 - 35 cycles of denaturation at 95 °C, 30 s; annealing (temperature dependent on primers), 30 s; elongation at 72 °C, 1 min/kb and final elongation at 72 °C, 7 min. For PCR from colonies, the initial denaturation was increased to 10 min. Denaturing temperature and elongation time varied according to manufacturer's instruction. Reactions without template (no-template controls) were included in all cases.

2.3.3 Site-directed mutagenesis

Site-directed mutations were introduced into *tdm* by PCR, using either overlap extension-PCR (SOE-PCR) (Choi and Schweizer 2005) (for mutants D198A, C263A, C279A, C343A, D39A) or the Q5 Site-Directed Mutagenesis kit for all other mutants according to manufacturer's instruction (New England Biolabs, Hertfordshire, UK). The mutagenic oligonucleotides used are shown in Appendix 1. The resulting mutant plasmids were verified by Sanger DNA sequencing. SOE-PCR was performed using Kapa high fidelity DNA Polymerase (Kapa). 10 ng of A and B regions were mixed and used as the template. SOE-PCR was run without primers for 5 cycles and then for 30 cycles after the upstream and downstream primers were added (Tdm_F, Tdm_R) to generate a ~2.3 kb PCR product.

2.3.4 Cloning

Plasmids and strains used for cloning and overexpression in *E. coli* are listed in **Table 2-2**.

Genes of interests were amplified by PCR and sub-cloned into the pGEM-T vector (Promega, Madison, USA) and transformed into *E. coli* JM109 competent cells (Promega, Madison, USA) according to manufacturer's instruction. Primers used to amplify genes and for sequencing are listed in Appendix 1. Where high fidelity DNA polymerase was used for PCR amplification, addition of a terminal 3' adenosine was accomplished by incubating purified PCR product with Kapa DNA polymerase (5 units), dATP (0.2 mM) and 1 × Taq buffer in a 20 µl final volume at 72 °C for 20 min.

Plasmids were extracted using the GeneJET kit (Fermentas) according to the manufacturer's instructions. Double restriction digest of DNA was carried out with FastDigest enzymes from Fermentas according to manufacturer's instruction. Desired fragments were separated by gel electrophoresis and recovered using the NucleoSpin Gel and PCR clean up kit (MACHEREY-NAGEL, Düren, Germany) according to manufacturer's instruction. Purified DNA fragments were ligated into the expression vector pET28a (Merck Biosciences, Darmstadt, Germany). Ligations were routinely carried out in 10 µl reactions with a vector: insert ratio of ~1:3. T4 DNA ligase from Promega or Fermentas was used according to the manufacturer's instructions.

2.3.5 Transformation

E. coli host strains JM109 (Promega, Madison, USA) and BLR(DE3) pLysS (Merck Biosciences, Darmstadt, Germany) were used for plasmid production and protein expression, respectively. Routinely, 20 µl competent cells were thawed on ice and DNA (~10 ng of plasmid DNA or ligation mix) added and gently mixed. Cells were subjected to heat shock at 42 °C for 45 s, and cooled on ice for 2 min. SOC medium (0.25 ml) was added and cells were then allowed to recover at 37 °C with shaking for one hour. Aliquots were spread on selective LB plates containing appropriate antibiotics where necessary and incubated at 37 °C for 18 – 24 h.

Table 2-2 Strains and plasmids used in this thesis

Strains/plasmids	**Relevant features**	**References**
Escherichia coli		
BLR(DE3)pLysS	Host for heterologous protein overexpression	Novagen
JM109	General cloning	Promega
Methylocella silvestris		
M.silvestris Δtmm::Kan	Wild-type with disrupted *tmm*	(Chen et al 2011)
M.silvestris Δtdm::Kan	Wild-type with disrupted *tdm*	This study
M.silvestris ΔtmoP::Kan	Wild-type with disrupted *tmoP*	This study
M.silvestris Δ3605::Kan	Wild-type with disrupted *Msil_3605*	This study
M.silvestris Δ3608::Kan	Wild-type with disrupted *Msil_3608*	This study
M.silvestris Δ3609::Kan	Wild-type with disrupted *Msil_3609*	This study

Strains/plasmids	Relevant features	References
Plasmids		
pGEM-T	Cloning vector	Promega
pET28a	Expression vector in *E. coli*	Novagen
pCOLADuet-1	Expression vector in *E. coli*	Novagen
pET28a-Tdm	Tdm expression, wild type	This study
pET28a-Tdm C263A	Tdm mutant C263A expression	This study
pET28a-Tdm C279A	Tdm mutant C279A expression	This study
pET28a-Tdm C343A	Tdm mutant C343A expression	This study
pET28a-Tdm H256A	Tdm mutant C256A expression	This study
pET28a-Tdm H276A	Tdm mutant H276A expression	This study
pET28a-Tdm D198A	Tdm mutant D198A expression	This study
pET28a-Tdm D198E	Tdm mutant D198E expression	This study
pET28a-Tdm D198N	Tdm mutant D198N expression	This study
pET28a-Tdm F259A	Tdm mutant F259A expression	This study
pET28a-Tdm Y305A	Tdm mutant Y305A expression	This study
pET28a-Tdm W321A	Tdm mutant W321A expression	This study

2.3.6 Induction and overexpression of recombinant protein in *E. coli*

E. coli cells containing the gene of interest in pET28a were grown in 250 ml LB with 25 μg ml^{-1} kanamycin at 37 °C with agitation at 250 $r \cdot min^{-1}$. When the cell density (OD_{600}) reached 0.4 - 0.6, isopropyl β-D-1-thiogalactopyranoside (IPTG) was added to a final concentration of 0.2 mM to induce protein expression and the cultures were then shifted to 25 °C for 18 h before cell harvesting by centrifugation (6,000×g, 10 min).

2.3.7 Protein purification

Cells were stored at -20 °C prior to cell lysis by passing three times through a chilled French press (American Instrument Co.) at 110 megapascals. The lysates were centrifuged at 100,000×g for 40 min, and the supernatants were saved as cell extracts for the purification of Tdm. Overexpressed Tdm was purified using a His-tag protein purification kit according to the manufacturer's instructions (Novagen, Germany) and eluted using an elution buffer, containing 1M imidazole, 0.5 M NaCl and 20 mM Tris-HCl (pH 7.8). Purified protein was desalted against 20 mM Tris-HCl (pH 7.8), 100 mM NaCl using the HiTrap desalting column (GE Healthcare, UK) for further analysis. Depending on the aim of the analysis, purified Tdm and mutant proteins were concentrated by ultrafiltration (Amicon Corporation) and further purified by size-exclusion chromatography.

2.3.8 Size-exclusion chromatography (Gel filtration)

A Superdex 200 10/300 GL gel filtration column (GE Healthcare, UK) was equilibrated with 20 mM Tris-HCl pH 7.8, 100 mM NaCl at 0.7 ml min^{-1} using an AKTA FPLC system (GE Healthcare, UK). A calibration curve of the partition coefficient (K_{av}) versus apparent molecular mass was determined by measuring the elution volume V_e for protein standards, consisting of dextran (2,000 kDa), ferritine (440 kDa), sweet potato β-amylase (200 kDa), aldolase2 (158 kDa), conalbumin (75 kDa), ovalbumin (43 kDa) (GE healthcare, UK), $K_{av} = (Ve–V_0)/(V_c–V_0)$ was used to calculate K_{av} for each standard, which was plotted vs. log M_r, where Mr is the known apparent molecular mass of the standard, V_e is the elution volume, and V_0 is the void volume which is 8.23 ml for the Superdex 200 10/300 GL column used (**Figure 2-1**).

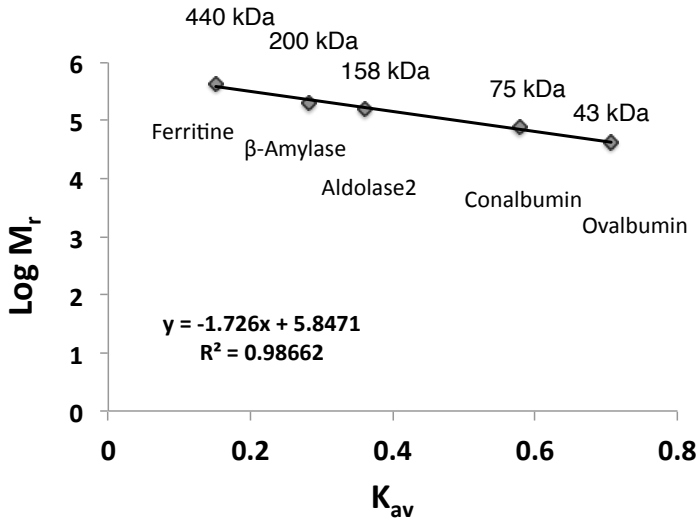


Figure 2-1 Calibration curve for protein molecular weight by gel filtration.

2.3.9 Protein quantification

Total protein concentration was determined at 595 nm by the Bradford assay (Bio-Rad, USA) using bovine serum albumin (BSA, 0.25, 0.5, 0.75, 1 mg ml^{-1}) (Bio-Rad, USA) as the standard according to the manufacturer's instruction.

2.3.10 Circular dichroism

Purified recombinant protein was buffer-exchanged to 20 mM sodium phosphate (pH 7.0) containing 50 mM NaCl using by ultrafiltration with Vivaspin centrifugal concentrator (cut-off of 10,000 Da) (GE healthcare, UK). Circular dichroism (CD) spectra were recorded in the range of 195–260 nm by a Jasco J-815 spectrometer (Jasco, UK) in a quartz cuvette of 1 mm path length at room temperature (~22 °C). Spectra were collected eight times per sample. Data were expressed as mean residue ellipticity (MRE) in degrees·cm^2·$dmol^{-1}$.

Spectra were deconvoluted using the online programme DICHROWEB (http://dichroweb.cryst.bbk.ac.uk/html/home.shtml) (Whitmore and Wallace 2004, Whitmore and Wallace 2008) and the CDSSTR algorithm was used to estimate the percentages of each secondary structure using the reference protein set 7 (Sreerama and Woody 2000).

2.3.11 6*His-tag cleavage

Removal of the 6*His-tag at the N-terminus of recombinant Tdm was carried out using thrombin (GE Healthcare, UK) according to the manufacturer's instruction. One mg purified recombinant Tdm was incubated with 10 units thrombin at 4 °C overnight (16 hrs). The 6*His-tag was efficiently removed as determined by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and its inability to bind to the nickel affinity column.

2.3.12 Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) analysis

Polypeptides were separated by SDS-PAGE using Mini-PROTEAN TGX precast polyacrylamide gels (12.5%, w/v) (Bio-rad, Hercules, CA, USA) conducted at 160-200 V during separation, using the running buffer containing glycine (72 g l^{-1}), Tris base (15 g l^{-1}) and SDS (5 g l^{-1}). On occasion, Bis-Tris Novex NuPAGE gels (Invitrogen, Carlsbad, CA, USA) were used with MOPS buffer following the manufacturer's instructions. Gels were stained with Fast Blue gel staining reagent (Expedeon, UK). Either the Bio-Rad dual colour protein ladder (Bio-Rad, USA) or the ColorBurst Marker (Sigma-Aldrich) was used as molecular mass markers for the estimation of protein molecular weight.

2.3.13 Native polyacrylamide gel electrophoresis

Native (non-denaturing) polyacrylamide gel electrophoresis was performed at a constant voltage of 150 V using an Invitrogen electrophoresis system on a NuPAGE Novex 3–8% Tris-Acetate (w/v) polyacrylamide gel. The gels were stained with the Fast Blue reagent (Expedeon, UK). The NativeMark™ Unstained Protein Standard from ThermoFisher Scientific was used as the standard.

2.3.14 Protein identification by mass-spectrometry

Bands of interest from SDS-PAGE were excised, digested with trypsin, and analysed to confirm their identity using the matrix-assisted laser desorption ionization–mass spectrometry (MALDI-MS) and tandem mass spectrometry at the Mass Spectrometry and Proteomics Facility Laboratory, School of Life Sciences, University of Warwick.

2.3.15 Crystallisation screening

Proteins were concentrated by ultrafiltration with Vivaspin centrifugal concentrator (cut-off of 10,000 Da) (GE healthcare) to ~10 mg mL^{-1} in 20 mM Tris-HCl pH 7.8, 100 mM NaCl for crystallisation screening unless otherwise stated.

Initial screenings were carried out in 96-well sitting drop crystallisation plates using a Mosquito robotic lipid handing system (TTP Labtech, UK). Each reservoir contains 70 µl of mother liquor. The protein drops constituted 200 nl of protein and

200 nl of crystallisation solution dispensed by the Mosquito robot. Screens used in this setup include JCSG+, PACT and Morpheus from Molecular Dimensions (Newmarket, UK). After setting up, all plates were sealed with a ThermalSeal RT sheet from AlphaLaboratories (AlphaLaboratories, UK) and incubated at either 4 °C or 18 °C with periodic examination for crystal formation. Plates were checked periodically using a SZ-PT Olympus microscope with an attached JVC colour video camera for imaging.

2.4 Analytical method

2.4.1 Ion-exchange liquid chromatography

Concentrations of methylated amines (carnitine, TMA, DMA, MMA *etc*) were determined by ion-exchange liquid chromatography. The ion chromatography system used consisted of a Metrohm 881 Compact IC Pro (Metrohm, UK) with a Metrosep C4/250 column. The eluent solution contained HNO_3 (1.5 mM), 2, 6-pyridinedicarboxylic acid (0.7 mM) and acetone (5%, v/v). All solutions were prepared using Milli-Q water (Millipore, USA).

2.4.2 Formaldehyde quantification

Enzyme assays for Tdm were carried out at room temperature (~22 °C) in a 96-well microplate (Bio-Rad, USA), containing 2.5 μg of purified Tdm in 50 μl of 10 mM 2-(N-morpholino)ethanesulfonic acid (MES) buffer unless otherwise stated. The

reactions were initiated by adding TMAO into the mixture (final concentration 10 mM) and incubated for 10 min. The time course experiment showed that enzyme activity slowed down after 30 min of incubation as assessed by formaldehyde release (**Figure 2-2**). Therefore, 10 min incubation time was applied as it was within the linear phase. Measurement of formaldehyde was performed by mixing 10 µl of the sample with 25 µl of 0.2% (w/v) Purpald reagent and 215 µl of Milli-Q water in a 96-well micro-plate (Bio-Rad, USA) (Quesenberry and Lee 1996). The Purpald solution was freshly prepared by dissolving in 1 M NaOH. Absorbance at 540 nm was determined after 20 min incubation at room temperature using a Bio-Rad iMark micro-plate reader. Calibration curves were prepared with 16% (v/v) formaldehyde of analytical grade purity (Thermo Scientific) from 20 µM to 180 µM (final concentration).

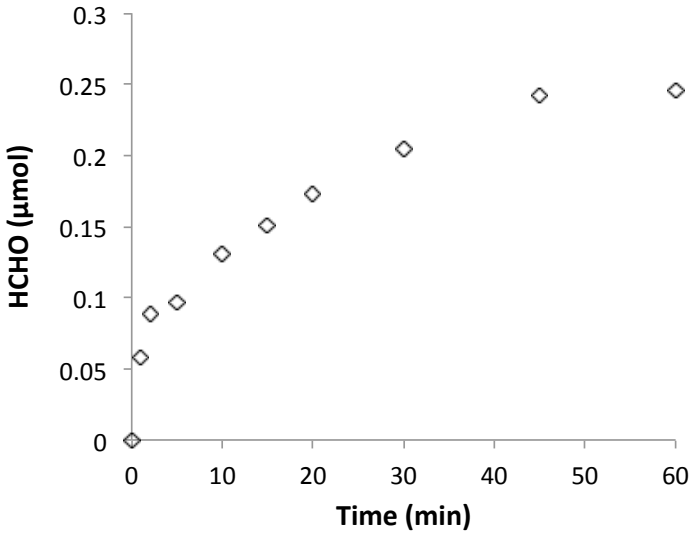


Figure 2-2 Time course of 10 µg recombinant 6*His-tag Tdm enzymatic reaction in 1 ml of 10 mM Tris-HCl (pH 7.8). 5 mM TMAO was used as the substrate.

2.5 Bioinformatics

Protein sequences were aligned, end-trimmed and analyzed using the MEGA5 package (Tamura et al 2011). All phylogenetic trees were constructed using the minimum evolution method (default settings) with 1,000 bootstrap replicates using the MEGA5 package. Analysis of conserved domains in proteins was carried out using the Pfam database (release 27.0) (Punta et al 2012).

Chapter 3 Identification and characterization of TMAO demethylase and TMAO permease in *Methylocella silvestris* BL2

3.1 Introduction

In recent years, the cycling of methylated amines (MAs) in the terrestrial environment has attracted great attention (Ge et al 2011). MAs, together with other aliphatic amines, have been identified as one of the important components of trace gases in the atmosphere, contributing to the growth of the so-called secondary organic aerosols and likely leading to the formation of cloud condensation nuclei (Cape et al 2011, Ge et al 2011). Since MAs are basic, they also play a role in neutralizing atmospheric acidity caused by organic and inorganic acids, including sulphuric acid, nitric acid and formic acid (Murphy et al 2007). MAs are produced both biologically and abiotically. Abiotic sources of MAs include biomass burning and emissions from vehicle exhaust (Ge et al 2011). The annual flux of MAs into the atmosphere is estimated to be in the order of 285 ± 78 Gg globally, a large proportion of which originates from animal husbandry and biomass burning although other anthropogenic activities such as agriculture also play a role (Ge et al 2011). The presence of such large quantities of MAs in the atmosphere can significantly affect human well-being. For example, MAs can be precursors for carcinogens such as *N*-nitrosodimethylamine, causing concerns for public health (Mitch et al 2003). Being an important component of organic nitrogen in the atmosphere, the cycling of MAs between land and oceans can also affect global biogeochemical cycles of

nitrogen through atmospheric deposition (Cape et al 2011). Therefore, understanding the sources and sinks of MAs in the environment will contribute to a better assessment of the MA cycle and subsequent impact on public health and ecosystem function.

Several processes contribute to biological MA production in the terrestrial environment, including degradation of herbicides and pesticides (Bhadbhade et al 2002, Dam et al 2005, Kamanavalli and Ninnekar 2000, Topp et al 1993), protein putrefaction (Kamiya and Ose 1984), anaerobic microbial respiration (Barrett and Kwan 1985), as well as degradation of quaternary amines (King 1987). Quaternary amines such as choline and carnitine are significant components of eukaryotic cells and are released to the environment due to normal cell turnover and programmed cell death. In agricultural and forest soils, MAs co-exist with quaternary amines, which represent a major pool of dissolved organic nitrogen, suggesting that quaternary amines are likely to be important MA precursors in these soils (Warren 2013a, Warren 2013b, Yu et al 2002).

It is known that many soil bacteria can sequester MAs from their environment as their carbon and nitrogen source (Anthony 1982). *Methylocella silvestris* BL2, a facultative one-carbon utilizing alphaproteobacterium isolated from a forest soil in Germany, can utilise MAs, including TMA, DMA and MMA, as carbon, nitrogen

and energy sources (Chen et al 2010a, Chen et al 2010c, Dunfield et al 2003). *M. silvestris* BL2 employs an indirect pathway involving TMAO as the key intermediate for the degradation of TMA to ammonium and formaldehyde (Chen et al 2011). The enzymes responsible for MMA degradation in this bacterium through γ-glutamylmethylamide (GMA) and *N*-methylglutamate (NMG) and the enzyme TMA monooxygenase responsible for the initial oxidation of TMA to TMAO have been identified previously (Chen et al 2010c, Chen et al 2011). It is hypothesized that TMAO can be further converted to MMA through a demethylation and an oxidation step, however, the genetics and biochemistry underpinning TMAO catabolism in this bacterium remains to be established. TMAO not only occurs in the natural environment but also is widely used as important industrial solvent (Yancey 2005). Studying the catalytic mechanisms of TMAO degradation by microorganisms and its subsequent conversion to methylated amines, such as DMA, will advance our understanding of the impact of TMAO release into the environment.

The main aims of this chapter are:

- To identify the gene encoding Tdm in the model bacterium *Methylocella silvestris*;
- To identify the gene encoding a TMAO transporter in this bacterium and
- To carry out initial biochemical characterisation of Tdm

3.2 Materials and methods

3.2.1 Growth of *Methylocella silvestris* and mutant strains

M.silvestris and mutants were grown in DNMS using MAs as nitrogen source as described in **Chapter 2**. Either methanol (10 mM) or succinate (5 mM) was used as the carbon source. Concentrations of MAs in the media were determined by ion-exchange liquid chromatography (see **Chapter 2**).

3.2.2 Construction of marker-exchange mutants in *Methylocella silvestris*

The marker-exchange mutants (**Table 2-2**) were generated by Dr. Yin Chen and already available when the project started. Mutants of *M. silvestris* were constructed as described previously (Chen et al 2010c). Briefly, a downstream region and an upstream region of the target gene were amplified by PCR and sub-cloned into the pGEM-T vector (Promega, Madison, USA) together with a kanamycin (*kan*) gene cassette amplified from the plasmid pCM184 (Marx and Lidstrom 2002), which was inserted between the two regions (primers used are listed in Appendix 1). The downstream and upstream regions together with the *kan* gene cassette were then released from the resulting plasmid and transformed into *M. silvestris* competent cells *via* electroporation as described previously (Chen et al 2010b). Mutants were selected on the solid DNMS medium containing kanamycin (25 μg ml^{-1}), which were then screened by diagnostic PCR and subsequent sequencing.

3.2.3 Cloning and heterologous expression of *tdm* in *Escherichia coli*

Cloning of *tdm* of *M. silvestris* into an expression vector, pET28a, was performed by Dr Yin Chen. Plasmids and strains used for cloning and overexpression of *tdm* in *E. coli* are listed in **Table 2-2**. Briefly, the *tdm* gene (Msil_3603) was amplified by PCR and sub-cloned into the pGEM-T vector (Promega, Madison, USA), which was then excised using the *Nde*I/*BamH*I sites and ligated into the expression vector pET28a (Merck Biosciences, Germany). The resulting plasmid was sequenced prior to being transformed into the expression host *E. coli* BLR(DE3) pLysS (Merck Biosciences, Germany).

3.2.4 Protein purification and enzymatic assays

Protein purification and 6*His-tag removal were carried out as described in **Chapter 2**. Several buffers with a range of pH were first compared in order to determine a suitable buffering system and the optimum pH for the purified Tdm. The data presented in **Figure 3-1** demonstrated that Tdm had highest activity at pH 6.0 in 10 mM MES (2-(*N*-morpholino) ethanesulfonic acid) buffer, which was then chosen for the following experiments. For enzymatic activity assays, either formaldehyde or DMA production from TMAO was quantified. DMA and formaldehyde quantification were carried out using the aforementioned ion-exchange chromatography and the Purpald assay as described in **Chapter 2**, respectively.

Steady-state kinetics were determined in triplicate. To determine the substrate specificity of Tdm, the compounds with similar structure as TMAO were tested. The assays were performed in triplicate and the compounds used were at a final concentration of 10 mM. To determine the stoichiometry of TMAO demethylation by Tdm, the enzyme reaction was initiated by adding TMAO at 2-8 mM and DMA and formaldehyde production was quantified after 60 min when TMAO was completely consumed.

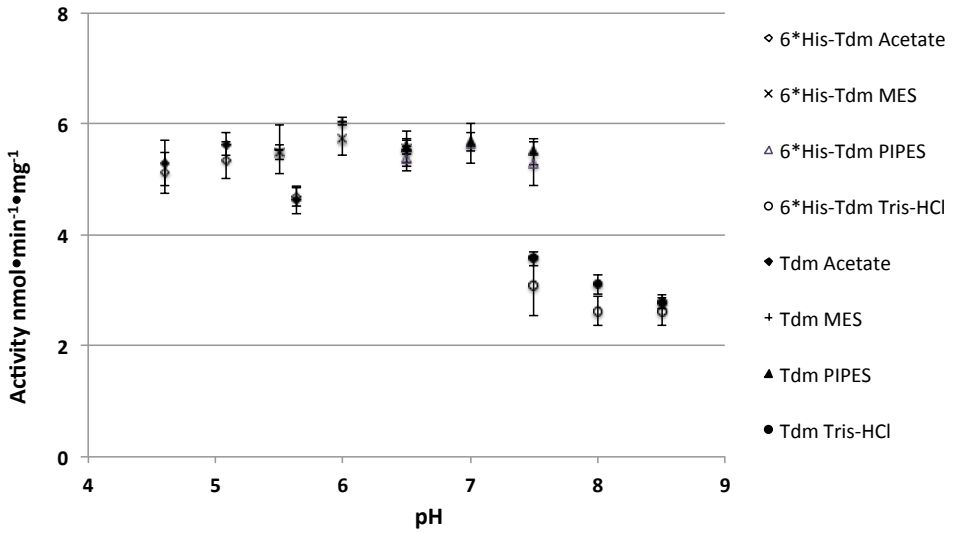


Figure 3-1 Optimal pH of purified recombinant 6*His-tag Tdm and 6*His-tag removed native Tdm in various pH buffers (final concentration, 10 mM). MES: 2-(N-morpholino) ethanesulfonic acid; PIPES: 1,4-piperazinediethanesulfonic acid; Tris-HCL: 2-amino-2-hydroxymethyl-propane-1,3-diol; Acetate: acetic acid and sodium acetate.

3.2.5 Bioinformatics

Homologous proteins were identified using the BLASTp programme using the Msil_3603 and Msil_3606 sequences of *M. silvestris* as the query. Protein sequences were aligned, end-trimmed and analyzed using the MEGA5 package (Tamura et al 2011). All phylogenetic trees were constructed using the minimum evolution method (default settings) with 1,000 bootstrap replicates. Accession numbers from the Uniprot database for all sequences used in phylogenetic analyses are listed in **Tables 3-2**, **3-3** and **3-4** for tetrahydrofolate (THF)-binding domains, DUF1989 domains and the amino acids-polyamine (APC) superfamily members, respectively. Analysis of conserved domains in protein was carried out using Pfam (release 27.0), (Punta et al 2012).

3.2.6 Analytical ultracentrifugation

Purified Tdm from recombinant *E. coli* was exhaustively dialyzed against 50 mM sodium phosphate (pH 7.0) containing 100 mM NaCl, 1 mM D, L-dithiothreitol. Tdm protein samples were centrifuged at 30,000 $r \cdot min^{-1}$ at 4 ºC for 16 h in an eight-cell An-50 Ti rotor in a Beckman XLI analytical ultracentrifuge (Beckman). Migration of the protein during centrifugation was monitored by measuring the distribution of absorbance at 280 nm across the sample in the centrepiece at 120 consecutive time points. Molecular masses were calculated by the SEDFIT package

using a c(s) model (Dam et al 2005, Schuck 2000). Protein partial specific volumes, buffer viscosities and densities were all calculated using SEDNTERP (http://sednterp.unh.edu/).

3.2.7 Native polyacrylamide gel electrophoresis

Native page was done as described in **Chapter 2**.

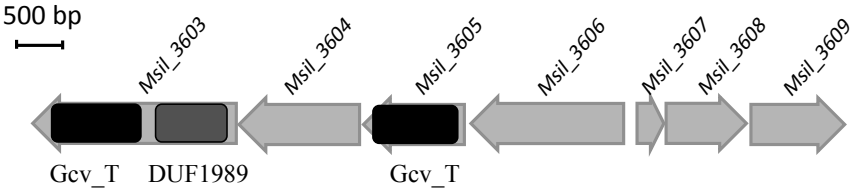
3.3 Results

3.3.1 Genes of Msil_3603 to Msil_3609 are involved in methylated amine metabolism in *Methylocella silvestris*

It is known that *M. silvestris* can utilize TMA as a sole carbon and nitrogen source through an indirect pathway, involving TMAO as the key intermediate (Chen et al., 2011). The gene encoding TMA monooxygenase (Tmm), the first enzyme in the TMA oxidation pathway in this bacterium, has already been identified (ORF *Msil_3604*) (Chen et al 2011). However the gene encoding the TMAO demethylase (Tdm) remains unknown. Comparative proteomics data have shown that peptides encoded by the genes in the neighbourhood of *tmm* (**Figure 3-2**) were also induced in the presence of TMA, suggesting a role in TMA oxidation (Chen et al 2011). In order to establish the role of the neighbouring gene in TMA metabolism and to identify the Tdm-encoding gene, several marker-exchange mutants in this bacterium have been generated. The mutants were then cultivated in a defined medium using

TMA, TMAO, DMA or MMA as the sole nitrogen source in order to establish their growth phenotype on MAs.

As it has been predicted previously (Chen et al 2011), the *tmm* mutant *(*Δ*Msil_3604*) could grow on TMAO, DMA and MMA, but not TMA (**Table 3-1**), confirming that it is only involved in the first step of TMA oxidation. Two mutants, Δ*Msil_3608* and Δ*Msil_3609*, could only grow on MMA (**Table 3-1**), suggesting that they are likely to encode subunits of the DMA monooxygenase (Dmm), which is the immediate upstream step for the oxidation of DMA to MMA. The gene *Msil_3607* is only 585 bp long and a mutant of *Msil_3607* is therefore not constructed in this study. However, Dmm has been previously purified from *Aminobacter aminovorans* and it is known to consist of 3 subunits (Alberta and Dawson 1987). Therefore, it is hypothesised that *Msil_3607-Msil_3609* may encode a Dmm in *M. silvestris*.



ORF	Length (AA)	Gene name	Annotation	Function	Reference
Msil_3603	761	*tdm*	Glycine cleavage system T protein, containing a N-terminal domain of unknown function (DUF1989) and a C-terminal tetrahydrofolate-binding domain (Gcv_T)	TMAO demethylase	This study
Msil_3604	451	*tmm*	A flavin-containing monooxygenase	TMA monooxygenase	Chen *et al*., 2011
Msil_3605	378	*dmmD*	Aminomethyltransferase, containing a tetrahydrofolate-binding domain (Gcv_T)	DMA monooxygenase subunit?	This study
Msil_3606	571	*tmoP*	Putative transmembrane amino acid transporter protein	TMAO permease	This study
Msil_3607	195	*dmmA*	Hypothetical protein	DMA monooxygenase subunit α	This study
Msil_3608	317	*dmmB*	Putative NADPH-flavodoxin reductase	DMA monooxygenase subunit β	This study
Msil_3609	351	*dmmC*	Protein of unknown function (DUF3445)	DMA monooxygenase subunit γ	This study

Figure 3-2 The genomic neighbourhood of TMA monooxygenase (*tmm*) in *Methylocella silvestris* BL2 and their putative functions. Conserved domains (highlighted in black) in Msil_3603 and Msil_3605 were identified using the conserved domain database (Marchler-Bauer et al., 2013).

Table 3-1 Growth of wild type and mutants of *Methylocella silvestris* on methylated amines*

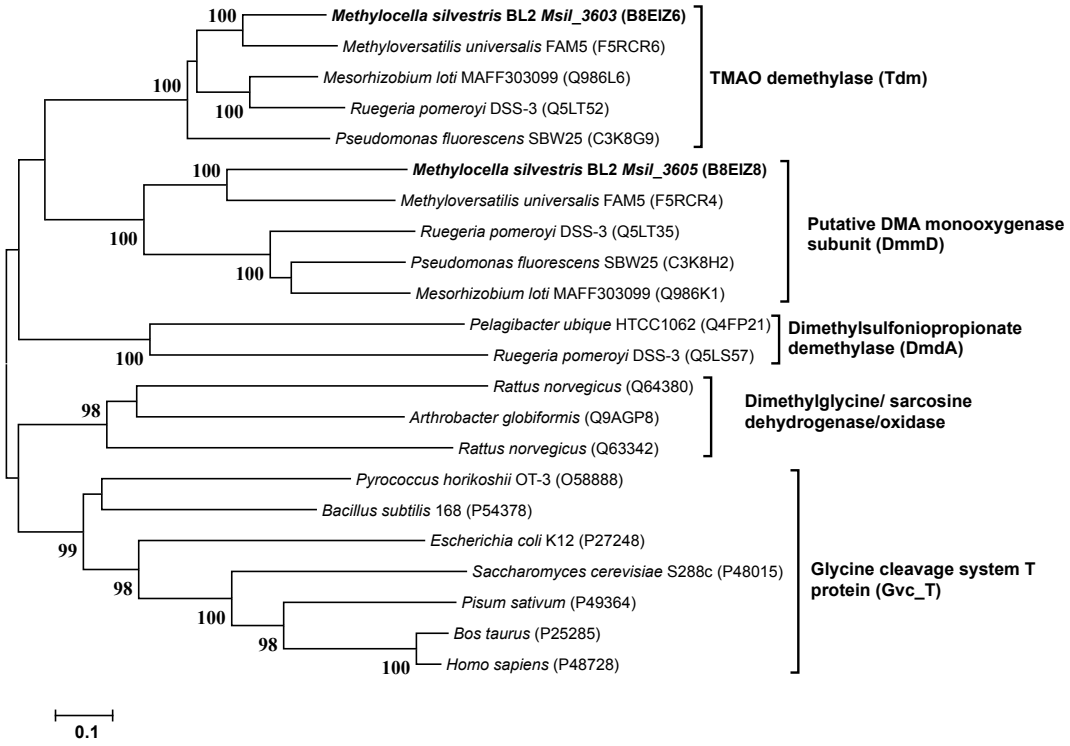
Substrate	**Wild type (h^{-1})**	**Δ*Msil_3603* (h^{-1})**	**Δ*tmm* (Δ*Msil_3604*) (h^{-1})**	**Δ*Msil_3605* (h^{-1})**	**Δ*Msil_3606* (h^{-1})**	**Δ*Msil_3608* (h^{-1})**	**Δ*Msil_3609* (h^{-1})**
TMA	0.030 ± 0.002	-	-	0.010 ± 0.001	0.025 ± 0.001	-	-
TMAO	0.045 ± 0.000	-	0.045 ± 0.003	0.010 ± 0.003	-	-	-
DMA	0.032 ± 0.002	0.027 ± 0.004	0.040 ± 0.006	0.024 ± 0.004	0.024 ± 0.001	-	-
MMA	0.031 ± 0.001	0.034 ± 0.005	0.047 ± 0.003	0.029 ± 0.000	0.025 ± 0.001	0.027 ± 0.003	0.027 ± 0.001

*, Methanol was used as the carbon source and methylated amines were used as the sole nitrogen source.

-, No growth. Values presented are average ± standard deviations of experiments run in triplicates.

The mutant Δ*Msil_3603* can utilize DMA and MMA, but not TMA or TMAO (**Table 3.1**). Msil_3603 is annotated as a glycine cleavage T protein (aminomethyl transferase) in the Genbank and Uniprot databases. It is composed of two domains, an uncharacterized N terminal domain (DUF1989) and a conserved THF-binding C terminal domain (Gcv_T), which is found in several well-characterized THF-dependent enzymes, such as glycine cleavage T protein (Okamura-Ikeda et al 2005) and dimethylsulfoniopropionate demethylase (Schuller et al 2012). Phylogenetic analysis of the THF-binding domain revealed that Msil_3603 formed a unique cluster in the family (**Figure 3-3 A**). Other sequences clustered within this clade include representatives that are known to metabolise MAs (Kalyuzhnaya et al 2006, Lidbury et al 2014). The N-terminus of Msil_3603 is an uncharacterized domain (DUF1989) with no known function. Phylogenetic analysis of the proteins of the DUF1989 superfamily showed the presence of four distinct clusters (**Figure 3-3 B**), none of which has been assigned function experimentally. DUF1989 in Msil_3603 shows modest sequence similarity (9-30%) to urea-carboxylase associated proteins, whose functions in urea catabolism are not yet known (Kanamori et al., 2004). It is hypothesized that the ORF *Msil_3603* may encode Tdm (E.C. 4.1.2.32) in this bacterium and this was tested further as shown below.

