

# Mechanistic insight into DsyB/DSYB, key enzymes in marine dimethylsulfoniopropionate synthesis

**Jonathan Todd** (✉ [jonathan.todd@uea.ac.uk](mailto:jonathan.todd@uea.ac.uk))

University of East Anglia <https://orcid.org/0000-0003-0777-9312>

**Chun-Yang Li**

Ocean University of China

**Jason Crack**

University of East Anglia

**Simone Newton-Payne**

University of East Anglia

**Andrew Murphy**

University of Warwick

**Xiu-Lan Chen**

Shandong University

**Benjamin Pinchbeck**

University of East Anglia

**Shun Zhou**

University of East Anglia

**Beth Williams**

University of East Anglia

**Ming Peng**

Shandong University

**Xiao-Hua Zhang**

Ocean University of China <https://orcid.org/0000-0002-7428-7775>

**Yin Chen**

University of Warwick <https://orcid.org/0000-0002-0367-4276>

**Nick Le Brun**

University of East Anglia <https://orcid.org/0000-0001-9780-4061>

**Yu-Zhong Zhang**

Shandong University

---

Article

**Keywords:** Marine Algae and Bacteria, Haptophyte and Dinoflagellate Algae, *Nisaea denitrificans*, Crystal Structure, Nucleophilic Attack, SAM Methyl Group

**Posted Date:** May 3rd, 2021

**DOI:** <https://doi.org/10.21203/rs.3.rs-454893/v1>

**License:** © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.  
[Read Full License](#)

---

# Abstract

Marine algae and bacteria produce eight billion tonnes of the organosulfur molecule dimethylsulfoniopropionate (DMSP) in Earth's surface oceans every year. DMSP is an anti-stress compound and, once released into the environment, a major nutrient, signalling molecule and source of climate-active gases. The methionine transamination pathway for DMSP synthesis is used by most known DMSP-producing algae and bacteria. The *S*-directed *S*-adenosylmethionine-dependent methyltransferase (SAM-MT) 4-methylthio-2-hydroxybutyrate (MTHB) *S*-methyltransferase, encoded by the *dsyB/DSYB* gene, is the key enzyme of this pathway, generating *S*-adenosylhomocysteine (SAH) and 4-dimethylsulfonio-2-hydroxybutyrate (DMSHB). *dsyB/DSYB*, present in most DMSP-producing bacteria and haptophyte and dinoflagellate algae with the highest known DMSP concentrations, is shown to be far more abundant and transcribed in marine environments than any other known DMSP synthesis pathway *S*-methyltransferase gene. Furthermore, we demonstrate *in vitro* activity of the bacterial DsyB enzyme from *Nisaea denitrificans*, and provide its crystal structure in complex with SAM and SAH-MTHB, which together provide the first mechanistic insights into a DMSP synthesis enzyme. Structural and mutational analyses imply that DsyB adopts a novel mechanism, distinct from any previously reported SAM-MT, in which the DsyB residue Tyr142 activates the sulfur atom of MTHB for nucleophilic attack on the SAM methyl group. Sequence analysis suggests that this mechanism is common to all bacterial DsyB enzymes and also, importantly, eukaryotic DSYB enzymes from e.g., algae that are the major DMSP producers in Earth's surface oceans.

## Introduction

Approximately eight billion tonnes of the compatible solute dimethylsulfoniopropionate (DMSP) are produced annually in Earth's surface waters (1), constituting up to 10% of surface ocean organic carbon (2). Many marine algae, bacteria, corals and some plants produce DMSP (3) for its proposed functions as e.g., a compatible solute (4), grazing deterrent (5), antioxidant (6) and protectant against hydrostatic pressure (7). Furthermore, DMSP is a major nutrient for marine microorganisms, and a precursor for climate-active volatiles such as dimethyl sulfide (DMS) (3, 8, 9). DMSP was thought to be mainly produced by marine algae in Earth's surface oceans, but recent studies suggest that bacteria, particularly in marine sediment, are also important DMSP producers (3, 10-12).

Three pathways for DMSP synthesis have been proposed based on the identification of intermediates and enzyme activities in various model DMSP producers: a methylation pathway in some plants and bacteria, a transamination pathway in algae and bacteria, and a decarboxylation pathway in one dinoflagellate (10, 12-15) (Fig. 1). Of these (Fig. 1), the transamination pathway is thought to be the most important in marine environments as it functions in the majority of DMSP producing algae (spanning dinoflagellates, haptophytes and diatoms) and bacteria (10, 12, 14). The committed enzyme of the transamination pathway (Fig. 1) is an *S*-adenosylmethionine (SAM)-dependent 4-methylthio-2-hydroxybutyrate (MTHB) *S*-methyltransferase that yields 4-dimethylsulfonio-2-hydroxybutyrate (DMSHB) (13, 14, 16). Recently, the key MTHB *S*-methyltransferase enzyme 'DsyB' (Fig. 1), was identified in many

DMSP-producing marine Alphaproteobacteria (10). Functional DsyB homologues (~33% amino acid identity), termed DSYB, exist in most DMSP-producing dinoflagellates, haptophytes, corals and ~ 20% of diatoms (12). This is consistent with the detection of DMSHB and its oxidative decarboxylation to DMSP in some prymnesiophytes, diatoms and prasinophytes (14). *dsyB* and *DSYB* genes are robust indicators of an organism's potential to produce DMSP (10, 12). The centric diatom *Thalassiosira pseudonana*, which lacks DSYB, contains an isoform MTHB *S*-methyltransferase enzyme termed TpMMT, but this enzyme has not been studied in any other organism (17). Published and new analysis here (see below) shows that *DsyB*/*DSYB* are far more abundant in known DMSP-producing microorganisms (phytoplankton and bacteria) and in marine environmental metagenome and metatranscriptome datasets than other identified DMSP synthesis genes (10-12). Furthermore, acquisition of *dsyB* is sufficient to enable some organisms to produce DMSP (10). Together these data suggest that transamination using DsyB/*DSYB* enzymes is the most important marine DMSP synthesis pathway.

