



University of Warwick institutional repository: <http://go.warwick.ac.uk/wrap>

This paper is made available online in accordance with publisher policies. Please scroll down to view the document itself. Please refer to the repository record for this item and our policy information available from the repository home page for further information.

To see the final version of this paper please visit the publisher's website. Access to the published version may require a subscription.

Author(s): Yin Chen, Nisha A. Patel, Andrew Crombie, James H. Scrivens, and J. Colin Murrell

Article Title: Bacterial flavin-containing monooxygenase is trimethylamine monooxygenase

Year of publication: 2011

Link to published article:

[http://dx.doi.org/ 10.1073/pnas.1112928108](http://dx.doi.org/10.1073/pnas.1112928108)

Publisher statement: Copyright ©2011 by the National Academy of Sciences

**Bacterial flavin-containing monooxygenase is trimethylamine monooxygenase, an overlooked enzyme of global importance**

Yin Chen<sup>1</sup>, Nisha A. Patel, Andrew Crombie, James H. Scrivens, J. Colin Murrell<sup>1</sup>

School of Life Sciences, the University of Warwick, CV4 7AL, United Kingdom

<sup>1</sup>Correspondence to Dr. Y. Chen, School of Life Sciences, the University of Warwick, CV4 7AL, United Kingdom. Phone: 00 44 24 7652 8976; Fax: 00 44 24 7652 3568; E-mail: [cheny98@gmail.com](mailto:cheny98@gmail.com)

or Prof. J.C. Murrell, School of Life Sciences, the University of Warwick, CV4 7AL. Phone: 00 44 24 7652 3553; Fax: 00 44 24 7652 3568; E-mail: [j.c.murrell@warwick.ac.uk](mailto:j.c.murrell@warwick.ac.uk)

Classification: Biological Sciences/ Microbiology

Flavin-containing monooxygenases (FMOs) are one of the most important monooxygenase systems in Eukaryotes and have many important physiological functions. FMOs have also been found in bacteria; however, their physiological function is not known. Here, we report the identification and characterization of trimethylamine (TMA) monooxygenase, termed *tmm*, from *Methylocella silvestris*, using a combination of proteomic, biochemical and genetic approaches. This bacterial FMO contains the FMO sequence motif (FXGXXXHXXXF/Y) and typical flavin adenine dinucleotide and nicotinamide adenine dinucleotide phosphate-binding domains. The enzyme was highly expressed in TMA-grown *M. silvestris* and absent during growth on methanol. The gene, *tmm*, was expressed in *Escherichia coli* and the purified recombinant protein had high TMA monooxygenase activity. Mutagenesis of this gene abolished the ability of *M. silvestris* to grow on TMA as a sole carbon and energy source. Close homologs of *tmm* occur in many *Alphaproteobacteria*, in particular *Rhodobacteraceae* (marine “*Roseobacter*” clade, MRC) and the marine SAR11 clade (*Pelagibacter ubique*). We showed that the ability for MRC to use TMA as sole carbon and/or nitrogen source is directly linked to the presence of *tmm* in their genomes and purified Tmm of MRC and SAR11 from recombinant *E. coli* showed TMA monooxygenase activities. The *tmm* gene is very abundant in the metagenomes of the Global Ocean Sampling Expedition and we estimate that 20% of the bacteria in the surface ocean contain *tmm*. Taken together, our results suggest that TMA monooxygenase, a bacterial FMO, plays an important, yet overlooked, role in the global carbon and nitrogen cycles.

flavin-containing monooxygenase | trimethylamine monooxygenase | *Methylocella silvestris* | SAR11 | marine *Roseobacter* clade

\body

## Introduction

Flavin-containing monooxygenases (FMOs) catalyze the NADPH-dependent N- or S-oxygenation of heteroatom-containing compounds (1). According to the definition by Berkel and colleagues, FMOs belong to class B flavoprotein monooxygenases, together with Baeyer-Villiger monooxygenases (BVMO) and *N*-hydroxylating monooxygenases (NMO) (2). This class of flavoprotein monooxygenases has characteristic conserved sequence motifs and contain enzyme-bound flavin adenine dinucleotide (FAD) (2, 3). FMOs in eukaryotes have been well-studied and diverse functions have been assigned to this group of enzymes (1, 4). The best known example of eukaryotic FMO is probably from mammals, where the predominant physiological function appears to be detoxification of a vast spectrum of xenobiotics, including trimethylamine (TMA) found in the diet of mammals. It has been shown that mutations in human FMO (isoform 3) cause the inheritable disease known as trimethylaminuria ( “fish-odor syndrome”) caused by the inability of the body to oxidize TMA to its non-odor oxygenated form trimethylamine *N*-oxide (TMAO), the TMA subsequently causing malodor in urine and the breath (5). FMOs have also been found in plants and fungi. Plant FMOs are involved in the biosynthesis of the plant-growth hormone auxin (6-7) and in pathogen defense (8-10). In fungi, FMOs are vital for the functional expression of proteins that contain disulfide bonds by controlling the redox potential within the endoplasmic reticulum (11-13). FMOs are also present in bacteria (14) and recently, the presence of FMO has been observed in many bacterial genomes. However, thus far the physiological roles of bacterial FMOs are unknown.

We hypothesized that since eukaryotic FMOs oxidize TMA to TMAO, bacterial FMOs will be involved in TMA metabolism by bacteria. An enzyme called “TMA monooxygenase” was suggested by Large and colleagues in the 1970s (15) to explain the

ability of *Aminobacter aminovorans* (previously known as *Pseudomonas aminovorans*), which lacked known pathways for TMA utilization (16), to grow on TMA. In this bacterium, TMA is converted to TMAO by this enzyme, and TMAO is then further converted to formaldehyde (assimilated as the carbon (C) source) and ammonium (assimilated as the nitrogen (N) source) via the intermediate monomethylamine (MMA) (15-17). In other bacteria such as *Pseudomonas* strain P, this enzyme is involved in utilization of TMA as a sole N source (17). However, the gene encoding TMA monooxygenase has never been identified.

Methylated amine compounds, including TMA, are ubiquitous in the environment, for example, as end products of protein putrefaction (18). In the marine environment, methylated amines are released due to the degradation of quaternary amine osmoregulators, such as glycine betaine, which are used by marine organisms to counteract water stress (19-21). Once released into the environment, methylated amines can be used by microorganisms as a C or N source. In fact, in the oceans, methylated amines represent a significant pool of C and N and standing concentrations up to hundreds of nM and  $\mu$ M have been reported in the water column (22-23) and sediment pore water (24-25), respectively. In addition to being involved in biogeochemical cycles of C and N, recent studies also suggest that methylated amines have the potential to affect global climate, being precursors in aerosol formation in the upper atmosphere (26-28).