A


Methylocella silvestris BL2 Msil_3603 (B8EIZ6)
Methyloversatilis universalis FAM5 (F5RCR6)
Mesorhizobium loti MAFF303099 (Q986L6)
Ruegeria pomeroyi DSS-3 (Q5LT52)
Pseudomonas fluorescens SBW25 (C3K8G9)
TMAO demethylase (Tdm)
Methylocella silvestris BL2 Msil_3605 (B8EIZ8)
Methyloversatilis universalis FAM5 (F5RCR4)
Ruegeria pomeroyi DSS-3 (Q5LT35)
Pseudomonas fluorescens SBW25 (C3K8H2)
Mesorhizobium loti MAFF303099 (Q986K1)
Putative DMA monooxygenase subunit (DmmD)
Pelagibacter ubique HTCC1062 (Q4FP21)
Ruegeria pomeroyi DSS-3 (Q5LS57)
Dimethylsulfoniopropionate demethylase (DmdA)
Rattus norvegicus (Q64380)
Arthrobacter globiformis (Q9AGP8)
Rattus norvegicus (Q63342)
Dimethylglycine/ sarcosine dehydrogenase/oxidase
Pyrococcus horikoshii OT-3 (O58888)
Bacillus subtilis 168 (P54378)
Escherichia coli K12 (P27248)
Saccharomyces cerevisiae S288c (P48015)
Pisum sativum (P49364)
Bos taurus (P25285)
Homo sapiens (P48728)
Glycine cleavage system T protein (Gvc_T)
0.1

B

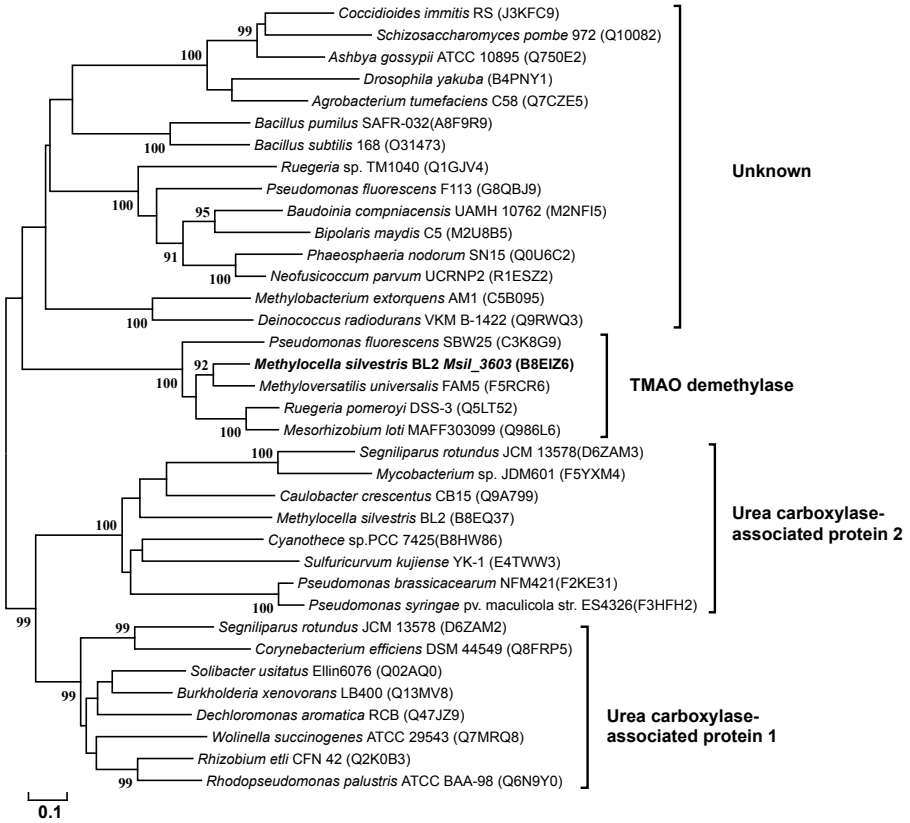

Coccidioides immitis RS (J3KFC9)
Schizosaccharomyces pombe 972 (Q10082)
Ashbya gossypii ATCC 10895 (Q750E2)
Drosophila yakuba (B4PNY1)
Agrobacterium tumefaciens C58 (Q7CZE5)
Bacillus pumilus SAFR-032(A8F9R9)
Bacillus subtilis 168 (O31473)
Ruegeria sp. TM1040 (Q1GJV4)
Pseudomonas fluorescens F113 (G8QBJ9)
Baudoinia compniacensis UAMH 10762 (M2NFI5)
Bipolaris maydis C5 (M2U8B5)
Phaeosphaeria nodorum SN15 (Q0U6C2)
Neofusicoccum parvum UCRNP2 (R1ESZ2)
Methylobacterium extorquens AM1 (C5B095)
Deinococcus radiodurans VKM B-1422 (Q9RWQ3)
Unknown
Pseudomonas fluorescens SBW25 (C3K8G9)
Methylocella silvestris BL2 Msil_3603 (B8EIZ6)
Methyloversatilis universalis FAM5 (F5RCR6)
Ruegeria pomeroyi DSS-3 (Q5LT52)
Mesorhizobium loti MAFF303099 (Q986L6)
TMAO demethylase
Segniliparus rotundus JCM 13578(D6ZAM3)
Mycobacterium sp. JDM601 (F5YXM4)
Caulobacter crescentus CB15 (Q9A799)
Methylocella silvestris BL2 (B8EQ37)
Cyanothece sp.PCC 7425(B8HW86)
Sulfuricurvum kujiense YK-1 (E4TWW3)
Pseudomonas brassicacearum NFM421(F2KE31)
Pseudomonas syringae pv. maculicola str. ES4326(F3HFH2)
Urea carboxylase-associated protein 2
Segniliparus rotundus JCM 13578 (D6ZAM2)
Corynebacterium efficiens DSM 44549 (Q8FRP5)
Solibacter usitatus Ellin6076 (Q02AQ0)
Burkholderia xenovorans LB400 (Q13MV8)
Dechloromonas aromatica RCB (Q47JZ9)
Wolinella succinogenes ATCC 29543 (Q7MRQ8)
Rhizobium etli CFN 42 (Q2K0B3)
Rhodopseudomonas palustris ATCC BAA-98 (Q6N9Y0)
Urea carboxylase-associated protein 1
0.1

Figure 3-3 Phylogenetic analyses of the tetrahydrofolate (THF)-binding domain (~358 amino acids) (A) and the DUF1989 domain (~197 amino acids) (B) of the putative TMAO demethylase (Tdm, encoded by Msil_3603). Bootstrap values (1,000 replicates) great than 90% are shown in percentage for each node. The bar represents 1 substitution per 10 amino acids in the aligned sequences. The Uniprot accession number of the sequences used for the phylogenetic analysis are listed in **Table 3-2, 3-3**.

The ORF *Msil_3606* encodes a membrane protein, consisting of 12 transmembrane helices. It is annotated as a putative transmembrane amino acid transporter protein in the Genbank and Uniprot databases. Our phylogenetic analyses suggest that Msil_3606 belongs to the amino acid-polyamine membrane transporter superfamily (APC family) (**Figure 3-4**). The APC family currently consists of 14 clades, 12 of which have been functionally assigned (Saier 2000). Msil_3606, together with sequences from known MA utilizers (*e.g. Methyloversatilis*) formed a distinct clade independent of the currently known APC family members. Marker-exchange mutagenesis experiments showed that the mutant (Δ*Msil_3606*) can grow on TMA, DMA and MMA, but not TMAO (**Table 3-1**), suggesting that it encodes a functional TMAO transporter. The role of *Msil_3605* in MA oxidation was not very clear. Msil_3605 has a single THF-binding domain (Gcv_T), which shows 32% identity to the C terminal Gcv_T domain in Msil_3603. The *ΔMsil_3605* mutant was able to grow on MMA, however its growth on TMA, TMAO and DMA was much slower compared to those of the wild type (**Table 3-1**).

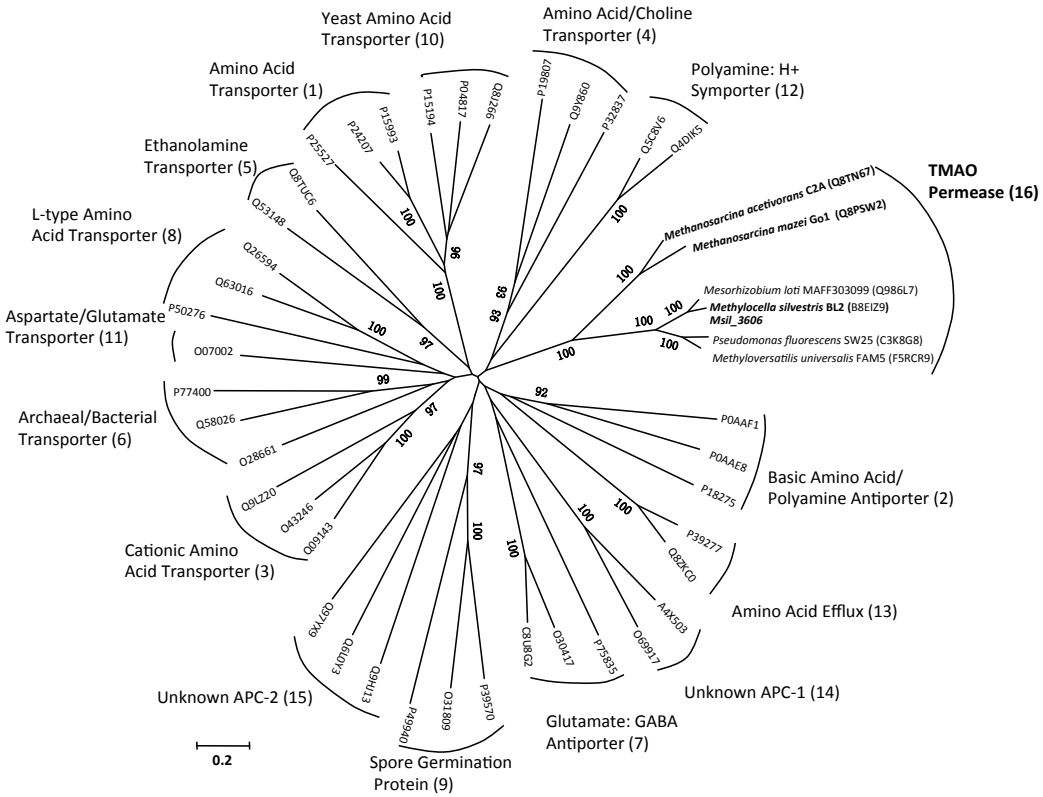


Figure 3-4 Phylogenetic analysis of the putative TMAO permease (TmoP) in the amino acid/polyamine/organocation (APC) superfamily (~375 amino acids). Bootstrap values (1,000 replicates) greater than 90% are shown for each node. The Uniprot accession number of the sequences used for the phylogenetic analysis are listed in **Table 3-4**.

3.3.2 Msil_3603 and Msil_3606 are required for *Methylocella silvestris* to grow on TMAO

In order to establish if *Msil_3603* and *Msil_3606* are indeed specifically required for TMAO metabolism, TMAO concentrations were quantified by ion-exchange chromatography in the culture medium of the wild type and the mutant strains of *M. silvestris*. Succinate was used as the sole carbon source instead of methanol because

methanol can damage the ion-exchange chromatography column used in this study. As shown in **Figure 3-5**, the wild type strain could grow on TMAO plus succinate and TMAO was completely depleted within 10 days (detection limit, 5 μM). However, the growth rates of the mutants (Δ*Msil*_3603, Δ*Msil_3606*) on TMAO as a sole nitrogen source were significantly reduced compare to that of the wild type, and TMAO concentrations in the medium remained unchanged throughout the experiment. Therefore, the data indicate that *Msil_3606* encodes a transporter required for TMAO uptake in *M. silvestris*, which was designated as TmoP.

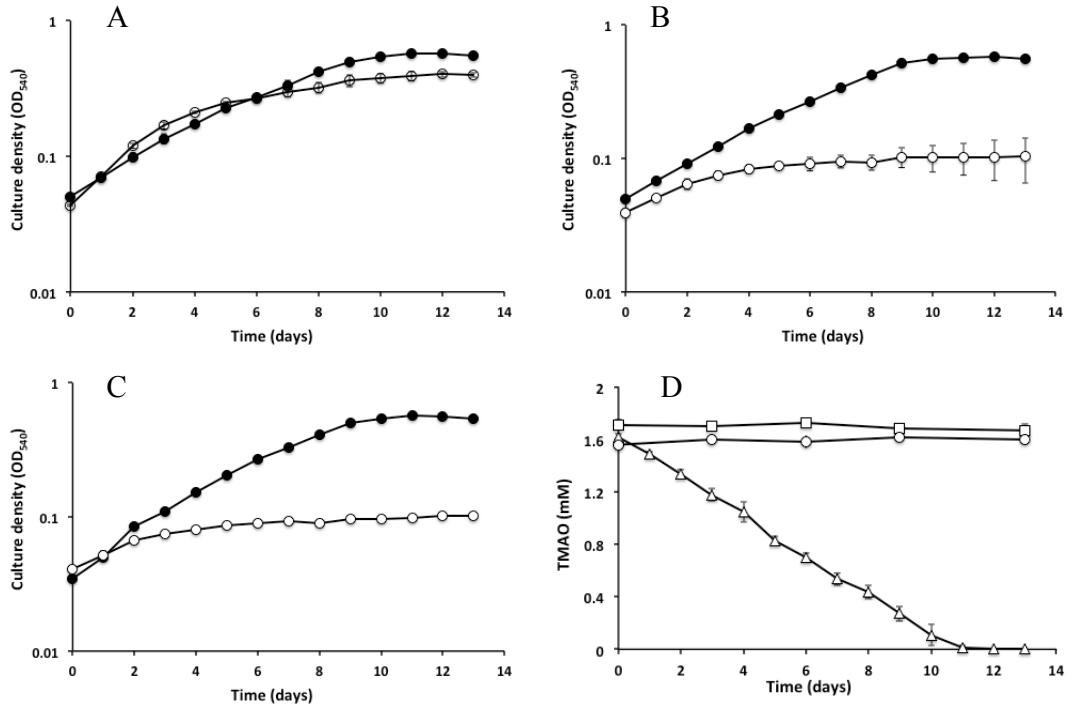


Figure 3-5 Growth of *Methylocella silvestris* of wild type (A), the Δ*Msil_3603* mutant (B) and the Δ*Msil_3606* mutant (C) on TMAO (—○—) or nitrate (—●—) as the sole nitrogen source. Quantification of TMAO consumption during growth of the wild type (—△—), Δ*Msil_3603* mutant (—○—) and the Δ*Msil_3606* mutant (—□—) (D). Nitrate was used as a sole nitrogen source as positive controls. Error bars indicate standard deviations of experiments run in triplicate.

3.3.3 *Msil_3603* encodes a bacterial TMAO demethylase

Gene *Msil_3603* was cloned from *M. silvestris* into an *Escherichia coli* host, overexpressed and further purified this protein with 6×His tag at its N-terminus by nickel (Ni^{2+}) affinity chromatography in order to establish whether it is a *bona fide* TMAO demethylase (Tdm). Eukaryotic Tdm has been purified previously (Fu et al 2006, Kimura et al 2000, Parkin and Hultin 1986, Takeuchi et al 2003), however, its microbial counterpart has only been partially purified from *Aminobacter aminovorans* and *Bacillus* sp. PM6 (Large 1971, Myers and Zatman 1971). The two products of TMAO demethylation are DMA and formaldehyde, and these two compounds were indeed detected when the purified protein was presented with TMAO (**Figure 3-6**). The stoichiometry of TMAO demethylation is determined to be 1 TMAO → 1 DMA + 1 HCHO.

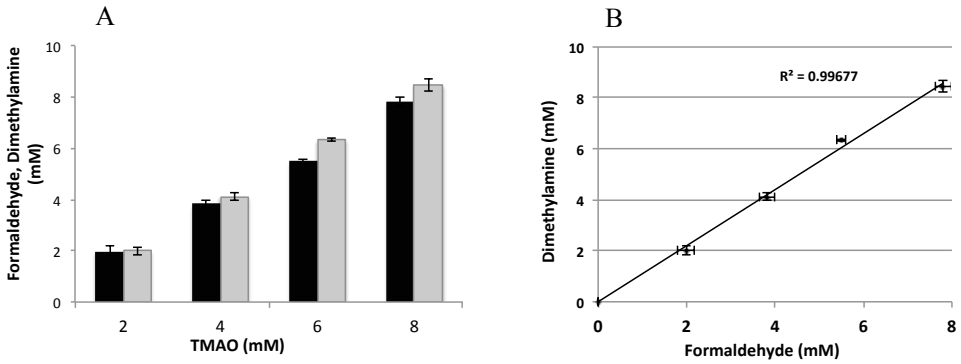


Figure 3-6 Stoichiometry of TMAO demethylation by Tdm (A). Black and grey bars represent formaldehyde and DMA concentrations, respectively. Ratio of DMA to formaldehyde (B). Error bars indicate standard deviations of triplicate experiments.

3.3.4 Initial characterization of Tdm of *M. silvestris* from recombinant *E. coli*.

The purified Tdm protein from recombinant *Escherichia coli* had a molecular weight of ~80 kDa under denaturing conditions (**Figure 3-7 A**), in good agreement with the calculated value from its amino acid sequence (82,547 Da). Its native molecular weight was estimated by two complementary methods, native gel electrophoresis and analytical ultracentrifugation, both of which suggested that the native Tdm was likely to be hexameric (**Figure 3-7 B, C**). The purified protein has an optimal pH at ~6.0 (**Figure 3-1**) and had no recognizable absorbance peak under UV-visible light (220 nm – 600 nm) besides the peak at 280 nm (**Figure 3-8**). Under optimum conditions, V_{max} and K_m of the recombinant Tdm were determined to be 21.7 ± 0.74 nmol min^{-1} mg^{-1} and 3.3 ± 0.64 mM, respectively by the Eadie-Hofstee plot (**Figure 3-9**). Its K_m value of Tdm of *M. silvestris* is in good agreement to that of *Aminobacter aminovorans* (2 mM) and *Bacillus* sp. (2.85 mM), respectively (Large 1971, Myers and Zatman 1971). The recombinant Tdm enzyme is specific for TMAO, among the compounds tested, it only showed ~50% activity to dimethyldodecylamine *N*-oxide (**Figure 3-10**). In contrast to the eukaryotic counterparts (Parkin and Hultin 1986), no enhancement of activity was found with additional ferrous iron or cysteine added to the *in vitro* enzyme assays (**Figure 3-11**),

however the impact of metals on Tdm activity was further investigated in detail in **Chapter 4**.

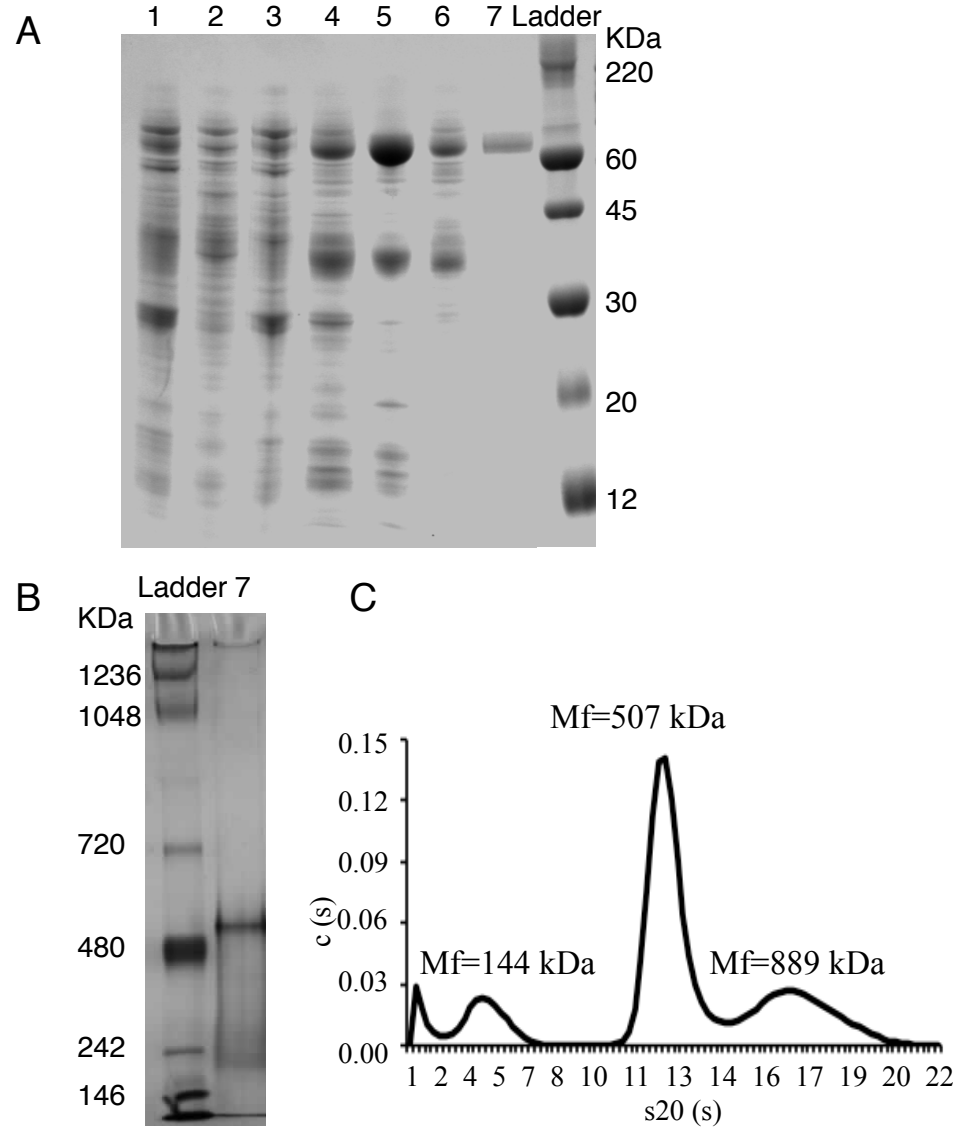


Figure 3-7 Estimation of molecular weight of purified Tdm by denaturing (A) and native (B) gel electrophoresis and analytical ultracentrifugation (C). 1, crude cell-free extract; 2, column flow through fraction of loaded cell-free extract; 3, column wash fraction with binding buffer containing 20 mM imidazole; 4-6, column wash fractions with 1, 3 and 6 bed volumns of washing buffer containing 60 mM imidazole; 7, elution fraction of the purified Tdm. Mf means the molar mass taking into account the current best-fit frictional ratio f/f0.

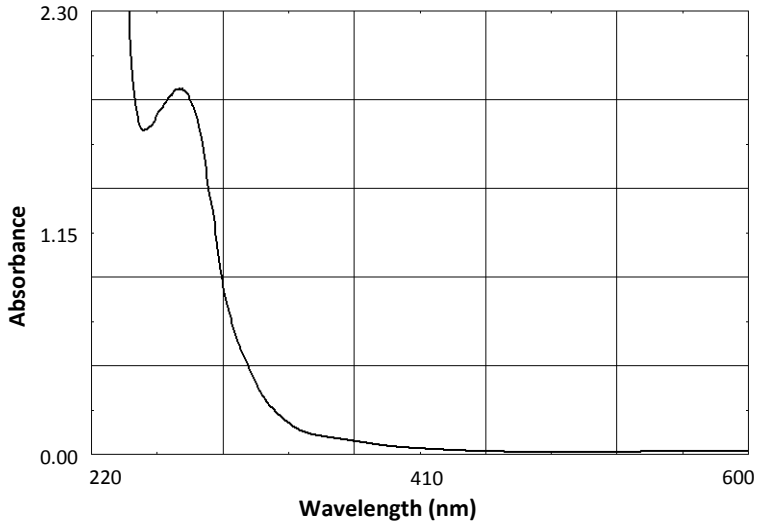


Figure 3-8 UV-visible absorbance spectrum (220 – 600 nm) of the purified recombinant Tdm of *Methylocella silvestris* BL2.

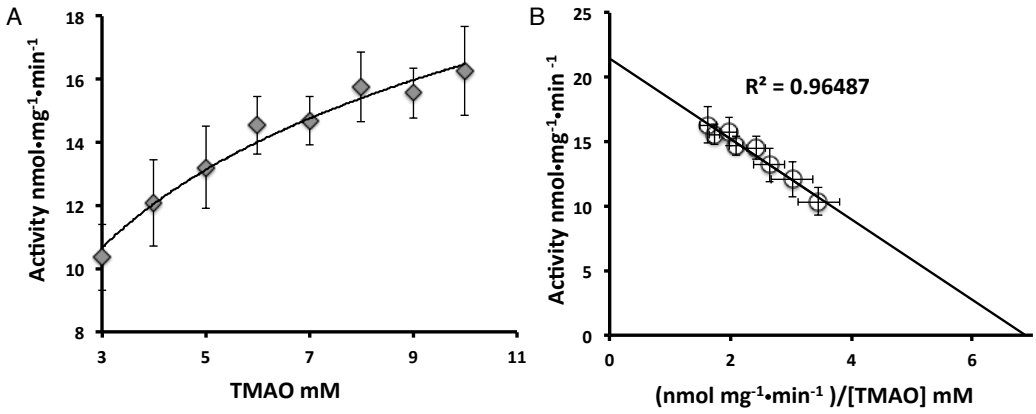


Figure 3-9 Steady-state kinetic parameters estimation. Michaelis-Menten representation of raw data (A). Eadie-Hofstee plot of kinetic data (B). Error bars indicate standard deviations of experiments run in triplicate.

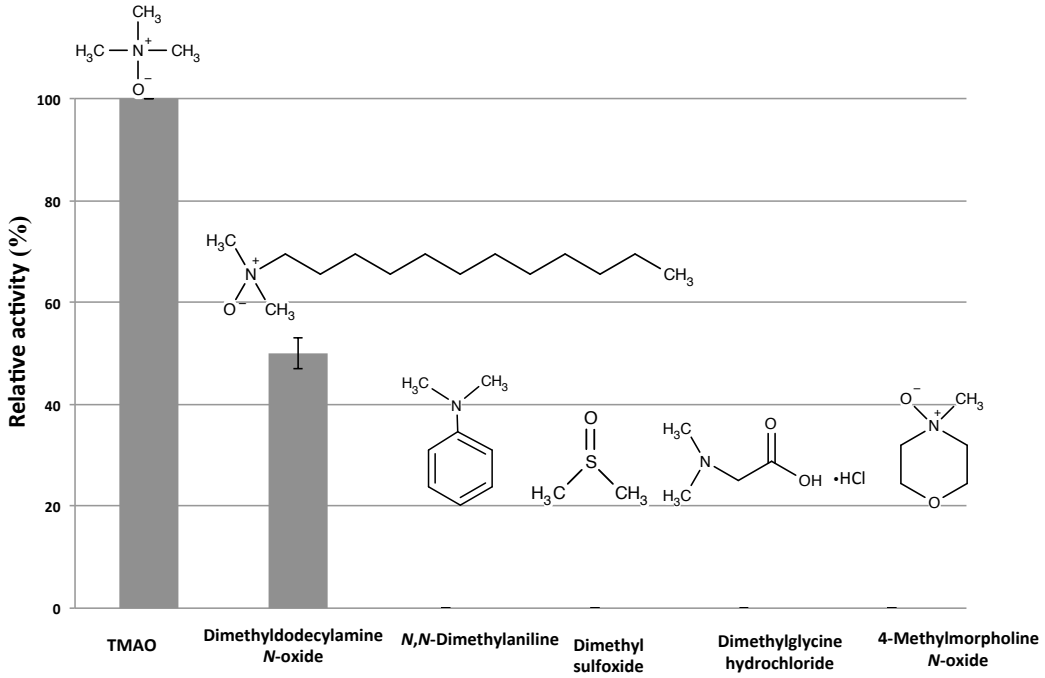


Figure 3-10 Relative activity of Tdm to selected structure analogues of TMAO. Error bars indicate standard deviations of experiments run in triplicate.

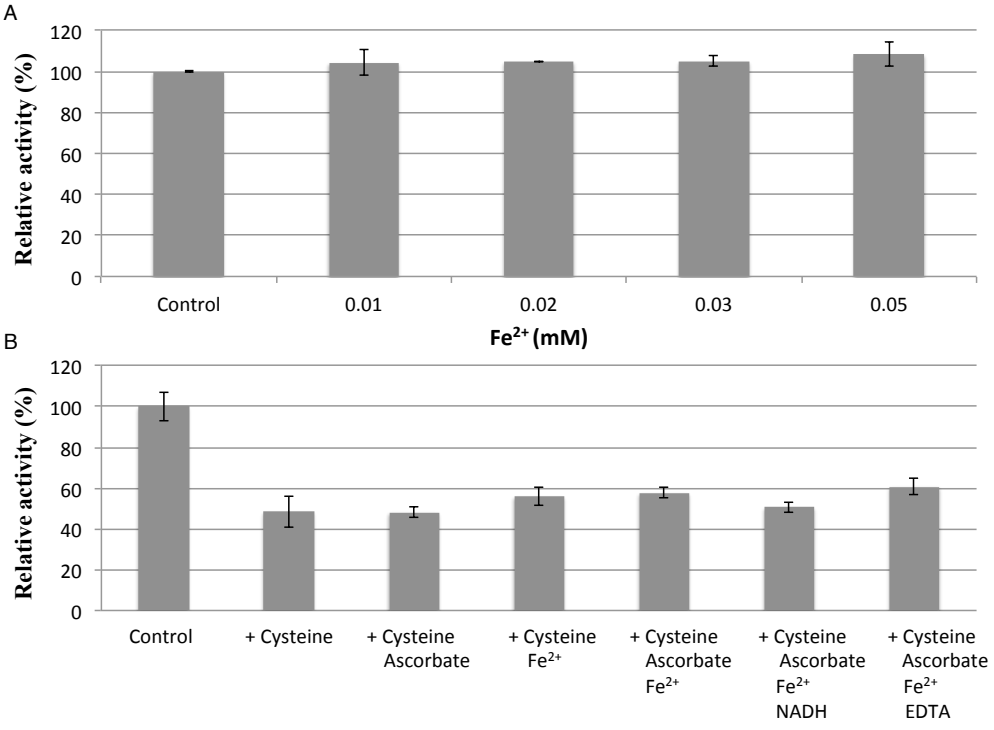


Figure 3-11 The effect of ferrous iron and cysteine on the activity of purified Tdm of *Methylocella silvestris*. Final concentration, cysteine, 2 mM; ascorbate, 2 mM; $FeCl_2$, 0.2 mM; NADH, 0.4 mM; EDTA, 10 mM.

3.3.5 Discussion

In this study, a TMAO membrane transporter (TmoP), the TMAO demethylase (Tdm) and the putative DMA monooxygenase have been identified in *Methylocella silvestris*. The presence of a specific transporter required for TMAO suggest that it can be taken up by *Methylocella silvestris* from the environment (Anthony 1982, Chen et al 2011). Although it is clear that TMAO can be used as a ubiquitous osmolyte by a range of marine biota (Gibb and Hatton 2004), the environmental sources of TMAO in soils and other terrestrial habitats are less clear. TMAO is a central metabolite involved in lipid metabolism in mammals and significant concentrations of TMAO have been detected in urine and other body fluids of humans (Smith et al 1994, Zhang et al 1992), rats (Smith et al 1994) and dogs (Richards et al 2013). It is therefore possible that the presence of TMAO in terrestrial environments, including soils, is the result of excretion from animals. However, it is also likely that TMAO is leaked out from microorganisms during the oxidation of TMA by microbial TMA monooxygenases (Chen et al 2011). Recent studies have shown that in agricultural and forest soils, precursors of TMA such as quaternary amines represent a major pool of dissolved organic nitrogen (Warren 2013a, Warren 2013b, Yu et al 2002). Microbial oxidation of TMA in soils may represent yet another source of TMAO in the environment.

The only known microbial TMAO transporter in the literature at the time when the project started was an ATP-dependent active transporter of the ABC superfamily found in *Aminobacter aminovorans* (Raymond and Plopper 2002). This ABC-type TMAO transporter was subsequently discovered by Ian Lidbury, a PhD student from Dr. Chen's group, in a range of marine bacteria, such as *Ruegeria pomeroyi* (Lidbury et al 2014). This study indicates that another type of microbial transporter for TMAO is present. This newly identified TMAO permease (TmoP) of *Methylocella silvestris* belongs to the APC superfamily but forms a distinct cluster (**Figure 3-4**). APC transporters are membrane permeases co-transporting another solute, acting as either a symporter or an antiporter (Saier 2000). It is not clear whether TmoP acts as a symporter or an antiporter and the co-transporting solute remains to be established. It is interesting to note that TmoP homologues are also found in some methanogenic *Archaea*, *e.g. Methanosarcina acetivorans*, *Methanosarcina mazei* (**Figure 3-4**) but it remains unclear whether TMAO can be directly used as a substrate for methanogenesis.

Microbial Tdm has been partially purified previously (Large 1971, Myers and Zatman 1971), and the gene encoding microbial Tdm remained unknown when the project was started. While my investigation on Tdm in *Methylocella silvestris* was underway, the *tdm* gene was also found in a range of marine bacteria (Lidbury et al

2014). Tdm from marine eukaryotes has also been purified, including those from the Alaskan Pollock (*Theragra chalcogramma*, Kimura et al 2000), the red hake (*Urophycis chuss*, Parkin and Hultin 1986) and the Humboldt squid (*Dosidicus gigas*, Fu et al 2006). Tdm sequences from bacteria and eukaryotes (Takeuchi et al 2003) have no sequence homology and have contrasting characteristics. For example, purified Tdm from *Dosidicus gigas* and *Theragra chalcogramma* have much smaller molecular mass, being 17.5 kDa and 25 kDa respectively. Their K_m values for TMAO (30 mM for *T. chalcogramma* and 26.2 mM for *D. gigas*) are significantly higher than those of the microbial Tdm (2–4 mM). Eukaryotic Tdm requires ferrous ion as an essential metal for activity whereas it has no obvious impact on microbial Tdm in *in vitro* assays (but also see **Chapter 4** for detailed characteriation of metal-dependent activity in Tdm). Tdm in bacteria and eukaryotes represent another example of convergent evolution where two forms of Tdm have evolved independently to catalyse the same biochemical reaction.

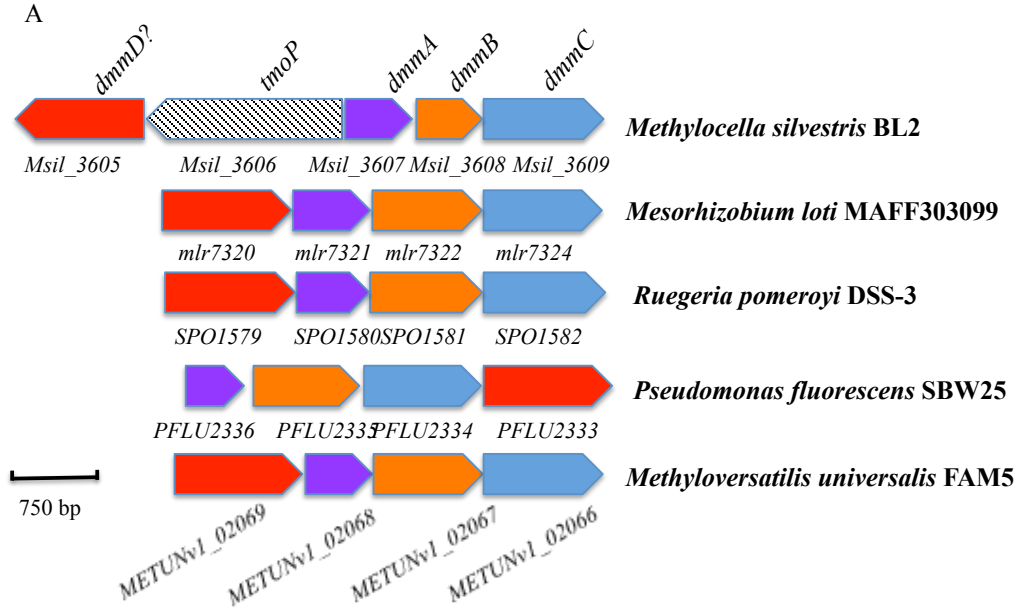
Another important finding from this study is the functional assignment of the DUF1989 domain as the N terminus of the microbial Tdm. Proteins having domains of unknown functions (DUF) currently represent more than a quarter of sequence entries in public databases such as Pfam (Punta et al 2012). Functional annotation of DUFs remains a great challenge for the scientific community since they not only

present a major knowledge gap between protein structure and functional relationship but also prevent the complete understanding of cellular functions from completed genomes (Galperin and Koonin 2010). Our phylogenetic analyses of DUF1989 representatives (1044 entries in Pfam in total) suggest the presence of at least four major clades, two of which are proteins associated with the urea carboxylase gene cluster (Kanamori et al 2004). However, the functions of the two DUF1989-containing proteins associated to this enzyme in microbial genomes remain unknown and warrant further experimental characterization.

The C-terminus of *M. silvestris* Tdm contains a highly conserved THF-binding domain, which is found in several enzymes catalysing the release of a formaldehyde molecule. Phylogenetic analyses of the THF-binding domain separate the sequences into five major clusters (**Figure 3-3 A**), three of which have been characterized previously, including the T protein of the glycine cleavage system, dimethylglycine and sarcosine dehydrogenase and dimethylsulfoniopropionate demethylase. The THF-binding domain of Tdm falls in to one of the previously recognized, but so far uncharacterized clades (Reisch et al 2008, Sun et al 2011). Comparative genomic analyses of the other group of THF-binding domain protein, represented by Msil_3605, revealed that they are located in the neighbourhood of the putative DMA

monooxygenases (Dmm) in other methylamine-utilizers (**Figure 3-12**), suggesting a role in DMA oxidation.