DSYB and DsyB belong to the SAM-dependent methyltransferase (SAM-MT) family (10, 12). SAM is the second most widely used enzyme substrate after ATP and is involved in many important biological processes (18). SAM-MTs are categorized based on the methyl accepting atom, usually *O*, *N*, *C* or *S* (19). The majority (54%) of known SAM-MTs are *O*-directed, whereas only 3% are *S*-directed, and up to now, all were in plants and mammals (19). SAM-MTs, which catalyse transmethylation via  $S_N2$  nucleophilic substitution (20), have evolved three distinct mechanisms: proximity and desolvation (PD), general acid/base-mediated catalysis, and a metal-dependent mechanism (19). To our knowledge, DsyB was the first reported *S*-directed SAM-MT identified in bacteria, and its mechanism is unknown.

Here, we investigate *Nisaea denitrificans* DR41\_21, a marine alphaproteobacterium (21) predicted to produce DMSP, and characterize its DsyB enzyme. With no reported protein crystallographic or mechanistic studies of DsyB/*DSYB* or any DMSP synthesis enzyme, X-ray crystallography and mutational analyses were employed to establish the structure and mechanistic properties of a functional DsyB enzyme and its interaction with SAM and MTHB substrates. Our results provide the first structural and mechanistic data for any DMSP synthesis enzyme. Furthermore, sequence alignment and structural analysis are used to elucidate mechanistic similarities between bacterial DsyB and algal DSYB enzymes. We propose the mechanism for DsyB/*DSYB*, which is the first for any bacterial/algal *S*-methyltransferase, and is distinct from all previously reported SAM-MT enzymes. We also probe marine microorganisms, metagenomes and metatranscriptomes for the presence of DsyB/*DSYB* and other key SAM-MT in DMSP synthesis pathways to investigate the importance of these proteins in the global oceans. Thus, this study identifies DsyB/*DSYB* as a new class of SAM-MT enzymes with a central role in driving marine microbial cycling of organosulfur.

## Results And Discussion

### *Nisaea denitrificans* DR41\_21 is a DMSP-producing bacterium

Isolated from coastal Mediterranean Sea surface waters, *Nisaea denitrificans* DR41\_21 (DSM 18348) is a marine alphaproteobacterium of the *Rhodospirillaceae* family which was not previously known to produce DMSP (21). *N. denitrificans* contains a DsyB homologue, 337 amino acid residues in length with 59% identity to *Labrenzia aggregata* DsyB and is thus predicted to make DMSP (10). Indeed, cloned *N. denitrificans dsyB* conferred onto *Rhizobium*, a heterologous host that lacks DsyB and makes no DMSP, MTHB S-methyltransferase activity. Furthermore, *N. denitrificans dsyB* fully restored DMSP production ( $105 \pm 3.4$  pmol DMSP  $\mu\text{g protein}^{-1}$ ) to an *L. aggregata dsyB* deletion mutant, which produces no DMSP without it (wild type *L. aggregata* produces  $99.8 \pm 1.2$  pmol DMSP  $\mu\text{g protein}^{-1}$  (10)). *N. denitrificans* itself produced DMSP when grown in the absence of methylated sulfur compounds, and production and *dsyB* transcription was enhanced by increased salinity and by nitrogen starvation (Fig 2A-B). This work further confirms that the presence of *dsyB* and its transcription in a bacterium reports on the ability of the strain to produce DMSP and the levels it makes, respectively.

### ***In vitro* characterization of *N. denitrificans* DsyB**

As shown above and in (10), DsyB has MTHB S-methyltransferase activity when expressed in alphaproteobacterial hosts. However, *L. aggregata* DsyB and *Chrysochromulina* DSYB enzymes (12), had no detectable MTHB S-methyltransferase activity when expressed in *E. coli*. The same was generally true of the recombinant *N. denitrificans* DsyB enzyme purified from *E. coli*, although variable MTHB S-methyltransferase activity was observed *in vitro* by hydrophilic interaction chromatography mass spectrometry (HILIC-MS), see below. The reason for this lack of activity upon isolation is unknown; one possibility is that the enzyme requires an essential co-factor or modification that was provided by an algal or alphaproteobacterial host, but not by *E. coli* (12). This hypothesis was initially supported by the fact that addition of heat denatured cell lysate fractions (from a PD10 desalting column) liberated from the *L. aggregata dsyB* deletion mutant, which produces no DMSP, recovered *N. denitrificans* DsyB MTHB S-methyltransferase activity (Fig. S1). Similar complementation was shown with the addition of heat-killed *Prymnesium parvum* extract to DSYB in (12). The activated DsyB protein was shown to have  $K_m$  and  $V_{max}$  values of 0.14 mM and 365 nmol min<sup>-1</sup> mg protein<sup>-1</sup>, respectively, for MTHB (Fig. 2C), which were similar to those previously established for *P. parvum* DSYB (0.09 mM and 294 nmol min<sup>-1</sup> mg protein<sup>-1</sup>) in (12). The activated DsyB had a  $K_m$  of 0.16 mM and  $V_{max}$  368.9 nmol min<sup>-1</sup> mg protein<sup>-1</sup> for the co-substrate SAM (Fig. 2D), which were also similar to those obtained with *P. parvum* DSYB (0.06 mM and 303 nmol min<sup>-1</sup> mg protein<sup>-1</sup>) in (12).