We report here identification and characterization of TMA monooxygenase (15, 17, 29) using a combination of genetic, proteomic and bioinformatic approaches. The corresponding gene is particularly abundant in marine “*Roseobacter*” clade (MRC) and the SAR11 clade, two of the most abundant marine bacteria in the surface ocean (30-33), as well as in bacterial metagenomic sequences in the Global Ocean Sampling Expedition (34). Our

results suggest that this overlooked enzyme may play an important role in the global C and N cycles.

## Results

### Identifying the function of a bacterial flavin-containing monooxygenase (FMO) in *M.*

*silvestris* as TMA monooxygenase. *M. silvestris* can grow on TMA as sole C and N source.

We detected inducible activities of three enzymes, TMA monooxygenase ( $20.4 \pm 1.9 \text{ nmol mg}^{-1} \text{ min}^{-1}$ ), TMAO demethylase ( $71.2 \pm 16.3 \text{ nmol mg}^{-1} \text{ min}^{-1}$ ) and dimethylamine (DMA) monooxygenase ( $21.6 \pm 0.7 \text{ nmol mg}^{-1} \text{ min}^{-1}$ ) in cell-free extracts of this bacterium when it was grown on methylated amines. Because the genes encoding these enzymes have not yet been identified, we carried out comparative proteomic analyses using soluble protein fractions from methanol-, monomethylamine (MMA)- and trimethylamine (TMA)-grown cultures. Quantitative proteomics data showed that polypeptides encoded by two gene clusters, orf Msil\_2632-Msil\_2639 and orf Msil\_3603-Msil\_3605 (**Figure 1a**), were particularly abundant in cell-free extracts when *M. silvestris* was grown on TMA, whereas they were not detectable in cell-free extracts from methanol-grown cultures (a full list of polypeptides detected in each sample is shown in **SI Table 1**). We have previously shown that the former gene cluster was involved in MMA metabolism in *M. silvestris* which encodes enzymes of the gamma-glutamylmethylamide and *N*-methylglutamate pathway, converting MMA to formaldehyde (which is used as C and energy source) and ammonium (which is used as N source) (**Figure 1b**) (35).

The second gene cluster (**Figure 1a**) encodes two proteins which are annotated as glycine cleavage T protein and a bacterial type FMO (Msil\_3604) which shows significant similarity to mammalian FMOs. Reverse transcriptase (RT)-PCR confirmed that transcription of the three genes of this second gene cluster was induced by TMA. These three genes were co-transcribed as an operon since primers targeting intergenic regions gave RT-PCR products

of the expected sizes (primers used are listed in **SI Table 2**). We hypothesized that orf Msil\_3603-Msil\_3605 encodes the three enzymes which metabolize TMA to MMA (**Figure 1b**), a pathway for TMA oxidation proposed by Large and colleagues in the 1970s (15, 17).

Sequence alignments with FMOs from mammals, plant and fungi identified conserved domains in Msil\_3604 (**SI Figure 1**). Msil\_3604 contains all of the three key characteristic sequence motifs of FMOs: the FAD-binding domain (GXGXXG), the FMO-identifying motif (FXGXXXHXXXF/Y) (3) and an NADPH-binding domain (GXSXXA). Phylogenetic analyses of Msil\_3604 and known FMOs from mammals, plant and fungi revealed that this bacterial FMO clustered within the known FMO family (**Figure 2**).

To identify the function of Msil\_3604, we cloned this gene from *M. silvestris* into pET28a, over-expressed it in *Escherichia coli* and purified the resulting polypeptide by affinity chromatography (**SI Figure 2**). Steady-state kinetic assays, performed at fixed NADPH concentrations with typical FMO substrates, showed Michaelis-Menten behavior (**Table 1**). As previously noted for other FMOs, FMO from *M. silvestris* had a rather broad substrate specificity, oxidizing many nitrogen and sulfur-containing compounds. However, unlike eukaryotic FMOs, thiol-containing molecules such as glutathione and cysteine were not oxidized by this bacterial FMO. Interestingly, this enzyme can also oxidize dimethylsulfide (DMS) and the  $k_m$  value is similar to that of TMA. Of all the substrates tested, this enzyme has the highest affinity for TMA and it is therefore likely to encode the TMA monooxygenase which are involved in TMA oxidation in this bacterium (**Figure 1b**). To further confirm its role in TMA oxidation in *M. silvestris*, a mutant of Msil\_3604 was constructed by marker-exchange mutagenesis. The growth rates of the mutant and wild type strain of *M. silvestris* on various methylamine compounds are presented in **Table 2**. Although the wild type strain grows on TMA as a sole C and N source, the mutant could no longer

grow on TMA. Overall, our data indicate that this bacterial type FMO encodes the so-called TMA monooxygenase in bacteria and is essential for bacterial TMA oxidation.

**Tmm in sequenced bacterial genomes.** The sequence of *M. silvestris* TMA monooxygenase (termed hereafter Tmm) was used to search microbial genome databases. No Archaeal genomes contain Tmm homologs. However, homologs were found in the genomes of some *Gammaproteobacteria*, including *Methylophaga* sp. SK1, *Pseudomonas putida* GB-1, *P. fluorescens* SBW25 and *P. mendocina* ymp. *Methylophaga* sp. SK1 and *P. mendocina* are known to use TMA as a sole C and N source (14, 36). Tmm homologs were also identified in the genomes of some terrestrial *Alphaproteobacteria*, including *Agrobacterium*, *Azorhizobium* and *Mesorhizobium*. Interestingly, Tmm homologues are also found in many marine *Alphaproteobacteria*, including *Rhodobacteraceae* (also known as the marine “*Roseobacter*” clade, MRC) and in three genomes of the SAR11 clade (*Pelagibacter ubique* sp. HTCC1002, HTCC1062 and HTCC7211), indicating a key role of TMA oxidation in marine C/N cycles.

Searching of MRC genome sequences revealed that 15 out of the 39 sequenced MRC genomes contain Tmm homologues (**SI Table 3**). All of these Tmm homologs have FAD and NADPH binding sites and the FMO sequence motif (FXGXXXHXXXF/Y) (**SI Figure 1**), and are clustered within the bacterial FMOs (**Figure 2**). We tested 11 out of these 39 MRC strains for their ability to use TMA as C and/or N source. Seven strains containing Tmm homologues in their genomes, grew on TMA as sole N source (**SI Table 3**). Of these, three strains (*Roseovarius* sp. 217, *Roseovarius* sp. TM1035 and *Roseovarius nubinhibens* ISM) can also use TMA as sole C and energy source. Four other strains tested, *Dinoroseobacter shibae*, *Oceanicola batsensis*, *Roseobacter* sp. SK209-2-6 and *Sagittula stellata* which lack *tmm* in their genomes did not grow on TMA as either C or N source. Representative growth curves of *Ruegeria pomeroyi* DSS-3 and *Roseovarius* sp. 217 growing on TMA is shown in

**Figure 3.** Inducible TMA monooxygenase activities were detected in the supernatant of TMA-grown cultures and the expression of Tmm in TMA-grown cultures was confirmed by mass spectrometry (**SI Figure 3**). To further prove the function of these Tmm homologues, we over-expressed *tmm* from *Roseovarius* sp. 217 and *Ruegeria pomeroyi* DSS-3 in *E. coli* and purified them. They both oxidize TMA (**Table 3**).