A



B

Organism	DmmA	DmmB	DmmC	DmmD
Methylocella silvestris BL2	Msil_3607 (100)	Msil_3608 (100)	Msil_3609 (100)	B8EIZ8 (100)
Mesorhizobium loti MAFF303099	mlr7321 (40)	mlr7322 (40)	mlr7324 (64)	mlr7320 (37)
Ruegeria pomeroyi DSS-3	SPO1580 (39)	SPO1581 (41)	SPO1582 (62)	SPO1579 (37)
Pseudomonas fluorescens SBW25	PFLU2336 (37)	PFLU2335 (40)	PFLU2334 (58)	PFLU2333 (39)
Methyloversatilis universalis FAM5	METUNv1_02068 (39)	METUNv1_02067 (39)	METUNv1_02066 (65)	METUNv1_02069 (54)

Figure 3-12 Putative DMA monooxygenase gene cluster in selected methylated amine utilizers (A). Purple: DMA monooxygenase α subunit (DmmA), Yellow: DMA monooxygenase β subunit (DmmB), Blue: DMA monooxygenase γ subunit (DmmC), Red: DMA monooxygenase δ subunit (DmmD). Gene locus tags in the genomes are shown. DMA monooxygenase homologues in selected methylated amine utilizers (B). Values in brackets are protein sequence identities to the respective homologues of *Methylocella silvestris* BL2. Identities to the respective homologues of *Methylocella silvestris* BL2 were given by IMG (The Integrated Microbial Genome) using Blastp program (https://img.jgi.doe.gov/cgi-in/w/main.cgi?section=FindGenesBlast&page= geneSearchBlast) (Markowitz et al., 2012).

Based on the present and previous studies (Chen et al 2010c, Chen et al 2011), it is now possible to propose a metabolic pathway of TMA metabolism in this bacterium. TMA is likely to be transported into the cell via an as-yet-unidentified transporter and is subsequently oxidised to release formaldehyde and ammonium (Figure 3-13). Previous genome analysis only identified the glutamine synthetase (GS)/glutamate synthase (GOGAT) as the pathway for ammonium assimilation in this bacterium (Chen et al 2010c). Formaldehyde released from MA oxidation can either be incorporated into biomass through the serine cycle or subjected to oxidation to CO_2 for generating energy and reducing equivalents. Because *M. silvestris* can grow on DMA and MMA, it is therefore likely that specific membrane transporters for these compounds are present in its genome. This study has suggested that the genes Msil_3607- Msil_3609 are likely to encode the Dmm whose activity has been confirmed previously in this bacterium (Chen et al 2011), and the knockout mutants can no longer grow on DMA. Dmm has previously been purified from *Aminobacter aminovorans* and shown to contain three subunits consisting of 24, 36 and 42 kDa respectively (Alberta and Dawson 1987), which are in good agreement with the predicted molecular mass of Msil_3607-Msil_3609, respectively. The role of the THF-containing ORF Msil_3605 in this pathway is not clear. The mutant had reduced growth rates when grown on DMA and TMAO, and it may encode a

subunit, which can be loosely associated with Dmm but facilitate the conjugation of formaldehyde released from TMAO demethylation, which may help to offset the toxicity effect of formaldehyde accumulation in the cell.

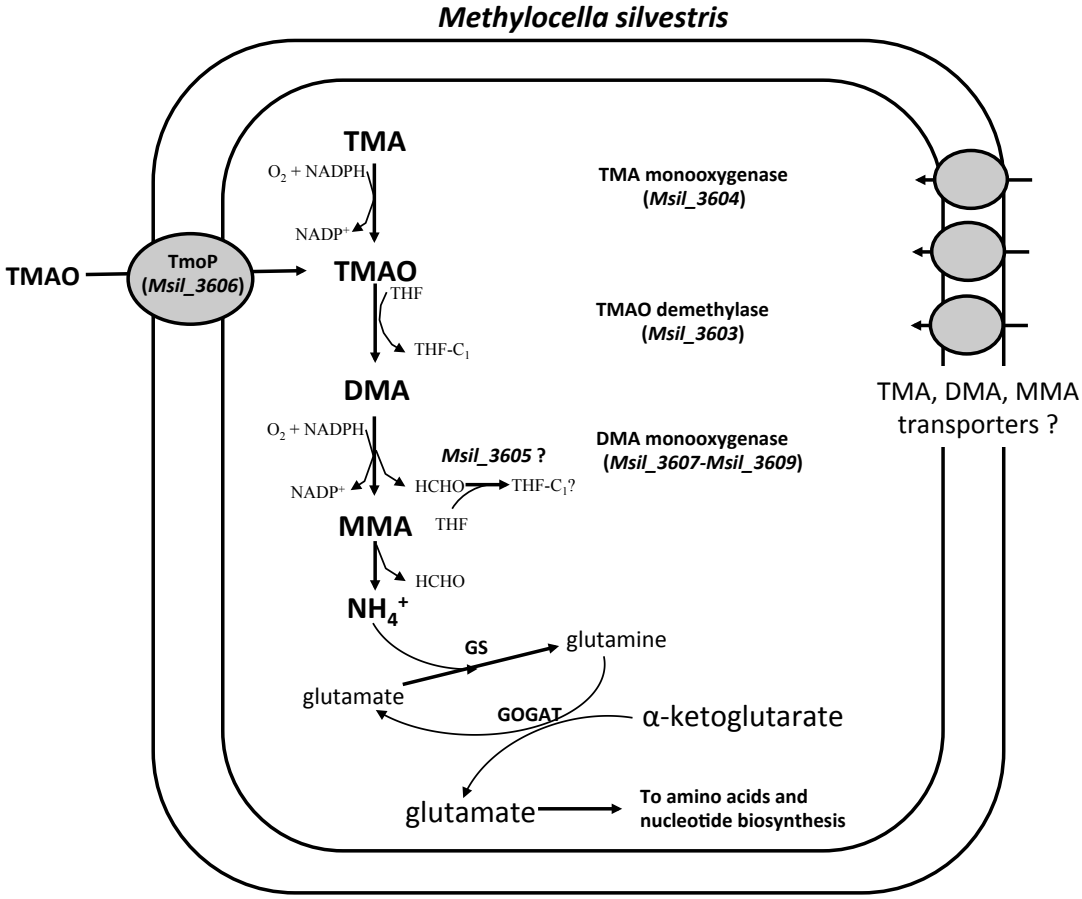


Figure 3-13 Proposed model of TMAO transport and metabolism in *Methylocella silvestris*. TMAO is either directly imported through the TmoP or resulted from the oxidation of TMA by Tmm. A membrane transporter for TMA in this bacterium is yet to be discovered. TMAO degradation by Tdm yields DMA and HCHO, which is likely to be conjugated to THF by the protein encoded by Msil_3605. MMA is further converted to ammonium through the γ-glutamylmethylamide/*N*-methylglutamate pathway, involving γ-glutamylmethylamide synthetase (Msil_2635), *N*-methylglutamate synthase (Msil_2632-Msil_2634) and *N*-methylglutamate dehydrogenase (Msil_2636-Msil_2639) (Chen et al 2010c). Ammonium is assimilated by *M. silvestris* as a nitrogen source through the glutamine synthetase (GS)/ glutamate synthase (GOGAT) pathway (Chen et al 2010c).

To conclude, genes encoding enzymes responsible for the uptake and catabolism of TMAO have been identified in *Methylocella silvestri*. The newly identified Tdm and TmoP proteins have not only furthered our understanding of TMA/TMAO degradation in this soil bacterium, but also expanded our knowledge about microbial cycling of Mas in terrestrial environments, functional assignment of the DUF1989 family and the expanding functions encoded in the APC superfamily.

Table 3-2 Representatives of the tetrahydrofolate (THF) binding domain-containing proteins

Cluster	Function	Uniprot accession No	Organism	Length (AA)	Identity (%)	Reference
Cluster 1	Glycine cleavage system, T protein (Gvc_T)	P49364	*Pisum sativum*	408	23	(Vauclare et al., 1998)
		P25285	*Bos taurus*	397	25	(Fujiwara et al., 1986)
		O58888	*Pyrococcus horikoshi*i OT-3	398	29	(Lokanath et al., 2004)
		P48728	*Homo sapiens*	403	25	(Okamura-Ikeda et al., 2005)
		P54378	*Bacillus subtilis* 168	362	32	(Kunst et al., 1997)
		P27248	*Escherichia coli* K12	364	27	(Stauffer et al., 1993)
		P48015	*Saccharomyces cerevisiae* S288c	400	19	(Sinclair et al., 1996)
Cluster 2	Dimethylglycine / sarcosine dehydrogenase/ oxidase	Q64380	*Rattus norvegicus*	911	26	(Cook et al., 1985)
		Q63342	*Rattus norvegicus*	848	27	(Porter et al., 1985)
		Q9AGP8	*Arthrobacter globiformis*	830	29	(Leys et al., 2003)
Cluster 3	Dimethylsulfoniopropionate demethylase (DmdA)	Q5LS57	*Ruegeria pomeroyi* DSS-3	364	26	(Reisch et al., 2008)
		Q4FP21	*Pelagibacter ubique* HTCC1062	369	24	(Reisch et al., 2008)

Cluster	Function	Uniprot accession No	Organism	Length (AA)	Identity (%)	Reference
Cluster 4	TMAO demethylase (Tdm)	B8EIZ6	*Methylocella silvestris* BL2	761	100	This study
		C3K8G9	*Pseudomonas fluorescens* SBW25	770	60	This study
		Q5LT52*	*Ruegeria pomeroyi* DSS-3	789	61	(Lidbury et al., 2014)
		Q986L6*	*Mesorhizobium loti* MAFF303099	789	64	This study
		F5RCR6	*Methyloversatilis universalis* FAM5	762	73	This study
Cluster 5	Putative DMA monooxygenase subunit (DmmD)	B8EIZ8	*Methylocella silvestris* BL2	378	32	This study
		Q5LT35	*Ruegeria pomeroyi* DSS-3	381	33	This study
		Q986K1	*Mesorhizobium loti* MAFF303099	377	34	This study
		C3K8H2	*Pseudomonas fluorescens* SBW25	376	35	This study
		F5RCR4	*Methyloversatilis universalis* FAM5	372	33	This study

AA: amino acids.

*Note that the start codon of these two genes has been mis-annotated in the genomes and a 408 AA upstream region are included for sequence alignment.

Table 3-3 Representatives of DUF1989 domain containing proteins

Cluster	Function	Uniprot accession No.	Organism	Length (AA)	Identity (%)	Reference
Cluster 1	TMAO demethylase	B8EIZ6	*Methylocella silvestris* BL2	761	100	This study
		C3K8G9	*Pseudomonas fluorescens* SBW25	770	57	This study
		Q5LT52*	*Ruegeria pomeroyi* DSS-3	789	63	(Lidbury et al., 2014)
		Q986L6*	*Mesorhizobium loti* MAFF303099	789	64	This study
		F5RCR6	*Methyloversatilis universalis* FAM5	762	78	This study
Cluster 2	Urea carboxylase associated protein 2	F2KE31	*Pseudomonas brassicacearum* NFM421	241	18	
		B8EQ37	*Methylocella silvestris* BL2	270	22	
		B8HW86	*Cyanothece* sp. PCC 7425	243	20	
		E4TWW3	*Sulfuricurvum kujiense* YK-1	251	13	
		D6ZAM3	*Segniliparus rotundus* JCM 13578	286	19	(Kanamori et al., 2004)
		F3HFH2	*Pseudomonas syringae* pv. Maculicola str. ES4326	241	16	
		Q9A799	*Caulobacter crescentus* CB15	279	22	
		F5YXM4	*Mycobacterium* sp. JDM601	244	9	

Cluster	Function	Uniprot accession No.	Organism	Length (AA)	Identity (%)	Reference
Cluster 3	Urea carboxylase associated protein 1	D6ZAM2	*Segniliparus rotundus* JCM 13578	220	30	
		Q8FRP5	*Corynebacterium efficiens* DSM 44549	238	26	
		Q02AQ0	*Solibacter usitatus* Ellin6076	216	25	
		Q13MV8	*Burkholderia xenovorans* LB400	218	27	
		Q47JZ9	*Dechloromonas 82romatic* RCB	219	26	
		Q7MRQ8	*Wolinella succinogenes* ATCC 29543	205	27	(Kanamori et al., 2004)
		Q2K0B3	*Rhizobium etli* CFN 42	216	27	
		Q6N9Y0	*Rhodopseudomonas palustris* ATCC BAA-98	211	26	
Cluster 4	Unknown	J3KFC9	*Coccidioides immitis* RS	292	16	This study
		Q10082	*Schizosaccharomyces pombe* 972	302	9	
		Q750E2	*Ashbya gossypii* ATCC 10895	304	14	
		B4PNY1	*Drosophila yakuba*	276	10	
		Q7CZE5	*Agrobacterium tumefaciens* C58	318	13	
		A8F9R9	*Bacillus pumilus* SAFR-032	198	27	
		O31473	*Bacillus subtilis* 168	198	28	
		Q1GJV4	*Ruegeria* sp. TM1040	233	31	
		G8QBJ9	*Pseudomonas fluorescens* F113	204	28	

Cluster	Function	Uniprot accession No.	Organism	Length (AA)	Identity (%)	Reference
Cluster 4	Unknown	M2NFI5	*Baudoinia compniacensis* UAMH 10762	527	20	This study
		M2U8B5	*Bipolaris maydis* C5	611	21	
		Q0U6C2	*Phaeosphaeria nodorum* SN15	256	20	
		R1ESZ2	*Neofusicoccum parvum* UCRNP2	255	24	
		C5B095	*Methylobacterium extorquens* AM1	231	32	
		Q9RWQ3	*Deinococcus radiodurans* VKM B-1422	216	29	

AA: amino acids.

*Note that the start codon of these two genes has been mis-annotated in the genomes and a 408 AA upstream region are included for sequence alignment.

Table 3-4 Representatives of the amino acid/polyamine/organocation (APC) superfamily membrane transporters

No	Family name (abbreviation)	Gene / locus number	Organism	Uniprot accession No	Description in database	Reference
1	The Amino Acid Transporter (AAT)	PheP	*Escherichia coli* K12	P24207	Phenylalanine:H^+ symporter	(Pi et al., 1991)
		AroP	*Escherichia coli* K12	P15993	Aromatic amino acid:H^+ symporter	(Honoré and Cole, 1990)
		GabP	*Escherichia coli* K12	P25527	γ-aminobutyrate:H^+ symporterTop of Form	(Niegemann et al., 1993)
2	Basic Amino Acid/Polyamine Antiporter (APA)	PotE	*Escherichia coli* K12	P0AAF1	Putrescine:ornithine antiporter	(Kashiwagi et al., 1991)
		CadB	*Escherichia coli* K12	P0AAE8	Cadaverine:lysine antiporter	(Soksawatmaekhin et al., 2004)
		ArcD	*Pseudomonas aeruginosa* LMG 12228	P18275	Arginine:ornithine antiporter	(Lüthi et al., 1990)

No	Family name (abbreviation)	Gene / locus number	Organism	Uniprot accession No	Description in database	Reference
3	Cationic Amino Acid Transporter (CAT)	CTR1	*Mus musculus*	Q09143	System Y^+ high affinity basic amino acid transporter	(Kim et al., 1991)
		CAT6	*Arabidopsis thaliana*	Q9LZ20	Neutral and cationic amino acids Transporter	(Hammes et al., 2006)
		CAT4	*Homo sapiens*	O43246	Cationic amino acid transporter	(Sperandeo et al., 1998)
4	Amino Acid/Choline Transporter (ACT)	Ctr	*Saccharomyces cerevisiae* S288c	P19807	Choline permease	(Nikawa et al., 1990)
		GabA	*Emericella nidulans*	Q9Y860	γ-aminobutyric acid (GABA) permease	(Hutchings et al., 1999)
		Uga4	*Saccharomyces cerevisiae* S288c	P32837	γ-aminobutyric acid (GABA) permease	(Uemura et al., 2007; Kashiwagi and Igarashi, 2011)
5	Ethanolamine Transporter (EAT)	EUTP	*Rhodococcus erythropolis*	Q53148	Ethanolamine import permease	(De Mot et al., 1994)
		MTTP	*Methanosarcina acetivorans*	Q8TUC6	Probable methylamine import permease	(Galagan et al., 2002)

No	Family name (abbreviation)	Gene / locus number	Organism	Uniprot accession No	Description in database	Reference
6	Archaeal/Bacterial Transporter (ABT)	Cat-1	*Archaeoglobus fulgidus*	O28661	Putative cationic amino acid permease	(Klenk et al., 1997)
		ApcT	*Methanocaldococcus jannaschii*	Q58026	Broad specificity amino acid transporter	(Bult et al., 1996)
		YbaT	*Escherichia coli* K12	P77400	Unknown	(Blattner et al., 1997)
7	Glutamate:GABA Antiporter (GGA)	GadC	*Lactococcus lactis subsp. Cremoris* MG1363	O30417	Glutamate:γ-aminobutyrate antiporter	(Sanders et al., 1998)
		YcaM	*Escherichia coli* K12	P75835	Unknown	(Serina et al., 2004)
		GadC	*Escherichia coli* O103:H2 12009	C8U8G2	Glutamate:GABA antiporter	(Ma et al., 2012)
8	L-type Amino Acid Transporter (LAT)	LAT1	*Rattus norvegicus*	Q63016	L-type neutral amino acid transporter	(Kanai et al., 1998)
		PRM1	*Schistosoma mansoni*	Q26594	Schistosome neutral and cationic amino acid transporter	(Krautz-Peterson et al., 2007)
		MUP1	*Saccharomyces cerevisiae* s288c	P50276	L-methionine and selenomethionine transporter	(Isnard et al., 1996)

No	Family name (abbreviation)	Gene / locus number	Organism	Uniprot accession No	Description in database	Reference
9	Spore Germination Protein (SGP)	GRBB	*Bacillus subtilis* 168	P39570	Spore germination protein B2	(Corfe et al., 1994)
		GerKB	*Bacillus subtilis* 168	P49940	Spore germination protein K2	(Kunst et al., 1997)
		YndE	*Bacillus subtilis* 168	O31809	Spore germination protein	(Kunst et al., 1997)
10	Yeast Amino Acid Transporter (YAT)	Gap1	*Saccharomyces cerevisiae* s288c	P19145	General amino acid permease	(Jauniaux and Grenson, 1990)
		Can1	*Saccharomyces cerevisiae* s288c	P04817	Arginine permease	(Hoffmann, 1985)
		GAP1	*Hebeloma cylindrosporum*	Q8J266	General amino acid uptake permease	(Wipf et al., 2002)
11	Aspartate/Glutamate Transporter (AGT)	YveA	*Bacillus subtilis* 168	O07002	Aspartate uptake permease	(Lorca et al., 2003)

No	Family name (abbreviation)	Gene / locus number	Organism	Uniprot accession No	Description in database	Reference
12	The Polyamine:H+ Symporter (PHS)	POT1	*Leishmania major*	Q5C8V6	Polyamine (putrescine, spermidine):H+ symporter	(Hasne and Ullman, 2005)
		POT1.1	*Trypanosoma cruzi*	Q4DIK5	Putriscene preferring polyamine:H+ symporter	(El-Sayed et al., 2005)
13	The Amino Acid Efflux (AAE)	YjeH	*Escherichia coli* K12	P39277	Probable hydrophobic amino acid efflux transporter	(Blattner et al., 1997)
		YjeH	*Salmonella typhimurium* LT2	Q8ZKC0	Ceftriaxone resistance porter	(McClelland et al., 2001)
14	Unknown APC-1 (U-APC1)	Strop_1487	*Salinispora tropica* CNB-440	A4X503	Unknown	Direct submission
		SCO6385	*Streptomyces coelicolor* A3(2)	O69917	Unknown	(Bentley et al., 2002)
15	Unknown APC-2 (U-APC2)	PTO0784	*Picrophilus torridus* JCM 10055	Q6L0Y3	Unknown	(Fuetterer et al., 2004)
		CAAT	*Thermoplasma acidophilum* JCM 9062	Q9HJ13	Cationic amino acid transporter related protein	(Ruepp et al., 2000)
		SSO1173	*Sulfolobus solfataricus* P2	Q97YX9	Unknown	(She et al., 2001)

No	Family name (abbreviation)	Gene / locus number	Organism	Uniprot accession No	Description in database	Reference
		Msil_3606	*Methylocella silvestris* BL2	B8EIZ9	TMAO transporter (100%)	This study
		MM_2964	*Methanosarcina mazei* Go1	Q8PSW2	Putative TMAO transporter? (26%)	This study
		PFLU_2329	*Pseudomonas fluorescens* SBW25	C3K8G8	Putative TMAO transporter? (53%)	This study
16	TMAO permease (TmoP)	mll7301	*Rhizobium loti* MAFF303099	Q986L7	Putative TMAO transporter? (76%)	This study
		METUNv1_02074	*Methyloversatilis universalis* FAM5	F5RCR9	Putative TMAO transporter? (56%)	This study
		Mtbp	*Methanosarcina acetivorans* C2A	Q8TN67	Putative TMAO transporter? (28%)	This study

Values in brackets are sequence identities to TmoP of *Methylocella silvestris* BL2

Chapter 4 O_2-independent demethylation of trimethylamine *N*-oxide by a TMAO demethylase of *Methylocella silvestris* BL2

4.1 Introduction

Bacterial Tdm is a key enzyme involved in bacterial degradation of TMA and TMAO (Chen et al 2011, Lidbury et al 2014, Zhu et al 2014) The enzyme was first proposed in the 1970s and has been partially purified from *Bacillus* sp. PM6 (Myers and Zatman 1971) and *Pseudomonas aminovorans* (now *Aminobacter aminovorans* (Large 1971)). Despite being purified from aerobic strains, Tdm can convert TMAO anaerobically to equimolar amounts of DMA and HCHO (1 TMAO → 1 DMA + 1 HCHO) (Large 1971, Lidbury et al 2014) (**Chapter 3**). The gene encoding Tdm has only been identified recently and it is now known that *tdm* is widely distributed in nature, particularly in heterotrophic bacteria of the *Roseobacter* clade and the SAR11 clade of marine bacterioplankton (Lidbury et al 2014).

Tdm is a homohexameric protein, each monomer comprising of two domains, an uncharacterized DUF1989-containing domain at its N-terminus and a tetrahydrofolate (THF)-binding domain (GCV_T) at its C-terminus. DUF1989 in Tdm shows modest sequence similarity (<30%) to urea-carboxylase associated proteins, whose functions in urea catabolism are as-yet unknown (Kanamori et al 2004). GCV_T domains, however, are found in several well-characterized THF-dependent enzymes, such as glycine cleavage T protein (Okamura-Ikeda et al 2005) and dimethylsulfoniopropionate demethylase (Schuller et al 2012), with a function of binding THF to accept formaldehyde. Therefore, it has been proposed previously that the N-terminal DUF1989 domain of Tdm may play a role in substrate binding and subsequent catalysis whereas its C-terminal GCV_T domain is responsible for HCHO conjugation with THF (**Chapter 3**).

It has been suggested that metal ions may play a role in Tdm catalysis (Myers and Zatman 1971). For example, the partially purified Tdm of *Bacillus* sp. PM6 was

strongly activated by ferrous iron (Myers and Zatman 1971). In agreement with the putative role of metals in catalysis, purified Tdm of *Methylocella silvestris* does not contain either FAD or NAD (**Chapter 3**). Although a crystal structure for Tdm has yet to be solved, structures of three DUF1989-domain containing proteins (3ORU, 3SIY, 3DI4) available in the PDB database all contain Zn^{2+}. However, the types of metal(s) present in Tdm are yet to be established experimentally, and the metal stoichiometry is not known.

Cytochrome P450 and horseradish peroxidase (HRP) oxidatively demethylate tertiary amine by inserting oxygen into the C_{α}−H bond by haem Fe(IV)-oxo (Cpd I) to generate a carbinolamine that spontaneously dealkylates forming the free amine and aldehyde (Roberts and Jones 2010). Two mechanisms of the initial C_{α}−H bond activation have been proposed, H-Atom Abstraction (HAT) that only forms C-centred substrate radical, and Single Electron Transfer (ET) pathway that forms both N- and C- centred substrate radicals during catalysis (**Figure 4-1**, (Chiavarino et al 2008)). Therefore, identification of the substrate radical intermediate will shed light on the reaction mechanism. 5,5-Dimethyl-1-Pyrroline-*N*-Oxide (DMPO) spin trap was employed to investigate the intermediate of HRP oxidation of *N*-substituted aromatic amines and found that N- or C-centred substrate radicals were detected depending on the structure of radical formed (Van der Zee et al 1989). In this chapter, DMPO spin trapping was employed to investigate substrate intermediate, hence to illuminate Tdm enzymatic mechanism.

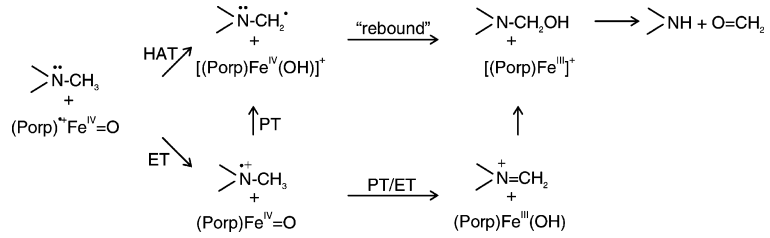


Figure 4-1 Proposed mechanisms in oxidative *N*-demethylation of tertiary amines catalyzed by $(Porp)^{\bullet+}Fe^{IV}=O$ (haem)(haem Fe(IV)-oxo), initiated by either H-Atom Abstraction (HAT) or Single Electron Transfer (ET). Adapted from (Chiavarino et al 2008).

The aims of the work described in this chapter are:

- To identify the metal cofactors and stoichiometry in Tdm
- To identify metal ion binding sites
- To investigate catalytic mechanism of Tdm

4.2 Materials and methods

4.2.1 Cloning, expression and purification of Tdm and mutants of *M. silvestris*

Plasmids and strains used for cloning and overexpression of Tdm and its mutants in *E. coli* are listed in **Table 2-2**. Tdm expression and purification were carried out as described in **Chapter 2**. Site-directed mutations in *tdm* were introduced by PCR and confirmed by DNA sequencing as described in **Chapter 2**. The oligonucleotides used in this study are shown in **Appendix 1**. For the Tdm reactivation and Fe^{2+} titration experiment, the 6*His-tag was removed from Tdm as described in **Chapter 2**. Protein sample preparation for various purposes is illustrated in **Figure 4-2**.

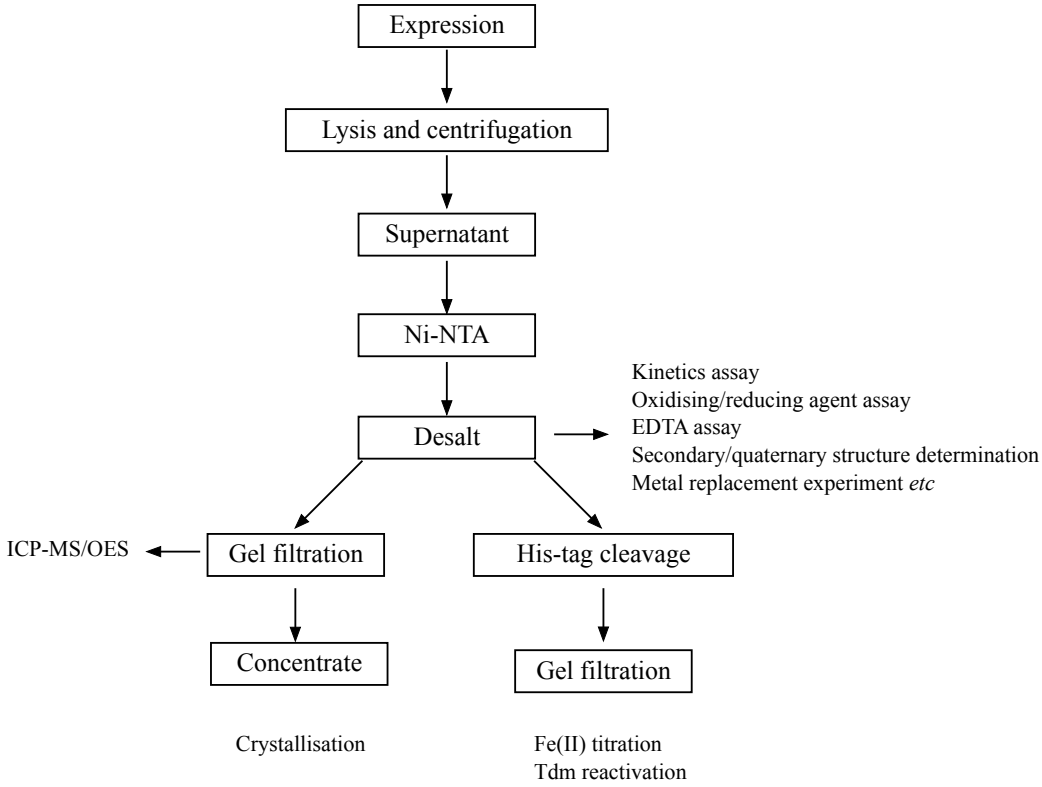


Figure 4-2 The workflow of Tdm sample preparation for biochemical analysis and crystallisation screening.

4.2.2 Enzyme activity assay

Enzyme activity was measured by quantifying HCHO production from TMAO degradation as described in **Chapter 2**. Enzyme kinetic parameters (V_{max}, K_m, k_{cat}) were calculated as described in **Chapter 2**. A range of concentrations of ethylenediaminetetraacetic acid (EDTA, 0 – 50 mM), ascorbic acid (0 – 10 mM) and H_2O_2 (0 – 100 μM) were added in the enzyme assays with 5 mM TMAO to a final volume of 50 μl to investigate their effects on Tdm activity.

For the metal replacement experiment, purified Tdm was diluted in a buffer containing 20 mM Tris-HCl (pH 7.8) and 100 mM NaCl to a final concentration of 1 mg ml^{-1} and incubated with various metal ions ($MgCl_2$, $MnCl_2$, $Fe(NH_4)_2(SO_4)_2$, $CoCl_2$, $NiSO_4$, $CuSO_4$, $ZnSO_4$) at an M^{2+}:Tdm ratio of 500 for 6 hours at 4 °C. 2 mM

ascorbic acid was used when $Fe(NH_4)_2(SO_4)_2$ was applied to maintain Fe^{2+} in its reduced state. 2.5 µg of Tdm and 5 mM TMAO were used for HCHO assay as described above.

Inhibition by dimethylethylamine (DMEA) was studied using assay solutions containing 1-4 mM TMAO and 0-4 mM DMEA. Inhibition constants (K_i) were determined on the basis of uncompetitive inhibition using **Equation 4-1** (Berg 2015),

$$V = \frac{V_{max}[S]}{K_m + [S](1 + \frac{[I]}{K_i})} \quad (4\text{-}1)$$

where V_{max} is the maximal activity, K_i is inhibition constant, [S] and [I] are the concentrations of substrate and inhibitor, respectively.

4.2.3 Fe^{2+} reactivation

6*His-tag-cleaved Tdm was diluted in buffer containing 20 mM Tris-HCl (pH 7.8) and 100 mM NaCl to a final concentration of 1 mg ml^{-1}, which was then incubated with 100 µM H_2O_2 for 20 min on ice. Wherever necessary, ascorbic acid and $Fe(NH_4)_2(SO_4)_2$ were added to a final concentration of 500 µM and 10–500 µM, respectively. The samples were then incubated for 20 min on ice prior to the quantification of Tdm activity using the aforementioned HCHO assay (**Chapter 2**).

4.2.4 Fe^{2+} titration

The Fe^{2+} titration, using $Fe(NH_4)_2(SO_4)_2$ was performed with the 6*His-tag-cleaved Tdm. To maintain Fe^{2+} in its reduced form, the titration assay was performed with 2 mM ascorbic acid with varying molar ratios of Fe^{2+}:Tdm in the Tris-NaCl buffer (20 mM Tris, 100 mM NaCl, pH 7.8) for 4 hrs before HCHO essay was carried out. The titration data for metal stoichiometry (n) determination were analysed by nonlinear curve fitting using **Equation 4-2** (Chai et al 2009, Drake and Klakamp 2007).

$$y = y_0 + (y_m - y_o) * \frac{(n+x+K_d) - \sqrt{(n+x+K_d)^2 - 4*n*x}}{2*n} \quad (4\text{-}2)$$

where x is the concentration of total metal ion and y is the activity percentage to the maximal activity, y_m is the maximal activity, y_0 is the activity percentage of as-isolated enzyme without added metal, K_d is the dissociation constant, and n is the number of binding sites (stoichiometry).

4.2.5 Inductively coupled plasma-mass spectrometry and optical emission spectrometry (ICP-MS/OES)

3% (v/v) trace metal grade nitric acid purified in house by sub-boiling point distillation was used as the sample matrix. ICP-MS analyses were carried out on an Agilent Technologies 7500 ICP-MS instrument. ICP-OES was performed on a Perkin Elmer Optical Emission Spectrometer Optima 5300DV instrument. The standards for calibration were freshly prepared by diluting Zn, Fe, S, Ni stock solutions (at 1000 mg l^{-1}, Sigma-Aldrich) with 3% (v/v) HNO_3 in doubly-deionized water with concentrations from 0.2–1 mg l^{-1} for Zn, Fe and Ni, and 4–20 mg l^{-1} for S. About 2.4 mg protein was diluted in 3% HNO_3 matrix for metal analysis. The content of S was quantified in order to determine the protein concentration. The contents of Zn, Fe, Ni and S were measured using the emission lines of 213.857 nm (Zn), 234.830 nm (Fe), 231.604 nm (Ni) and 180.669 nm (S), respectively.

4.2.6 Homology modelling

A homology model for the DUF1989 domain of Tdm (residues 1 to 383) was created using the DUF1989 domain family protein (PDB: 3ORU) as the template using the SWISS-MODEL webserver (http://swissmodel.expasy.org) (Arnold et al 2006). 3-D structures were visualized using Chimera (Pettersen et al 2004).

4.2.7 Secondary and quaternary structure determination

Secondary and quaternary structure was determined by CD and size-exclusion chromatography respectively as described in **Chapter 2**.

4.2.8 Gas chromatography–mass spectrometry (GC-MS) determination of secondary amines

Fifty μg purified Tdm was incubated at room temperature (~20 °C) in a final volume of 1 ml in 10 mM MES buffer (pH 6.0) with either 50 mM TMAO or 25 mM dimethylethylamine (DMEA), or both for 30 min. Secondary amines (*i.e.* DMA, MEA) were derivatized using benzenesulfonyl chloride (BSC) (Sigma-Aldrich) and determined by GC-MS as described previously (Zhang et al 2012). The mixture was basified with 2 mL of 10 mol l^{-1} aqueous sodium hydroxide solution and 0.2 mL BSC were added and left for 30 min at room temperature with occasional gentle shaking. The residue derivatives were hydrolysed at 80°C for 20 min. Subsequently the solution was cooled down to room temperature and acidified with 18.5% aqueous solution of hydrochloric acid to pH 5.5. The mixture was extracted with 1 ml dichloromethane. The organic phase was used for the analysis of secondary amines as benzenesulfonyl derivatives by GC-MS. An Agilent 6890/5973 GC-MS platform equipped with an automatic liquid sampler was used for amine analysis. An aliquot (1 μl) of the dichloromethane layer was injected into GC-MS. GC conditions were as follows: column, Agilent HP-5ms capillary column (30 m × 0.25 mm i.d.; film thickness, 0.25 μm); column temperature, 100°C for 3 min, then the temperature was increased at 5°C min^{-1} rate up to 180°C, after that 10°C min^{-1} rate up to 250°C, followed by 5 min at 290°C; carrier gas, helium; flow rate, 1 ml min^{-1}; split ratio, 100:1.

4.2.9 Isotope labelling studies with $TMA^{18}O$

$TMA^{18}O$ was synthesized by oxidizing TMA by $H_2{}^{18}O_2$ (Dustan et al 1899). 100 μg TMA solution in water (≥99%, Sigma-Aldrich) was added to 250 μg of 2-3% (v/v) $H_2{}^{18}O_2$ (Sigma-Aldrich), which was then left overnight at room temperature.

Excessive TMA was evaporated by heating up to 90˚C for 15 min. The fish odour of TMA gradually disappeared. $H_2{}^{18}O$ was removed by freeze-drying. The white crystal that remained was dissolved in 1 ml distilled water. Purity of synthesized $TMA^{18}O$ was then assessed by direct infusion to mass spectrometer without any previous separation. An ion trap mass spectrometry (Thermo Scientific, Waltham, MA, USA) equipped with an electrospray ionization (ESI) source in the positive ion mode was used for analysis. ESI settings were as follows: capillary temperature 200 ºC, capillary voltage 4 kV. $TMA^{18}O$ concentration was quantified by ion chromatography as described in **Chapter 2**.

HCHO adduction and detection were carried out according to (Jiang et al 2013) with the following modification. HCHO was adducted by sodium bisulfite ($NaSO_3$) (Sigma-Aldrich) to form hydroxymethanesulfonate (HMS^-) $CH_2(OH)SO_3^-$. The reaction mixture for HCHO activity assays contained 10 mM $TMA^{18}O$ or TMAO, 10 mM $NaSO_3$ in 1ml of 10 mM MES buffer, pH 6.0. The reaction was initiated by adding 10 µg of purified recombinant Tdm and incubated for 20 min at room temperature.