Liquid chromatography with mass spectrometry (LC-MS) and/or native mass spectrometry was used in an attempt to identify the activation factor in the *L. aggregata dsyB* extract (Fig. S2). A prominent peak at 37,084 Da was observed in the LC-MS spectrum for both the as-isolated and activated samples, which corresponds to DsyB with its N-terminal Met residue cleaved (commonly observed for proteins over-expressed in *E. coli*) (22). A lower intensity peak at +131 Da, corresponding to the full-length protein (37,215 Da), was also observed in the as-isolated sample, indicating that the Met cleavage was not complete (Fig. S2A). There was an additional minor peak at +269 Da of unknown origin in the activated

sample (Fig. S2B). Under non-denaturing conditions, both monomeric and dimeric forms of DsyB were detected in the as-isolated sample, a feature commonly observed in non-denaturing mass spectra of solution dimers (23-25). In the monomeric region, the main protein peaks (due to cleaved and non-cleaved proteins) were again observed, but, in addition, a number of adduct species were present in the spectrum. Two of these, at +36 and +98 Da, correspond to chloride and (most likely) phosphate adducts. An additional adduct at +63 Da was also observed, possibly due to metal ion binding. In general, the spectrum of the dimeric form of DsyB was less well-resolved, but the main protein peak (at 74,168 Da) along with chloride and possible metal ion adducts were all detected (Fig. S3). The non-denaturing mass spectrum of the activated monomeric DsyB (Fig. S2B) revealed a number of adducts, including those most likely due to chloride, phosphate and metal ion binding (all common within the as-isolated DsyB), along with an additional adduct at +122 Da (and at +244 Da), which is likely due to Tris buffer. Thus, we have no data to support there being a cofactor or modification of DsyB caused by the addition of the heat-killed *L. aggregata dsyB* extract to as-isolated DsyB, and further work is required to understand the variable nature of DsyB activity (see below).

The association of metal ions with DsyB was investigated further. Inductively coupled plasma mass spectrometry (ICP-MS) analysis revealed variable metal ion content with some preparations of as-isolated DsyB containing up to 0.85 Cu per protein, with other metals such as Ni (up to 0.5 per protein), Zn (0.4) and Fe (0.14) also detected. However, there was no correlation between metal ion content and activity of as-isolated samples.

Despite mostly lacking consistent *in vitro* MTHB S-methyltransferase activity (see below), native MS showed that the as-isolated *N. denitrificans* DsyB enzyme binds to SAM. The deconvoluted mass spectrum of a DsyB sample under non-denaturing conditions and containing 25 equivalents of SAM contained a peak in the DsyB dimer region at +870 Da (predicted mass of a  $(\text{DsyB})_2\text{-(SAM)}_2$  is 74,966 Da), indicative of a  $(\text{DsyB})_2\text{-(SAM-Cl)}_2$  adduct, that was not observed in the absence of SAM (Fig. S4A). Evidence for SAM binding was also apparent in the monomer region, though the presence of chloride adducts spreads out the intensity in this region (Fig. S4B). Evidence for an MTHB-bound form of DsyB was also observed (although, again, the presence of chloride adducts spreads out intensity, Fig. S4B). Although care is needed in interpreting intensities of peaks in the non-denaturing mass spectrum, the low intensity of the SAM- and MTHB-bound forms of DsyB suggests relatively low affinities when these substrates are present individually.

Small molecule HILIC-MS analysis of reactions following addition of MTHB to DsyB-SAM resulted in the detection of substrates SAM and MTHB, and products DMSHB and SAH (Fig. 3). Non-denaturing mass spectrometry of similarly generated samples resulted in the loss of SAM- and MTHB-bound forms of DsyB (Fig. S4B). Together, the data are consistent with DsyB being a SAM-dependent MTHB S-methyltransferase.

## Overall structure of DsyB

To analyse the catalytic mechanism of DsyB, we solved the crystal structures of complexes of DsyB with SAM and with SAH-MTHB. The crystal structure of the DsyB-SAM complex was determined by the single-wavelength anomalous dispersion (SAD) method using a selenomethionine derivative (Se derivative) (Table S1).

Crystals of the DsyB-SAM complex belonged to the  $P2_12_12_1$  space group, with four molecules arranged as a tetramer in the asymmetric unit. Each DsyB molecule contains two domains, an N-terminal domain (N-domain, Met1-Ala125) and a C-terminal domain (C-domain, Thr126-Glu337), which can be seen binding to the SAM molecule (Fig. 4A). The DsyB C-domain contains seven  $\beta$ -strands surrounded by six  $\alpha$ -helices, which together adopt the typical Rossmann-like  $\alpha/\beta$  fold of Class I SAM-MTs (Fig. 4A). Structural analysis showed that two DsyB monomers are tightly intertwined, mainly through interactions of residues from the N-domains of two adjacent monomers (Fig. 4A). Analysis of DsyB using the PISA server ([http://www.ebi.ac.uk/msd-srv/prot\\_int/pistart.html](http://www.ebi.ac.uk/msd-srv/prot_int/pistart.html)) predicted the DsyB dimer to be stable in solution. Indeed, gel filtration analysis indicated that DsyB is likely a dimer in solution (Fig. 4B), consistent with the non-denaturing mass spectrometry data above (Fig. S3, S4A). These results indicate that DsyB functions as a dimer in the same way as other SAM-MTs, whose N-domains are also responsible for dimerization (19, 26, 27). ICP-MS and LC-MS analyses showed that the as-isolated DsyB contained variable metals. However, in the crystal structure of DsyB-SAM complex, no explicit electron density associated with metals was observed, suggesting that the binding site of metals may not be specific in DsyB.