The three genomes of members of the SAR11 clade also contain Tmm homologues. Due to the lack of available isolates to test their growth on TMA, we chemically synthesized two *tmm* genes (sp. HTCC1002 and HTCC7211), over-expressed these genes in *E. coli* and further purified the recombinant proteins. Kinetics analyses also demonstrated that these Tmm homologues can also oxidize TMA (**Table 3**).

**Tmm is abundant in the marine bacterial metagenome.** The wide occurrence of Tmm homologs in the MRC and the SAR11 clade bacteria suggested that Tmm may play an important role in the marine C and N cycles. We therefore estimated the abundance of Tmm homologs in the Global Ocean Sampling (GOS) Expeditions database (34). This database contains ~ 6.1 million deduced gene products from bacteria sampled directly from surface waters at various sites in the North-west Atlantic Ocean, the Sargasso Sea and the Eastern Pacific Ocean. Using a relatively stringent BLASTP search ( $< e^{-80}$ ), we identified 570 Tmm homologs (**Table 4**). These homologs occur at nearly all GOS sites, except for two non-marine sites (Lake Gatun and Punta Cormorant) and one marine site which was sampled with larger-sized fraction filters (Dirty Rock Cocos Island) (34). The retrieved Tmm homologs from the GOS dataset all contain the FMO sequence motif. The abundance of Tmm homologs is comparable to other enzymes carrying out key environmental functions in the marine environment, such as DmdA which is involved in dimethylsulfoniopropionate (DMSP) catabolism (37), and is about 10 times more frequent than DddQ, DddP and DddL, which are involved in DMSP-dependent DMS production (38). On average, we estimate that ~ 20% of

the microbial genomes in the surface ocean contain Tmm homologs (**Table 4**). Phylogenetic analysis indicated that the Tmm homologs from the GOS database could be placed into three groups (**Figure 4**), the majority of which were closely related to Tmm homologs in SAR11 clade bacteria (94.1%) and the MRC bacteria (4.7%) respectively (**SI Figures 4**). The predominance of the MRC and the SAR11 clade bacteria in marine TMA metabolism is supported by the analysis of GmaS, gamma-glutamylmethanamide synthetase, an enzyme involved in a subsequent step in the pathway of TMA metabolism (**Figure 1b**) (35, 39). GmaS homologs are also found at nearly all GOS sites and are as abundant as the Tmm enzyme (**Table 4**). Phylogenetic analysis also indicated the sequences retrieved from the GOS dataset are from either the SAR11 clade (78%) or the MRC (22%) (**SI Figure 5**).

## Discussion

In this study we have identified the gene encoding TMA monooxygenase (29, 40). The wide occurrence of Tmm homologs in the MRC and the SAR11 clade bacteria indicates that methylated amine compounds may play an important role in marine C and N cycles. These two groups of marine bacteria are particularly abundant in the surface oceans and previous studies have shown that they play a key role in marine carbon and sulfur cycles (30-33, 41-42). Although few studies have measured the concentrations of methylated amines in the oceans, existing data suggest that these compounds represent a significant pool of C and N in the marine water column (22-23) and in marine and coastal sediments (24-25). We tested representatives of the MRC bacteria and the ability to use TMA is directly correlated to the presence of Tmm in their genomes (**SI Table 3**). The purified Tmm enzyme of MRC and SAR11 bacteria all oxidize TMA (**Table 3**), indicating that this enzyme is likely to contribute to the removal of methylated amines in the marine environment. Further evidence comes from analysis of the GOS data set, where it is estimated that on average 20% of the bacteria in the surface ocean contain Tmm, the majority of which originated from the SAR11 clade

bacteria (**Table 4**). Indeed, previous metaproteome analysis of bacteria collected from surface seawater has identified Tmm polypeptides from SAR11 clade bacteria (43). Therefore, it is tempting to speculate that the MRC and the SAR11 clade bacteria may play key roles in sequestering N from methylated amines. This may therefore give them a selective advantage in the oceans over microorganisms unable to use such compounds, where they face severe competition for N.

The purified Tmm has broad substrate specificity and can also oxidize DMA, although the affinity for DMA is lower than for TMA (**Table 1, SI Table 4**). This ability to oxidise secondary amines has been shown previously with purified TMA monooxygenase from *A. aminovorans* and it has been shown that oxidizing DMA has no physiological function (40). Non-specific oxidation of DMA produces formamide, which is toxic (40) and this may have inhibited the growth of the wild type *M. silvestris* on DMA alone (**Table 2**). This is supported by the fact that wild type *M. silvestris* can grow on TMA in the presence of DMA (**Table 2**), since TMA is the preferred substrate of TMA monooxygenase. Although FMO from *M. silvestris* can oxidize DMA, its function as a DMA monooxygenase can be ruled out since FMOs are not sensitive to carbon monoxide, a characteristic of secondary amine monooxygenases (44). In fact, DMA monooxygenase would release formaldehyde from the oxidation of DMA (**Figure 1b**) and the presence of tetrahydrofolate-binding domains in orfs Msil\_3603 and Msil\_3605 tends to suggest that one of these genes may encode the DMA monooxygenase (the other is likely to encode TMA *N*-oxide demethylase, which also releases formaldehyde).

The ability of bacterial FMOs to oxidize DMS and dimethylsulfoxide (DMSO) is surprising. DMSO was initially used as a solvent to dissolve some of the substrates. It was noticed previously that the FMO from *Methylophaga* sp. SK1 also oxidized DMSO (45). The ability of bacterial FMOs to oxidize DMS may be of special interest. It has been previously

noticed that there was an uncharacterized route for oxidation of DMS to DMSO by heterotrophic bacteria, such as *Delftia acidovorans* (46). In this bacterium, the oxidation of DMS to DMSO was dependent on NADPH, which could not be replaced by NADH. A search of the genome of *D. acidovorans* sp. SPH-1 does indeed reveal FMO-like protein encoding genes. Since FMO of the MRC and the SAR 11 clade bacteria can also oxidize DMS, they may also represent a previously-uncharacterized sink of DMS from the marine environment. Whether or not FMO is responsible for this uncharacterized DMS conversion in these bacteria, and to what extent they reduce DMS emission from the marine environment certainly warrants further investigation.

To conclude, our results revealed the function of bacterial FMOs as an enzyme involved in bacterial TMA metabolism, TMA monooxygenase. The wide occurrence of this enzyme in marine bacteria, together with its high abundance in the surface ocean, points to a key role for methylated amines in marine C and N cycling.