A high performance liquid chromatography LC System (Waters, Milford, MA, USA) equipped with a Xbridge BEH amide analytical column (150 × 3.0 mm × 2.5 µm; Warters, Irland) was used for separation. The column was thermostated at 30 °C, and a flow rate of 0.30 ml min^{-1} was employed. The following 20 min gradient program was used: 5 min hold at 90% mobile phase A (100% acetonitrile) and 10% mobile phase B (10 mM ammonium acetate in H_2O, pH 4.5), gradient to 50% A in 10 min, 5 min hold at 50% A, All samples and standards utilized a 5 µL injection volume.

Aforementioned ESI-MS in the negative ion mode was used for analysis. ESI settings were as follows: capillary temperature 200 ºC, capillary voltage 2 kV. The

deprotonated molecular ions $[M-H]^-$ with *m/z*=111 and 113 were identified as the unique peak for HMS and ^{18}O-HMS, respectively. Fragmentation of HMS and ^{18}O-HMS form product *m/z*=81 (HSO_3^-) after dissociation of $-CH_2O$=30 and $CH_2{}^{18}O$=32.

4.2.10 DMPO spin trapping

The spin trap 5,5-dimethyl- pyrroline-*N*-oxide (DMPO) (for EPR-spectroscopy) was obtained from Sigma-Aldrich. To trap radicals formed during Tdm catalysis, the reaction solutions contained 100 mM DMPO, 100 mM TMAO, 60 μM Tdm which was prepared in 20 mM Tris-HCl, pH 7.8 and 100 mM NaCl at room temperature. Controls were prepared without Tdm or TMAO.

All EPR spectra were recorded at ambient temperature (ca. 291 K) on a Bruker EMX (X-band) spectrometer. 1.0 mm quartz tubes with inner diameter (I.D.) of 1.0 mm and outer diameter of 1.2 mm (Wilmad Labglass) were used. Typical key EPR spectrometer settings were modulation amplitude 2.0 G and microwave power 0.63 mW, 2 mT modulation amplitude, 1.0 $\times 10^5$ receiver gain, sweep gain 100 s with repeated number of 8 X-scans.

4.2.11 Multiple sequence alignments

Multiple sequence alignments were performed using the iterative alignment program MUSCLE (Edgar 2004).

4.2.12 Statistical analyses

Analysis of variance (ANOVA) and Tukey HSD post-hoc tests were performed using the R software package version 3.2. (2015). Data are expressed as means ± standard deviations.

4.3 Results

4.3.1 Tdm is a novel zinc-iron dependent protein

Early studies in the 1970s have suggested that bacterial Tdm is a metal-dependent enzyme. The *Bacillus* Tdm is strongly stimulated by ferrous iron and reducing agents such as ascorbate and glutathione (Myers and Zatman 1971). In agreement with these previous studies, inhibition of Tdm activity was observed when the purified enzyme was incubated with the metal chelator EDTA (**Figure 4-3**). To characterize the metal ion(s) in Tdm, I carried out ICP-MS metal scan analyses of purified recombinant Tdm of *Methylocella silvestris*, which detected the presence of Zn, Fe and Ni above background levels (data not shown).

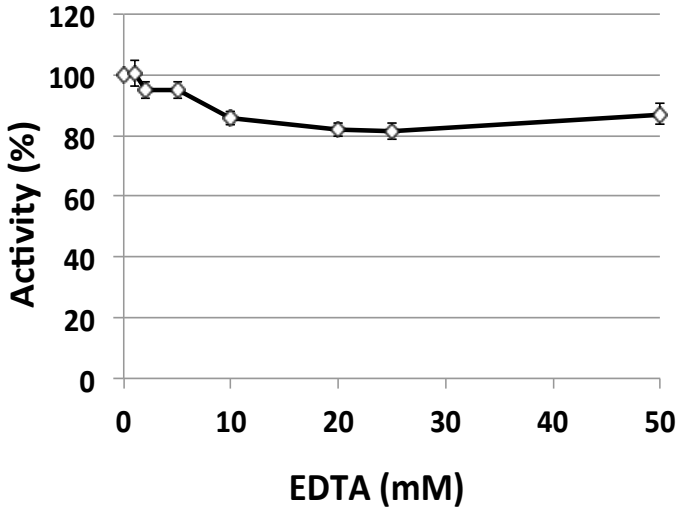


Figure 4-3 Tdm activity is inhibited by EDTA. Activities were determined by quantifying HCHO production using 5 mM TMAO. Activity%: The activity percentage of Tdm with EDTA to the one without EDTA. Error bars indicate standard deviations of experiments run in triplicate.

To obtain a more accurate estimation of metal contents in Tdm, Zn, Fe, Ni as well as S (for accurate determination of protein concentrations) were quantified by ICP-OES. The results showed that 1 monomer of Tdm contained 0.97 ± 0.03 Zn^{2+} and 0.35 ±

0.02 iron in the as-isolated Tdm (CK-Tdm, **Table 4-1**). Trace amounts of Ni^{2+} (<0.1 molar equivalents per Tdm monomer) were also found.

To address whether Zn or Fe could be replaced by each other, Fe and Zn enriched Tdm (Fe-Tdm, Zn-Tdm) were purified from *E. coli* cultivated in media supplemented with either $Fe(NH_4)_2(SO_4)_2$ or $ZnCl_2$ (0.5 mM final concentration). Tdm expressed in Fe-supplemented media showed slightly yet significantly higher Fe^{2+} content (0.38 ± 0.02 mol per mol monomer) (p<0.05), whereas Tdm expressed in Zn-supplemented media had reduced Fe^{2+} content (0.12 ± 0.01 mol per mol monomer) in purified Tdm, in coincidence with an increase of Zn^{2+} content (1.33 ± 0.03 mol per mol monomer) (p<0.05), suggesting a replacement of Fe by Zn in purified Tdm. Additionally, Fe-enriched Tdm had a higher catalytic activity compared to that of CK-Tdm or Zn-enriched Tdm (**Table 4-1**), suggesting a role of Fe in catalysis.

Table 4-1 Steady-state kinetics parameters of Zn and Fe-enriched Tdm

		CK-Tdm	Fe-Tdm	Zn-Tdm
K_m (mM)		3.88 ± 0.19[a]	4.09 ± 0.21[a]	4.29 ± 0.12[a]
V_{max} (nmol min^{-1} mg^{-1})		14.61 ± 1.13[a]	16.99 ± 0.70[b]	12.41 ± 0.52[c]
k_{cat} (s^{-1})		19.49 ± 1.51[a]	22.65 ± 0.94[b]	17.30 ± 1.37[c]
k_{cat}/K_m $M^{-1}s^{-1}*10^3$		5.03 ± 0.34[a]	5.54 ± 0.36[a]	4.02 ± 0.22[b]
Metal equiv	Zn^{2+}	0.97 ± 0.03[a]	0.92 ± 0.04[a]	1.33 ± 0.03[b]
	Fe^{2+}	0.33 ± 0.02[a]	0.38 ± 0.02[b]	0.12 ± 0.01[c]

CK-Tdm: Tdm purified recombinant *E. coli* cultivated using the LB medium. Fe-Tdm: The culture medium was supplemented with $Fe(NH_4)_2(SO_4)_2$ at a final concentration of 0.5 mM. Zn-Tdm: The culture medium was supplemented with $ZnCl_2$ at a final concentration of 0.5 mM. Values are means ± standard deviations. Different superscript letters in the same row between samples denote significant differences between different metal enrichment (p< 0.05).

To further probe the iron species in as-isolated Tdm, activity assays were performed with the addition of a reducing agent (ascorbic acid, Asc) or an oxidizing agent

(hydrogen peroxide, H_2O_2) at varying concentrations. The results demonstrated Asc did not show any inhibition until a final concentration of 8 mM (**Figure 4-4**) whilst H_2O_2 inhibited Tdm activity effectively. More than 80% activity was lost upon incubation with 100 μM H_2O_2 (**Figure 4-5 A)**. Furthermore, when isolated Tdm was incubated with various divalent metal ions, it was observed that Fe^{2+} significantly enhanced Tdm activity, particularly in the presence of Asc (**Figure 4-5 B**). Interestingly, the activity of H_2O_2-pretreated Tdm can at least be partially restored by incubating with Asc and Fe^{2+} (**Figure 4-5 C**). Taken together, the data suggest a role of ferrous iron in the native as-isolated Tdm during catalysis.

Due to the traditional protein overexpression and purification procedures, oxygen sensitive ferrous iron is prone to loss (Gantt et al 2006, Zhu et al 2003). The ICP-OES analyses of as-isolated Tdm may have underestimated iron contents in this protein (**Table 4-1**). To determine Fe^{2+} stoichiometry, a Fe^{2+} titration experiment was carried out. To eliminate the non-specific binding of metal contamination to the 6*His-tag, 6*His-tag-free Tdm was used. The data presented in **Figure 4-5 D** gave a stoichiometry number of $n = 0.91$. Together, the data suggest that Tdm is a Zn^{2+} and Fe^{2+} dependent protein with Zn^{2+}:Fe^{2+}:Tdm monomer ratios of 1:1:1.

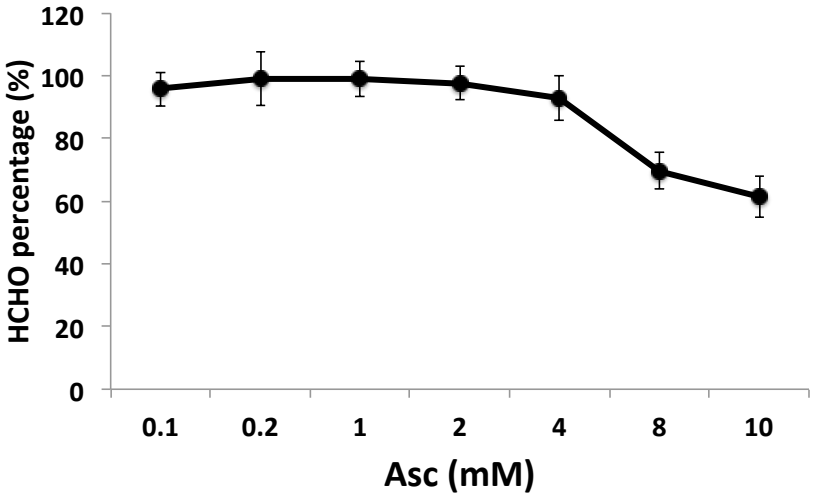


Figure 4-4 The effect of ascorbic acid on Tdm activity. Activity%: The activity percentage of Tdm with ascorbic acid to the one without ascorbic acid. The error bars represent standard deviation from experiments run in triplicate

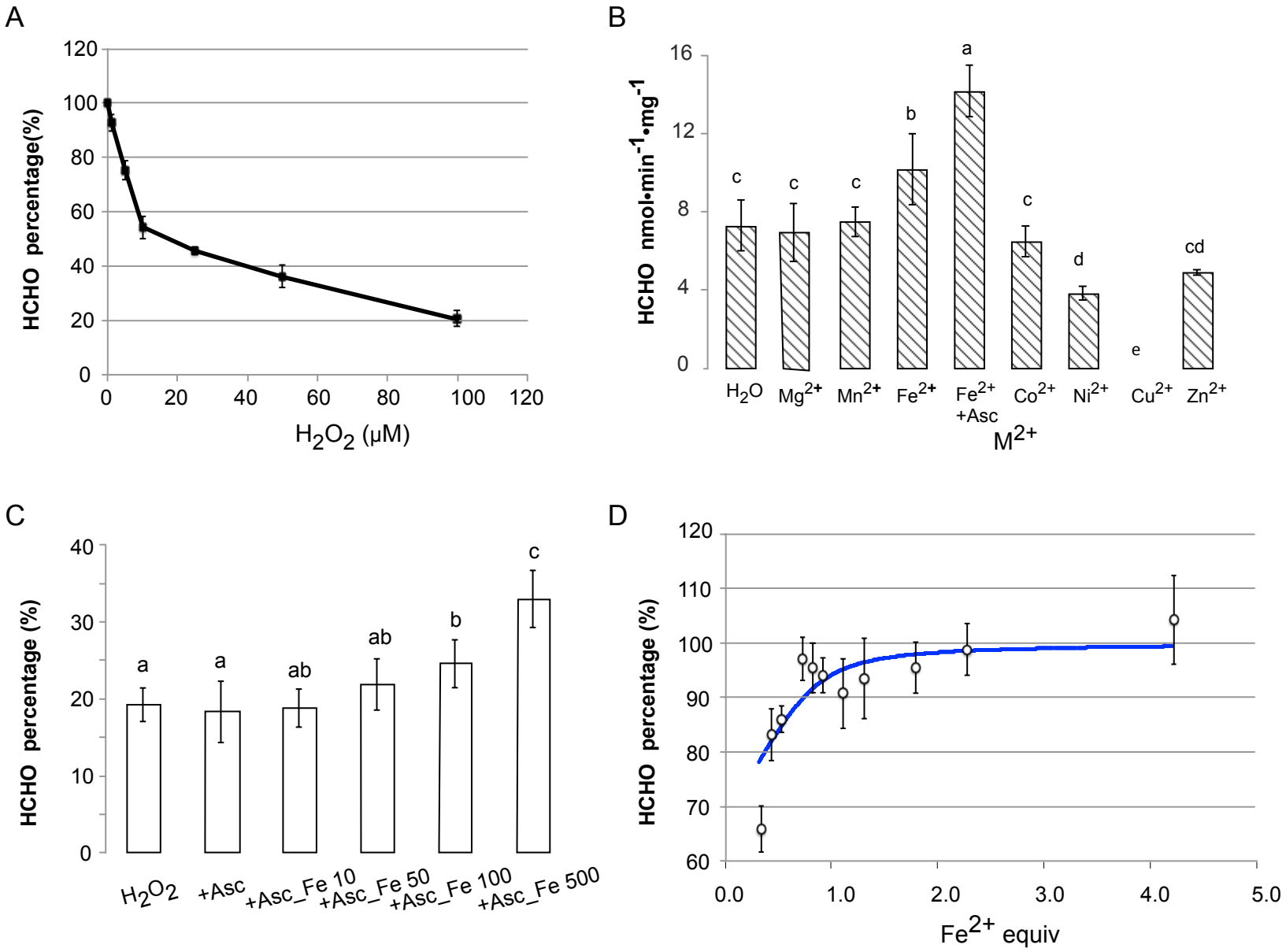

A
HCHO percentage(%)
120
100
80
60
40
20
0
0
20
40
60
80
100
120
H2O2 (μM)
B
HCHO nmol•min-1•mg-1
16
12
8
4
0
a
b
c
c
c
c
d
e
cd
H2O
Mg2+
Mn2+
Fe2+
Fe2+ +Asc
Co2+
Ni2+
Cu2+
Zn2+
M2+
C
HCHO percentage (%)
40
30
20
10
0
a
a
ab
ab
b
c
H2O2
+Asc
+Asc_Fe 10
+Asc_Fe 50
+Asc_Fe 100
+Asc_Fe 500
D
HCHO percentage (%)
120
110
100
90
80
70
60
0.0
1.0
2.0
3.0
4.0
5.0
Fe2+ equiv

Figure 4-5 (Previous page) Fe^{2+} is a native cofactor for Tdm. The addition of H_2O_2 significantly affects Tdm activity (A). Tdm activity was measured by quantifying the formation of HCHO. The effects of metal ion additions on Tdm activity (B). Different letters denote significant differences between different M^{2+} metal ions reconstitution ($p < 0.05$). Mg^{2+}: $MgCl_2$. Mn^{2+}: $MnCl_2$. Fe^{2+}: $Fe(NH_4)_2(SO_4)_2$. Co^{2+}: $CoCl_2$. Ni^{2+}: $NiSO_4$. Cu^{2+}: $CuSO_4$. Zn^{2+}: $ZnSO_4$. The loss of Tdm activity by the addition of H_2O_2 can be partially restored by adding reducing agents in the enzyme assay (C). Tdm was pretreated with 100 μM H_2O_2 for 15 min. Inactivated Tdm was then reactivated by adding 500 μM Asc and Fe^{2+} at varying concentration for 20 min before enzyme activity assay was carried out. H_2O_2: Tdm pretreated with 100 μM H_2O_2; +Asc: H_2O_2-pretreated Tdm incubated with 500 μM ascorbic acid; +Asc_Fe 10, +Asc_Fe 50, +Asc_Fe 100, +Asc_Fe 500: H_2O_2-pretreated Tdm incubated with 500 μM ascorbic acid and 10, 50, 100 or 500 μM of $Fe(NH_4)_2(SO_4)_2$ respectively. Different letters denote significant differences between different metal reconstitution ($p< 0.05$). Titration of Tdm with increasing concentrations of Fe^{2+} and the activity was measured by quantifying HCHO formation (D). Titration data were plotted using a non-linear fitting using the multiple independent binding sites model (Chai et al 2009, Drake and Klakamp 2007), giving $n = 0.91$ per Tdm monomer. The error bars represent standard deviation from experiments run in triplicate.

4.3.2 Three cysteine residues (C263, C279, C343) contribute to Zn^{2+} coordination in Tdm

A 3D structure for any Tdm protein is yet to be determined. Tdm is composed of two domains (see **Chapter 3**). The C-terminal GCV_T domain is best characterized in the T protein of the glycine cleavage complex, which is required for glycine catabolism (Okamura-Ikeda et al 2005) and is not known to contain a metal cofactor. It is therefore postulated that metal-binding sites for Zn^{2+} and Fe^{2+} are likely located in the N-terminal uncharacterized DUF1989 domain.

The PDB database contains structures of 3 members of the DUF1989 protein family (3SIY, 3ORU, 3DI4) from genus *Rugeria* (previously known as *Silicibacter*), which all contain Zn^{2+}. To predict the Zn^{2+}-binding sites in Tdm, homology modelling was therefore applied. The SWISS-MODEL template library was searched with Blast and HHBlits (HMM-HMM–based lightning-fast iterative sequence search). Both algorithms indicated that the N-terminal domain of Tdm (residues 128 to 352) gave the highest sequence identity (34%) to the sequence of 3ORU (resolution of 1.1 Å), which was therefore chosen as the reference structure for modelling.

The established model predicted a conserved Zn^{2+} coordination motif in Tdm, despite its poor global and per-residue model quality assessed using the QMEAN scoring function (Qualitative Model Energy Analysis) (Benkert et al 2011). In agreement with the existing structures of DUF1989 family proteins in the PDB database, homology modelling predicted that Zn^{2+} was coordinated by three cysteine residues in Tdm (Cys263, Cys279, Cys343) with the thiol S-Zn^{2+} distance around 2.3 Å (**Figure 4-6 A**). Multiple sequence alignment of Tdm proteins from a range of microbes of terrestrial and oceanic origins covering α-, β-, γ-proteobacteria revealed strict conservation of this 3-Cys Zn^{2+} binding motif in Tdm proteins (**Figure 4-10**).

To probe the role of these residues in Tdm activity, site directed mutants were constructed. These mutant proteins were purified (**Figure 4-6 B**) and their activities and metal contents were characterized. Tdm C343A showed 2 contaminants around 42 and 22 kDa. To remove the contaminant protein, I overloaded the column to compete away the other proteins. After optimisation, the purified recombinant Tdm C343A was >95% pure shown by SDS-PAGE (lane 5, **Figure 4-6 B**). The results showed that all three single mutants were inactive, along with significantly reduced Zn^{2+} and Fe^{2+} contents (p<0.05) (**Table 4-2**). Although CD spectroscopy revealed only very minor changes in overall secondary structure of these three mutants (**Figure 4-6 C**), the native homohexamer, which dominates in wild-type Tdm, was virtually absent in the mutants (**Figure 4-6 D**). Together, the results suggested that C263, C279 and C343 were crucial for maintaining structural integrity in native Tdm.

Table 4-2 Activity and metal quantification of wild type Tdm and mutants

	Oligomeric status	Activity	M^{2+}:Tdm (monomer) ratio	
		(nmol min^{-1} mg^{-1})	Zn^{2+}	Fe^{2+}
WT	Hexamer	14.61±1.13	0.97 ± 0.03 [a]	0.33 ± 0.02[a]
C263A	Monomer	0	0.44 ± 0.01[g]	0.04 ± 0.05[de]
C279A	Monomer	0	0.26 ± 0.14[h]	0.03 ± 0.01[e]
C343A	Aggregate/monomer	0	0.23 ± 0.09[h]	0.04 ± 0.03[de]
D198A	Unknown	0	0.66 ± 0.05[e]	0.12 ± 0.01[c]
D198N	Aggregate/hexamer	0	0.67 ± 0.02[e]	0.18 ± 0.02[b]
D198E	Aggregate/hexamer	0	0.53 ± 0.01[f]	0.33 ± 0.01[a]
H256A	Monomer	0	0.72 ± 0.01[d]	0.04 ± 0.01[e]
H276A	Hexamer	0	0.91 ± 0.01[b]	0.09 ± 0.02[d]
F259A	Hexamer	0	0.87 ± 0.03[c]	0.05 ± 0.00[e]
W321A	Hexamer	0	0.83 ± 0.03[c]	0.11 ± 0.00[c]
Y305A	Hexamer	0	0.94 ± 0.03[a]	0.15 ± 0.02[b]

Values are means ± standard deviations. Different superscript letters in the same column between samples denote significant differences between different metal reconstitution (p< 0.05).

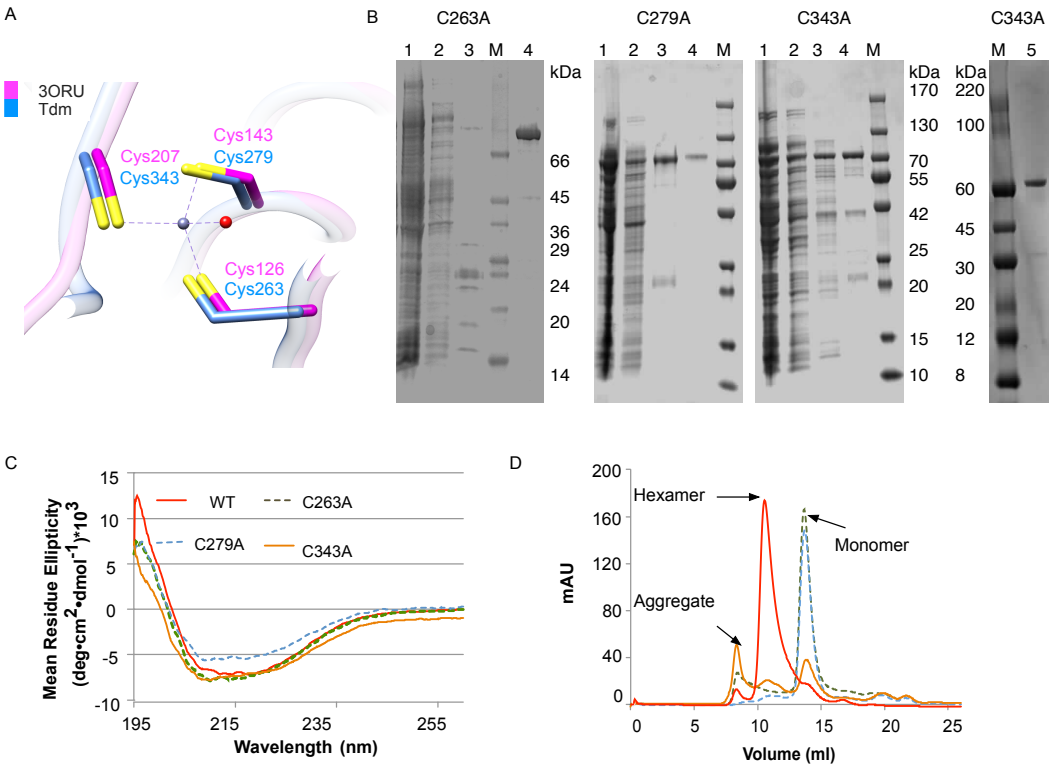


Figure 4-6 Determination of Zn^{2+} binding site in Tdm. Homology modelling of the DUF1989 domain in Tdm (blue) suggests that C263, C279 and C343 are potential Zn^{2+} binding sites (A). SDS-PAGE of Tdm variants (B). 1, crude cell-free extract; 2, column wash fraction with binding buffer containing 20 mM imidazole; 3, column wash fraction with washing buffer containing 60 mM imidazole; 4, elution fraction of the purified protein; 5, elution fraction of Tdm C343A after optimisation of protein purification. Far-UV CD spectra of wild-type Tdm (WT) and site-directed mutants (C263A, C279A, C343A) (C). Oligomeric states of variants (D). Red: WT; dark green: C263A; blue: C279A; orange: C343A. The native Tdm exists predominantly as a homohexamer (**Chapter 3**).

4.3.3 H276 is a potential Fe^{2+} binding ligand in Tdm

Due to the lack of a crystal structure of Tdm and the absence of Fe^{2+} in existing crystal structures of DUF1989 family proteins, I performed multiple sequence alignment of Tdm sequences in order to gain insight into conserved residues, which may shed light on the residues involved in Fe^{2+} coordination (**Figure 4-10 A**). A 2-His-1-carboxylate facial triad is a common motif in a number of non-haem Fe^{2+}-containing enzymes, where two histidine residues and one carboxylate-containing side chain are arranged at one face of an octahedron whereas the opposite face of the octahedron is available to coordinate a variety of exogenous ligands (Hegg and Que 1997, Koehntop et al 2005). The sequence alignment revealed the presence of two strictly conserved histidines (His256, His276) in all Tdm analysed (**Figure 4-10 A**).

In order to test whether H256 and H276 were indeed Fe^{2+} binding sites, point mutants were constructed, and the mutants were subsequently purified (**Figure 4-7**) and characterized. Enzyme assays reveal that the two mutants (H256A, H276A) completely lost activity (**Table 4-2**). Metal analysis showed that the H276A mutant indeed had lost Fe^{2+} but its Zn^{2+} content remained unchanged compared to that of the wild type Tdm. Furthermore, the loss of Fe^{2+} in the H276A mutant cannot be attributed to a structural alteration caused by site-directed mutagenesis since its secondary and quaternary structure is comparable to that of the wild type (**Figure 4-7 B, C**).

The H256A mutant was also inactive and had almost completely lost its Fe^{2+}. It also had a significantly lower Zn^{2+} content (Zn^{2+}:Tdm monomer ratio was 0.72 ± 0.01, compared to 0.97 ± 0.03 for wild type) (**Table 4-2**). Although its secondary structure remained largely unchanged as revealed by CD spectroscopy, this mutant did not form native hexamer (**Figure 4-7 B, C**). Therefore, it is concluded that H276 is a

potential Fe^{2+}-binding site whilst H256 plays a role in maintaining overall structure and may also be a Fe^{2+}-binding ligand.

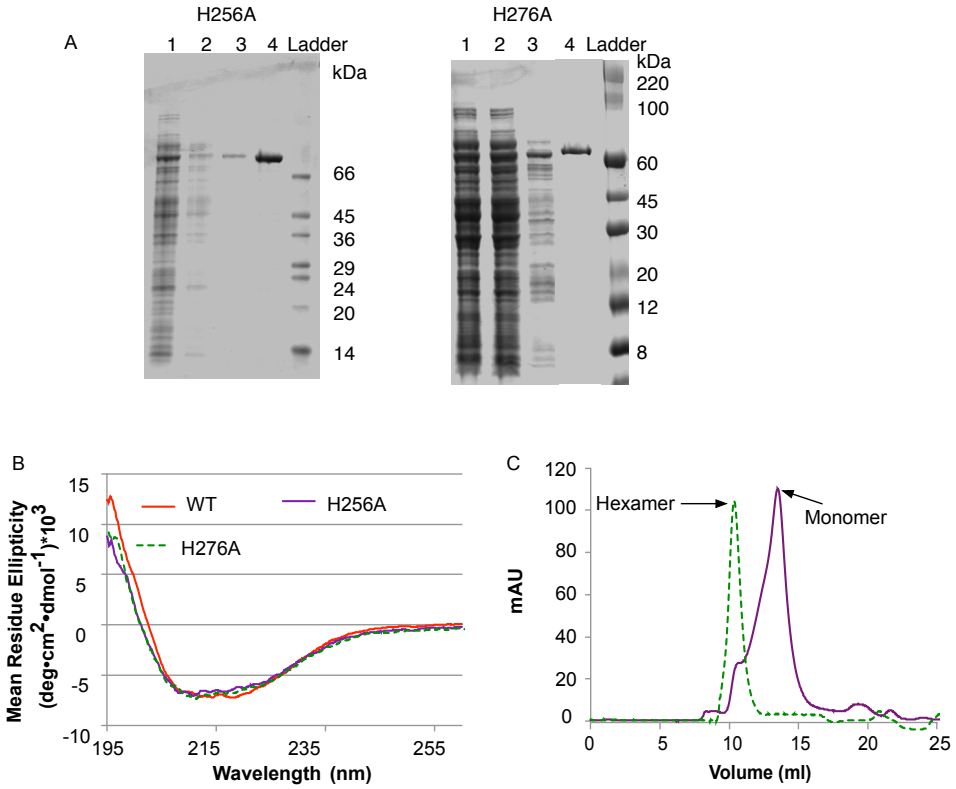


Figure 4-7 H276 is a potential Fe^{2+}-binding site. SDS-PAGE of Tdm H256A and H276A (A). 1, crude cell-free extract; 2, column wash fraction with binding buffer containing 20 mM imidazole; 3, column wash fraction with washing buffer containing 60 mM imidazole; 4, elution fraction of the purified protein. Far-UV CD spectra of wild type Tdm (WT) and variants (H256A, H276A) (B). Oligomeric states of variants (C). purple: H256A, blue: H276A.

4.3.4 The "bridging" nature of D198 in Tdm

When Zn^{2+} is involved in maintaining protein structure, it is commonly coordinated by four protein ligands in the order of Cys>His>Asp/Glu in a tetrahedral geometry (Auld 2001). Intriguingly, the three existing crystal structures of DUF1989 family proteins (*i.e.* 3ORU) employ a 3-Cys-OH_2 Zn^{2+} binding motif with the fourth ligand being a water molecule in the crystal structure (**Figure 4-8 A)**. Searching the surrounding zone (<2.5 Å) (Harding 2001) of the Zn^{2+} ion did not reveal any other potential binding ligand.

Although cysteines are commonly found in structural Zn sites, the 3-Cys-OH_2 tetrahedral coordination is found in the catalytic sites of several enzymes, *e.g.* cytidine deaminase (CDA) from *Mycobacterium tuberculosis* (3IJF) (Sánchez-Quitian et al 2010), human (1MQ0) (Chung et al 2005) and mouse (1ZAB) (Teh et al 2006), and CDA-related enzyme, Blasticidin S deaminase from *Aspergillus terreus* (2Z3G, 1WN6) (Kumasaka et al 2007). In these structures, the catalytic H_2O forms H-bonding to the side chain oxygen of Glu and main chain nitrogen of Cys in the vicinity to maintain a tetrahedral arrangement. Zinc ion appears to act as electrophilic catalyst through an accompanying water molecule. Structure superimposition revealed that a 3-Cys-OH_2 motif of 3ORU is similar to CDA (**Figure 4-8 A**). Asp66 of 3ORU is in good agreement with the conserved Glu in CDA, *i.e.* Glu58 of 3IJF and Glu56 of 2Z3G (**Figure 4-8 A**), indicating that D66 may also H-bond to catalytic H_2O.

Multiple sequence alignment of various Tdm proteins and DUF1989 family proteins demonstrated strict conservation of this aspartate residue (corresponding to D198) in Tdm (**Figure 4-10 A**). To investigate whether D198 is indeed important in TMAO demethylation by Tdm, three site-directed mutants were made, D198A (no oxygen atom), D198N (one oxygen atom, neutral side chain) and D198E (two oxygen atoms,

negatively charged). Variants were purified by Ni-NTA and purities were examined by SDS-PAGE (**Figure 4-8 B**). All three mutants were inactive, and their Zn^{2+} contents (Zn^{2+}:Tdm ratios between 0.5-0.7) were significantly lower than that of the WT (0.97 ± 0.03) (p<0.05) (**Table 4-2**). Although their overall secondary structure remained largely unchanged, alteration of quaternary structure was observed and the mutants were prone to aggregation (**Figure 4-8 C, D**).

Incidentally, the Fe^{2+} content of D198 mutants varied too, with a trend of D198A (0.12 ± 0.01) < D198N (0.18 ± 0.02) < D198E (0.33 ± 0.01). The D198E mutant retained its binding capacity to Fe^{2+}, similar to that of the wild type enzyme (0.33 ± 0.02) and a large proportion of this mutant still retained hexameric state as observed in the wild type Tdm (**Table 4-2**). Therefore, the data suggest that D198 also likely provides a carboxyl group for Fe^{2+} binding.

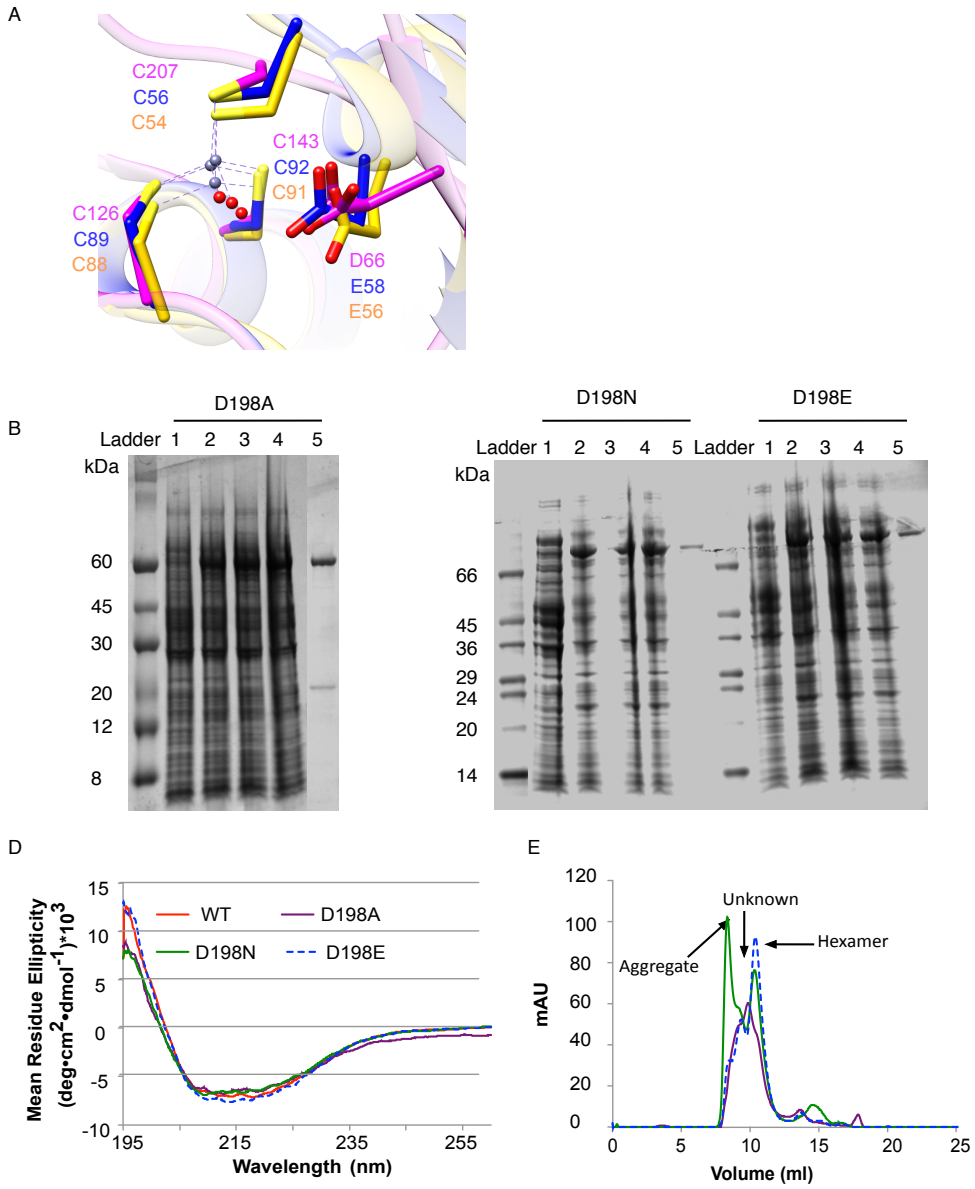


Figure 4-8 A unique "bridging" aspartate residue D198. Superimposition of Zn^{2+}-binding motif of 3ORU (magenta) with catalytic Zn^{2+}-site of CDA of *Mycobacterium tuberculosis* (3IJF, blue) and *Aspergillus terreus* (2Z3G, golden) (A). Zn^{2+} ion is shown as grey sphere. Water is shown as red sphere. SDS-PAGE of Tdm D198A, D198N and D198E (B). 1, uninduced crude cell-free extract; 2, cell-free extract induced by 0.2 mM IPTG; 3, cell-free extract induced by 0.5 mM IPTG; 4, cell-free extract induced by 1 mM IPTG; 5, purified recombinant protein. Far-UV CD spectra of WT and variants (D198A, D198N, D198E) (C). Oligomeric states of variants (D). purple: D198A, green: D198N, blue: D198E.