The crystals of the DsyB-SAH-MTHB complex belong to the  $P2_1$  space group (Table S1) and the resulting structure has a similar overall structure to that of the DsyB-SAM complex (Fig. 4C). Interestingly, in this case the MTHB co-substrate molecule is located between the C-domain and the N-domain of one DsyB monomer (Fig. 4C).

### The conformational change of DsyB in binding SAM

During structural refinement of the DsyB-SAM complex, we found that three monomers (chain A, B and C) of the asymmetric unit contained SAM molecules. The structures of these three monomers are similar, with root mean square deviations (RMSDs) of  $\sim 0.5$  Å between any two monomers. The chain D of the DsyB-SAM complex is not bound to a SAM molecule. The conformation of chain D is different to the other monomers bound to SAM, with a RMSD of  $\sim 2.4$  Å between chain D and chain A. Moreover, residues Asp123 to Tyr143 in chain D exhibited weak electron density, suggesting that this region is highly flexible. By superposing molecules of chain A and chain D, we observed that the N-domains of chain A and chain D are almost completely aligned, whereas the C-domain rotates  $\sim 10^\circ$  as a rigid body (Fig. 4D). These structural differences indicate that DsyB possesses two conformations: an “open” form and a “closed” form. Although DsyB can bind SAM and MTHB individually (Fig. S4B), the binding of SAM triggers the conformational change of DsyB from the “open” form to the “closed” form, shrinking the cavity between the N-domain and the C-domain of DsyB and possibly promoting the subsequent binding of MTHB.



The structure of the DsyB-SAM complex is similar to that of the *Streptosporangium sibiricum* SibL protein (PDB code: 4U1Q), a C-directed Class I SAM-MT, with an RMSD of  $\sim 1.3$  Å between these two structures. SAM binding also triggers the conformational change of SibL from an “open” form to a “closed” form to complete the formation of a binding site for its methyl acceptor 3-hydroxykynurenine (26). Similar conformational changes have also been observed in other C/O-directed SAM-MTs, despite their low sequence identities (19, 27).

### Binding sites of SAM and MTHB

The SAM molecule within the DsyB-SAM complex (Fig. 4A) is bound mainly by hydrogen bonds with residues in the DsyB C-terminal domain. DsyB residues Asp223 and Ala224 participate in binding the adenine ring of SAM; Asp196 forms hydrogen bonds with the ribose moiety of SAM; and Ser150, Gly173 and Ser239 interacts with the terminal amino acid moiety of SAM (Fig. 5A). A similar binding mode was observed between the same DsyB C-terminal residues for SAH in the DsyB-SAH-MTHB complex.

In the DsyB-SAH-MTHB complex, the electron density of the MTHB molecule is relatively poor and the distance between SAH and MTHB is more than 6 Å, which is too far to enable the methyl transfer reaction. We speculate that the position of MTHB observed in the structure is not the exact location of MTHB when the reaction occurs under physiological conditions, and that the observed structure represents a state where the MTHB molecule has not completely entered into the active site. Nevertheless, the location of MTHB clearly implies its initial binding site (Fig. 5B), and several residues likely involved in the binding of MTHB were identified, including Tyr97 and Gln101 from the N-domain and Tyr129, Tyr142, Gln146 and His291 from the C-domain of DsyB (Fig. 5C).

To determine the importance of the Tyr97, Gln101, Tyr129, Tyr142, Gln146 and His291 in DsyB binding to MTHB, we performed site-directed substitutions of these residues and assayed the *in vivo* MTHB S-methyltransferase activity of the resultant variant DsyB derivatives in *R. leguminosarum*. *In vivo* assays were performed in *Rhizobium* because of the sensitivity of *in vitro* assays, see above. Site-directed mutations of Tyr97, Gln101, Tyr129, Tyr142, Gln146 or His291 to alanine severely decreased the enzymatic activity of DsyB (Fig. 5D), indicating the potentially important roles of these residues in binding MTHB. In particular, activity of the Tyr97Ala and Tyr142Ala variants was completely abolished (Fig. 5D), suggesting that these two residues play key roles in the methyl transfer reaction.

Based on this mutational analysis, residues Tyr97 and Tyr142 of DsyB are the best candidates for catalytic residues. In the structure of the DsyB-SAH-MTHB complex, the distance between Tyr97 and MTHB is 3.9 Å, which is similar to the distance between Tyr142 and MTHB (3.7 Å). However, as discussed above, the DsyB-SAH-MTHB complex structure observed is likely an intermediate state, since the MTHB molecule needs to be closer to SAM to attack the SAM methyl group. Because the phenolic sidechain of Tyr97 is pointing in the opposite direction to SAM, and that of Tyr142 is located in between MTHB and SAM (Fig. 5C), movement of MTHB from its initial binding site (observed in the structure) towards SAM would lengthen the distance between Tyr97 and MTHB, and shorten the distance between Tyr142 and MTHB. Therefore, Tyr142 is the most probable catalytic residue of DsyB. Interactions with Gln101, Tyr129



and Gln146 might be important in lowering the  $pK_a$  of the phenolic proton of Tyr142, which may be essential for the activation of MTHB. The residue Tyr97 may play a vital role in MTHB binding based on mutational and structural analyses.