## **Materials and Methods**

**Cultivation of *Methylocella silvestris*.** *M. silvestris* was grown at 25 °C either in a 4-liter fermentor or in 125 ml serum vials using diluted mineral salt medium as described previously (35). To test if the trimethylamine monooxygenase (*tmm*) mutant could grow on methylated amines, growth experiments were set up in triplicate using 120 ml serum vials, containing 20 ml medium with an inoculum size of 10%. Details of growth conditions and reagents are described in *SI Material and Methods*.

**Quantitative comparative proteomics.** A total of 700 µg of soluble protein extract from MMA, TMA and methanol grown *M. silvestris* cells were used for proteomic analyses (47). The details of the method are shown in *SI Material and Methods*.

**Marker-exchange mutagenesis of *tmm* in *M. silvestris*.** To construct a *tmm* mutant of *M. silvestris*, a downstream region (containing a *KpnI* site) and an upstream region (containing

an *Mlu*I site) of the target were amplified by PCR and sub-cloned into pGEMT together with a kanamycin gene cassette (amplified from pCM184) inserted between the two regions (primers used are listed in **SI Table 2**). The resulting plasmid was then cut with *Kpn*I and *Mlu*I and the 2.6 kb fragment containing the kanamycin gene cassette was electroporated into *M. silvestris* as described previously (35).

**Over-expression of *tmm* in *E. coli*.** The *tmm* genes from *M. silvestris*, *Roseovarius* sp. 217 and *Ruegeria pomeroyi* DSS-3 were amplified by PCR (primers used are listed in **SI Table 2**) and sub-cloned into pGEMT vector (Promega), which were then excised using *Nde*I/*Hind*III and ligated into the expression vector pET28a (Merck Biosciences). The *tmm* homologues in the *Pelagibacter ubique* genomes (strain HTCC1002 and HTCC7211) were synthesized commercially with *E. coli* codon usage (GenScript Corporation). The synthesized genes were inserted into the expression vector pET28a using the *Nde*I/*Hind*III restriction sites. The resulting plasmids were then transformed into the expression host *E. coli* BLR(DE3) pLysS (Merck Biosciences). To over-express TMA monooxygenases, *E. coli* cells were grown at 37 °C to an OD<sub>540</sub> of 0.6 and isopropyl β-D-1-thiogalactopyranoside (IPTG) was then added to a final concentration of 0.1 mM.

**Enzyme assays and kinetics.** Cells were broken for protein analyses and enzyme assays by passing three times through a French pressure cell (American Instrument Co.) at 110 MPa. Cell debris was removed by centrifugation at 6,000 g for 15 min. One-dimensional protein analyses were carried out using a pre-cast NuPAGE Bis-Tris gel (10%, Invitrogen). All enzyme assays were carried out in triplicate in 80 mM PIPES buffer (pH 7.6) at room temperature (22 °C) unless otherwise stated. Details are shown in *SI Material and Methods*.

**Cultivation of MRC bacteria on methylamines.** Growth tests were carried out using a defined medium in triplicate, containing artificial sea salts from Sigma-Aldrich (S9883) 40 g l<sup>-1</sup>, sodium phosphate 0.2 mM (pH 8.0), 4-(2-Hydroxyethyl)piperazine-1-ethanesulfonic acid

sodium (HEPES) 10 mM (pH 8.0), methylated amines 2 mM, FeCl<sub>3</sub> 50  $\mu$ M, and a mixture of C sources containing glucose (5.6 mM), fructose (5.6 mM), succinate (8.5 mM), pyruvate (11.4 mM), glycerol (10.9 mM) and acetate (17 mM). The vitamins were added as described in *SI Materials and Methods*. Positive growth controls with ammonium chloride (1.0 mM) and negative controls with no added N compounds were also set up. To test if these isolates could use TMA as a sole C source, the above medium was used, except that the carbon mixture was omitted and TMA was added to a final concentration of 5 mM.

**Identification of Tmm and GmaS homologs in the GOS metagenome.** Tmm sequence of *M. silvestris* was used as the query sequence for a BLASTP search of the GOS peptides at CAMERA (<https://portal.camera.calit2.net/gridsphere/gridsphere?cid=>) ('GOS': all ORF peptides (P) database, e<sup>-80</sup>). This resulted in 568 unique sequences. An additional search using the same settings was performed using the FMO of *Methylophaga* sp. SK1 (JC7986) as the query and two additional sequences were retrieved. These 570 sequences were further grouped into 127 unique groups (identity > 90% within each group) using the CD-HIT program (48). Representative sequences from each group were aligned using the ARB program (49) and each of the sequence was manually checked for the FMO sequence motif. To estimate the frequency of Tmm-containing cells, the numbers of Tmm homologs were normalized against the average numbers of the six essential single-copy genes (*atpD*, *rpoB*, *gyrB*, *hsp70*, *tufA* and *recA*) as described previously (37). The identification of GmaS homologs were estimated as described above except that the GmaS of *M. silvestris* (ACK51558) was used as the query, which resulted in 682 unique sequences. Additional queries using the GmaS sequences of *M. universalis* FAM5 (ADH10360), *M. mays* (BAF99006) and *A. tumefaciens* C58 (AAK89209) did not yield additional new sequences. These 682 sequences were further grouped by CD-HIT and analyzed by ARB.

## Acknowledgement

This work was supported by NERC through grants NE/H016236/1 to YC and NE/E016855/1 to JCM. We thank Ms. S. Slade for assistance with proteomics analysis, Dr. K. Thalassinos for bioinformatics and Mr. W. Yang for protein over-expression. We are grateful to Drs. H. Schäfer (University of Warwick), A. Buchan (University of Tennessee), R. Belas (University of Maryland Biotechnology Institute), B. Ward (Princeton University) and I. Wagner-Döbler (Helmholtz Centre for Infection Research) for providing bacterial strains.

## References

1. Ziegler DM (1993) Recent studies on the structure and function of multisubstrate flavin-containing monooxygenases. *Annu Rev Pharmacol Toxicol* 33: 179-199.
2. van Berkel WJH, Kamerbeek NM, Fraaije MW (2006) Flavoprotein monooxygenases, a diverse class of oxidative biocatalysts. *J Biotechnol* 124: 670-689.
3. Fraaije MW, Kamerbeek NM, van Berkel WJH, Janssen DB. (2002) Identification of a Baeyer-Villiger monooxygenase sequence motif. *FEBS Lett* 518:43-47.
4. Ziegler DM (2002) An overview of the mechanism, substrate specificities, and structure of FMOs. *Drug Metab Rev* 34: 503-511.
5. Mitchell SC, Smith RL (2001) Trimethylaminuria: the fish malodor syndrome. *Drug Metab Dispos* 29: 517-521.
6. Schlaich NL (2007) Flavin-containing monooxygenases in plants: looking beyond detox. *Trends Plant Sci* 12: 412-418.
7. Zhao Y, Christensen SK, Fankhauser C, Cashman JR, Cohen JD, et al. (2001) A role for flavin monooxygenase-like enzymes in auxin biosynthesis. *Science* 291: 306-309
8. Mishina TE, Zeier J (2006) The *Arabidopsis* flavin-dependent monooxygenase FMO1 is an essential component of biologically induced systemic acquired resistance. *Plant Physiol* 141: 1666-1675.