4.3.5 The substrate-binding pocket

TMAO-binding pockets have been studied previously in two enzymes, namely the substrate-binding protein (TmoX) of the TMAO ABC transporter and TMAO reductase (TorT) (Li et al 2015, Moore and Hendrickson 2012). In both proteins, a hydrophobic substrate-binding pocket, composed of three to four aromatic residues, was found to recognize and bind TMAO via cation-π interaction (Gallivan and Dougherty 1999). Sequence alignment of Tdm proteins indeed revealed the presence of several conserved phenylalanine (Phe), tyrosine (Tyr) and tryptophan (Trp) residues (**Figure 4-10 A**). In an attempt to identify TMAO-binding sites in Tdm, site-directed mutants were generated by individually replacing these Phe, Tyr and Trp residues with Ala. Variants were purified by Ni-NTA and purities were examined by SDS-PAGE (**Figure 4-9 A**). Three of these mutants, F259A, Y305A and W321A, have completely lost activity (**Table 4-2**), while the overall secondary and structure was retained (**Figure 4-9 B, C**), suggesting a role of these aromatic residues in substrate binding. The remainder of the mutants (Y185A, Y237A, Y267A, Y273A, W298A, W327A, Y363A) were, however, still active (**Table 4-3**), thus these residues are unlikely to contribute to substrate binding.

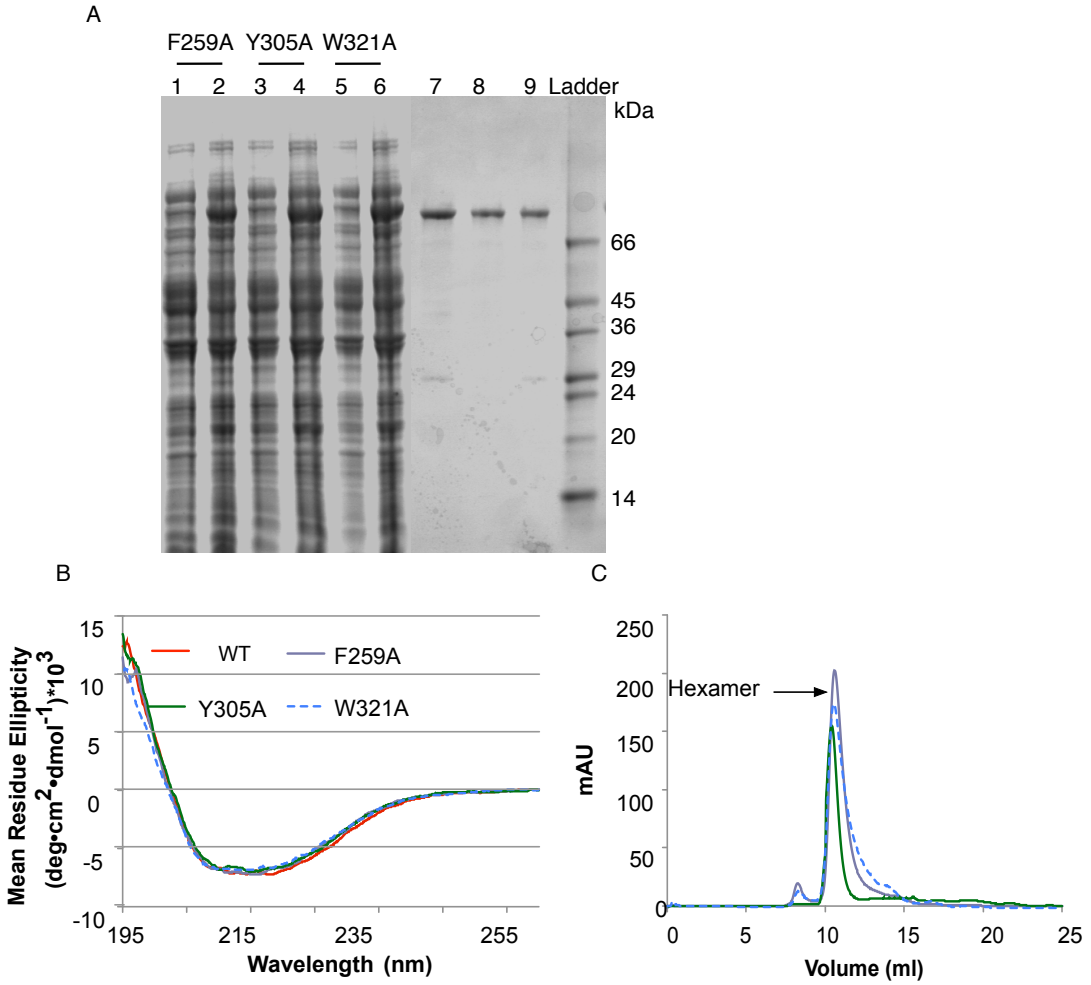


Figure 4-9 F259, Y305 and W321 are potential substrate pocket residues. SDS-PAGE of Tdm F259A, Y305A and W321A (A). 1, 3, 5, uninduced crude cell-free extract; 2, 4, 6, Induced crude cell extract; 7, 8, 9, purified F259A, Y305A, W321A, respectively. Far-UV CD spectra of wild type (WT) Tdm and variants (F259A, Y305A, W321A) (B). Oligomeric states of variants (C). grey: F259A, green: Y305A, blue: W321A

Table 4-3 Site-directed mutants of conserved aromatic residues in Tdm

Mutants	Activity
Y178A	n.d.
Y185A	+
Y237A	+
F238A	n.d.
F259A	-
Y267A	-
Y268A	n.d.
Y273A	+
W298A	+
F304A	n.d.
Y305A	-
W321A	-
Y327A	+
W353A	n.d.
Y363A	+

n.d.: mutants were either not constructed or constructed but can not be induced by IPTG, therefore the activities were not determined; -: inactive; +: active.

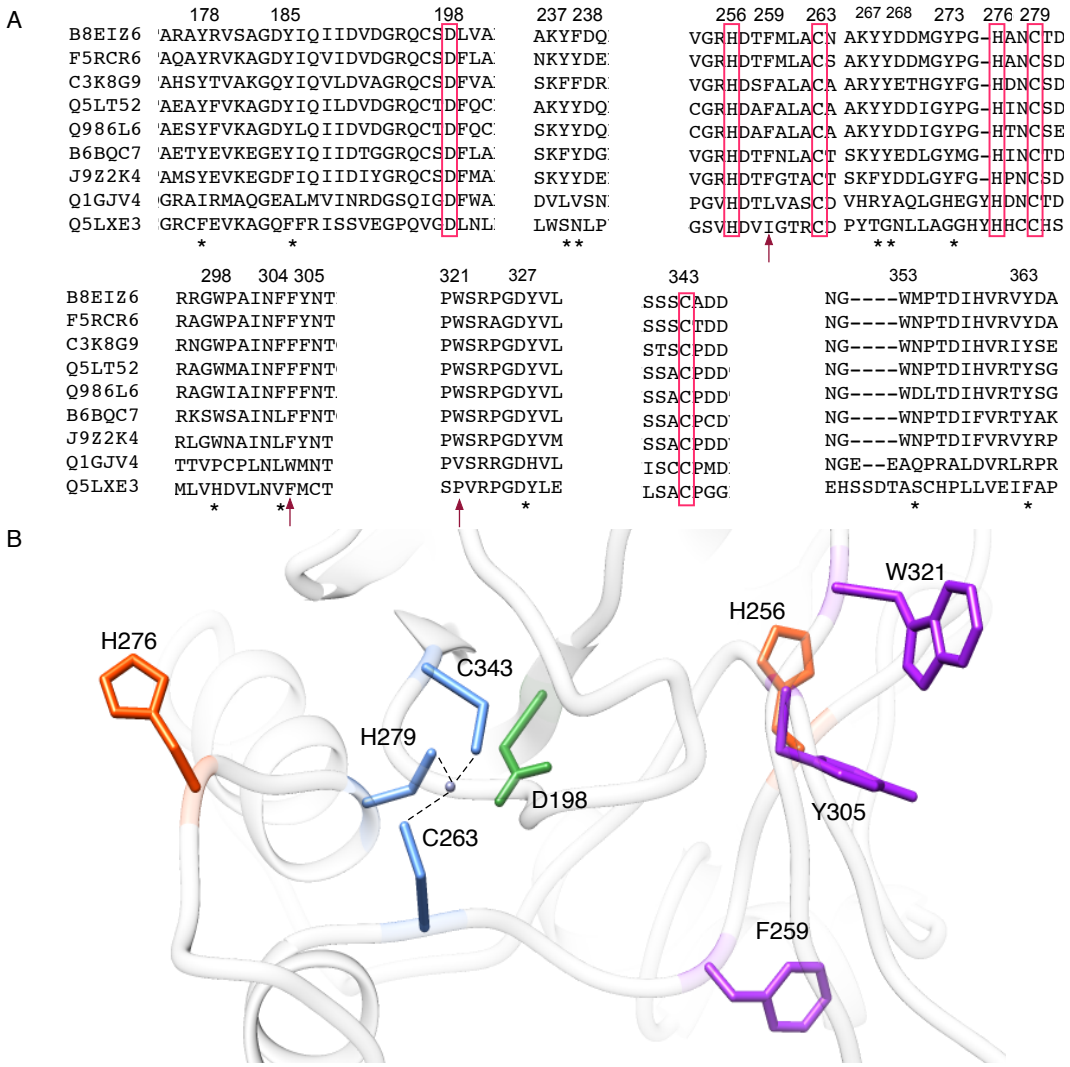


Figure 4-10 Investigation and visualisation of the residues involved in metal coordination and substrate binding. Multiple sequence alignment of bacterial Tdm proteins (Uniprot accession numbers are shown) and two DUF1989 family protein entries whose structures are available from the PDB database (**Q5LXE3**; **Q1GJV4**) (A). **B8EIZ6**: *Methylocella silvestris* BL2; **C3K8G9**: *Pseudomonas fluorescens* SBW25; **Q5LT52**: *Ruegeria pomeroyi* DSS-3; **Q986L6**: *Mesorhizobium loti* MAFF303099; **F5RCR6**: *Methyloversatilis universalis* FAM5; **B6BQC7**: *Candidatus* Pelagibacter sp. HTCC7211; **J9Z2K4**: *alphaproteobacterium* strain HIMB59. The conserved putative Zn^{2+} (C263, C279, C343) and Fe^{2+} (H256 and H276) coordination centres are highlighted in red boxes. Conserved aromatic amino acid residues in Tdm but not **Q5LXE3** or **Q1GJV4** are indicated by asterisk (*). Potential residues contributing to the formation of the substrate pocket (F259, Y305, W321) are indicated by arrows. Visualisation of the identified residues forming metal binding centers and substrate pocket (B). Blue: Zn^{2+} binding sites; orange: potential Fe^{2+} binding site; green: "bridging" D198; purple: putative substrate-binding pocket. Cautions should be taken that, due to the low sequence identity and absence of Fe^{2+} in the template (**3ORU**), the rotamer of H256, H276, F259, Y305 and W321 are not of high confidence according to the homology model.

4.3.6 The catalytic mechanism of TMAO degradation by Tdm

TMAO degradation catalysed by Tdm resembles *N*-dealkylation by other non-haem Fe^{2+}-containing enzymes, such as DNA dealkylase (AlkB) (Mishina and He 2006), histone demethylase (JHDM1, JMJD6) (Chang et al 2007, Tsukada et al 2006) and Rieske-type demethylase (Daughtry et al 2012, Summers et al 2012). A high-valent Fe(IV)-oxo complex is a common active species for attacking the C-H bond of saturated carbon centres by non-haem Fe^{2+}-containing proteins. O_2 is required as oxygen donor for all the aforementioned reactions to form the Fe(IV)-oxo complex. However, a previous study on bacterial Tdm suggested that TMAO demethylation is O_2-independent (Large 1971). The enzyme assays performed using purified Tdm showed no difference in kinetics aerobically and anaerobically, supporting that TMAO demethylation is indeed O_2-independent (**Figure 4-11 A**). The O_2-independency therefore suggests that TMAO functions as both the substrate and the oxygen donor. It is therefore speculated that the substrate TMAO can donate its oxygen atom to Fe^{2+} and act as a surrogate oxygen donor. Although such a mechanism of surrogate oxygen donation has not been observed in non-haem iron enzymes, it has been found in haem-containing P450 enzymes in the absence of NAD(P)H and oxygen (reviewed in ref. (Guengerich et al 1996, Hrycay and Bandiera 2015)). Based on the well-studied P450 enzymes, it is proposed that a high-valent oxidant (*e.g.* Fe(IV)-oxo) is required and a tertiary amine intermediate (*i.e.* TMA) is formed (**Figure 4-11 B**). Firstly, TMAO donates oxygen to activate the Fe^{2+} centre ② . Fe^{2+} shared two electrons with O atom therefore forms high valent Fe(IV)-oxo or Fe(V)-oxo complex ③, which then hydroxylates the *N*-methyl group to give an

intermediate ④ that decomposes in water to form dimethylamine (DMA) and formaldehyde (HCHO). Meanwhile, Tdm returns to its ferrous resting state ①.

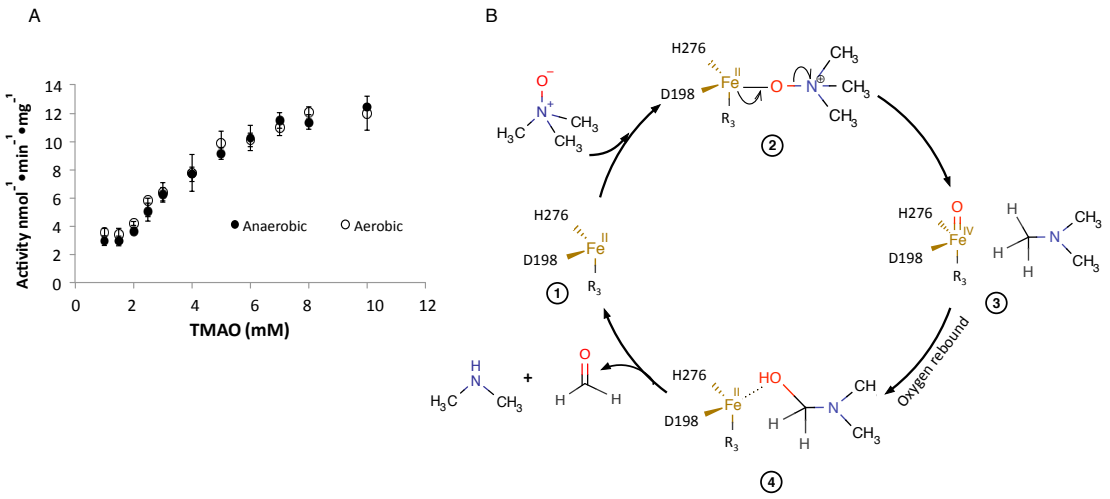


Figure 4-11 Proposed mechanism of Tdm. Tdm catalyzes an O_2-independent demethylation of TMAO (A). The error bars represent standard deviation from experiments run in triplicate. Schematic diagram of the proposed mechanism of Tdm catalytic cycle (B). The substrate is proposed to act as the oxygen donor to activate the Fe centre ② to form a putative high valent iron-oxo intermediate (*e.g.* Fe(IV)-oxo) ③, which then hydroxylates the methyl to give an intermediate ④ that decomposes in water to form DMA and HCHO. Meanwhile, Tdm returns to its ferrous resting state ①.

To test this hypothesis, a crossover experiment was performed using a TMA analogue in order to trap the formation of putative secondary amine species that can be released during the Tdm catalytic cycle (Back and Dyck 1997). DMEA, a structural homologue of the postulated TMA intermediate, was added to the enzyme assays with and without the substrate, TMAO. If the tertiary amine intermediate (*i.e.* TMA) is formed and acts as a substrate during the catalytic cycle, the high-valent Fe species, *e.g.* Fe(IV)-oxo, may also abstract H from its analogue DMEA, hence forming the corresponding secondary amine species, MEA (**Figure 4-12 A**). The secondary amine products (DMA, MEA) were derivatized by BSC and quantified by GC-MS. In the absence of TMAO, Tdm does not catalyse the demethylation of DMEA. However, in the presence of TMAO, both DMA (*m/z*: 77.1, 141.1, 185.1) and MEA (*m/z*: 77.1, 141.1, 184.1) were detected, and DMA formation was competitively reduced in the presence of DMEA (**Figure 4-12 B,C,D**). The results therefore supported the postulated mechanism, both confirming that TMAO is required as oxygen donor, and that the resulting tertiary amine is a substrate for oxidative demethylation during the catalytic cycle of Tdm.

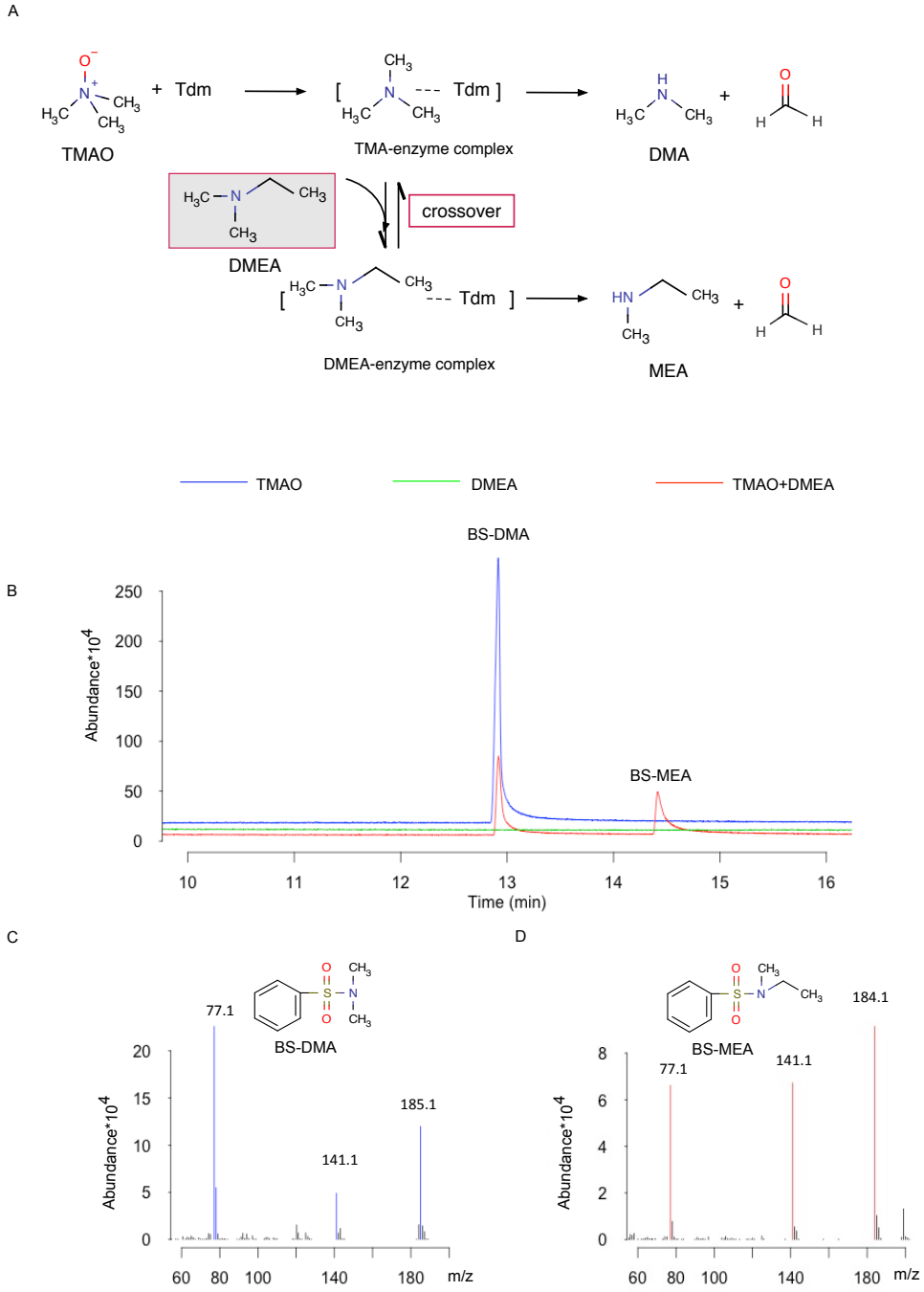


Figure 4-12 Formation of a secondary amine product, MEA by Tdm using a TMA-analogue, DMEA. Proposed diagram of MEA formation from DMEA in the presence of the natural substrate TMAO (A). Gas chromatography – mass spectrometry (GC-MS) analysis of secondary amines produced during Tdm reaction in the presence or absence of DMEA (B). The enzyme reaction contains either 50 mM TMAO (blue line), or 25 mM DMEA (green line) or both (red line). Only in the presence of the natural substrate TMAO, MEA was formed (red line). Mass spectra of benzenesulfonyl chloride derivatized secondary amines, BS-DMA (C, blue) and BS-MEA (D, red).

To examine the effect of DMEA on Tdm activity, the HCHO assay was performed with varying concentrations of DMEA and TMAO. Double reciprocal plots revealed that DMEA acts as an uncompetitive inhibitor of Tdm shown by constant slope at different concentration of DMEA. K_m and V_{max} are both reduced (**Figure 4-13 A**), with inhibition constant (K_i) of 2.7 mM (**Figure 4-13 B**). V_{max} of Tdm was nearly 0 when DMEA at 5 mM final concentration.

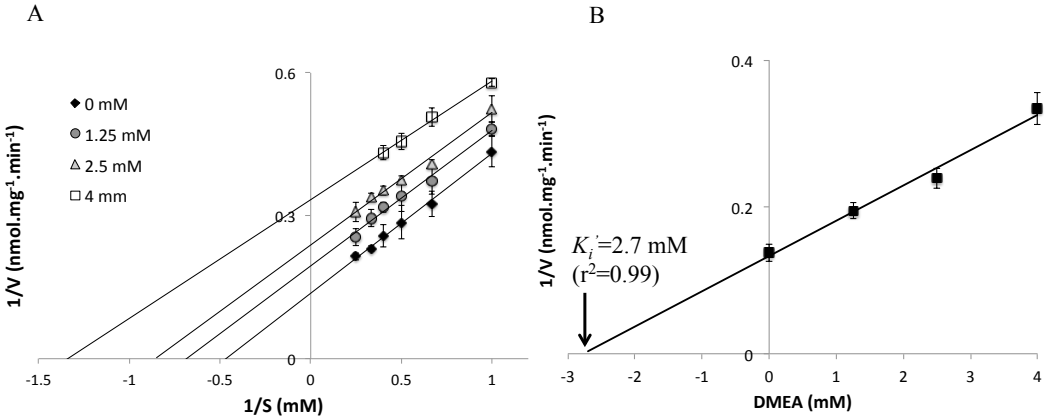


Figure 4-13 DMEA inhibits Tdm activity. Lineweaver-Burk plot of Tdm in the absence and presence of DMEA at varying concentrations (A). A secondary plot of the intercept derived from the primary Linewever-Burk plot versus DMEA concentrations (B). With both decreased V_{max} and K_m, DMEA is a uncompetitive inhibitor of Tdm (K_i = 2.7 mM).

4.3.7 Unlabelled formaldehyde was produced using TMA^{18}O as substrate

To further confirm that TMAO, rather than O_2, is the oxygen donor for Tdm, TMA^{18}O was synthesized and the ^{18}O label was followed during Tdm catalysis. Synthesized TMA^{18}O was verified by mass spectrometry (**Figure 4-14 A**). Formaldehyde produced during Tdm catalysis was adducted by sodium bisulfite and a subsequent ion with *m/z*=111 (HMS^-) would be detected by mass spectrum. If the hypothesis is true that the oxygen atom of formaldehyde is derived from the substrate TMA^{18}O, formaldehyde should be ^{18}O-labeled, hence an *m/z*=113 (^{18}O-HMS^-) would

be observed. However, only the *m/z*=111 ion was found in both TMAO and TMA^{18}O reaction mixture (**Figure 4-14 B**) at the retention time of 6.2 min with similar intensity values. MS/MS analysis of products from bombardment of the *m/z*= 111 ion selected from the initial MS ion separations resulting in detection of HSO_3^-(*m/z*=81), a product derived from bombardment of HMS^- (*m/z*=111) and the release of HCHO (*m/z*=32) (**Figure 4-14 C**). Extracted ion chromatogram at *m/z*=113 was observed from both reactions with ^{18}O-labeled and TMA^{16}O, but did not display a distinct peak (**Figure 4-14 D**). MS/MS analysis of ion *m/z*=113 at 6.2 min resulted in an ion at *m/z*=95.8 (**Figure 4-14 E**), suggesting that *m/z*=113 is from reaction matrix rather than ^{18}O-HMS^-.

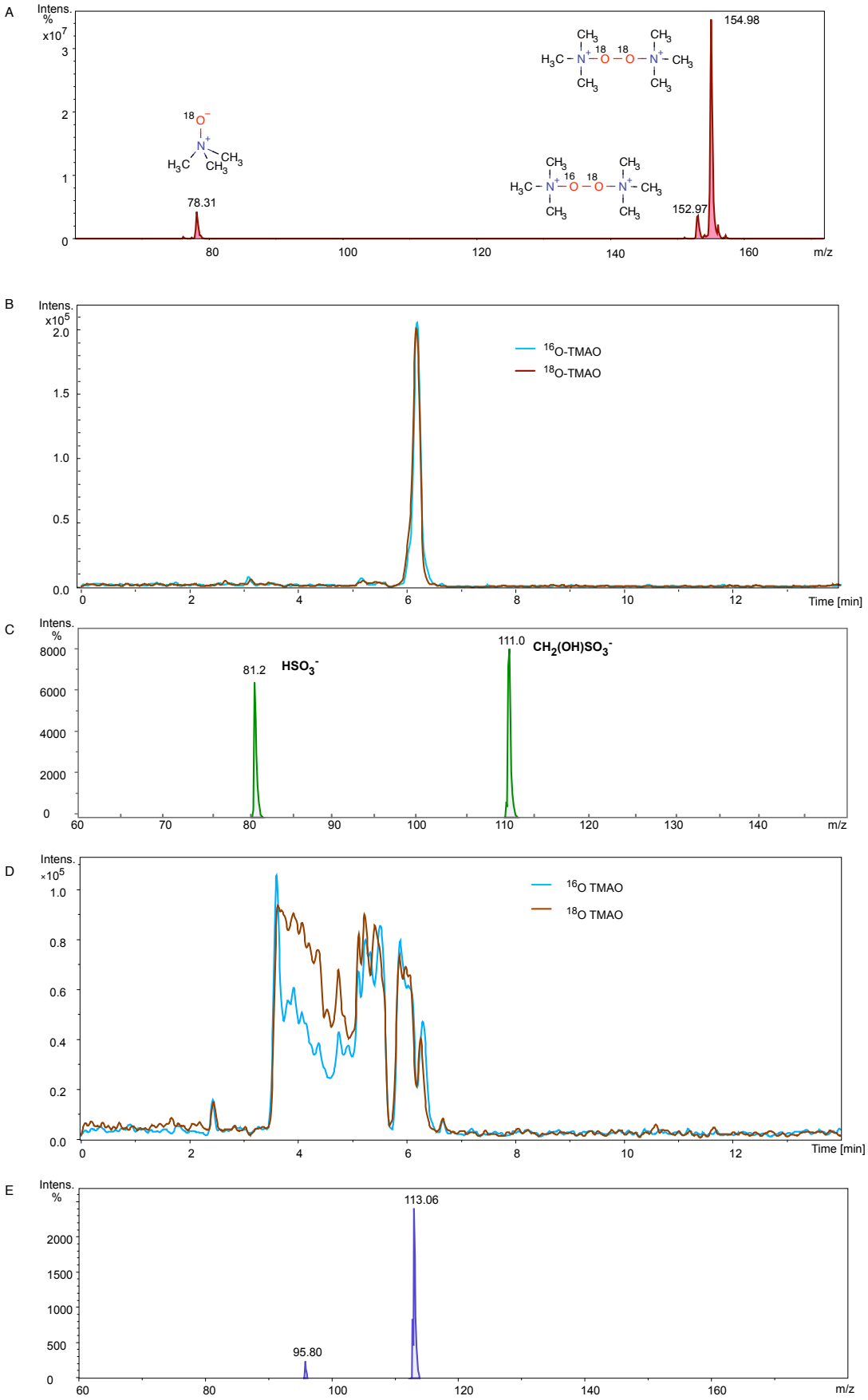

A
Intens.
%
x10^7
3
2
1
0
154.98
78.31
152.97
80
100
120
140
160
m/z
B
Intens.
x10^5
2.0
1.5
1.0
0.5
0.0
16O-TMAO
18O-TMAO
0
2
4
6
8
10
12
Time [min]
C
Intens.
%
8000
6000
4000
2000
0
81.2
HSO3-
111.0
CH2(OH)SO3-
60
70
80
90
100
110
120
130
140
m/z
D
Intens.
x10^5
1.0
0.8
0.6
0.4
0.2
0.0
16O TMAO
18O TMAO
0
2
4
6
8
10
12
Time [min]
E
Intens.
%
2000
1500
1000
500
0
113.06
95.80
60
80
100
120
140
160
m/z

Figure 4-14 Determination of O-transfer using $TMA^{18}O$. Purity of the chemically synthesized TMAO, as assessed by mass spectrometry (A). $TMA^{18}O$ displayed a M+1 peak at *m/z*=78.31 and 2M+1 peak at *m/z*=154.98. The small peak of 152.98 *m/z* represents 2M+1 with one ^{16}O and one ^{18}O. H_2O_2 used for $TMA^{18}O$ synthesis contained 90% $H_2{}^{18}O_2$ and 10% $H_2{}^{16}O_2$ Extracted ion chromatogram of the HMS ion *m/z*=111 from Tdm reaction using $TMA^{16}O$ (light blue) and $TMA^{18}O$ (brown) (B). MS/MS analyses of the *m/z*=111 ion resulted in the detection of HSO_3^- (*m/z* = 81) after releasing of HCHO (*m/z*= 32) (C). Extracted ion chromatogram the unidentified ion *m/z*=113 from Tdm reaction using $TMA^{16}O$ (light blue) and $TMA^{18}O$ (brown) (D). MS/MS analysis of the *m/z*=113 from both reactions produced an unidentified ion of *m/z*=95.8. HSO_3^- (*m/z* = 81) was not detected from the parent ion (E).

4.3.8 DMPO-OH but not substrate radical was detected

A DMPO spin trap experiment was designed to reveal substrate radicals formed during Tdm catalysis. However, neither a N- nor C-centred substrate intermediate radical was detected. Interestingly, DMPO-OH adduct was detected (**Figure 4-15**, marked by *). When TMAO was omitted from the incubation mixture, weak DMPO-OH adduct was also detected, possibly due to slow hydrolysis of DMPO. There were some small radical adduct signals detected from both with and without TMAO (**Figure 4-15**, marked by arrows), suggesting that these radicals were either from unspecific catalysis of DMPO by Tdm or DMPO spontaneously.

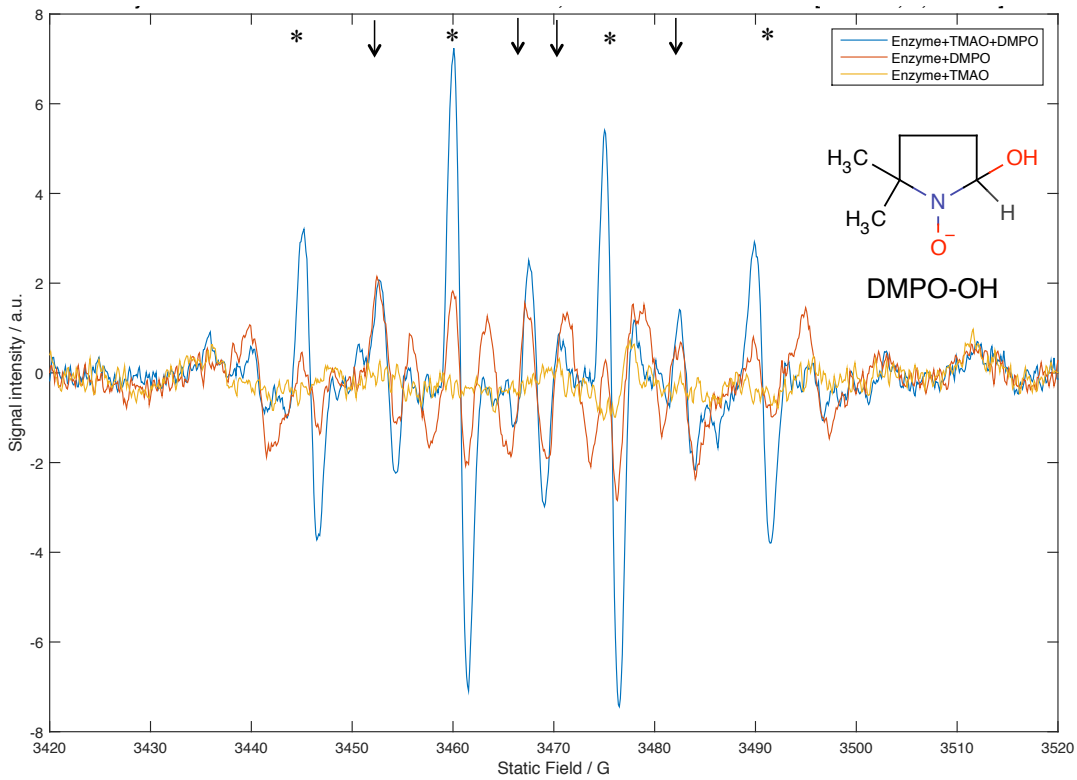


Figure 4-15 DMPO spin trap of intermediates during TMAO demethylation. *: DMPO-OH, arrows: DMPO-unknown species.

4.4 Discussion

4.4.1 Tdm is a novel Zn^{2+} and Fe^{2+} containing metalloprotein

Data presented in this chapter suggest that Tdm is a Zn^{2+} and Fe^{2+}-dependent metalloenzyme and the Zn^{2+}:Fe^{2+}:Tdm monomer ratio is likely to be 1:1:1. In Tdm, Zn^{2+} is coordinated by three cysteine thiolates (C263, C279 and C343) and 1 water molecule, which resembles catalytic Zn^{2+} site of CDA (Chung et al 2005, Kumasaka et al 2007, Sánchez-Quitian et al 2010, I et al 2006). Meanwhile, I also demonstrated that Fe^{2+} plays a catalytic role. It is noteworthy that both Zn^{2+} and Fe^{2+} are in the close vicinity of Asp198. Therefore, it is possible that Tdm has a co-catalytic binuclear Zn^{2+}-Fe^{2+} centre. Binuclear Zn^{2+}-Fe^{2+} sites have been found in many enzymes, such as glycerophosphodiesterase (Daumann et al 2013), glyoxalase II (Zang et al 2001), enamidase (Kress et al 2008) and protein phosphatase 2B (Namgaladze et al 2002). The resolved structures of existing binuclear Zn^{2+}-Fe^{2+} proteins revealed that Zn^{2+} is predominantly coordinated by a histidine side chain. However, caution should be applied when interpreting the Zn^{2+} site in Tdm due to the low quality of the overall model obtained through homology modelling. We cannot rule out the possibility that the fourth Zn^{2+} ligand may be provided by the D198 in Tdm.

Indeed, it is propose that a D198 has a dual role of stabilizing both Zn^{2+} and catalytic Fe^{2+}. Mutagenesis study of the conserved D198 in Tdm supported the idea that D198 is crucial for maintaining structural integrity through interaction with Zn^{2+}, either directly or indirectly through H-bonding with Zn^{2+}-bound water (**Figure 4-16**). Replacing D198 with Ala, Asn and Glu resulted in quaternary structural alteration as well as reduced Zn^{2+} content in the Tdm mutants (0.5~0.7 vs 1 in the wild type).

Figure 4-16 Dual role of D198 by coordinating Zn^{2+} and Fe^{2+} ion simultaneously. D198 and H276 are Fe^{2+}-binding sites in Tdm. D198 is involved in Zn^{2+} coordination by either directly (A) or indirectly through H-bonding with Zn^{2+}-bound water (B).

D198 also likely contributes to Fe^{2+}-binding together with H276. The mutagenesis results demonstrate that a carboxyl group is important in maintaining Fe^{2+} stoichiometry in Tdm and the Fe^{2+}:Tdm ratio of the D198E mutant (0.33 ± 0.01) was comparable to that of the WT (0.33 ± 0.02) while the D198A mutant (0.12 ± 0.01) lost Fe^{2+} significantly (p<0.05). So far, it is unclear what other residue could contribute to Fe^{2+} coordination in Tdm. It is very likely that Tdm employs a non-classic 2-His-1-carboxylate triad because the conserved H256 also seems to have a structure maintaining role. Variations of the classical 2-His-1-carboxylate triad motif for Fe^{2+} coordination have recently been found in a number of enzymes. For example, in the halogenase SyrB2 (uniprot entry: Q9RBY6), the carboxylate is absent and instead, a halogen ion takes its place in the coordination sphere (Blasiak et al 2006). In diketone-cleaving dioxygenase, Dke1 (Uniprot entry: Q8GNT2) (Diebold et al 2010) and cysteine dioxygenase, CDO (Uniprot entry: Q16878, P60334), a three-histidine triad is found (Gardner et al 2010, McCoy et al 2006) whereas in carotenoid oxygenase (Uniprot entry: P74334), a four-histidine motif is present (Kloer et al 2005). Although H256 and H276 are far apart from homology modelling, we should take into caution that the prediction of these two residues may not be reliable due to low sequence identidy and absence of Fe^{2+} in the template (**Figure 4-10 B**). Further

structural and biochemical investigations are certainly warranted to conclusively map the ligands involved in Zn^{2+} and Fe^{2+}-coordination in Tdm.