Three distinct catalytic mechanisms have been reported for SAM-MTs, including the PD mechanism, the general acid/base-mediated mechanism and the metal-dependent mechanism (19). Structural and biochemical analyses indicate that the activity of DsyB is not metal-dependent, but is catalytic residue-dependent. Furthermore, MTHB, the substrate of DsyB, does not contain a thiol group, and so activation of MTHB for attack on SAM cannot occur via deprotonation. This implies that DsyB does not follow the typical acid/base mechanism where the substrate is activated by deprotonation. Thus, the catalytic mechanism of DsyB likely represents a novel catalytic strategy for SAM-MTs. Rather than acting as a catalytic base, the catalytic residue Tyr142 most likely activates MTHB by making the sulfur atom more nucleophilic, which is essential for the subsequent attack of MTHB on SAM.

### **The catalytic mechanism of DsyB**

Based on our structural and biochemical results, we propose DsyB first binds a SAM molecule to generate a conformational change from “open” state to “closed” state, which promotes the binding of MTHB (Fig. 6A). When an MTHB molecule enters the active site, the deprotonated phenolic group of Tyr142 in DsyB activates the sulfur atom of MTHB for nucleophilic attack on the methyl group of SAM (Fig. 6B). Subsequently, the generated DMSHB and SAH are released, and DsyB can re-bind a SAM molecule from the intracellular environment in preparation for the next reaction.

Currently DsyB is the sole known bacterial SAM-dependent *S*-methyltransferase with a crystal structure, although there are *S*-directed SAM-MT structures available from mammal and plant enzymes (19). For example, thiopurine *S*-methyltransferase (TPMT) is a murine *S*-directed SAM-MT that methylates 6-mercaptopurine (28). Unlike DsyB, TPMT does not contain an N-terminal domain likely involved in dimerization, as its N-terminus only constitutes 40 residues (28). Although Arg147 and Arg221 are possible participants in 6-mercaptopurine deprotonation, the modest decrease in the enzymatic activities of the corresponding mutants suggests that TPMT may possess the PD strategy for catalysis (28). *Catharanthus roseus* CrSMT1 is another *S*-directed SAM-MT that methylates a broad range of substrates including benzene thiol and furfuryl thiol (29). Homology modelling suggested that CrSMT1 contains an N-domain for dimerization (29), which is similar to DsyB. However, CrSMT1 is thought to use a histidine residue as a general base to deprotonate the thiol group of the substrate (29). Thus, although the *S*-directed SAM-MTs only constitute a small portion of the reported SAM-MTs (19), their catalytic mechanisms appear diverse. Moreover, the catalytic mechanism of DsyB proposed here represents a novel catalytic strategy of SAM-MTs.

### **Universality of the catalytic mechanism of DsyB**

The majority of bacteria containing DsyB are Rhodobacterales, which are abundant in marine environments, but this enzyme is also found in some Rhizobiales and Rhodospirillales (including *N.*

*denitrificans*) (10, 30). To investigate the ubiquity of the DsyB catalytic mechanism, we performed sequence alignment of DsyB proteins from different Rhodobacterales, Rhizobiales and Rhodospirillales bacteria (Fig. S5). Most residues involved in initial MTHB binding (Tyr97, Gln101, Tyr129, Gln146 and His291) and SAM binding (Ser150, Gly173, Asp196, Asp223, Ala224 and Ser239), and the catalytic residue (Tyr142) are highly conserved in DsyB proteins from these marine bacteria, indicating that mechanistic insight gained here for *N. denitrificans* DsyB has universal significance in bacteria containing DsyB.

Eukaryotic DSYB, which may originate from bacterial DsyB, is a key enzyme for DMSP synthesis in many phytoplankton, such as marine haptophytes, dinoflagellates and some diatoms (12). DSYB shares 33% sequence identity to DsyB, and we predicted the structure of DSYB from *Chrysochromulina tobin* CCMP291 by homologous modelling using SWISS-MODEL (<https://swissmodel.expasy.org/>) (31). Structural alignment of DSYB and DsyB indicated that residues involved in binding MTHB and the catalytic residue are perfectly superposed (Fig. S6). Moreover, sequence alignment of DsyB and DSYB from different marine algae showed that residues which play important roles in DsyB are highly conserved in different DSYB proteins (Fig. S7), suggesting that DSYB proteins adopt a similar catalytic mechanism to DsyB.

### **DsyB/DSYB are the most abundant and transcribed S-methyltransferase enzymes of known DMSP synthesis pathways in marine microorganisms and environments**

Having the identity of the key S-methyltransferases in diverse DMSP synthesis pathways (DsyB/DSYB and TpMMT in the bacterial and algal transamination pathway, and MmtN and BurB in bacterial methylation pathways, Fig. 1), we carefully analysed their presence in marine microorganisms and their abundance and transcript levels in published global 'omics datasets to quantify the potential environmental importance of these pathways.

Of the known DMSP synthesis enzymes DsyB is by far the most abundant in sequenced and/or isolated bacteria (65.8 % of cultured isolates) (Table S2)(7, 10, 11, 32, 33). DsyB is mostly found in alphaproteobacterial Rhodobacterales, Rhizobiales and Rhodospirillales, but is also sporadically found in e.g. an actinobacterial *Ponticoccus* isolate (7), and in some Betaproteobacteria and Bacteroidetes metagenome assembled genomes (34). MmtN has much fewer (14.4%), but equally diverse, host bacteria, being found in Alphaproteobacteria, Gammaproteobacteria and Actinobacteria (Table S2). Finally, BurB is confined to very closely related *Burkholderia* spp. that likely use DMSP as an intermediate in toxin production (32).