9. Koch M, Vorwerk S, Masur C, Sharifi-Sirchi G, Olivieri N, et al. (2006) A role for a flavin-containing mono-oxygenase in resistance against microbial pathogens in *Arabidopsis*. *Plant J* 47: 629-639.
10. Bartsch M, Gobbato E, Bednarek P, Debey S, Schultze JL, et al. (2006) Salicylic acid-independent ENHANCED DISEASE SUSCEPTIBILITY1 signaling in *Arabidopsis* immunity and cell death is regulated by the monooxygenase FMO1 and the Nudix hydrolase NUDT7. *Plant Cell* 18: 1038-1051.
11. Suh JK, Poulsen LL, Ziegler DM, Robertus JD (1996) Molecular cloning and kinetic characterization of a flavin-containing monooxygenase from *Saccharomyces cerevisiae*. *Arch Biochem Biophys* 336: 268-274.
12. Suh JK, Poulsen LL, Ziegler DM, Robertus JD. (1999) Yeast flavin-containing monooxygenase generates oxidizing equivalents that control protein folding in the endoplasmic reticulum. *Proc Natl Acad Sci USA* 96: 2687-2691.
13. Suh JK, Robertus JD. (2000) Yeast flavin-containing monooxygenase is induced by the unfolded protein response. *Proc Natl Acad Sci USA* 97: 121-126.
14. Choi HS, Kim JK, Cho Eh, Kim YC, Kim JI, et al. (2003) A novel flavin-containing monooxygenase from *Methylophaga* sp. strain SK1 and its indigo synthesis in *Escherichia coli*. *Biochem Biophys Res Commun* 306: 930-936.
15. Large PJ, Boulton CA, Crabbe MJC. (1972) The reduced nicotinamide-adenine dinucleotide phosphate- and oxygen-dependent *N*-oxygenation of trimethylamine by *Pseudomonas aminovorans*. *Biochem J* 128: 137p-138p.
16. Anthony C (1982). The biochemistry of methylotrophs. Academic Press, London, United Kingdom.
17. Chandler SR (1983) The utilization of methylamine-nitrogen by the methazotrophic bacterium *Pseudomonas* P. PhD thesis. University of Reading.
18. Kamiya A, Youki O (1984) Study of odorous compounds produced by putrefaction of foods. *J Chromatogr A* 292:329-331.

19. King GM (1988) *In* Nitrogen cycling in coastal marine environments. Eds. TH Blackburn, J Sorensen. P143-173.
20. Oren A (1990) Formation and breakdown of glycine betaine and trimethylamine in hypersaline environments. *Antonie van Leeuwenhoek* 58:291-298.
21. Welsh DT (2000) Ecological significance of compatible solute accumulation by micro-organisms: from single cells to global climate. *FEMS Microbiol Rev* 24:263-290.
22. Gibb SW and Hatton AD (2004) The occurrence and distribution of trimethylamine-N-oxide in Antarctic coastal waters. *Mar Chem* 91:65-75.
23. Van Neste A, Duce RA, Lee C (1987) Methylamines in the marine atmosphere. *Geophy Res Lett* 14:711-714.
24. Fitzsimons MF, Kahni-Danon B, Dawitt M (2001) Distributions and adsorption of the methylamines in the inter-tidal sediments of an East Anglian estuary. *Environ Exp Bot* 46: 225-236.
25. Lee C, Olsen BL (1984) Dissolved, exchangeable and bound aliphatic amines in marine sediments: initial results. *Org Geochem* 6:259-263.
26. Facchini MC, Decesari S, Rinaldi M, Carbone C, Finessi E, et al. (2008) Important source of marine secondary organic aerosol from biogenic amines. *Environ Sci Technol* 42:9116-9121.
27. Müller C, Iinuma Y, Karstensen J, van Pinxteren D, Lehmann S, et al. (2009) Seasonal variation of aliphatic amines in marine sub-micrometer particles at the Cape Verde islands. *Atmos Chem Phys* 9:9587-9897.
28. Rinaldi M, Decesari S, Finessi E, Giulianelli L, Carbone C, et al. (2010) Primary and secondary organic marine aerosol and oceanic biological activity: recent results and new perspectives for future studies. *Adv Meteorol* ID 310682.
29. Boulton CA, Crabbe MJ, Large PJ (1974) Microbial oxidation of amines: partial purification of a trimethylamine mono-oxygenase from *Pseudomonas aminovorans* and its role in growth on trimethylamine. *Biochem J* 140: 253-263.

30. Buchan A, González JM, Moran MA. (2005) Overview of the marine *Roseobacter* lineage. *Appl Environ Microbiol* 71:5665-5667.
31. Brinkhoff T, Giebel H-A, Simon M. (2008) Diversity, ecology, and genomics of the *Roseobacter* clade: a short overview. *Arch Microbiol* 189:531-539.
32. Wagner-Döbler I, Biebl H. (2006) Environmental biology of the marine *Roseobacter* lineage. *Annu Rev Microbiol* 60:255-280.
33. Morris RM, Rappé MS, Cannon SA, Vergin KL, Siebold WA, et al (2002) SAR11 clade dominates ocean surface bacterioplankton communities. *Nature* 420:806-810.
34. Rusch DB, Halpern AL, Sutton G, Heidelberg KB, Williamson S, et al (2007) The *Sorcerer II* Global Ocean Sampling Expedition: Northwest Atlantic through Eastern Tropical Pacific. *PLOS Biol.* 5:e77.
35. Chen Y, Scanlan J, Song L, Crombie A, Rahman MT, et al. (2010) Gamma-glutamylmethylamide is an essential intermediate in the metabolism of methylamine by *Methylocella silvestris*. *Appl Environ Microbiol* 76: 4530-4537.
36. Boden R, Thomas E, Savani P, Kelly DP, Wood AP (2008) Novel methylotrophic bacteria isolated from the River Thames (London, UK). *Environ Microbiol* 10: 3225-3236.
37. Howard EC, Sun S, Biers EJ, Moran MA (2008) Abundant and diverse bacteria involved in DMSP degradation in marine surface water. *Environ Microbiol* 10:2379-2410.
38. Todd JD, Curson ARJ, Dupont CL, Nicholson P, Johnston AWB (2009) The *dddP* gene, encoding a novel enzyme that converts dimethylsulfoniopropionate into dimethyl sulfide, is widespread in ocean metagenomes and marine bacterial and also occurs in some Ascomycete fungi. *Environ Microbiol* 11:1376-1385. Corrigendum *Environ Microbiol* 11:1624-1625.
39. Chen Y, McAleer KL, Murrell JC (2010) Monomethylamine as a nitrogen source from a nonmethylotrophic bacterium *Agrobacterium tumefaciens*. *Appl Environ Microbiol* 76:4102-4104.