4.4.2 Substrate recognition and binding by a hydrophobic pocket

The identification of F259, Y305 and W321 indicated that Tdm likely recognizes and binds TMAO by cation-π interaction, which has also been found in other TMAO-binding proteins, such as TmoX and TorT (Li et al 2015, Moore and Hendrickson 2012). The cation-π interaction has long been recognized as an important non-covalent binding interaction relevant to structural biology (Gallivan and Dougherty 1999). The aromatic rings of Phe, Tyr and Trp provide negative electrostatic potential allowing interaction with cations (Dougherty 1996). Such cation-π interaction also seems common in proteins involved in quaternary amine transport and metabolism, *e.g.* choline-TMA lyase (CutC) from *Desulfovibrio alaskensis* G20 (Craciun et al 2014), acetylcholine esterase from *Tetronarce californica* (Harel et al 1993, Zhong et al 1998), substrate-binding protein (ChoX) of choline/acetylcholine from *Sinorhizobium meliloti* (Oswald et al 2008), and substrate-binding protein (ProX) of glycine betaine from *Archaeoglobus fulgidus* (Schiefner et al 2004). Interestingly, the Fe^{2+}:Tdm ratios of these mutants were also significantly reduced compared to that of the wide type. It is therefore likely that these hydrophobic residues are located adjacent to the Fe^{2+} centre and thus may directly or indirectly influence Fe^{2+} coordination in Tdm. Same as the prediction of Fe^{2+}-coordination sites, the structure of substrate pocket may not be reliable as a result of low sequence identity and non-conservative hydrophobic amino acids (**Figure 10 A, B**).

4.4.3 The O_2-independent *N*-dealkylation of Tdm

Non-haem iron enzymes participate in many metabolically important reactions by activating dioxygen (Bugg 2001, Que and Ho 1996). Tdm carries out an oxidative *N*-

dealkylation which resembles other well characterized α-ketoglutarate-dependent mononuclear non-haem enzymes, such as DNA dealkylase (AlkB) (Mishina and He 2006), and histone demethylase (*e.g.* JHDM1, JMJD6) (Chang et al 2007, Tsukada et al 2006). A-ketoglutarate-dependent oxygenases activate dioxygen to form high-valent Fe(IV)-oxo species as the oxidant. Fe(IV)-oxo complex subsequently transfers an oxygen atom to the substrate to generate a carbinolamine intermediate, followed by formaldehyde release from the carbinolamine intermediate, producing demethylated adenine or lysine. However, Tdm is different from α-ketoglutarate-dependent oxygenases in that the Tdm-catalysed reaction is O_2-independent (**Figure 4-11 A**).

Chemically, many single-oxygen atom donors are capable of generating Fe(IV)-oxo species (*e.g.* NaOX, X = Cl or Br, iodosylbenzene (PhIO)) (Balland et al 2004, Rohde et al 2003). Although surrogate single-oxygen atom donors have been found in biological systems, such as P450 enzymes, they havn't been found in non-haem iron containing enzymes. *N, N*-dimethylaniline *N*-oxide (DMAO) is an intermediate formed during P450-mediated demethylation of *N, N*-dimethylaniline. Recently studies have demonstrated that DMAO can serve as a surrogate single oxygen donor because DMAO gave identical isotope effects as natural system using NADPH and O_2 (Dowers et al 2004, Roberts and Jones 2010). It is proposed that, akin to DMAO for P450 enzymes, TMAO can act as a surrogate oxygen donor for Tdm. A potential mechanism for TMAO degradation by Tdm is shown in **Figure 4-11 B**. Our cross-over experiments did support the presence of a TMA-alike intermediate during the catalytic cycle as proposed in **Figure 4-12**. The TMA analogue, in this study, DMEA, uncompetitively inhibited TMAO demethylation by stabilised enzyme-substrate complex.

To validate this hypothesis that TMAO has a dual role of being an oxygen atom donor and the substrate, TMA^{18}O was used and the formation of ^{18}O-HCHO was trapped using sodium bisulfite (Jiang et al 2013). The formation of ^{18}O-labelled HCHO-bisulfite adduct would give a mass peak at *m/z*=113 whereas ^{16}O-unlabelled HCHO-bisulfite adduct would give a peak at *m/z*=111. However, only ^{16}O-unlabelled-adduct was observed (**Figure 4-14**) in reactions when Tdm was incubated with TMA^{18}O. Nevertheless, the data do not rule out the possibility that rapid exchange of oxygen atom may have occurred between the highly active Fe species (*e.g.* Fe(IV)-oxo, Fe(V)-oxo) and the solvent (water) prior to oxygen atom transfer to the $-CH_3$ group of TMAO. Similar rapid exchange of oxygen in high-valent Fe has been reported previously in biomimetic systems (Seo et al 2004).

The nature of the high-valent Fe species during Tdm catalysis remains unclear. The Fe(IV)-oxo species is common in haem and non-haem Fe^{2+}-containing enzymes (Costas et al 2004, Krebs et al 2007, Makris et al 2006, Nehru et al 2007). In addition, another high-valent iron-oxo complex, Fe(V)-oxo has been postulated as an active oxidant in Rieske dioxygenase enzyme (Kovaleva and Lipscomb 2008, Shan and Que 2006) and authenticated in non-haem iron biomimetic systems (Tiago de Oliveira et al 2007). Clearly further investigation is required to confirm the active Fe species during Tdm catalysis.

4.4.4 C-H bond activation mechanism remains unclear

According to Tdm mechanism hypothesis (**Figure 4-11 B**), a C- or N-centred substrate radical is formed before oxygen rebound to substrate intermediate depending on the mechanism of C-H bond activation. However, no substrate radical was detected in this DMPO spin-trapping experiment. One possible explanation might

be that, although DMPO was added immediately after reaction started, substrate intermediate may have turned over too fast to be trapped.

Although the detection of hydroxyl radical fits the hypothesis of cocatalytic Zn^{2+}, the DMPO-OH signals should be viewed cautiously because, quite often, they are artefacts rather than free hydroxyl radicals derived from the systems under study. Possible causes of DMPO-OH signals are 1) spontaneous hydrolysis of DMPO, which are usually weak; 2) non-specific reaction with other chemicals, such as H_2O_2, or nucleophilic addition of water (Ranguelova and Mason 2011). As a result, it is necessary to perform kinetic-based competition experiment with hydroxyl radical scavengers (*e.g.* ethanol, formate, DMSO) in order to establish the existence of free hydroxyl radical in spin-trapping experiment (Ranguelova and Mason 2011).

4.4.5 Hypothetical roles of co-catalytic Zn

The exact role of Zn^{2+} in the progression of TMAO demethylation remains unclear. Although the site-directed mutagenesis data largely supported the role of Zn^{2+} in maintaining Tdm structure, its involvement as a co-catalytic Zn^{2+}-Fe^{2+} centre is also possible. For example, one can envisage that Zn^{2+} may facilitate oxygen atom transfer from TMAO to Fe^{2+} through Lewis acid-base interaction with the O atom at step ② (**Figure 4-17 A**). Similarly, Zn^{2+} may bind to the O atom of the complex formed at step ④ to mediate C-N bond cleavage (**Figure 4-17 C**). A similar role of Zn^{2+} in catalytic site of alcohol dehydrogenase (ADH) for C-H bond cleavage is well known (Bugg 2012). Alternatively, Zn^{2+} may also stabilize the reactive high valent Fe species during catalysis through electrostatic interaction via a solvent oxygen, *e.g* Fe(IV)-oxo (step ③, **Figure 4-17 B**). It has been shown that the conserved positively charged arginine residue can help to stabilize and polarize the negative charge on the

reactive Fe(III)-OOH complex in peroxidases (Neri et al 1998, Schiødt et al 2007) (Singh et al 2012).

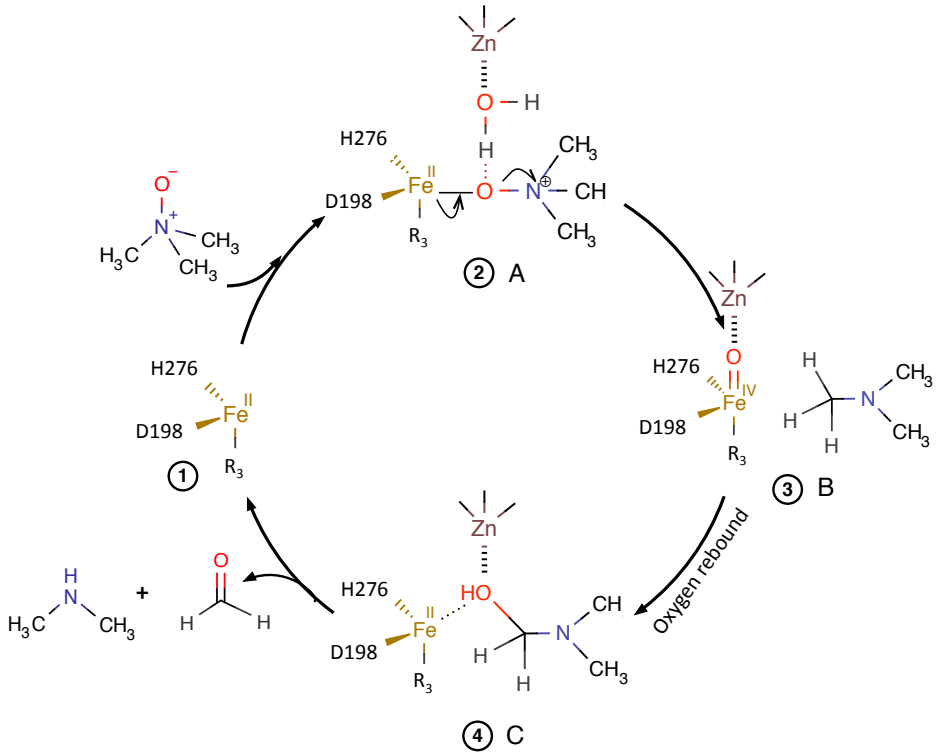


Figure 4-17 Possible function of Zn^{2+} in TMAO demethylation of polarizing TMAO (A); stabilizing Fe(IV)-oxo (B) and mediating C-N bond breakage (C).

In summary, the combination of site-directed mutagenesis, homology modelling and analytical chemistry have provided insight into the structure- function relationship of a novel Zn^{2+} and Fe^{2+} metalloprotein, Tdm. It carries out an unusual O_2-independent oxidative demethylation utilizing the substrate as the oxygen donor. Determination of the three-dimensional structure is now required to validate the model proposed in this study.

Chapter 5 Crystallisation and biochemical characterization a Tdm homologue from a marine bacterium *Ruegeria pomeroyi* DSS-3

5.1 Introduction

TMAO is a ubiquitous organic osmolyte that occurs in a wild variety of marine biota, including algae, zooplankton and fish. It is postulated that in these Eukaryotes, TMAO is derived from hydrolysis of phosphatidylcholine to TMA followed by oxidation by host flavin-containing monooxygenases (Seibel and Walsh 2002, Yancey 2005). TMAO is a potent protein stabiliser to counteract the effect of destabilizers (*i.e.* temperature, urea), and can enhance protein folding (Singh et al 2005, Street et al 2006, Yancey 2005).

TMAO increases with depth in bony fish (Kelly and Yancey 1999, Samerotte et al 2007). Levels of TMAO in the Kermadec snail fish captured at 7000 m revealed they had the highest recorded TMAO content of ~400 mmol kg^{-1} (wet mass), giving the evidence that TMAO maybe biochemically restrain marines fish from inhabiting deepest ocean depth (Yancey et al 2014).

Based on the numerous studies that proved important biological and physiological roles TMAO plays in marine organism, it is reasonable to assume that TMAO presents in seawater. However, the study of TMAO concentration of marine surface water has been rare due to technical limit at nanomolar level (Gibb and Hatton 2004). Gibb & Hatton (Gibb and Hatton 2004) used a coupled flow injection-ion chromatographic technique to determine TMAO in natural seawater and found that TMAO ranged from below the analytical detection limit (1.65 nmol l^{-1}) to 76.9 nmol l^{-1} in the coastal waters off the Antarctic Peninsula.

Despite the difficulty to quantify TMAO content in seawater, it has been recognized to form part of the MAs pool that are important nutrients for heterotrophic marine bacterioplankton (Anthony 1982, King 1987, Lidbury et al 2014, Lidbury et al 2016). Recent study has identified genes encoding Tdm and a TMAO ABC transporter in

numerically dominant marine bacteria of the *Roseobacter* clade (MRC) and the SAR11 clade (Lidbury et al 2014, Lidbury et al 2015). MRC accounts for around 10% of bacterioplankton in the open ocean (Buchan et al 2005, Labbe and Rettmer 1989, Rappé and Giovannoni 2003) and up to 25% in coastal waters (DeLong 2005, Giebel et al 2011, Suzuki et al 2001). SAR11 clade of the *Alphaproteobacteria* comprises about 25% cells of coastal, estuary and open sea habitats (Malmstrom et al 2004). MRC and SAR11 clade are key players in the ocean carbon cycle (Rusch et al 2007). The metabolism of TMAO as energy source has a substantial ecological significance which may explain the dominance of MRC and the SAR11 clade. Efficient convertion of organic substrate into biomass stimulates growth, which provides the ecological advantage of these bacteria. Meanwhile, N-remineralisation to ammonium, and export to support other microorganisms in the environment (Halsey et al 2012, Lidbury et al 2015, McClelland et al 2001).

In **Chapter 4**, recombinant Tdm of *Methylocella silvestris* (Tdm_BL2 hereafter) has been identified as a Zn^{2+} and Fe^{2+}-dependent enzyme, in which Zn^{2+} and Fe^{2+} constitute a unique cocatalytic centre. Metal replacement experiment of Tdm_BL2 *in vitro* (**Chapter 4**) has shown that Fe^{2+} is essential for the maximum activity. However, the native metal cofactor in microbial Tdms of different origins (*i.e.* terrestrial *versus* marine) may differ as a result of metal availability. For example, Zinc replaced iron to form protoporphyrin (ZnPP) during haem synthesis (Labbe and Rettmer 1989); phytoplankton substituted Fe-containing ferredoxin by flavodoxin (LaRoche et al 1996). Substitution of cobalt and cadmium for Zn^{2+} at active sites of zinc enzymes such as carbonic anhydrase in *Thalassiosira weissflogii* (Lane and Morel 2000, Price and Morel 1990) has also been observed when Zn^{2+} was limited.

The dependency of Zn^{2+} and Fe^{2+} raises the question that would Zn^{2+} or Fe^{2+} be replaced by other ions when Zn^{2+} or Fe^{2+} is in short supply, *e.g.* High Nutrient–Low Chlorophyll (HNLC) region including the Southern Ocean, the equatorial Pacific, and parts of the Antarctic Ocean (Carol and Timothy 1997, Moore et al 2013). HNLC oceans typically have subnanomolar concentrations of dissolved iron (Martin et al 1994, Mawji et al 2015). Such low concentration of dissolved iron has been identified as the limiting nutrient for both primary productivity and bacteria growth (Behrenfeld and Kolber 1999, Church et al 2000, Moore et al 2013, Tagliabue et al 2014), thus iron fertilisation triggered massive primary production in HNLC region (Coale et al 1996). Apart from "Fe-hypothesis", Morel (Morel et al 1994) proposed a "Zn-hypothesis", suggesting Zn as well as Fe may be limiting phytoplankton growth. Dissolved Zn concentrations in the surface waters of the open Pacific, Atlantic Oceans or Southern Ocean are in the subnanomolar range (Lohan et al 2002, Zhao et al 2014). ~98% of Zn is bound to strong and uncharacterized organic ligands such that the concentration bioavailable Zn, is in the picomolar range (Andersen et al 2011, Lohan et al 2002).

Given the wide distribution of Tdm genes in marine bacterioplankton (Lidbury et al 2014), one can envisage that Tdm may use other metals to replace cocatalytic metal centre where there is a limited supply of Fe or Zn in the seawater. Manganese (Mn) is of particular interest in this regard, because dissolved Mn is present in relatively high concentrations compared with dissolved Fe in surface waters of the sea (Chester 1990, Mawji et al 2015, Saager et al 1989). Mn has a reduction potential that overlaps with that of Fe (da Silva and Williams 2001). It has been reported that Mn and Fe are exchangeable by replacing metal ion at the active centre, for example swapping metals in Fe- and Mn-dependent homoprotocatechuate 2,3-dioxygenases homologues

from *Brevibacterium fuscum* (Fe-dependent) and *Arthrobacter globiformis* (Mn-dependent) *in vitro* did not interfere enzyme activity (Emerson et al 2008). Alternatively, Mn or Fe are incorporated in similar active site of homologues enzymes, but exhibit significant activity with only their native cofactor, for example Mn-, Fe-superoxide dismutases (SOD) (Ken et al 2005, Stallings et al 1984, Vance and Miller 1998).

Therefore it is hypothesised that Tdm homologues from marine environment may have different metal cofactors, compared to the Zn^{2+} and Fe^{2+} found in Tdm_BL2. To test the hypothesis, two oceanic Tdm homologues were characterised, the pelagic *α*-proteobacterium SAR11 strain HIMB59 (uniprot entry number: J9Z2K4, Tdm_HIMB59 hereafter) and the coastal MRC strain *Ruegeria pomeroyi* DSS-3 (uniprot entry number: Q5LT52, Tdm_DSS3 hereafter).

The aims of this chapter are:

- To compare the enzymatic activity of Tdm homologues present in the marine bacteria, SAR11 strain HIMB59 and *Ruegeria pomeroyi* DSS-3
- To identify the metal cofactors of Tdm homologues from marine bacterium
- To set up crystallization trials for the determination of Tdm structure in order to reveal its structure-function relationship

5.2 Materials and methods

5.2.1 Cloning and heterologous expression of *tdm* wild type, mutants and C-terminal truncated Tdm in *Escherichia coli*

Tdm_HIMB59 and Tdm_DSS3 overexpression constructs were generated by Dr. Ian Lidbury (Lidbury et al 2014). Briefly, The *tdm* gene from *R. pomeroyi* DSS-3 was amplified by PCR and cloned into the expression vector pET28a (Merck Biosciences, Germany). The *tdm* gene from the SAR11 strain HIMB59 was chemically synthesized

(GenScript Corporation) and cloned into pET28a. The resulting plasmids were transformed into the expression host *E. coli* BLR(DE3) pLysS (Merck Biosciences, Germany). Tdm_BL2 wild type and mutant overexpression constructs (C263A, C279A) were generated as described in **Chapter 2**. The C-terminal truncated Tdm_BL2 ranging from amino acid position 1 to 383 (Tdm_383) and 1-389 (Tdm_389) were amplified by PCR using plasmid pET28a-Tdm_BL2 as the template with the primers Tdm_F/Tdm_383_R and Tdm_F/Tdm_389_R, respectively (**Appendix 1**). The C-terminal truncated Tdm_BL2 genes were then excised using the *Nde*I/*BamH*I sites and ligated into the expression vector pET28a (Merck Biosciences, Germany) as aforementioned. Plasmids were sequenced for confirmation.

5.2.2 Protein purification and enzymatic assays

Protein induction, purification using the Ni-NTA column followed by desalting, and subsequent enzymatic assays were performed as described in **Chapter 2**.

Inhibition assay of DMSO on Tdm_BL2 was determined using assay solutions containing 1.5-6 mM TMAO and 0.1-1 mM DMSO. Inhibition constants (K_i) were determined as described in **Chapter 4**.

5.2.3 Size-exclusive chromatography

As described in **Chapter 2**.

5.2.4 Inductively coupled plasma - mass spectrometer and optical emission spectrometer (ICP-MS/OES)

As described in **Chapter 4**.

5.2.5 Crystallisation screening

Crystallisation screening was performed as described in **Chapter 2**. Tdm_BL2 and Tdm_DSS3 were screened in the Structure Biology Lab (C10) at the University of Warwick in collaboration with Prof. Vilmos Fulop and Dr. Alexander Cameron.

Crystallisation screening of Tdm_BL2 variants C263A and C279A were carried out in Prof. Yuzhong Zhang's lab, State Key Lab of Microbial Technology, Shandong University, China. Protein concentration was ~4 mg mL^{-1} in 20 mM Tris-HCl pH 7.8, 100 mM NaCl. Screens used include Procomplex suit (Qiagen), PEG/Ion HR2-126 (Qiagen), PEG/Ion 2 Screen (Hampton), PEGRx 1 HR2-082 (Hampton), PEGRx 2 HR2-084 (Hampton), Crystal Screen Lite HR2-128 (Hampton), Crystal Screen 2 HR2-112 (Hampton), Crystal Screen 2 HR2-110 (Hampton), Index (Rigaku), Wizard 1/2/3/4 (Rigaku).

5.2.6 Expansion of successful hits in a 24-well format

To improve crystal morphology and diffraction, conditions identified as supporting either nucleation or crystal formation in the 96-well screens were further expanded in a 24-well format. This allowed for changes in crystal formation upon slight adjustments in the pH and concentration of the precipitant to be explored. In this case, 1 mL of mother liquor was dispensed into the well. Typically, protein to mother liquor ratios of 1 µl:1 µl and 2 µl:1 µl were set up on the same plastic coverslip, which was then placed over the well and sealed with vacuum grease to produce a hanging drop set up. As before, plates were left at 18^{o} C and periodically checked for better crystal formation.

5.2.7 Seeding

For seeding experiments, a horsehair was dipped into the drop of parent crystals and streaked across six drops in consequence, resulting in serial dilution of the crystal seed across the plate.

5.2.8 Additive screen

To further refine crystallisation conditions, additional screens were exploited using the Crystal Screen Lite HR2-128 (Hampton Research) supplemented with various

cations. Sitting drops were comprised of 1 µl of protein samples, 1 µl of well solution and 0.3 µl of additives including, $BaCl_2$, $CaCl_2$, $CoCl_2$, $CuCl_2$, $MgCl_2$, $MnCl_2$, $SrCl_2$, YCl_3, $ZnCl_2$

5.3 Results

5.3.1 Tdm homologue from *R. pomeroyi* DSS-3 is a trimer

Denatured molecular weight of recombinant Tdm_HIMB59 and Tdm_DSS3 are ~80 kDa (**Figure 5-1 A**), which are consistent with the molecular weight calculated from amino acid sequence (88.33 kDa and 86.20 kDa, respectively). The yield of Tdm_HIMB59 was so low that no further characterisation was performed. Gel filtration revealed that the native molecular weight of Tdm_DSS3 was ~260 kDa (**Figure 5-1 B),** indicating that Tdm_DSS3 is a trimer.

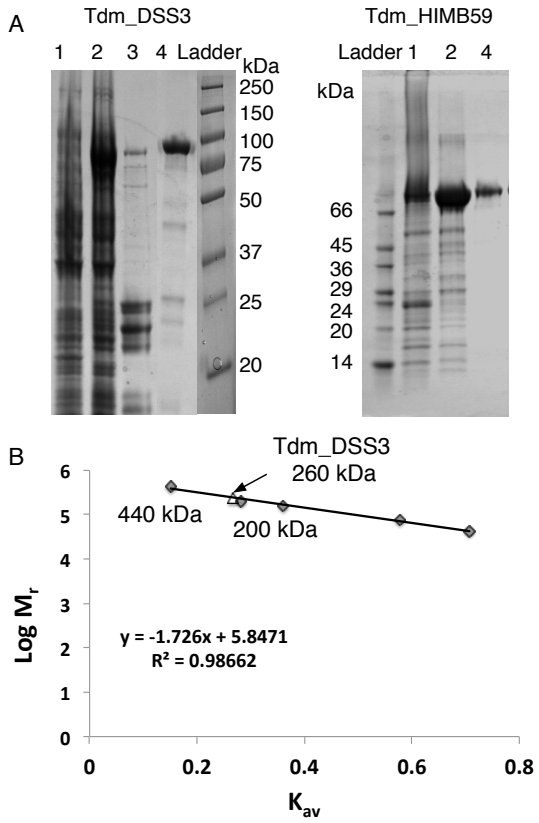


Figure 5-1 Molecular weight determination of the recombinant Tdm_DSS3 and Tdm_HIMB59. SDS-PAGE of Tdm_DSS3 and Tdm_HIMB59 showed a denatured molecular weight ~ 80 kDa (A). 1, crude cell-free extract; 2, column wash fraction with binding buffer containing 20 mM imidazole; 3, column wash fraction with washing buffer containing 60 mM imidazole; 4, elution fraction of the purified truncated protein. Native molecular weight estimation by gel filtration (B). Tdm from the α-proteobacterium HIMB59 gave very low yield so that gel filtration and subsequent biochemical characterisation were not performed. Standard curve for the molecular mass determination was established using Ferritin (440 kDa), β-Amylase from sweet potato (200 kDa), Aldolase2 (158 kDa), Conalbumin (75 kDa) and Ovalbumin (43 kDa) as described in **Chapter 2**. The calculated size of Tdm_DSS3 was 260 kDa.

5.3.2 The activity of Tdm_DSS3 is comparable to Tdm_BL2

The purified recombinant Tdm_DSS3 has an optimum pH at around 6 (**Figure 5-2 A**). Under optimum conditions, V_{max} and K_m of the recombinant Tdm_DSS3 were determined to be 13.71 ± 0.62 nmol min^{-1} mg^{-1} and 3.19 ± 0.15 mM, respectively by the Eadie-Hofstee plot (**Figure 5-2 B**). Its K_m value of Tdm_DSS3 is in good agreement with that of *Methylocella silvestris* (3.3 ± 0.64 mM), *Aminobacter aminovorans* (2 mM) and *Bacillus* sp. (2.85 mM), respectively (Large 1971, Myers and Zatman 1971, Zhu et al 2014).

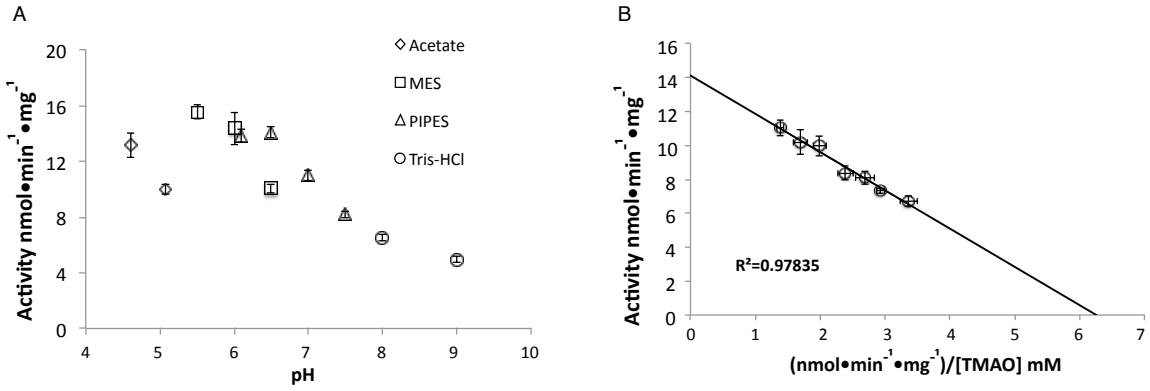


Figure 5-2 Optimum pH (A) and steady-state kinetic parameters (B) of recombinant Tdm_DSS3 by the Eadie-Hofstee plot. Error bars indicate standard deviations of experiments run in triplicate.

5.3.3 Tdm_DSS3 does not use DMSO as a substrate

Previous study has shown that bacterial Tmm, which is responsible for the oxidation of TMA into TMAO, can also oxidise TMA analogue, dimethylsulphide (DMS) at comparable rates to that of TMA (Chen et al 2011). Lidbury and Kröber (Lidbury et al 2016) proposed that DMS oxidation by Tmm may be a significant route for DMSO production in the oceans. However, the subsequent Tdm_DSS3 that catalyse TMAO can not degrade DMSO nor DMS **(Figure 5-3)**. The same substrate specificity has been observed for Tdm_BL2 (**Figure 3-10**) and DMSO acts as an effective uncompetitive inhibitor of Tdm_BL2 with inhibition constant (K_i) of 0.66 mM (**Figure 5-4**).

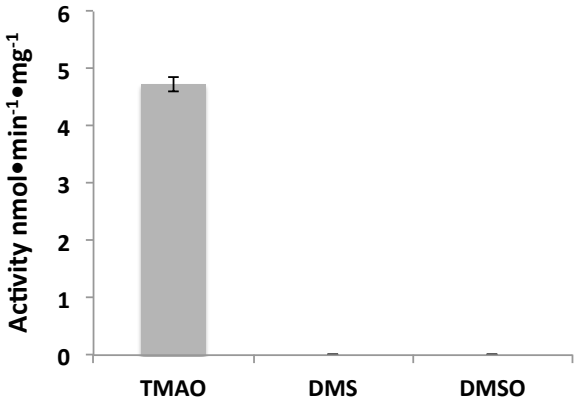


Figure 5-3 Tdm_DSS3 does not degrade DMS nor DMSO

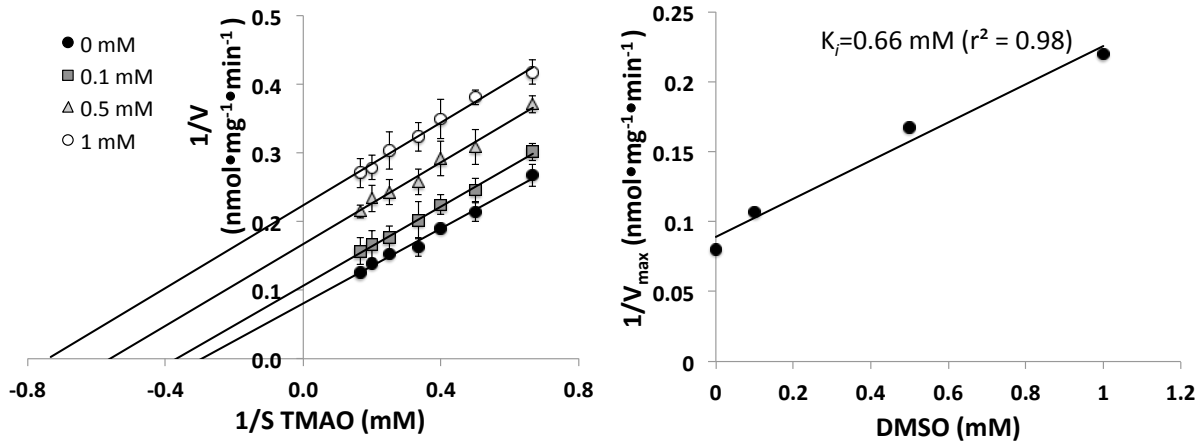


Figure 5-4 DMSO inhibits Tdm activity. Lineweaver-Burk plot of Tdm_BL2 in the absence or presence of DMSO at varying concentrations (A). A secondary plot of the intercept is derived from the primary Lineweaver-Burk plot versus DMSO concentration (B). With both decreased V_{max} and K_m, DMSO is therefore a uncompetitive inhibitor of TMAO demethylation by Tdm_BL2, with K_i of 0.66 mM.

5.3.4 Tdm from DSS-3 is more susceptible to EDTA

EDTA inhibition assay performed previously on Tdm_BL2 showed that less than 20% activity was inhibited upon incubation with 10 mM EDTA. Tdm_DSS3, however, is more sensitive to EDTA. At a concentration of 10 mM, EDTA caused ~40% reduction of the enzyme activity (**Figure 5-5**).

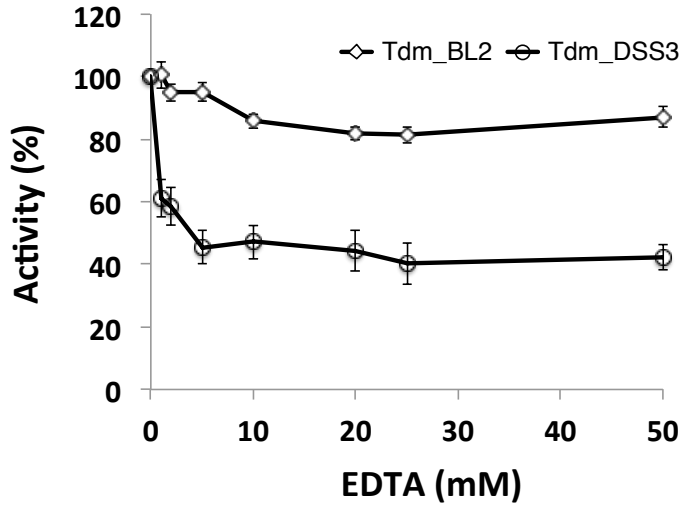


Figure 5-5 Inhibition of Tdm activities by EDTA.

5.3.5 No other metals other than Zn and Fe was detected in Tdm from *R. pomeroyi*

ICP-MS scan revealed the presence of Fe, Zn and Ni in recombinant Tdm_DSS3. Fe, Zn, Ni and S were therefore quantified by ICP-OES. ICP-OES data showed that Ni was negligible (<0.1), which was probably due to the leaching from the Ni-NTA column. Zn content was 0.1-0.2 equivalent per monomer of Tdm_DSS3, much lower than that of Tdm_BL2 (1 equivalent per monomer of Tdm_BL2). The Fe content was more difficult to ascertain. The first batch of purified Tdm_DSS3 showed Fe:Tdm ratio of 1.15 ± 0.23 per monomer, whilst the second batch was 0.29 ± 0.06 per

monomer (**Table 5-1**). As a result, further experiments are required to confirm Fe:Tdm_DSS3 ratio.

Table 5-1 Metal ion quantification of recombinant Tdm_DSS3 by ICP_OES

	M^{2+}:Tdm (monomer) ratio		
	Zn	Fe	Ni
Batch 1	0.10±0.09	1.15±0.23	<0.1
Batch 2	0.16±0.04	0.29±0.06	<0.1

Each batch of enzyme was prepared in triplicate using 3 separate Ni-NTA columns.

5.3.6 96-well crystallisation screening trials

Purified wild type Tdm from *M. silvestris* and *R. pomeroyi*, and two mutants of Tdm_BL2 were screened (**Table 5-2**). In order to set up crystallization screening, a highly concentrated recombinant Tdm proteins (4-10 mg ml^{-1}) are required. However, Tdm_BL2 and Tdm_DSS3, tend to aggregate during concentration using the spin columns (Vivaspin, GE healthcare, UK) as shown in **Figure 5-6 A**. Therefore, the hexamer fraction of Tdm_BL2 and the trimer fraction of Tdm_DSS3 were collected, which were then used for crystallisation screenings directly without further concentration in order to avoid aggregation. The collected fractions represent the Tdm enzymes in its native form with the concentrations between 2-4 mg ml^{-1}, lower than the recommended concentration recommended for crystallization trials.

Table 5-2 Proteins and polypeptides used for crystallisation screening

Protein	**Organism**	**Oligomeric status**	**Comments**
Tdm_BL2	*Methylocella silvestris* BL2	Hexamer	Wild type
Tdm_DSS3	*Ruegeria pomeroyi* DSS-3	Trimer	Wild type
Tdm_BL2 C263A	*Methylocella silvestris* BL2	Monomer	C263A mutant
Tdm_BL2 C279A	*Methylocella silvestris* BL2	Monomer	C279A mutant

I also intended to set up crystallization trials using the Tdm_BL2 mutants, C263A and C279A. These two mutant proteins form monomers, which might be easier to form

crystals due to smaller molecular weight. However, it was observed that these two mutants formed multiple oligomeric aggregates after concentrating using the spin columns (**Figure 5-6 B**). Similarly, only the monomers collected through gel filtration columns were used for crystallisation screening (concentrations ~3mg ml^{-1}), however, no promising crystals were observed throughout the experiments despite that a range of screening matrices were used.