This hierarchy of DMSP synthesis gene abundance in bacteria (DsyB > MmtN > BurB) was mirrored in marine environmental data. In the *Tara* Oceans prokaryotic database, both *dsyB* and *mmtN* were found throughout the water column (Fig. 7A), but no close homologues of *BurB* (e-value <1e-40) were detected. This is consistent with BurB-mediated DMSP production in *Burkholderia* spp., possibly for toxin production, not being an important process in marine systems. *DsyB* was significantly more abundant

than *mmtN* in both the metagenomic (Kruskal-Wallis  $X^2 = 83.781$ ,  $p < .001$ ) and metatranscriptomic (Kruskal-Wallis  $X^2 = 33.64$ ,  $p < .001$ ) *Tara* Oceans datasets (Fig. 7A). Additionally, *dsyB* was found at 172 and 153 sampling sites (treating each depth as a separate site) in the metagenomes and metatranscriptomes, respectively, whereas *mmtN* was found at only 74 and 63 sites, respectively. Given this, our analysis of median abundance overestimates the contribution of *mmtN* to DMSP production in the global ocean. We therefore determined the relative abundance of *dsyB:mmtN* across depths at each sampling site in both metagenomes and metatranscriptomes (Fig. S8). *DsyB* was more abundant at almost all sites in the metagenomes, and was more highly expressed across most, though there were a number of locations in the South Atlantic and South Pacific where *mmtN* was predominant (Fig. S8). Taxonomic examination of both *dsyB* and *mmtN* sequences in the *Tara* database (Fig. 7B) showed that both genes were exclusively from Alphaproteobacteria, primarily within the Orders Rhodobacterales and Rhizobiales for *dsyB*, and the genus *Thalassospira* for *mmtN* (Fig. 7B). These data highlight *DsyB* as the most abundant, transcribed and, likely, important of the known bacterial DMSP synthesis enzymes in marine waters, which likely play a significant role in the global production of DMSP.

Moving to eukaryotic DMSP synthesis, we carefully analysed available transcriptome data from marine eukaryotes in the marine microbial eukaryote transcriptome sequencing project (MMETSP) (35). The TpMMT MTHB S-methyltransferase has only been characterised in the centric diatom *Thalassiosira pseudonana* CCMP1335, and close homologues (~70% protein identity) with the same singular domain structure only exist in 17/82 diatom transcriptomes (7 of which also contain *DSYB*), and no other phytoplankton (Table S3 and S4) (36-51). The next most homologous TpMMT-like proteins, present in e.g. *Thalassiosira oceanica* (EJK59074) and *Fistulifera solaris* (GAX25165) are more diverse (the methyltransferase domain being <50% identical to TpMMT), contain extra protein domains and, thus, are much larger proteins whose function is unknown. These TpMMT-like proteins cannot be considered as functional MTHB S-methyltransferase enzymes and were omitted from this study. In contrast, *DSYB* is found in the transcriptomes of 47/61 dinoflagellates and 24/30 haptophytes, organisms known to produce the highest levels of DMSP per cell (>50mM) (37, 44). Furthermore, 15/82 diatom transcriptomes, typically known to produce lower cellular DMSP levels (generally < 50 mM) (37), and some Ochrophyta, Cnidaria and Cilophora transcriptomes also contained *DSYB*. These data show *DSYB* to be the most abundant and widespread DMSP synthesis enzyme known in eukaryotic DMSP-producing organisms.

Within the eukaryotic Marine Atlas of Tara Ocean Unigenes (MATOU) we found both *DSYB* and *TpMMT* within epipelagic (surface, SRF and deep chlorophyll maximum, DCM) waters. Initial examination showed *DSYB* to be more abundant in  $\leq 3 \mu\text{m}$  fractions than in larger fractions (Fig S9). Data from these smaller  $\leq 3 \mu\text{m}$  fractions that likely contain picoeukaryotes (CoP) were considered together. Likewise, fractions with a minimum filter size of  $\geq 3 \mu\text{m}$  that likely exclude picoeukaryotes (ExP) were also considered together. Abundance was not significantly different between SRF and DCM sampling depths for either CoP or ExP *DSYB* (Kruskal-Wallis  $X^2 = 0.113$ ,  $p = .74$ , and Kruskal-Wallis  $X^2 = 0.004$ ,  $p = .95$ , respectively), or for CoP or ExP *TpMMT* (Kruskal-Wallis  $X^2 = 0.102$ ,  $p = .75$ , and Kruskal-Wallis  $X^2 = 0.194$ ,  $p = .66$ , respectively), and as