40. Boulton CA, Large PJ. (1975) Oxidation of N-alkyl- and NN-dialkylhydroxylamines by partially purified preparations of trimethylamine mono-oxygenase from *Pseudomonas aminovorans*. *FEBS Lett* 55: 286-290.
41. González JM, Kiene RP, Moran MA. (1999) Transformation of sulfur compounds by an abundant lineage of marine bacteria in the  $\alpha$ -subclass of the class *Proteobacteria*. *Appl Environ Microbiol* 65: 3810-3819.
42. Tripp HJ, Kitner JB, Schwalbach MS, Dacey JWH, Wilhelm LJ, et al. (2008) SAR11 marine bacteria require exogenous reduced sulphur for growth. *Nature* 452:741-744.
43. Sowell SM, Wilhelm LJ, Norbeck AD, Lipton MS, Nicora CD, et al (2009) Transport functions dominate the SAR11 metaproteome at low-nutrient extremes in the Sargasso Sea. *ISME J* 3:93-105.
44. Alberta JA, Dawson JH. (1987) Purification to homogeneity and initial physical characterization of secondary amine monooxygenase. *J Biol Chem* 262: 11857-11863.
45. Alfieri A, Malito E, Orru R, Fraaije MW, Mattevi A. (2008) Revealing the moonlighting role of NADP in the structure of a flavin-containing monooxygenase. *Proc Natl Acad Sci USA* 105: 6572-6577.
46. Zhang L, Kuniyoshi I, Hirai M, Shoda M. (1991) Oxidation of dimethyl sulfide by *Pseudomonas acidovorans* DMR-11 isolated from peat biofilter. *Biotechnol Lett* 13: 223-228.
47. Patel VJ, Thalassinou K, Slade SE, Connolly JB, Crombie A, et al. (2009) A comparison of labelling and label-free mass spectrometry-based proteomics approaches. *J Proteome Res* 8: 3752-3759.
48. Li W, Godzik A (2006) Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* 22:1658-1659.
49. Ludwig W, Strunk O, Westram R, Richter L, Meier H, et al (2004) ARB: a software environment for sequence data. *Nucleic Acids Res* 32:1363-1371.
50. Tamura K, Dudley J, Nei M, Kumar S. (2007) MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. *Mol Biol Evol* 24:1596-1599.

## Figure legends

**Figure 1 a)** A summary of comparative proteomics and transcriptional analyses of the three gene cluster containing a bacterial FMO in *M. silvestris*. The function was based on the annotation from the genome sequence using BLASTP search; the abundance of each of the polypeptide is shown in percentages of the total soluble proteome in each condition and the presence (+) or absence (-) of transcription of each gene were confirmed by reverse transcription PCR. MMA, monomethylamine; TMA, trimethylamine; MeOH, methanol. **b)** The proposed pathway for TMA oxidation involving TMA monooxygenase.

**Figure 2** A neighbour-joining phylogenetic tree showing the relationship of TMA monooxygenase (Tmm) to other FMOs. The tree was drawn using the MEGA4 (50) based on an alignment of ~ 450 amino acids of FMOs. Baeyer-Villiger monooxygenases were used as an out-group. Bootstrap values of 100 replicates are shown.

**Figure 3** A representative growth curve of *Ruegeria pomeroyi* DSS-3 and *Roseovarius* sp. 217 growing on trimethylamine (TMA) as sole nitrogen source. Controls were set up with no added nitrogen. The inner panels show the activity of TMA monooxygenase from the supernatant of *R. pomeroyi* DSS-3 and *Roseovarius* sp. 217 cultures, respectively. No enzyme activities were detected when ammonium was used as the sole nitrogen source. Error bars indicate standard deviation of each experiment (n=3).

**Figure 4** An unrooted tree showing Tmm homologs retrieved from sequenced bacterial genomes and the Global Ocean Sampling Expedition data set. The neighbor-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 out-group.

**Table 1** Steady state kinetic assays on purified trimethylamine monooxygenase of *M. silvestris* from recombinant *Escherichia coli*\*

Substrate	$k_m$ ( $\mu\text{M}$ )	$V_{\max}$ ( $\text{nmol mg}^{-1} \text{min}^{-1}$ )
TMA	$9.4 \pm 2.1$	$29.4 \pm 3.2$
DMA	$89.7 \pm 13.2$	$6.9 \pm 0.1$
MMA	-	-
DMSO	$3575 \pm 151$	$4.8 \pm 1.5$
DMS	$10.3 \pm 0.7$	$34.6 \pm 0.2$
Cysteamine	$3139 \pm 534$	$76.2 \pm 9.1$
Methimazole	$28.2 \pm 5.1$	$14.4 \pm 0.8$
Dimethylaniline	$35.7 \pm 2.5$	$29.2 \pm 0.9$
Glutathione	-	-
Cysteine	-	-

\* - : Substrate not oxidized.

**Table 2** Specific growth rates of wild type and the *tmm* mutant of *Methylocella silvestris* grown on methylated amines\*

Growth substrate	Wild type (h <sup>-1</sup> )	Mutant (h <sup>-1</sup> )
TMA	0.034 ± 0.003	-
DMA	-	0.039 ± 0.003
TMA+DMA	0.042 ± 0.008	0.037 ± 0.002
MMA	0.029 ± 0.003	0.029 ± 0.001

\* - , no growth.

**Table 3** Kinetic parameters with trimethylamine for purified recombinant Tmm enzymes

<b>Substrate</b>	<b><math>k_m</math> (<math>\mu\text{M}</math>)</b>	<b><math>V_{\text{max}}</math> (<math>\text{nmol mg}^{-1} \text{min}^{-1}</math>)</b>
<i>Roseovarius</i> sp. 217	$21.6 \pm 1.9$	$1133.6 \pm 58.6$
<i>Ruegeria pomeroyi</i> DSS-3	$20.8 \pm 2.9$	$267.7 \pm 52.2$
<i>Pelagibacter ubiquus</i> HTCC1002	$27.5 \pm 4.2$	$70.8 \pm 7.7$
<i>Pelagibacter ubiquus</i> HTCC7211	$28.5 \pm 4.4$	$67.3 \pm 3.2$

**Table 4** Distribution of Tmm homologues at different sampling sites in the Global Ocean Sampling Expedition metagenomic data set.