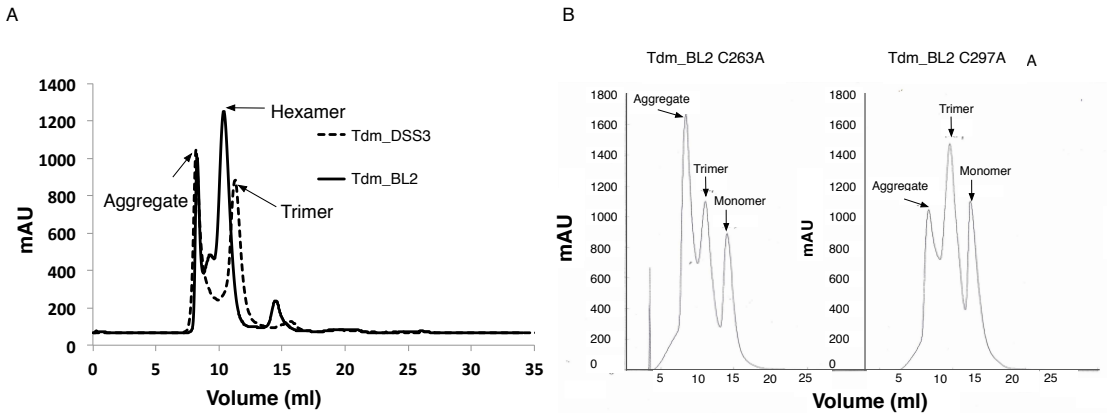
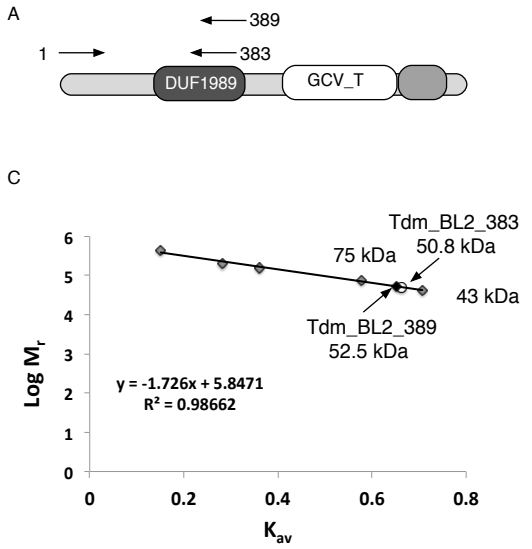


Figure 5-6 Gel filtration purification of wild-type Tdm_BL2 and Tdm_DSS3 (A) and variants Tdm_BL2 C263A and C279A (B).

Given that Tdm_BL2 has a massive native molecular weight of 480 kDa, it may be difficult to crystallise. I have previously hypothesized that the N-terminal DUF1989 domain contains the metal cofactors and is where TMAO bound and catalysed to produce HCHO (see **Chapter 4** for more information), thus DUF1989 is of more interest than the well studied C-terminal GCV_T domain. As a result, two versions of C-terminal truncated Tdm_BL2 were constructed, Tdm_BL2_383 ranging from amino acid positions 1 to 383 where the DUF1989 domain ends, and Tdm_BL2_389 ranging from amino acid positions 1 to 389 (**Figure 5-7 A**). Tdm_BL2_383 and Tdm_BL2_389 were induced for expression in recombinant *E. coli* and purified using an identical approach used for the full-length Tdm_BL2. Purity and molecular weight

of the truncated Tdms were determined by SDS-PAGE (**Figure 5-7 B**). Gel filtration data showed that, in contrast to the wild-type Tdm, Tdm_BL2_383 and Tdm_BL2_389 existed as monomers (**Figure 5-7 C**). Activity assays using truncated Tdm_BL2_383 and Tdm_BL2_389 by quantifying formaldehyde formation demonstrated that they were inactive. Metal quantification by ICP-OES revealed that C-terminal truncated mutants did not contain Fe, and the Zn^{2+}:protein ratio was reduced to 0.5-0.6 per monomer (**Figure 5-7 D**), suggesting that C-terminal truncated mutants did not fold properly. Therefore, Tdm_BL2_383 and Tdm_BL2_389 were not included in the crystallisation screening trials.

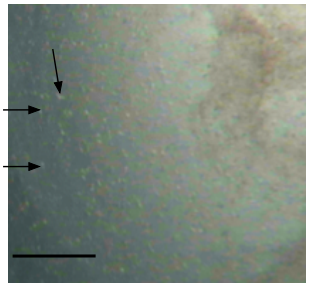


D

	M^{2+}: protein (monomer)	
	Zn	Fe
Tdm_BL2_383	0.58±0.02	0
Tdm_BL2_389	0.64±0.01	0

Figure 5-7 Construction and characterisation of C-terminal truncated Tdm_BL2 mutants. Scheme of C-terminal truncated Tdm_BL2 mutants (A). SDS-PAGE (12.5% w/v, Bio-Rad, USA) analysis of C-terminal truncated Tdm_BL2 mutants (B). Ladder, PageRuler Prestained Protein Ladder (left, Thermofisher Scientific), Sigma-Aldrich colour burst (right, Sigma-Aldrich); 1, uninduced crude cell-free extract of *E.coli* containing Tdm_BL2_383; 2, induced crude cell-free extract of *E.coli* containing Tdm_BL2_383; 3, uninduced crude cell extract of *E.coli* containing Tdm_BL2_389; 4, induced crude cell-free extract of *E.coli* containing Tdm_BL2_389; 5, elution fraction of Tdm_BL2_383; 6, elution fraction of Tdm_BL2_389. Native molecular weight estimation by gel filtration (C). Standards were the same as **Figure 5-2**. The calculated sizes of Tdm_BL2_383 and Tdm_BL2_389 were 50.8 and 52.5 kDa, respectively.

Initial screenings were prepared using the sitting drop vapour diffusion method. The initial crystallisation trails indicated that Tdm_DSS3 was more promising with small plates observed under pH 7-8 with PEG3350 or PEG6000 as precipitant at 20% concentration (PACT premier, Molecular Dimensions) after 3 days at 18ºC (example shown **Figure 5-8**).



0.1M Tris-HCl (pH 8.0), 0.2 M NaCl, 20% PEG6000

Figure 5-8 Microcrystal formation of Tdm_DSS3. The sitting drop method of vapour diffusion was used with commercial sparse matrix screens. Small crystals are indicated by arrows. Scale bar=50 μm.

Although Tdm_BL2 WT and variants formed promising precipitates such as light/heavy amorphous precipitates under certain conditions, no crystal was observed. As a result, Tdm_DSS3 was processed for subsequent fine-screening.

5.3.7 24-well refinement of suitable crystallisation conditions for Tdm_DSS3

Refinement of Tdm_DSS3 crystalisation condition was carried using the 24-well plate setup as described in materials and methods. Tris-HCl (pH 8.0) and HEPES (pH 7.5) buffers were tested with either PEG3350 or PEG6000, varying from 16%-26% (w/v) in a total of six 2% increments. NaCl remained to be 0.2 M throughout all conditions. The results of refinement, after incubation for 3 days at 18ºC, were shown in **Figure 5-9**. As indicated in **Figure 5-9**, HEPES (pH 7.5) buffer with 16% PEG6000 gave the biggest crystals, two-dimensional plates, about 15-30 μm in length and 5-20 μm in width. No well-formed three-dimensional crystals were observed. Diffraction quality is affected by crystal size and regularity (Glaeser et al 2000), as a result, seeding and

additive screening were exploited to optimise crystallisation. The crystals grew smaller as PEG6000 concentration increased to 26%. Therefore, 0.1 M HEPES (pH 7.5), 0.2 M NaCl, 16% (v/v) PEG 6000 was adopted for further seeding and additive refinements.

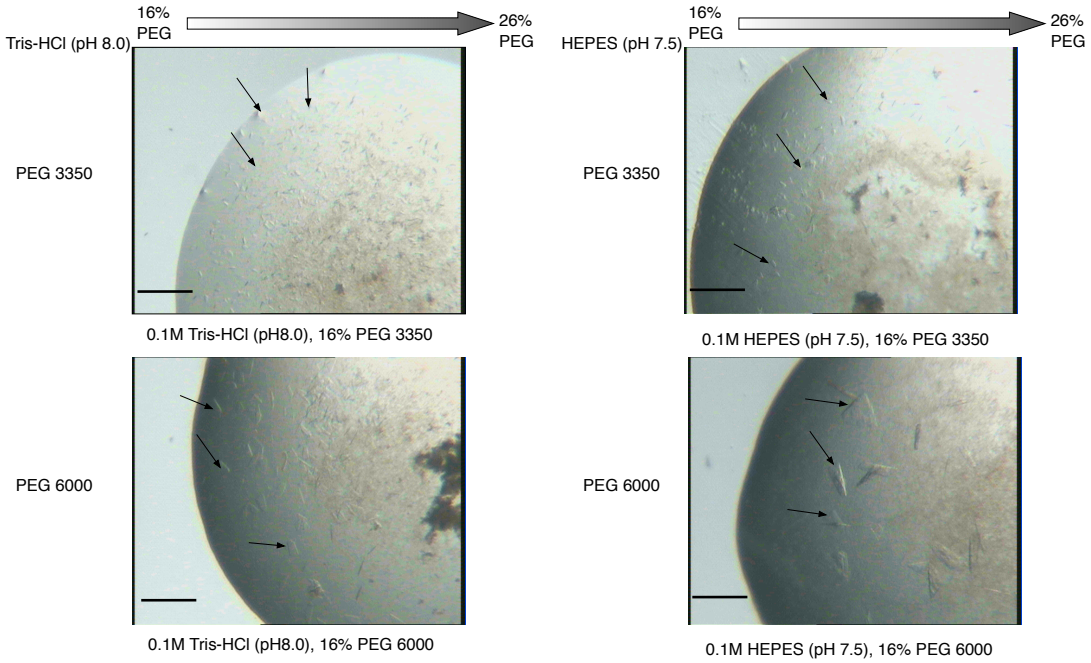


Figure 5-9 Optimisation of crystallisation conditions for Tdm_DSS3. Crystals are indicated by arrows. Scale bar=50 μm.

5.3.8 Refinement of Tdm_DSS3 crystal morphology and diffraction by seeding and additive screen

Horsehair crystal seeding was carried out in the 24-well format under the optimised condition after 1 day incubation. However, neither seeding nor additive screening using additional multivalent metal ions showed any obvious improvement in terms of crystal shape or size.

5.4 Discussion

5.4.1 Tdm_DSS3 shows comparable activity to Tdm_BL2 but has different quaternary structure

Tdm_DSS3 had an optimal pH of 6 and displayed comparable activity to that of *M. silvestris*. Tdm_DSS3 is more sensitive to EDTA inhibition comparing to Tdm_BL2, which might be due to different oligomeric status of the enzymes. Comparing to the homotrimeric Tdm from *R. pomeroyi*, Tdm of *M. silvestris* is homohexameric that may have restricted access of EDTA to the metal centre.

Although the structure of DMSO resembles TMAO, it is neither a substrate nor a competitive inhibitor of Tdm. In fact, DMSO acts as an effective uncompetitive inhibitor of Tdm, indicating that DMSO binds to the Tdm-TMAO complex. The other uncompetitive inhibitor for Tdm that has been found (**Chapter 4**) is a TMA analogue, DMEA (**Figure 4-12**). Interestingly, TMA itself did not show any inhibition of TMAO demethylation by Tdm_BL2 even at the concentration as high as 100 mM (data unshown). It remains unclear where DMSO or DMEA binds and how they inhibited Tdm demethylation. Resolving Tdm structure will certainly help to illuminate the mechanism of uncompetitive inhibition by DMSO and DMEA.

5.4.2 Tdm_DSS3 is an Fe-dependent enzyme

The presence of Fe and the absence of Mn suggested that Tdm_DSS3 employs Fe to mediate electron transfer. Further study is required to accurately determine Fe: Tdm_DSS3 ratio. Interestingly, it displayed 0.1-0.2 Zn^{2+}:Tdm_DSS3 (monomer) ratio, much lower than that of Tdm_BL2. In **Chapter 4**, it was concluded that, in Tdm_BL2, Zn^{2+} is a cocatalytic centre and contributes to the maintenance of overall structure of Tdm_BL2. So far, it is not clear whether the lower amount of Zn^{2+} found in Tdm of *R. pomeroyi* was due to zinc contamination or in deed, majority of Zn^{2+} in

this enzyme was lost during purification. The latter explanation is less likely because my previous study on Tdm_BL2 (**Chapter 4**) revealed that Zn^{2+} is essential for Tdm_BL2 and a loss of Zn^{2+} leads to structure alteration. Besides, the strong binding affinity of Zn^{2+} to the thiol group of the cysteines (C263, C279, C343), which are also conserved in Tdm of *R. pomeroyi* (C287, C303, C367) (**Figure 4-6 B**), suggests that a lower Zn^{2+} content in Tdm_DSS3 is not likely due to its loss during purification. However, additional experiments are required to investigate whether Zn addition to the purified Tdm_DSS3 will enhance its activity. Further study is required to determine the precise metal composition and metal centres in Tdm_DSS3.

5.4.3 Crystal obtained from Tdm_DSS3 grew as small plates

Although I have managed to observe Tdm_DSS3 crystals, they were two-dimensional plates with small sizes, not good enough to get good quality X-ray diffraction data. Crystallisation optimization strategies, such as micro-seeding and additive screening, did not significantly improve the crystals. Further optimization can be performed, such as using different temperatures, different protein concentrations, or using various Tdm mutants. Given that Tdm is a ferrous iron containing protein, which is sensitive to oxidation by air, anaerobic crystallisation should also be considered. Alternatively, one can try to create a C-terminal truncated Tdm with longer flanking region in order to keep the mutant properly folded. Since the N-terminus contain the metal binding sites, such a smaller truncated protein may be relatively easier for crystallization trials and further optimization. Alternatively, Cryo-electron microscopy can be used to investigate Tdm 3D structure without crystallisation (Baumeister and Steven 2000, Saibil 2000).

Chapter 6 Conclusions and future perspective

This Chapter discusses the principal findings of this project in the context of the main objectives outlined in Section 1.9, and highlights areas where further investigation is necessary.

The identification of *tdm* and TMAO transporter gene (*tmoP*) in *Methylocella silvestris* BL2 (**Aim 1**) was sought in **Chapter 3**. Marker-exchange mutagenesis enabled the identification of several genes involved in TMAO metabolism, including *Msil_3606*, a permease of the amino acids-polyamine (APC) superfamily, and *Msil_3603*, consisting a N-terminal domain of unknown function (DUF1989) and a C-terminal tetrahydrofolate-binding domain. Null mutants of *Msil_3603* and *Msil_3606* can no longer grow on TMAO. Thus it is hypothesised that *Msil_3603* and *Msil_3606* encode *tdm* and *tmoP*, respectively. *In vitro* enzyme assay of recombinant Msil_3603 further supported the idea that *Msil_3603* encodes Tdm. Purified Msil_3603 from recombinant *Escherichia coli* can convert TMAO to dimethylamine and formaldehyde (1 TMAO → 1 dimethylamine + 1 formaldehyde). Phylogenetic analysis of TmoP showed that it forms a distinct group from the other APC transporters. APC transporters are membrane permeases co-transporting another solute, acting as either a symporter or an antiporter (Saier 2000). It is not clear whether TmoP acts as a symporter or an antiporter and the co-transporting solute remains to be established.

TmoP homologues are also found in some methanogenic *Archaea*, *e.g. Methanosarcina acetivorans*, *Methanosarcina mazei*, and annotated as DMA transporter. It remained unclear whether these TmoP homologues are true DMA transporter. If not, it might suggest that these methanogenic *Archaea* can utilise TMAO directly as well. Methanogenic *Archaea* are able to degrade TMA, DMA and

MMA anaerobically by respective corrinoid-dependent methyltransferase (Burke et al 1998, Paul et al 2000, Wassenaar et al 1998).

TMAO has been shown to be an alternative electron acceptor among diverse bacteria when they encounter microaerobic or anoxic conditions (Barrett and Kwan 1985, Czjzek et al 1998, Sellars et al 2002) and a TMAO reductase operon (*torCAD*) has been identified about two decades ago (Méjean et al 1994). Recent study has shown that haloarchaea, *Halobacterium* sp. Strain NRC-1, can also use DMSO/TMAO as alternative electron acceptors but with a different reductase operon, *dmsREABCD* (Müller and DasSarma 2005). Therefore it is reasonable to expect that methanogenic *Archaea* can also use TMAO anaerobically as an electron acceptor. However, blasting *Methanosarcina acetivorans*, *Methanosarcina mazei* with *dmsREABCD* or *torCAD* as entries did not reveal any hit. As a result, it remains elusive whether TMAO can be used directly by methanogenic *Archaea*.

Genetic characterisation of Tdm (**Aim 2**) revealed that it consists of a N-terminal DUF1989 domain without assigned function and a C-terminal GCV_T domain. GCV_T domain has been found to play an important role in one-carbon unit metabolism in that is responsible for HCHO conjugation with THF. As a result, GCV_T domain is found in several enzymes catalysing the release of HCHO, including T protein of the glycine cleavage system, dimethylglycine and sarcosine dehydrogenase and dimethylsulfoniopropionate demethylase. Phylogenetic analysis of GCV_T domain of Tdm revealed that it falls into one of the previously recognized, but so far uncharacterized clades (Reisch et al 2008, Sun et al 2011). Initial biochemical and biophysical characterisation (**Aim 2**) demonstrated that Tdm is a hexamer, displays high substrate specificity that only catalyses demethylation of TMAO and a structural homologue, dimethyldodecylamine *N*-oxide, and each

monomer incorporates 1 equivalent of Zn^{2+} and 1 equivalent of Fe^{2+} as metal cofactors.

To elucidate the structure and functional mechanism of Tdm (**Aim 3**), site-directed mutagenesis, homology modelling and metal analyses by inorganic mass spectrometry have been applied in combination. Zn^{2+} is coordinated by a 3-sulfur-1-O motif. H276, and maybe H256, contribute to Fe^{2+} binding. An aspartate residue (D198) likely bridges Fe^{2+} and Zn^{2+} centres, suggesting that co-catalytic dinuclear centre. Site-directed mutagenesis of Tdm also led to the identification of three hydrophobic aromatic residues likely involved in substrate coordination (F259, Y305, W321), potentially through a cation-π interaction.

Tdm catalyses a unique O_2-independent demethylation. Therefore, it is hypothesised that TMAO has a dual role of being both a substrate and an oxygen donor for formaldehyde formation. The cross-over experiment using a substrate intermediate analogue supported the hypothesis by giving direct evidence of a TMA-alike intermediate produced during the Tdm catalytic cycle.

With the previous studies on haem and non-haem Fe^{2+}-containing enzymes, one can envisage that a high-valent iron-oxo complex (*e.g.* Fe(IV)-oxo, Fe(V)-oxo) is produced and responsible for C-H activation. As a result, it is postulated that TMAO acts as the oxygen donor to activate the Fe centre to form a putative high valent iron-oxo intermediate (*e.g.* Fe(IV)-oxo, Fe(V)-oxo) , which then hydroxylates the methyl group to give a carbinolamine intermediate that decomposes in water to form DMA and HCHO. Meanwhile, Tdm returns to its ferrous resting state.

In future, advanced spectroscopic methods such as electron paramagnetic resonance and Mössbauer spectroscopy (Rohde et al 2003, Tiago de Oliveira et al 2007) could

be used to determine the nature of the iron-containing active site and its redox state during catalysis.

Methylocella species show remarkable metabolic diversity, from one carbon compounds, methane, methanol, methylamines, to multicarbon compounds, acetate, succinate, malate, ethanol, and pyruvate (Dedysh et al., 2005). Crombie and Murrell (2014) recently described the simultaneous growth of the *M. silvestris* on methane and short-chain propane that has never been documented over a century of research into methanotroph. Identification and mechanism investigation of key enzymes involved in methylamine metabolism from *M. silvestris*, such as Tdm in this study, gives a valuable insight into its metabolic versatility.

By studying the mechanism of Tdm, it may shed some light on the function of DUF1989 domain (**Aim 4**). Since the well-studied GCV_T domain does not have any metal cofactor, it is believed that Zn^{2+} and Fe^{2+} are located in the DUF1989 domain. This is consistent with the hypothesis that DUF1989 cleaves HCHO, which was then transferred to the GCV_T domain where conjugation to THF takes place. The aforementioned homology modelling was performed using single DUF1989 domain protein structure available in PDB to give some insight into the structural conformation of DUF1989 domain of Tdm. Although it helped to identify 3 Cys that coordinate Zn^{2+} and a conserved Asp that probably links Zn^{2+} and Fe^{2+}, the overall quality of the established homology model was poor. In order to gain more accurate 3D-structure information of DUF1989 domain (**Aim 3**), C-terminal truncated mutants were constructed. However, the structure was disrupted indicated by the loss of metal cofactors and altered quaternary structure. Tdm of *Methylocella silvestris* and *Ruegeria pomeroyi* DSS-3 were screened for crystallisation. So far, the best Tdm crystals I have obtained are from Tdm of *Ruegeria pomeroyi* DSS-3, which are 2-

dimensional small plates. Further optimisation is required to achieve high quality crystals. Alternatively, Cryo-electron microscopy can be used to investigate Tdm 3D structure without crystallisation (Baumeister and Steven 2000).

Tdm homologues are widely distributed in marine bacterioplankton, from nutrient-enriched coastal water to oligotrophic open sea. Fe^{2+} and Zn^{2+} are regarded as the limiting micronutrients for the low primary productivity in some areas of the seawater (Anderson et al 1978, Behrenfeld and Kolber 1999, Carol and Timothy 1997, Moore et al 2013, Morel et al 1994). To investigate metal dependency of Tdm homologues from Zn^{2+} and/or Fe^{2+}-depleted environment (**Aim 5**), Tdm homologues from *R. pomeroyi* DSS-3 isolated from coastal water and oligotrophic *α*-proteobacterium HIMB59 were heterologously expressed in *E. coli*, purified and characterised. Tdm from *R. pomeroyi* DSS-3 displayed comparable activity as the one from *M. silvestris* implied by close K_m (3.19 ± 0.15 mM and 3.3 ± 0.64 mM, respectively) and V_{max} (13.71 ± 0.62 nmol min^{-1} mg^{-1} and 14.61 ± 1.13 nmol min^{-1} mg^{-1}, respectively). Preliminary metal determination suggested that Tdm from *R. pomeroyi* DSS-3 may still incorporate Fe^{2+}, but not Zn^{2+}. Further study is required to determine the precise metal composition and metal centres in Tdm of *R. pomeroyi* DSS-3. The yield of Tdm from the *α*-proteobacterium HIMB59 was too low to allow subsequent characterisation. Since TMAO is an important nutrient (C, N and energy source) for marine bacteria (Lidbury et al 2014, Lidbury et al 2015, Sun et al 2011), and the capability of using TMAO as energy source may explain the predominance of MRC and SAR11 clade, identification of metal cofactors in Tdm from marine bacteria will illuminate how bacteria adapt to micronutrient stress and establish the link between micronutrient cycling (*e.g.* trace metals) and C/N cycling.

In summary, this thesis makes a contribution to the understanding of indirect MAs metabolic pathway through the identification and mechanistic investigation of Tdm. The O_2-independent demethylation performed by Tdm gives a new insight into the generation of active high-valent iron-oxo complex from non-haem Fe^{2+}-containing enzyme. Furthermore, a novel TMAO transport, TmoP, has been identified, which is different to the previously known ABC-type TMAO transporter (Li et al 2015, Lidbury et al 2015, Raymond and Plopper 2002) and a putative Dmm has also been proposed.

References

Alberta JA, Dawson JH (1987). Purification to homogeneity and initial physical characterization of secondary amine monooxygenase. *J Biol Chem* **262:** 11857-11863.

Almeida J, Schobesberger S, Kurten A, Ortega I, Kupiainen-Maatta O, Praplan A *et al* (2013). Molecular understanding of sulphuric acid-amine particle nucleation in the atmosphere. *Nature* **502:** 359-+.

Andersen M, Vance D, Archer C, Anderson R, Ellwood M, Allen C (2011). The Zn abundance and isotopic composition of diatom frustules, a proxy for Zn availability in ocean surface seawater. *Earth Planet Sci Lett* **301:** 137-145.

Anderson M, Morel F, Guillard R (1978). Growth limitation of a coastal diatom by low zinc ion activity. *Nature* **276:** 70-71.

Anthony C (1982). *The biochemistry of methylotrophs*. Academic Press: London, United Kingdom.

Arumugam M, Raes J, Pelletier E, Le Paslier D, Yamada T, Mende DR *et al* (2011). Enterotypes of the human gut microbiome. *Nature* **473:** 174-180.

Arnold K, Bordoli L, Kopp J, Schwede T (2006). The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. *Bioinformatics* **22:** 195-201.

Auld DS (2001). Zinc coordination sphere in biochemical zinc sites. *Biometals* **14:** 271-313.

Back TG, Dyck BP (1997). A novel camphor-derived selenenamide that acts as a glutathione peroxidase mimetic. *J Am Chem Soc* **119:** 2079-2083.

Balland V, Charlot M, Banse F, Girerd J, Mattioli T, Bill E *et al* (2004). Spectroscopic characterization of an Fe-IV intermediate generated by reaction of XO- (X = Cl, Br) with an Fe-II complex bearing a pentadentate non-porphyrinic ligand - Hydroxylation and epoxidation activity. *Eur J Inorg Chem* **2004:** 301-308.

Barrett EL, Kwan HS (1985). Bacterial reduction of trimethylamine oxide. *Annu Rev Microbiol* **39:** 131-149.

Bateman A, Coggill P, Finn R (2010). DUFs: families in search of function. *Acta Crystallographica Section F-Structural Biology and Crystallization Communications* **66:** 1148-1152.

Baumeister W, Steven A (2000). Macromolecular electron microscopy in the era of structural genomics. *Trends Biochem Sci* **25:** 624-631.

Behrenfeld M, Kolber Z (1999). Widespread iron limitation of phytoplankton in the South Pacific Ocean. *Science* **283:** 840-843.

Berntsson RP, Smits SH, Schmitt L, Slotboom DJ, Poolman B (2010). A structural classification of substrate-binding proteins. *FEBS Lett* **584:** 2606-2617.

Benkert P, Biasini M, Schwede T (2011). Toward the estimation of the absolute quality of individual protein structure models. *Bioinformatics* **27:** 343-350.

Berg JM (2015). *Biochemistry*, Eighth edn. W.H. Freeman and Company: New York.

Bhadbhade BJ, Sarnaik SS, Kanekar PP (2002). Biomineralization of an organophosphorus pesticide, Monocrotophos, by soil bacteria. *J Appl Microbiol* **93:** 224-234.

Blasiak LC, Vaillancourt FH, Walsh CT, Drennan CL (2006). Crystal structure of the non-haem iron halogenase SyrB2 in syringomycin biosynthesis. *Nature* **440:** 368-371.

Bockus A, Seibel B (2016). Trimethylamine oxide accumulation as a function of depth in Hawaiian mid-water fishes. *Deep-Sea Research Deep-Sea Rea Pt I* **112:** 37-44.

Boyd G, Mathews F, Packman L, Scrutton N (1992). Trimethylamine dehydrogenase of bacterium W(3)A(1): molecular-cloning, sequence determination and over-expression of the gene. *FEBS Letters* **308:** 271-276.

Brand L, Sunda W, Guillard R (1983). Limitation of marine-phytoplankton reproductive rates by zinc, manganese, and iron. *Limnol Oceanogr* **28:** 1182-1198.

Brown AD (1976). Microbial water stress. *Bacteriol Rev* **40:** 803-846.

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

Bugg T (2012). *Introduction to enzyme and coenzyme chemistry*, 3rd edn. Wiley: Chichester, West Sussex.

Bugg TD (2001). Oxygenases: mechanisms and structural motifs for O_2 activation. *Curr Opin Chem Biol* **5:** 550-555.

Canchi DR, Jayasimha P, Rau DC, Makhatadze GI, Garcia AE (2012). Molecular mechanism for the preferential exclusion of TMAO from protein surfaces. *J Phys Chem B* **116:** 12095-12104.

Canchi DR, García AE (2013). Cosolvent effects on protein stability. *Annu Rev Phys Chem* **64:** 273-293.

Cape JN, Cornell SE, Jickells TD, Nemitz E (2011). Organic nitrogen in the atmosphere - Where does it come from? A review of sources and methods. *Atmos Res* **102:** 30-48.

Carol LM, Timothy PR (1997). *Biological Oceanography: An Introduction*, 2nd ed. Oxford : Butterworth-Heinemann.

Chai SC, Lu JP, Ye QZ (2009). Determination of binding affinity of metal cofactor to the active site of methionine aminopeptidase based on quantitation of functional enzyme. *Anal Biochem* **395:** 263-264.

Chang B, Chen Y, Zhao Y, Bruick RK (2007). JMJD6 is a histone arginine demethylase. *Science* **318:** 444-447.

Chen Y, Crombie A, Rahman MT, Dedysh SN, Liesack W, Stott MB *et al* (2010a). Complete genome sequence of the aerobic facultative methanotroph *Methylocella silvestris* BL2. *J Bacteriol* **192:** 3840-3841.

Chen Y, McAleer KL, Murrell JC (2010b). Monomethylamine as a nitrogen source for a nonmethylotrophic bacterium, *Agrobacterium tumefaciens*. *Appl Environ Microbiol* **76:** 4102-4104.

Chen Y, Scanlan J, Song L, Crombie A, Rahman MT, Schäfer H *et al* (2010c). {gamma}-Glutamylmethylamide is an essential intermediate in the metabolism of methylamine by Methylocella silvestris. *Appl Environ Microbiol* **76:** 4530-4537.

Chen Y, Patel NA, Crombie A, Scrivens JH, Murrell JC (2011). Bacterial flavin-containing monooxygenase is trimethylamine monooxygenase. *Proc Natl Acad Sci U S A* **108:** 17791-17796.

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.

Chester R (1990). *Marine Geochemistry*. Springer Netherlands.

Chiavarino B, Cipollini R, Crestoni ME, Fornarini S, Lanucara F, Lapi A (2008). Probing the Compound I-like reactivity of a bare high-valent oxo iron porphyrin complex: the oxidation of tertiary amines. *J Am Chem Soc* **130:** 3208-3217.

Chistoserdov A, Chistoserdov L, Mcintire W, Lidstrom M (1994). Genetic organisation of the mau gene-cluster in Methylobacterium extorquens AM1: complete nucleotide sequence and generation and characterisation of mau mutants. *J Bacteriol* **176:** 4052-4065.

Chistoserdova L, Kalyuzhnaya M, Lidstrom M (2009). The Expanding World of Methylotrophic Metabolism. *Annu Rev Microbiol* **63:** 477-499.

Choi KH, Schweizer HP (2005). An improved method for rapid generation of unmarked *Pseudomonas aeruginosa* deletion mutants. *BMC Microbiol* **5:** 30.

Chung SJ, Fromme JC, Verdine GL (2005). Structure of human cytidine deaminase bound to a potent inhibitor. *J Med Chem* **48:** 658-660.

Church M, Hutchins D, Ducklow H (2000). Limitation of bacterial growth by dissolved organic matter and iron in the Southern Ocean. *Appl Environ Microbiol* **66:** 455-466.

Coale K, Wang X, Tanner S, Johnson K (2003). Phytoplankton growth and biological response to iron and zinc addition in the Ross Sea and Antarctic Circumpolar Current along 170 degrees W. *Deep-Sea Research Part Ii-Topical Studies in Oceanography* **50:** 635-653.

Coale KH, Johnson KS, Fitzwater SE, Gordon RM, Tanner S, Chavez FP *et al* (1996). A massive phytoplankton bloom induced by an ecosystem-scale iron fertilization experiment in the equatorial Pacific Ocean. *Nature* **383:** 495-501.

Costas M, Mehn MP, Jensen MP, Que L (2004). Dioxygen activation at mononuclear nonheme iron active sites: enzymes, models, and intermediates. *Chem Rev* **104:** 939-986.

Craciun S, Marks JA, Balskus EP (2014). Characterization of choline trimethylamine-lyase expands the chemistry of glycyl radical enzymes. *ACS Chem Biol* **9:** 1408-1413.

Crawford D, Lipsen M, Purdie D, Lohan M, Statham P, Whitney F *et al* (2003). Influence of zinc and iron enrichments on phytoplankton growth in the northeastern subarctic Pacific. *Limnol Oceanogr* **48:** 1583-1600.

Crombie A (2011). Metabolism of methane and propane and the role of the glyoxylate bypass enzymes in *Methylocella silvestris* BL2. **Doctor of Philosophy** thesis, University of Warwick

Crombie A, & Murrell C. (2014). Trace-gas metabolic versatility of the facultative methanotroph Methylocella silvestris. *Nature*, *510*, 148-151.

da Silva FJJR, Williams RJP (2001). *The Biological Chemistry of the Elements: The Inorganic Chemistry Of Life*, 2nd ed. Oxford: Oxford University Press.

Dam J, Velikovsky CA, Mariuzza RA, Urbanke C, Schuck P (2005). Sedimentation velocity analysis of heterogeneous protein-protein interactions: Lamm equation modelling and sedimentation coefficient distributions c(s). *Biophys J* **89:** 619-634.

Daughtry KD, Xiao Y, Stoner-Ma D, Cho E, Orville AM, Liu P *et al* (2012). Quaternary Ammonium Oxidative Demethylation: X-ray Crystallographic, Resonance Raman, and UV-Visible Spectroscopic Analysis of a Rieske-Type Demethylase. *J Am Chem Soc* **134:** 2823-2834.

Daumann LJ, McCarthy BY, Hadler KS, Murray TP, Gahan LR, Larrabee JA *et al* (2013). Promiscuity comes at a price: catalytic versatility vs efficiency in different metal ion derivatives of the potential bioremediator GpdQ. *Biochim Biophys Acta* **1834:** 425-432.

Davidson AL, Chen J (2004). ATP-binding cassette transporters in bacteria. *Annu Rev Biochem* **73:** 241-268.

Dedysh S, Liesack W, Khmelenina V, Suzina N, Trotsenko Y, Semrau J *et al* (2000). *Methylocella palustris* gen. nov., sp nov., a new methane-oxidizing acidophilic bacterium from peat bags, representing a novel subtype of serine-pathway methanotrophs. *Int J Syst Evol Microbiol* **50:** 955-969.

Dedysh S, Berestovskaya Y, Vasylieva L, Belova S, Khmelenina V, Suzina N *et al* (2004). *Methylocella tundrae* sp nov., a novel methanotrophic bacterium from acidic tundra peatlands. *Int J Syst Evol Microbiol* **54:** 151-156.

Dedysh S, Knief C, Dunfield P (2005). *Methylocella* species are facultatively methanotrophic. *J Bacteriol* **187:** 4665-4670.

DeLong EF (2005). Microbial community genomics in the ocean. *Nat Rev Microbiol* **3:** 459-469.

Delorme C, Huisman TT, Reijnders WN, Chan YL, Harms N, Stouthamer AH *et al* (1997). Expression of the *mau* gene cluster of *Paracoccus denitrificans* is controlled by MauR and a second transcription regulator. *Microbiology* **143 (Pt 3):** 793-801.

Diebold AR, Neidig ML, Moran GR, Straganz GD, Solomon EI (2010). The three-his triad in Dke1: comparisons to the classical facial triad. *Biochemistry* **49:** 6945-6952.

Dougherty DA (1996). Cation-pi interactions in chemistry and biology: a new view of benzene, Phe, Tyr, and Trp. *Science* **271:** 163-168.

Dowers TS, Rock DA, Jones JP (2004). Kinetic isotope effects implicate the iron-oxene as the sole oxidant in P450-catalyzed *N*-dealkylation. *J Am Chem Soc* **126:** 8868-8869.

Drake AW, Klakamp SL (2007). A rigorous multiple independent binding site model for determining cell-based equilibrium dissociation constants. *J Immunol Methods* **318:** 147-152.

Dunfield PF, Khmelenina VN, Suzina NE, Trotsenko YA, Dedysh SN (2003). *Methylocella silvestris* sp. nov., a novel methanotroph isolated from an acidic forest cambisol. *Int J Syst Evol Microbiol* **53:** 1231-1239.

Edgar R (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* **32:** 1792-1797.

Emerson JP, Kovaleva EG, Farquhar ER, Lipscomb JD, Que L (2008). Swapping metals in Fe- and Mn-dependent dioxygenases: evidence for oxygen activation without a change in metal redox state. *Proc Natl Acad Sci U S A* **105:** 7347-7352.

Ellwood M, Hunter K (2000). The incorporation of zinc and iron into the frustule of the marine diatom Thalassiosira pseudonana. *Limnol Oceanogr* **45:** 1517-1524.

Erupe M, Viggiano A, Lee S (2011). The effect of trimethylamine on atmospheric nucleation involving H_2SO_4. *Atmos Chem Phys* **11:** 4767-4775.

Facchini M, Decesari S, Rinaldi M, Carbone C, Finessi E, Mircea M et al (2008). Important source of marine secondary organic aerosol from biogenic amines. *Environ Sci Technol* **42:** 9116-9121.

Fu XY, Xue CH, Miao BC, Liang JN, Li ZJ, Cui FX (2006). Purification and characterization of trimethylamine-*N*-oxide demethylase from jumbo squid (*Dosidicus gigas*). *J Agric Food Chem* **54:** 968-972.

Gallivan JP, Dougherty DA (1999). Cation-pi interactions in structural biology. *Proc Natl Acad Sci U S A* **96:** 9459-9464.

Galperin MY, Koonin EV (2010). From complete genome sequence to 'complete' understanding? *Trends Biotechnol* **28:** 398-406.

Gantt SL, Gattis SG, Fierke CA (2006). Catalytic activity and inhibition of human histone deacetylase 8 is dependent on the identity of the active site metal ion. *Biochemistry* **45:** 6170-6178.

Gardner JD, Pierce BS, Fox BG, Brunold TC (2010). Spectroscopic and computational characterization of substrate-bound mouse cysteine dioxygenase: nature of the ferrous and ferric cysteine adducts and mechanistic implications. *Biochemistry* **49:** 6033-6041.

Ge X, Wexler AS, Clegg SL (2011). Atmospheric amines - Part I. A review. *Atmos Environ* **45:** 524-546.