such these sampling depths were combined for the purposes of comparative analyses between *DSYB* and *TpMMT* abundance. *DSYB* was significantly more abundant than *TpMMT* in both CoP and ExP fractions of both the MATOU metagenome and metatranscriptome (Fig. 8A) (*DSYB* vs *TpMMT* CoP post-hoc Dunn's test  $z = 16.22$ ,  $p < .001$ , ExP post-hoc Dunn's test  $z = 6.97$ ,  $p < .001$ , and CoP post-hoc Dunn's test  $z = 15.16$ ,  $p < .001$ , ExP post-hoc Dunn's test  $z = 7.33$ ,  $p < .001$ , respectively). *DSYB* was significantly more abundant within the CoP fraction than within the ExP fraction, again for both the MATOU metagenome and metatranscriptomes (Fig. 8A), (post-hoc Dunn's test  $z = 11.30$ ,  $p < .001$ , and post-hoc Dunn's test  $z = 12.20$ ,  $p < .001$ , respectively). In contrast, *TpMMT* abundance was not significantly different between CoP and ExP fractions (Fig. 8A) (post-hoc Dunn's test  $z = 1.61$ ,  $p = .11$ , and post-hoc Dunn's test  $z = 0.92$ ,  $p = .35$ , respectively). Again, these analyses likely overestimated the abundance of *DSYB* in the ExP fraction, and *TpMMT* in general, because *DSYB* was detected at 138/140 CoP fraction sites, but was only found at 178/272 ExP fraction sites ( $\chi^2$  (1,  $N = 412$ ) = 56.77,  $p < .001$ ), whereas *TpMMT* was detected at 90/140 CoP fraction sites, and at 39/272 ExP fraction sites ( $\chi^2$  (1,  $N = 412$ ) = 107.21,  $p < .001$ ). As such, *DSYB* was detected significantly more sites than *TpMMT*, in both the CoP ( $\chi^2$  (1,  $N = 280$ ) = 54.41,  $p < .001$ ) and ExP ( $\chi^2$  (1,  $N = 544$ ) = 148.12,  $p < .001$ ) fractions. Similarly, within MATOU\_v1\_metaT, *DSYB* was detected at 139/140 CoP fraction sites and at 251/272 ExP fraction sites ( $\chi^2$  (1,  $N = 412$ ) = 8.98,  $p < .01$ ) whereas *TpMMT* was detected at 94/140 CoP fraction sites, and at 99/272 ExP fraction sites ( $\chi^2$  (1,  $N = 412$ ) = 35.09,  $p < .001$ ). Again, *DSYB* was detected at significantly more CoP ( $\chi^2$  (1,  $N = 280$ ) = 51.78,  $p < .001$ ) and ExP ( $\chi^2$  (1,  $N = 544$ ) = 185.10,  $p < .001$ ) fraction sites than *TpMMT* in the metatranscriptome data. Given the greater abundance of *DSYB* over *TpMMT* in the environmental data, and that the majority of environmental *DSYB* sequences are likely from dinoflagellates and/or haptophytes, known to be high producers (37, 44) of DMSP compared to *TpMMT* in the generally low-producing diatoms (37), *DSYB* is currently the most important known DMSP synthesis enzyme (Fig. 8B).

## Conclusion

DMSP is an abundant and ecologically important organosulfur compound. DsyB/DSYB enzymes catalyse the committed S-methylation of MTHB to generate DMSHB, which is the key step of the transamination pathway for DMSP synthesis in most bacteria and algae (10-12). Furthermore, DsyB/DSYB enzymes are present in the most prodigious DMSP-producing haptophyte and dinoflagellate phytoplankton, and represent the most abundant and transcribed S-methylase genes of known DMSP synthesis pathways in marine waters. In this study, we solved the first crystal structures of bacterial DsyB-SAM and DsyB-SAH-MTHB complexes and demonstrated conversion of SAM and MTHB into SAH and DMSHB. Based on structural and mutational analyses, the catalytic mechanism of DsyB is proposed, which represents a novel catalytic strategy of SAM-MTs, and has universal significance in bacteria containing DsyB, and in marine algae containing DSYB. Our results provide novel insights into DMSP synthesis, shedding light on the global sulfur cycling. This study also offers a better understanding of the catalysis of SAM-MTs, especially the S-directed SAM-MTs.

# Materials And Methods

## Bacterial strains and growth conditions

*E. coli* BL21 (DE3) was grown in Lysogeny Broth (LB) medium at 37°C. *R. leguminosarum* J391 was grown in TY (52) complete medium or Y (52) minimal medium (10 mM NH<sub>4</sub>Cl as nitrogen source) at 28 °C. *L. aggregata* J571 (*dsyB*) was grown in YTSS (53) complete medium or MBM (54) minimal medium (0.5 mM NH<sub>4</sub>Cl as nitrogen source) at 30 °C. *N. denitrificans* DR41\_21 (DSM 18348), purchased from DSMZ, Germany, was cultured in the Difco 2216 medium at 30°C (<http://www.dsmz.de/>) or MBM medium under different salinity and nitrogen levels for differential DMSP production experiments. Standard conditions were 10 mM NH<sub>4</sub>Cl and 35 PSU compared to 5 PSU (low salinity) or 50 PSU (high salinity). Cultures were sampled for DMSP and for RT-qPCR work in exponential-phase growth (after ~7 days). For nitrogen starvation experiments exponential phase cells grown under standard conditions were harvested and incubated overnight in standard MBM media with no added NH<sub>4</sub>Cl. Where necessary, 10 mM succinate was used as carbon source and antibiotics were added at the following concentrations: streptomycin (400 µg ml<sup>-1</sup>), kanamycin (20 µg ml<sup>-1</sup>), spectinomycin (200 µg ml<sup>-1</sup>), gentamicin (20 µg ml<sup>-1</sup>), ampicillin (100 µg ml<sup>-1</sup>).

## General *in vivo* and *in vitro* genetic manipulations

Plasmids (Table S5) were transferred to *E. coli* by transformation, and *R. leguminosarum* J391 or *L. aggregata* J571 by conjugation, in a triparental mating using the helper plasmid pRK2013 (55). Routine restriction digestions and ligations for cloning were performed essentially as in Carrion et al., (56). The oligonucleotide primers used for molecular cloning were synthesized by Eurofins Genomics and are detailed in Table S6. Sequencing of plasmids and PCR products was performed by Eurofins Genomics.