Sample number	Sampling site	Description	Tmm*	GmaS*	DddP*	Dddd*	DddL*	DmdA*	RecA*	% Genomes with Tmm†	% Genomes with GmaS†
2	Gulf of Maine	Coastal	12	16	1	2	0	21	35	29	39
3	Brown's Bank-Gulf of Maine	Coastal	13	11	0	0	0	39	52	24	21
4	Outside Halifax Nova Scotia	Coastal	30	21	0	0	0	34	40	55	38
5	Bedford Basin Nova Scotia	Emabymment	7	7	2	3	0	26	34	18	18
6	Bay of Fundy Nova Scotia	Estuary	23	25	0	0	0	50	80	39	42
7	Northern Gulf of Maine	Coastal	8	14	2	0	0	20	53	13	23
8	Newport Harbor	Coastal	2	7	4	0	0	9	43	3	12
9	Block Island	Coastal	15	15	0	0	0	28	33	39	39
10	Cape May	Coastal	18	19	0	0	0	36	63	27	29
11	Delaware Bay	Estuary	5	2	0	0	0	7	43	10	3
12	Chesapeake Bay	Estuary	5	4	0	0	0	17	39	10	8
13	Off Nag's Head	Coastal	7	9	1	0	0	17	43	18	24
14	South of Charleston	Coastal	18	14	0	0	0	42	67	29	22
15	Off KeyWest	Coastal	5	15	2	0	0	35	51	8	26
16	Gulf of Mexico	Coastal	14	7	1	0	0	41	50	26	13
17	Yucatan Channel	Open ocean	9	13	1	0	0	32	61	15	23
18	Rosario Bank	Open ocean	9	13	3	0	0	43	83	14	20
19	Northeast of Colon	Coastal	6	16	2	0	0	33	72	10	27
20	Lake Gatun	Fresh water	0	1	0	0	0	3	61	0	1
21	Gulf of Panama	Coastal	5	14	4	0	0	31	53	10	27
22	250 miles from Panama City	Open ocean	12	16	1	0	0	35	50	22	28
23	30 miles from Cocos	Open ocean	11	16	0	0	0	35	54	18	26
25	Dirty Rock Cocos Island	Fringing reef	0	1	1	0	0	5	13	0	6
26	134 miles NE of Galapagos	Open ocean	7	11	1	0	0	41	55	13	20
27	Devil's Crown	Coastal	16	13	4	0	0	41	66	26	20
28	Coastal Floreana	Coastal	10	12	2	0	0	48	70	16	20
29	North James Bay	Coastal	9	24	2	0	0	46	50	20	52
30	Warm Seep Roca Redunda	Warm seep	19	17	1	0	0	52	55	36	33

31	Upwelling Fernandina Island	Coastal	26	27	1	0	0	57	61	46	49
32	Mangrove on Isabella	Mangrove	8	5	1	1	0	16	43	20	14
33	Punta Cormorant	Hypersaline	0	2	4	0	6	2	43	0	4
34	North Seamore Island	Coastal	9	17	2	0	0	42	51	17	32
35	Wolf Island	Coastal	11	13	2	0	0	38	51	21	26
36	Cabo Marshall	Coastal	14	19	0	0	0	45	54	24	33
37	Equatorial TAO	Open ocean	14	6	5	0	0	48	38	30	13
47	201 miles from French	Open ocean	9	18	5	0	0	48	62	15	29
51	Rangiora Atoll	Coral reef	12	14	2	0	0	28	53	25	29
Average			11	13	1.5	0.2	0.2	32	52	20	23

\* The sampling sites are as detailed in Howard and colleagues (2008). Numbers of Tmm and GmaS homologues, per 100,000 reads, with BLASTP values of  $< e^{-80}$ . Values of DddD, DddL, DmdA and RecA were from Howard and colleagues (2008) and values of DddP were from Todd and colleagues (2009).

† Frequencies of Tmm and GmaS-containing strains, expressed as genome equivalents, as described in Howard and colleagues (2008).

	Msil_3603	Msil_3604	Msil_3605
Annotation	Glycine cleavage T protein	Bacterial flavin-containing monooxygenase	Glycine cleavage T protein
Polypeptide abundance in			
TMA	0.72%	1.60%	4.16%
MMA	1.35%	3.88%	6.20%
MeOH	Not detectable	Not detectable	Not detectable
Transcription in			
TMA	+	+	+
MeOH	-	-	-