Gibb SW, Hatton AD (2004). The occurrence and distribution of trimethylamine-*N*-oxide in Antarctic coastal waters. *Mar Chem* **91:** 65-75.

Giebel HA, Kalhoefer D, Lemke A, Thole S, Gahl-Janssen R, Simon M *et al* (2011). Distribution of *Roseobacter* RCA and SAR11 lineages in the North Sea and characteristics of an abundant RCA isolate. *ISME J* **5:** 8-19.

Gillett M, Suko J, Santoso F, Yancey P (1997). Elevated levels of trimethylamine oxide in muscles of deep-sea gadiform teleosts: A high-pressure adaptation? *J Exp Zool* **279:** 386-391.

Glaeser R, Facciotti M, Walian P, Rouhani S, Holton J, MacDowell A *et al* (2000). Characterization of conditions required for x-ray diffraction experiments with protein microcrystals. *Biophys J* **78:** 3178-3185.

Gruffaz C, Muller E, Louhichi-Jelail Y, Nelli Y, Guichard G, Bringel F (2014). Genes of the *N*-methylglutamate pathway are essential for growth of *methylobacterium extorquens* DM4 with Monomethylamine. *Appl Environ Microbiol* **80:** 3541-3550.

Guengerich FP, Yun CH, Macdonald TL (1996). Evidence for a 1-electron oxidation mechanism in *N*-dealkylation of *N,N*-dialkylanilines by cytochrome P450 2B1 - Kinetic hydrogen isotope effects, linear free energy relationships, comparisons with horseradish peroxidase, and studies with oxygen surrogates. *J Biol Chem* **271:** 27321-27329.

Halsey KH, Carter AE, Giovannoni SJ (2012). Synergistic metabolism of a broad range of C1 compounds in the marine methylotrophic bacterium HTCC2181. *Environ Microbiol* **14:** 630-640.

Harding MM (2001). Geometry of metal-ligand interactions in proteins. *Acta Crystallogr D Biol Crystallogr* **57:** 401-411.

Harel M, Schalk I, Ehret-Sabatier L, Bouet F, Goeldner M, Hirth C *et al* (1993). Quaternary ligand binding to aromatic residues in the active-site gorge of acetylcholinesterase. *Proc Natl Acad Sci U S A* **90:** 9031-9035.

Hegg EL, Que L (1997). The 2-His-1-carboxylate facial triad--an emerging structural motif in mononuclear non-heme iron(II) enzymes. *Eur J Biochem* **250:** 625-629.

Hrycay EG, Bandiera SM (2015). Monooxygenase, peroxidase and peroxygenase properties and reaction mechanisms of cytochrome P450 enzymes. *Adv Exp Med Biol* **851:** 1-61.

Hua L, Zhou R, Thirumalai D, Berne BJ (2008). Urea denaturation by stronger dispersion interactions with proteins than water implies a 2-stage unfolding. *Proc Natl Acad Sci U S A* **105:** 16928-16933.

Husain M, Davidson VL (1987). Purification and properties of methylamine dehydrogenase from Paracoccus denitrificans. *J Bacteriol* **169:** 1712-1717.

Jakuba R, Saito M, Moffett J, Xu Y (2012). Dissolved zinc in the subarctic North Pacific and Bering Sea: Its distribution, speciation, and importance to primary producers. *Global Biogeochem Cy* **26**.

Jiang W, Wilson MA, Weeks DP (2013). O-Demethylations catalyzed by Rieske nonheme iron monooxygenases involve the difficult oxidation of a saturated C-H bond. *ACS Chem Biol* **8:** 1687-1691.

Kalyuzhnaya MG, De Marco P, Bowerman S, Pacheco CC, Lara JC, Lidstrom ME *et al* (2006). *Methyloversatilis universalis* gen. nov., sp. nov., a novel taxon within the Betaproteobacteria represented by three methylotrophic isolates. *Int J Syst Evol Microbiol* **56:** 2517-2522.

Kamanavalli CM, Ninnekar HZ (2000). Biodegradation of propoxur by *Pseudomonas* species. *World J Microb Biot* **16:** 329-331.

Kamiya A, Ose Y (1984). Study of odorous compounds produced by putrefaction of foods: V. Fatty acids, sulfur compounds and amines. *J Chromatogr* **292:** 383-391.

Kanamori T, Kanou N, Atomi H, Imanaka T (2004). Enzymatic characterization of a prokaryotic urea carboxylase. *J Bacteriol* **186:** 2532-2539.

Kelly RH, Yancey PH (1999). High contents of trimethylamine oxide correlating with depth in deep-sea teleost fishes, skates, and decapod crustaceans. *Biol Bull* **196:** 18-25.

Ken CF, Hsiung TM, Huang ZX, Juang RH, Lin CT (2005). Characterization of Fe/Mn-superoxide dismutase from diatom Thallassiosira weissflogii: cloning, expression, and property. *J Agric Food Chem* **53:** 1470-1474.

Kimura M, Seki N, Kimura I (2000). Purification and characterization of trimethylamine-*N*-oxide demethylase from walleye pollack muscle. *Fish Sci* **66:** 967-973.

King GM (1987). Methane production from methylated amines in hypersaline sediments. *Abstr Pap Am Chem S* **193:** 17-GEOC.

Kloer DP, Ruch S, Al-Babili S, Beyer P, Schulz GE (2005). The structure of a retinal-forming carotenoid oxygenase. *Science* **308:** 267-269.

Koehntop KD, Emerson JP, Que L (2005). The 2-His-1-carboxylate facial triad: a versatile platform for dioxygen activation by mononuclear non-heme iron(II) enzymes. *J Biol Inorg Chem* **10:** 87-93.

Kovaleva EG, Lipscomb JD (2008). Versatility of biological non-heme Fe(II) centers in oxygen activation reactions. *Nat Chem Biol* **4:** 186-193.

Krebs C, Galonić Fujimori D, Walsh CT, Bollinger JM (2007). Non-heme Fe(IV)-oxo intermediates. *Acc Chem Res* **40:** 484-492.

Kress D, Alhapel A, Pierik AJ, Essen LO (2008). The crystal structure of enamidase: a bifunctional enzyme of the nicotinate catabolism. *J Mol Biol* **384:** 837-847.

Kumasaka T, Yamamoto M, Furuichi M, Nakasako M, Teh AH, Kimura M *et al* (2007). Crystal structures of blasticidin S deaminase (BSD): implications for dynamic properties of catalytic zinc. *J Biol Chem* **282:** 37103-37111.

Labbe RF, Rettmer RL (1989). Zinc protoporphyrin: a product of iron-deficient erythropoiesis. *Semin Hematol* **26:** 40-46.

Landry M, Barber R, Bidigare R, Chai F, Coale K, Dam H *et al* (1997). Iron and grazing constraints on primary production in the central equatorial Pacific: An EqPac synthesis. *Limnol Oceanogr* **42:** 405-418.

Lane T, Morel F (2000). Regulation of carbonic anhydrase expression by zinc, cobalt, and carbon dioxide in the marine diatom Thalassiosira weissflogii. *Plant Physiol* **123:** 345-352.

Large PJ (1971). Non-oxidative demethylation of trimethylamine *N*-oxide by *Pseudomonas aminovorans*. *FEBS Lett* **18:** 297-300.

LaRoche J, Boyd P, McKay R, Geider R (1996). Flavodoxin as an in situ marker for iron stress in phytoplankton. *Nature* **382:** 802-805.

Latypova E, Yang S, Wang Y, Wang T, Chavkin T, Hackett M *et al* (2010). Genetics of the glutamate-mediated methylamine utilization pathway in the facultative methylotrophic beta-proteobacterium *Methyloversatilis universalis* FAM5. *Mol Microbiol* **75:** 426-439.

Laxson C, Condon N, Drazen J, Yancey P (2011). Decreasing urea: trimethylamine N-Oxide ratios with depth in chondrichthyes: a physiological depth limit? *Physiol Biochem Zool* **84:** 494-505.

Levitch ME (1976). The demonstration of two discrete enzymes catalyzing the synthesis of glutamine and gamma-glutamylmethylamide in *Pseudomonas* MS. *Biochem Biophys Res Commun* **76:** 609-614.

Li CY, Chen XL, Shao X, Wei TD, Wang P, Xie BB *et al* (2015). Mechanistic insight into trimethylamine *N*-oxide recognition by the marine bacterium *Ruegeria pomeroyi* DSS-3. *J Bacteriol* **197:** 3378-3387.

Lidbury I, Murrell JC, Chen Y (2014). Trimethylamine *N*-oxide metabolism by abundant marine heterotrophic bacteria. *Proc Natl Acad Sci U S A* **111:** 2710-2715.

Lidbury I, Kröber E, Zhang Z, Zhu Y, Murrell JC, Chen Y *et al* (2016). A mechanism for bacterial transformations of DMS to DMSO: A missing link in the marine organic sulfur cycle. *Environ Microbiol* doi:10.1111/1462-2920.13354

Lidbury ID, Murrell JC, Chen Y (2015). Trimethylamine and trimethylamine *N*-oxide are supplementary energy sources for a marine heterotrophic bacterium: implications for marine carbon and nitrogen cycling. *ISME J* **9:** 760-769.

Lohan M, Statham P, Crawford D (2002). Total dissolved zinc in the upper water column of the subarctic North East Pacific. *Deep-Sea Res Pt II* **49:** 5793-5808.

Ma J, Pazos I, Gai F (2014). Microscopic insights into the protein-stabilizing effect of trimethylamine *N*-oxide (TMAO). *Proc Natl Acad Sci U S A* **111:** 8476-8481.

Makris TM, von Koenig K, Schlichting I, Sligar SG (2006). The status of high-valent metal oxo complexes in the P450 cytochromes. *J Inorg Biochem* **100:** 507-518.

Malmstrom R, Kiene R, Cottrell M, Kirchman D (2004). Contribution of SAR11 bacteria to dissolved dimethylsulfoniopropionate and amino acid uptake in the North Atlantic ocean. *Appl Environ Microbiol* **70:** 4129-4135.

Martin D, Bartlett D, Roberts M (2002). Solute accumulation in the deep-sea bacterium *Photobacterium profundum*. *Extremophiles* **6:** 507-514.

Martin J, Coale K, Johnson K, Fitzwater S, Gordon R, Tanner S *et al* (1994). Testing the iron hypothesis in ecosystems of the equatorial pacific-ocean. *Nature* **371:** 123-129.

Marx CJ, Lidstrom ME (2002). Broad-host-range cre-lox system for antibiotic marker recycling in gram-negative bacteria. *Biotechniques* **33:** 1062-1067.

Mawji E, Schlitzer R, Dodas E, Abadie C, Abouchami W, Anderson R *et al* (2015). The GEOTRACES Intermediate Data Product 2014. *Mar Chem* **177:** 1-8.

McClelland M, Sanderson KE, Spieth J, Clifton SW, Latreille P, Courtney L *et al* (2001). Complete genome sequence of *Salmonella enterica* serovar typhimurium LT2. *Nature* **413:** 852-856.

McCoy JG, Bailey LJ, Bitto E, Bingman CA, Aceti DJ, Fox BG *et al* (2006). Structure and mechanism of mouse cysteine dioxygenase. *Proc Natl Acad Sci U S A* **103:** 3084-3089.

Mello C, Barrick D (2003). Measuring the stability of partly folded proteins using TMAO. *Protein Sci* **12:** 1522-1529.

Meyer M, Granderath K, Andreesen JR (1995). Purification and characterization of protein PB of betaine reductase and its relationship to the corresponding proteins glycine reductase and sarcosine reductase from *Eubacterium acidaminophilum*. *Eur J Biochem* **234:** 184-191.

Mishina Y, He C (2006). Oxidative dealkylation DNA repair mediated by the mononuclear non-heme iron AlkB proteins. *J Inorg Biochem* **100:** 670-678.

Mitch WA, Sharp JO, Trussell RR, Valentine RL, Alvarez-Cohen L, Sedlak DL (2003). *N*-nitrosodimethylamine (NDMA) as a drinking water contaminant: A review. *Environ Eng Sci* **20:** 389-404.

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

Moore JO, Hendrickson WA (2012). An asymmetry-to-symmetry switch in signal transmission by the histidine kinase receptor for TMAO. *Structure* **20:** 729-741.

Morel F, Reinfelder J, Roberts S, Chamberlain C, Lee J, Yee D (1994). Zinc and carbon co-limitation of marine-phytoplankton. *Nature* **369:** 740-742.

Morel F, Price N (2003). The biogeochemical cycles of trace metals in the oceans. *Science* **300:** 944-947.

Muller C, Iinuma Y, Karstensen J, van Pinxteren D, Lehmann S, Gnauk T *et al* (2009). Seasonal variation of aliphatic amines in marine sub-micrometer particles at the Cape Verde islands. *Atmos Chem Phys* **9:** 9587-9597.

Murphy SM, Sorooshian A, Kroll JH, Ng NL, Chhabra P, Tong C *et al* (2007). Secondary aerosol formation from atmospheric reactions of aliphatic amines. *Atmos Chem Phys* **7:** 2313-2337.

Myers PA, Zatman LJ (1971). The metabolism of trimethylamine *N*-oxide by *Bacillus* PM6. *Biochem J* **121:** 10P.

Namgaladze D, Hofer HW, Ullrich V (2002). Redox control of calcineurin by targeting the binuclear Fe^{2+}-Zn^{2+} center at the enzyme active site. *J Biol Chem* **277:** 5962-5969.

Nayak D, Marx C (2014). Methylamine utilization via the *N*-methylglutamate pathway in *Methylobacterium extorquens* PA1 involves a novel flow of carbon through C-1 assimilation and dissimilation pathways. *J Bacteriol* **196:** 4130-4139.

Nehru K, Seo MS, Kim J, Nam W (2007). Oxidative *N*-dealkylation reactions by oxoiron(IV) complexes of nonheme and heme ligands. *Inorg Chem* **46:** 293-298.

Neri F, Indiani C, Welinder KG, Smulevich G (1998). Mutation of the distal arginine in *Coprinus cinereus* peroxidase--structural implications. *Eur J Biochem* **251:** 830-838.

Okamura-Ikeda K, Hosaka H, Yoshimura M, Yamashita E, Toma S, Nakagawa A *et al* (2005). Crystal structure of human T-protein of glycine cleavage system at 2.0 Å resolution and its implication for understanding non-ketotic hyperglycinemia. *J Mol Biol* **351:** 1146-1159.

Oswald C, Smits SH, Höing M, Sohn-Bösser L, Dupont L, Le Rudulier D *et al* (2008). Crystal structures of the choline/acetylcholine substrate-binding protein ChoX from *Sinorhizobium meliloti* in the liganded and unliganded-closed states. *J Biol Chem* **283:** 32848-32859.

Parkin KL, Hultin HO (1986). Characterization of trimethylamine-*N*-oxide (TMAO) demethylase activity from fish muscle microsomes. *J Biochem* **100:** 77-86.

Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC *et al* (2004). UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem* **25:** 1605-1612.

Price N, Morel F (1990). Cadmium and cobalt substitution for zinc in a marine diatom. *Nature* **344:** 658-660.

Punta M, Coggill PC, Eberhardt RY, Mistry J, Tate J, Boursnell C *et al* (2012). The Pfam protein families database. *Nucleic Acids Res* **40:** D290-301.

Que L, Ho RY (1996). Dioxygen activation by enzymes with mononuclear non-heme iron active sites. *Chem Rev* **96:** 2607-2624.

Quesenberry MS, Lee YC (1996). A rapid formaldehyde assay using purpald reagent: application under periodation conditions. *Anal Biochem* **234:** 50-55.

Ramanathan V, Crutzen P, Kiehl J, Rosenfeld D (2001). Atmosphere - Aerosols, climate, and the hydrological cycle. *Science* **294:** 2119-2124.

Ranguelova K, Mason RP (2011). The fidelity of spin trapping with DMPO in biological systems. *Magn Reson Chem* **49:** 152-158.

Rappé MS, Giovannoni SJ (2003). The uncultured microbial majority. *Annu Rev Microbiol* **57:** 369-394.

Raymond JA, Plopper GE (2002). A bacterial TMAO transporter. *Comp Biochem Physiol B Biochem Mol Biol* **133:** 29-34.

R Core Team (2015). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Reisch CR, Moran MA, Whitman WB (2008). Dimethylsulfoniopropionate-dependent demethylase (DmdA) from *Pelagibacter ubique* and *Silicibacter pomeroyi*. *J Bacteriol* **190:** 8018-8024.

Richards SE, Wang Y, Claus SP, Lawler D, Kochhar S, Holmes E *et al* (2013). Metabolic phenotype modulation by caloric restriction in a lifelong dog study. *J Proteome Res* **12:** 3117-3127.

Roberts KM, Jones JP (2010). Anilinic *N*-oxides support cytochrome P450-mediated *N*-dealkylation through hydrogen-atom transfer. *Chemistry* **16:** 8096-8107.

Roesser M, Müller V (2001). Osmoadaptation in bacteria and archaea: common principles and differences. *Environ Microbiol* **3:** 743-754.

Rohde JU, In JH, Lim MH, Brennessel WW, Bukowski MR, Stubna A *et al* (2003). Crystallographic and spectroscopic characterization of a nonheme Fe(IV)-O complex. *Science* **299:** 1037-1039.

Rosgen J, Jackson-Atogi R (2012). Volume exclusion and H-bonding dominate the thermodynamics and solvation of trimethylamine-*N*-oxide in aqueous urea. *J Am Chem Soc* **134:** 3590-3597.

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

Saager P, Debaar H, Burkill P (1989). Manganese and iron in indian-ocean waters. *Geochim Cosmochim Acta* **53:** 2259-2267.

Saibil H (2000). Macromolecular structure determination by cryo-electron microscopy. *Acta Crystallographica Section D-Biological Crystallography* **56:** 1215-1222.

Saier MH (2000). Families of transmembrane transporters selective for amino acids and their derivatives. *Microbiology* **146 (Pt 8):** 1775-1795.

Saitou N, Nei M (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* **4:** 406-425.

Samerotte AL, Drazen JC, Brand GL, Seibel BA, Yancey PH (2007). Correlation of trimethylamine oxide and habitat depth within and among species of teleost fish: an analysis of causation. *Physiol Biochem Zool* **80:** 197-208.

Schiefner A, Holtmann G, Diederichs K, Welte W, Bremer E (2004). Structural basis for the binding of compatible solutes by ProX from the hyperthermophilic archaeon *Archaeoglobus fulgidus*. *J Biol Chem* **279:** 48270-48281.

Schiødt CB, Veitch NC, Welinder KG (2007). Roles of distal arginine in activity and stability of *Coprinus cinereus* peroxidase elucidated by kinetic and NMR analysis of the Arg51Gln, -Asn, -Leu, and -Lys mutants. *J Inorg Biochem* **101:** 336-347.

Schuck P (2000). Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modelling. *Biophys J* **78:** 1606-1619.

Schuller DJ, Reisch CR, Moran MA, Whitman WB, Lanzilotta WN (2012). Structures of dimethylsulfoniopropionate-dependent demethylase from the marine organism *Pelagabacter ubique*. *Protein Sci* **21:** 289-298.

Seibel BA, Walsh PJ (2002). Trimethylamine oxide accumulation in marine animals: relationship to acylglycerol storage. *J Exp Biol* **205:** 297-306.

Seo MS, In JH, Kim SO, Oh NY, Hong J, Kim J *et al* (2004). Direct evidence for oxygen-atom exchange between nonheme oxoiron(IV) complexes and isotopically labeled water. *Angew Chem Int Ed Engl* **43:** 2417-2420.

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

Shan X, Que L (2006). High-valent nonheme iron-oxo species in biomimetic oxidations. *J Inorg Biochem* **100:** 421-433.

Shaw WV, Tsai L, Stadtman ER (1966). The enzymatic synthesis of N-methylglutamic acid. *J Biol Chem* **241:** 935-945.

Singh R, Haque I, Ahmad F (2005). Counteracting osmolyte trimethylamine *N*-oxide destabilizes proteins at pH below its pKa. Measurements of thermodynamic parameters of proteins in the presence and absence of trimethylamine N-oxide. *J Biol Chem* **280:** 11035-11042.

Singh R, Grigg JC, Armstrong Z, Murphy ME, Eltis LD (2012). Distal heme pocket residues of B-type dye-decolorizing peroxidase: arginine but not aspartate is essential for peroxidase activity. *J Biol Chem* **287:** 10623-10630.

Sousa S, Fernandes P, Ramos M (2007). The carboxylate shift in zinc enzymes: A computational study. *J Am Chem Soc* **129:** 1378-1385.

Smith JL, Wishnok JS, Deen WM (1994). Metabolism and excretion of methylamines in rats. *Toxicol Appl Pharmacol* **125:** 296-308.

Sreerama N, Woody RW (2000). Estimation of protein secondary structure from circular dichroism spectra: comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set. *Anal Biochem* **287:** 252-260.

Stallings WC, Pattridge KA, Strong RK, Ludwig ML (1984). Manganese and iron superoxide dismutases are structural homologues. *J Biol Chem* **259:** 10695-10699.

Street TO, Bolen DW, Rose GD (2006). A molecular mechanism for osmolyte-induced protein stability. *Proc Natl Acad Sci U S A* **103:** 13997-14002.

Summers RM, Louie TM, Yu CL, Gakhar L, Louie KC, Subramanian M (2012). Novel, highly specific *N*-demethylases enable bacteria to live on caffeine and related purine alkaloids. *J Bacteriol* **194:** 2041-2049.

Sun J, Steindler L, Thrash JC, Halsey KH, Smith DP, Carter AE *et al* (2011). One carbon metabolism in SAR11 pelagic marine bacteria. *PLoS One* **6:** e23973.

Suzuki M, Preston C, Chavez F, DeLong E (2001). Quantitative mapping of bacterioplankton populations in seawater: field tests across an upwelling plume in Monterey Bay. *Aquat Microb Ecol* **24:** 117-127.

Sánchez-Quitian ZA, Schneider CZ, Ducati RG, de Azevedo WF, Bloch C, Basso LA *et al* (2010). Structural and functional analyses of *Mycobacterium tuberculosis* Rv3315c-encoded metal-dependent homotetrameric cytidine deaminase. *J Struct Biol* **169:** 413-423.

Tagliabue A, Sallee J, Bowie A, Levy M, Swart S, Boyd P (2014). Surface-water iron supplies in the Southern Ocean sustained by deep winter mixing. *Nat Geosci* **7:** 314-320.

Takeuchi K, Hatanaka A, Kimura M, Seki N, Kimura I, Yamada S *et al* (2003). Aspolin, a novel extremely aspartic acid-rich protein in fish muscle, promotes iron-mediated demethylation of trimethylamine-*N*-oxide. *J Biol Chem* **278:** 47416-47422.

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S (2011). MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* **28:** 2731-2739.

Tang WH, Wang Z, Fan Y, Levison B, Hazen JE, Donahue LM *et al* (2014). Prognostic value of elevated levels of intestinal microbe-generated metabolite trimethylamine-N-oxide in patients with heart failure: refining the gut hypothesis. *J Am Coll Cardiol* **64:** 1908-1914.

Tang WH, Wang Z, Kennedy DJ, Wu Y, Buffa JA, Agatisa-Boyle B *et al* (2015). Gut microbiota-dependent trimethylamine N-oxide (TMAO) pathway contributes to both development of renal insufficiency and mortality risk in chronic kidney disease. *Circ Res* **116:** 448-455.

Teh AH, Kimura M, Yamamoto M, Tanaka N, Yamaguchi I, Kumasaka T (2006). The 1.48 Å resolution crystal structure of the homotetrameric cytidine deaminase from mouse. *Biochemistry* **45:** 7825-7833.

Thomas GH (2010). Homes for the orphans: utilization of multiple substrate-binding proteins by ABC transporters. *Mol Microbiol* **75:** 6-9.

Tiago de Oliveira F, Chanda A, Banerjee D, Shan X, Mondal S, Que L *et al* (2007). Chemical and spectroscopic evidence for an FeV-oxo complex. *Science* **315:** 835-838.

Topp E, Hanson RS, Ringelberg DB, White DC, Wheatcroft R (1993). Isolation and characterization of an *N*-methylcarbamate insecticide-degrading methylotrophic bacterium. *Appl Environ Microbiol* **59:** 3339-3349.

Tortell P, Rau G, Morel F (2000). Inorganic carbon acquisition in coastal Pacific phytoplankton communities. *Limnology and Oceanography* **45:** 1485-1500.

Tsukada Y, Fang J, Erdjument-Bromage H, Warren ME, Borchers CH, Tempst P *et al* (2006). Histone demethylation by a family of JmjC domain-containing proteins. *Nature* **439:** 811-816.

Van der Zee J, Duling DR, Mason RP, Eling TE (1989). The oxidation of *N*-substituted aromatic amines by horseradish peroxidase. *J Biol Chem* **264:** 19828-19836.

Vance C, Miller A (1998). Spectroscopic comparisons of the pH dependencies of Fe-substituted (Mn)superoxide dismutase and Fe-superoxide dismutase. *Biochemistry* **37:** 5518-5527.

Wang Z, Klipfell E, Bennett BJ, Koeth R, Levison BS, Dugar B *et al* (2011). Gut flora metabolism of phosphatidylcholine promotes cardiovascular disease. *Nature* **472:** 57-63.

Warren CR (2013a). High diversity of small organic N observed in soil water. *Soil Biol Biochem* **57:** 444-450.

Warren CR (2013b). Quaternary ammonium compounds can be abundant in some soils and are taken up as intact molecules by plants. *New Phytol* **198:** 476-485.

Whitmore L, Wallace BA (2004). DICHROWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data. *Nucleic Acids Res* **32:** W668-673.

Whitmore L, Wallace BA (2008). Protein secondary structure analyses from circular dichroism spectroscopy: methods and reference databases. *Biopolymers* **89:** 392-400.

Wischer D, Kumaresan D, Johnston A, El Khawand M, Stephenson J, Hillebrand-Voiculescu AM *et al* (2015). Bacterial metabolism of methylated amines and identification of novel methylotrophs in Movile Cave. *ISME J* **9:** 195-206.

Yancey P, Clark M, Hand S, Bowlus R, Somero G (1982). Living with water stress: evolution of osmolyte systems. *Science* **217:** 1214-1222.

Yancey P, Fyfe-Johnson A, Kelly R, Walker V, Aunon M (2001). Trimethylamine oxide counteracts effects of hydrostatic pressure on proteins of deep-sea teleosts. *Journal of Experimental Zoology* **289:** 172-176.

Yancey P, Rhea M, Kemp K, Bailey D (2004). Trimethylamine oxide, betaine and other osmolytes in deep-sea animals: Depth trends and effects on enzymes under hydrostatic pressure. *Cellular and Molecular Biology* **50:** 371-376.

Yancey PH (2005). Organic osmolytes as compatible, metabolic and counteracting cytoprotectants in high osmolarity and other stresses. *J Exp Biol* **208:** 2819-2830.

Yancey P, Siebenaller J (2015). Co-evolution of proteins and solutions: protein adaptation versus cytoprotective micromolecules and their roles in marine organisms. *Journal of Experimental Biology* **218:** 1880-1896.

Yancey PH, Gerringer ME, Drazen JC, Rowden AA, Jamieson A (2014). Marine fish may be biochemically constrained from inhabiting the deepest ocean depths. *Proc Natl Acad Sci U S A* **111:** 4461-4465.

Yang CC, Packman LC, Scrutton NS (1995). The primary structure of Hyphomicrobium X dimethylamine dehydrogenase. Relationship to trimethylamine dehydrogenase and implications for substrate recognition. *Eur J Biochem* **232:** 264-271.

Youn J, Crosbie E, Maudlin L, Wang Z, Sorooshian A (2015). Dimethylamine as a major alkyl amine species in particles and cloud water: Observations in semi-arid and coastal regions. *Atmospheric Environment* **122:** 250-258.

Yu Z, Zhang Q, Kraus TEC, Dahlgren RA, Anastasio C, Zasoski RJ (2002). Contribution of amino compounds to dissolved organic nitrogen in forest soils. *Biogeochemistry* **61:** 173-198.

Zang TM, Hollman DA, Crawford PA, Crowder MW, Makaroff CA (2001). Arabidopsis glyoxalase II contains a zinc/iron binuclear metal center that is essential for substrate binding and catalysis. *J Biol Chem* **276:** 4788-4795.

Zhang AQ, Mitchell SC, Ayesh R, Smith RL (1992). Determination of trimethylamine and related aliphatic-amines in human urine by headspace gas-chromatography. *J Chromatogr-biomed* **584:** 141-145.

Zhang H, Ren S, Yu J, Yang M (2012). Occurrence of selected aliphatic amines in source water of major cities in China. *J Environ Sci (China)* **24:** 1885-1890.

Zhang X, Fuller J, Mcintire W (1993). Cloning sequencing, expression, and regulation of the structural gene for the copper/topa quinone-containing methylamine oxidase from Arthrobacter strain P1, a gram-positive facultative methylotroph. *Journal of Bacteriology* **175:** 5617-5627.

Zhao Y, Vance D, Abouchami W, de Baar H (2014). Biogeochemical cycling of zinc and its isotopes in the Southern Ocean. *Geochimica Et Cosmochimica Acta* **125:** 653-672.

Zhong W, Gallivan JP, Zhang Y, Li L, Lester HA, Dougherty DA (1998). From ab initio quantum mechanics to molecular neurobiology: a cation-pi binding site in the nicotinic receptor. *Proc Natl Acad Sci U S A* **95:** 12088-12093.

Zhu J, Dizin E, Hu X, Wavreille AS, Park J, Pei D (2003). S-Ribosylhomocysteinase (LuxS) is a mononuclear iron protein. *Biochemistry* **42:** 4717-4726.

Zhu W, Gregory J, Org E, Buffa J, Gupta N, Wang Z *et al* (2016). Gut Microbial Metabolite TMAO Enhances Platelet Hyperreactivity and Thrombosis Risk. *Cell* **165:** 111-124.

Zhu Y, Jameson E, Parslow RA, Lidbury I, Fu T, Dafforn TR *et al* (2014). Identification and characterization of trimethylamine *N*-oxide (TMAO) demethylase and TMAO permease in *Methylocella silvestris* BL2. *Environ Microbiol* **16:** 3318-3330.

Appendix 1-1 List of oligos used in this thesis

Primer	Sequence(5'-3')	Note	Reference
Primers for general use			
M13_F	GTAAAACGACGGCCAG		(Vieira and Messing 1982)
M13_R	CAGGAAACAGCTATGAC		
T7 promoter	TAATACGACTCACTATAGGG		
T7 terminator	GCTAGTTATTGCTCAGCGG		
Primers used for *tdm*			
Tdm_F	CATATGAATGTTAATGCGGCGGCCCCAAAGGGGT	Amplification of *tdm* from *Methylocella silvestris*	Thesis
Tdm_R	GGATCCCTATGCCCTGACTCGCGCCT		
3603_A_F	CGACGAAGCGATCATTACAA	Construction of *Msil_3603*::*kan* mutant of *M. silvestris*	Thesis
3603_A_R	CCACGTCGATGATCTGGATA		
3603_B_F	GTTTGAGGGCCTTTGTCAGA		
3603_B_R	GGTGACGACTCCCACCTG		
3604_A_F	CCCTCTTGTTCTTCCCCTTC	Construction of *tmm*::*kan* mutant of *M. silvestris*	Thesis
3604_A_R	TTGACCGTCACCGTAAACAA		

Primer	Sequence(5'-3')	Note	Reference
3604_B_F	CGCCAGCCAATGCTATAAAT		
3604_B_R	GCGATATCCCATGATGCTCT		
3605_A_F	TCTTGAATGATCGCCACAGA	Construction of *3605*::*kan* mutant mutant of *M. silvestris*	Thesis
3605_A_R	ATGCTCGAAGTAGCGCAGAT		
3605_B_F	CGTGTTCCTGTGGGATCAG		
3605_B_R	ATCGTAGAACGGCGTCTTGA		
3606_A_F	CGTCCTGATGCTTCTGGTTT	Construction of *3606*::*kan* mutant of *M. silvestris*	Thesis
3606_A_R	CACAAGGATGCTTTCGATGA		
3606_B_F	TACCTCAGCCATGTCAACGA		
3606_B_R	AGCAACAAAACGACCAATCC		
3608_A_F	GAAGAGGCCTTTCCATGAGC	Construction of *3608*::*kan* mutant of *M. silvestris*	Thesis
3608_A_R	TGGTCTTGAGATAGGCGAGAA		
3608_B_F	GCAAGATCACGATCCATCC		
3608_B_R	CATGATCGAGCGATTGGAC		
3609_A_F	ATGAACCAGCGATTCAAGC	Construction of *3609*::*kan* mutant of *M. silvestris*	Thesis
3609_A_R	TATATTCGAACGGCTCACAGG		

Primer	Sequence(5'-3')	Note	Reference
3609_B_F	ATTCAACATCGGCATGACCT		
3609_B_R	GCGGGTCAATCCCTTATAGTC		
3603_con_F	ATCCGACAGGCATTGAAAAG	Confirmation of *Msil_3603*::*kan* mutant	Thesis
3603_con_R	GCGCCTTTTCTGGATCATAG		
3604_con_F	ATCGCGCTACTTCGAGCAT	Confirmation of *tmm*::*kan* mutant	Thesis
3604_con_R	GAGCCTTTTCAATGCCTGTC		
3605_con_F	GGATTGGTCGTTTTGTTGCT	Confirmation of *Msil_3605*::*kan* mutant	Thesis
3605_con_R	GTCGGACTGTTTTTCGAAGC		
3606_con_F	GTTCGGACGCCGATAAGAC	Confirmation of *Msil_3606*::*kan* mutant	Thesis
3606_con_R	GTCGACGATGTTGCTGATTG		
3608_con_F	GTTTGTCCGCGATCATTTCT	Confirmation of *Msil_3608*::*kan* mutant	Thesis
3608_con_R	GTCGCCAAACACAAAGGACT		
3609_con_F	CACCGCCATCTTGAAGAAAT	Confirmation of *Msil_3609*::*kan* mutant	Thesis
3609_con_R	GAAGACGGGATAACCGATCA		
Kan_F	GGTAGGTCGACGCATGCGAGCTCGGAAAGCCACGTTGTGTCTC	Amplification of kanamycin gene cassette	Thesis
Kan_R	GGTAGGTCGACGCATGCGAGCTCAAGGTGTTGCTGACTCATAC		

Primer	Sequence(5'-3')	Note	Reference
Tdm1187_F	GCGTGAATCCGGTTTCCATT	Sequencing Confirmation of *tdm* mutants	Thesis
Tdm1363_R	CGCTTTGCAACAGGATTTCG		
D198A_F	5' GCAATGCTCCGCGCTCGTCGCCTTC 3'	Construction of Tdm D198A	Thesis
D198A_R	5' AAGGCGACGAGCGCGGAGCATTGCC 3'		
C263A_F	5' ACACATTCATGCTCGCCGCGAACGCGAAATATTACG 3'	Construction of Tdm C263A	Thesis
C263A_R	5' TCGTAATATTTCGCGTTCGCGGCGAGCATGAATGTG 3'		
C279A_F	5' CCGGCCACGCCAATGCGACCGACAATTTCAAC 3'	Construction of Tdm C279A	Thesis
C279A_R	5' TTGAAATTGTCGGTCGCATTGGCGTGGCCGGG 3'		
C343A_F	5' TCCTCGTCCGCGGCCGACGACATTGATC 3'	Construction of Tdm C343A	Thesis
C343A_R	5' GTCGTCGGCCGCGGACGAGGACGCGCAG 3'		
D39A_F	5' GTTCGACCTTTCGGCCGGCGCGGAAGCGATCATTACAAATC 3'	Construction of Tdm D39A	Thesis
D39A_R	5'GATTTGTAATGATCGCTTCCGCGCCGGCCGAAAGGTCGAAC 3'		
E47A_F	5' TACAAATCTCGCGGGGGGCCAGG 3'	Construction of Tdm E47A	Thesis
E47A_R	5' ATGATCGCTTCGTCGCCG 3'		
D167A_F	5' AGTAAAGCTCGCTCTCAGGGTCGACGC 3'	Construction of Tdm D167A	Thesis
D167A_R	5' GGGGCAAGCGAAGGCGGC 3'		

Primer	Sequence(5'-3')	Note	Reference
D190A_F	5'CCAGATCATCGCCGTGGACGGCC 3'	Construction of Tdm D190A	Thesis
D190A_R	5'ATATAATCGCCGGCGCTGAC 3'		
H256A_F	5' CGTCGGGCGGGCAGACACATTCATGC 3'	Construction of Tdm H256A	Thesis
H256A_R	5' GTGTCGCGGACGATCTCA 3'		
H276A_F	5' CTATCCCGGCGCAGCCAATTGCAC 3'	Construction of Tdm H276A	Thesis
H276A_R	5' CCCATGTCGTCGTAATATTTC 3'		
Tdm_383_R	5' GGATCCTTAGGGGGCGGCGTCAGCGGTCA 3'	Construction of C-terminal truncate Tdm	Thesis
Tdm_389_R	5' GGATCCTTACAGCGCGGAAGTCCTTGAATGGAAACC 3'	Construction of C-terminal truncate Tdm	Thesis