Figure 1a

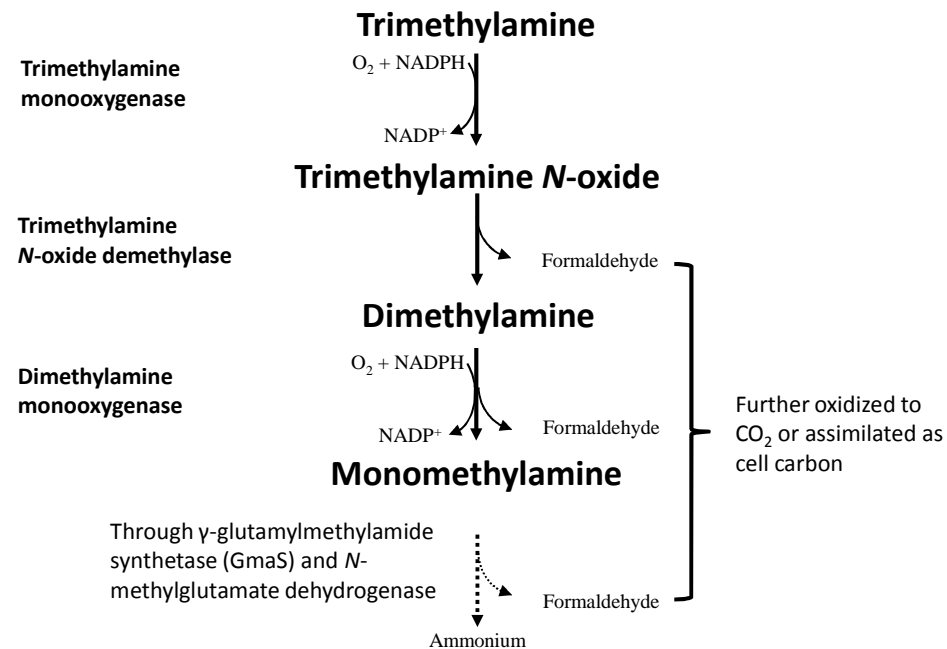


Figure 1b

Figure 2

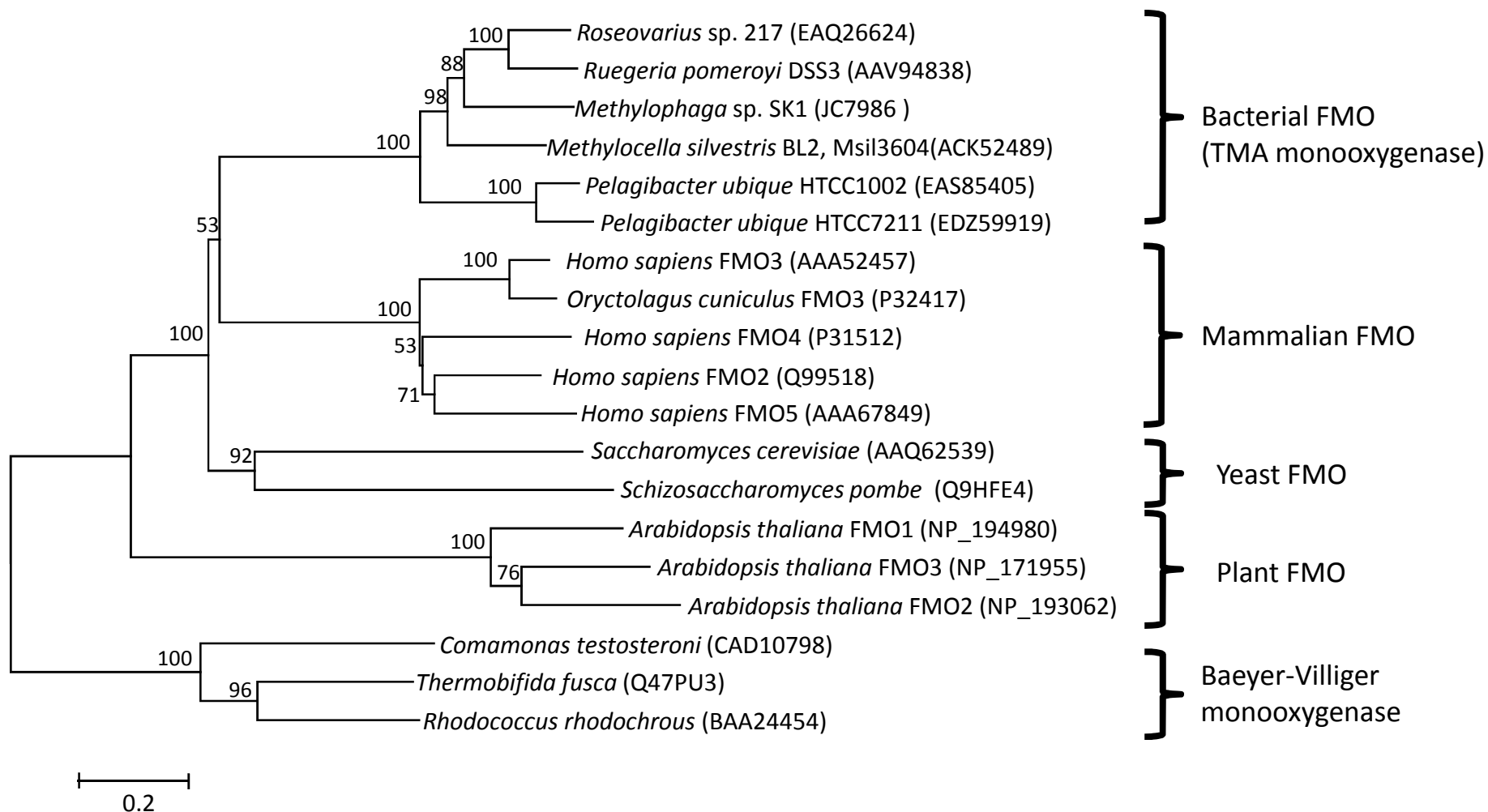


Figure 3

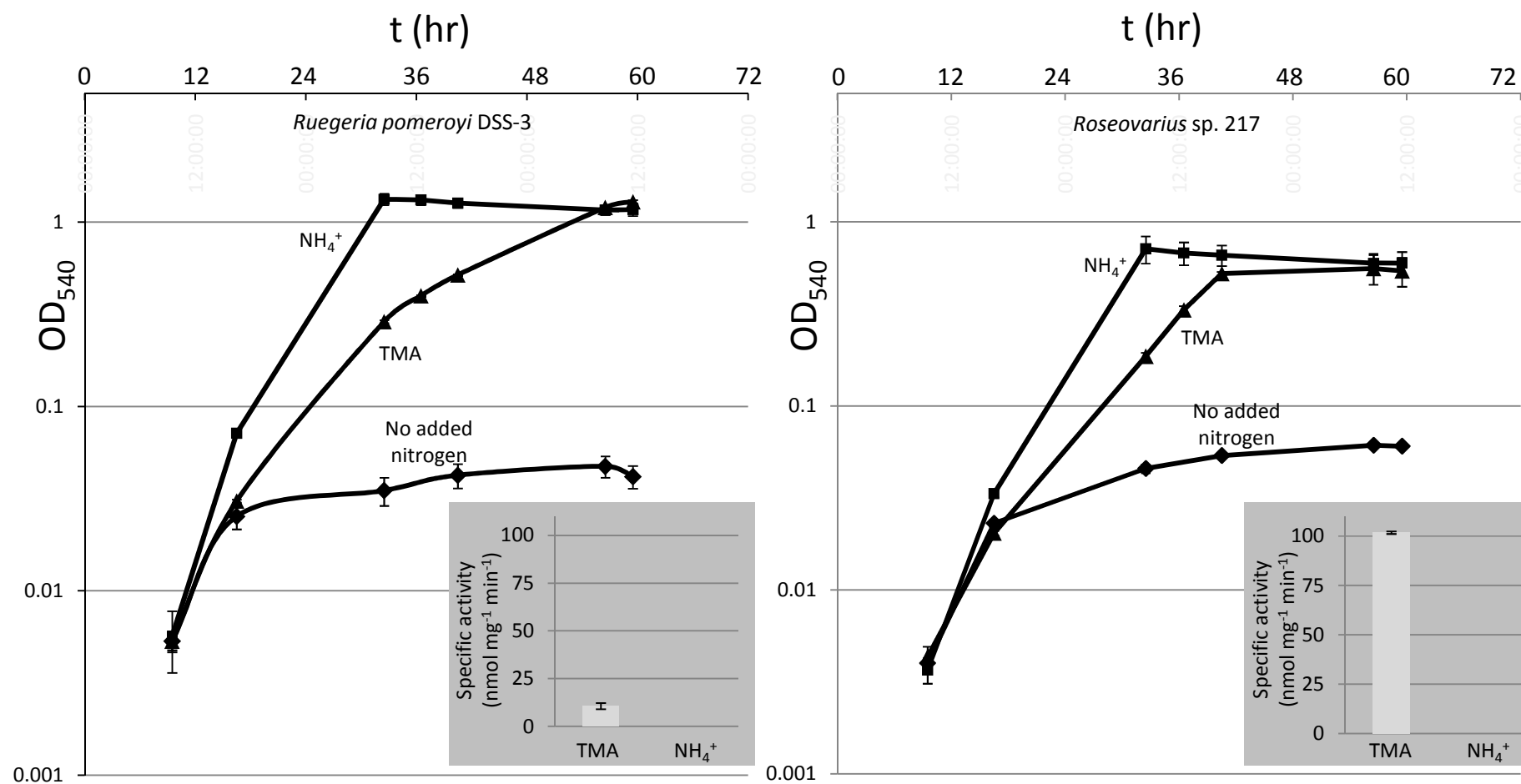


Figure 4