The *dsyB* gene was amplified from *N. denitrificans* genomic DNA and cloned into the pET22b (Novagen, America) for the expression of DsyB with a C-terminal His-tag or into pLMB509, for expression in *Rhizobium* and *Labrenzia* (57). Amino acid substitution mutations in DsyB were generated using QuikChange® mutagenesis kit (Agilent Technologies) and the primers in Table S6. All site directed mutant (SDM) variant plasmids were verified by DNA sequencing.

## Reverse transcription quantitative polymerase chain reaction (RT-qPCR)

RNA was isolated from 100 ml *N. denitrificans* cultures using RNeasy Mini Kit (Qiagen) according to the manufacturer's protocol with some modifications. On-Column DNase digestion was performed with RNase-Free DNase Set (Qiagen). Reverse transcription of 1 µg of DNA-free RNA per sample was done using the QuantiTect Reverse Transcription Kit (Qiagen). PCR on RNA and cDNA samples confirmed that RNA samples were DNA-free.

Primers for RT-qPCR for *N. denitrificans* *dsyB* and housekeeping genes *recA* and *gyrB*, were designed using Primer5 (58) (Table S6). Primer efficiencies were all 90–110% and within recommended limits.



qPCR was performed on a StepOnePlus instrument (Applied Biosystems). Quantification was performed using the SYBR<sup>®</sup> Green JumpStart<sup>™</sup> Taq ReadyMix<sup>™</sup> (Sigma-Aldrich) following the manufacturer's instructions. Reactions (20 µl) contained 2 µl cDNA, 0.8 µl of 10 µM primers and with an annealing temperature of 55 °C. For each condition and gene, the cycle threshold (Ct) values of triplicate technical and biological replicates were averaged. Relative expression levels were determined with efficiency correction (59). *dsyB* transcription was displayed as normalized fold change to the standard condition.

### **DsyB enzyme assays *in vivo***

To measure MTHB *S*-methyltransferase activity from pLMB509 clones (and SDM derivatives) in *R. leguminosarum* J391, cultures were grown (in triplicate) overnight in TY complete medium; 1 ml of culture was centrifuged at 20,000 g for 2 min, resuspended in the same volume of Y medium and then diluted 1:100 into 5 ml Y with 10 mM taurine (to induce expression, Sigma-Aldrich, T0625), 0.5 mM DL-MTHB (Sigma-Aldrich, 55875) and gentamicin, and incubated at 28 °C for 60 h before sampling for gas chromatography (GC) analysis to determine the amount of DMSP product (see below).

To measure MTHB *S*-methyltransferase activity from pLMB509 clones expressing the *dsyB* gene in the *L. aggregata dsyB<sup>-</sup>* mutant strain J571, cultures were grown (in triplicate) overnight in YTSS complete medium. Following incubation, 1 ml of culture was then centrifuged at 20,000 g for 2 min, resuspended in the same volume of MBM medium and then diluted 1:50 into 5 ml MBM with 10 mM taurine (to induce expression; Sigma-Aldrich), rifampicin and gentamicin, and incubated at 30 °C for 24 h. DL-MTHB (0.5 mM) was added as substrate to the cultures and these were incubated for 4 h at 30 °C before sampling for GC and protein estimation by the Bradford assay.

To measure DMSHB/DMSP in *Rhizobium* or *Labrenzia* assay mixtures, 200 µl of culture was added to a 2 ml glass serum vial then 100 µl 10 M NaOH was added and vials were crimped with PTFE/rubber crimp caps immediately. Vials were incubated at 80°C for 10 minutes (to capture DMSHB as well as DMSP) and then for 24 h at room temperature in the dark before being monitored by GC assay. All GC assays involved measurement of headspace DMS using a flame photometric detector (Agilent 7890 A GC fitted with a 7693 autosampler) and an HP-INNOWax 30 m × 0.320 mm capillary column (Agilent Technologies J&W Scientific). Calibration curves were produced by alkaline lysis of DMSP standards in water. The detection limit for headspace DMS from DMSP was 0.015 nmol and from DMSHB was 0.3 nmol. DsyB activity is expressed as pmol DMSHB/DMSP mg protein<sup>-1</sup> min<sup>-1</sup>. Protein concentrations were determined using the Bradford method (BioRad). Control assays of *Rhizobium* or *Labrenzia* J571 containing pLMB509 were carried out, as above, and gave no detectable DsyB activity.

### **Protein expression and purification**

*E. coli* BL21 (DE3) containing pET22b::*dsyB* clones were cultured aerobically at 37°C in LB medium containing ampicillin. At mid-exponential growth (OD<sub>600</sub> 0.5–0.7) 0.5 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) was added and the cells incubated at 20°C for 16 h. Cells were harvested by centrifugation (20 min, 7500 × g, 4°C), washed and resuspended in 25 mM Tris-HCl, pH 8.0, 150 mM



NaCl. Cell lysis was performed by three passes through a French Press (16,000 psi), cell debris and unbroken cells were pelleted by centrifugation (30 min, 5,500 × g, 4°C) and the supernatant was recovered and subjected to centrifugation (60 min, 185,000 × g, 4°C) to pellet the membrane fraction. Soluble cell lysate was applied to a slurry of Ni-NTA resin (Qiagen) at a 3:1 ratio for 90 minutes with shaking at 4°C. The lysate/slurry mix was loaded into Econo-Pac polypropylene columns, washed with 50 mM Tris-HCl, 250 mM NaCl, 20 mM imidazole, pH 8.0 and DsyB was eluted in 5 × 1 ml fractions using 50 mM Tris-HCl, 250 mM NaCl, 250 mM imidazole, pH 8.0. Fractions containing DsyB were concentrated and buffer exchanged into 2 ml of 50 mM Tris-HCl, 100 mM NaCl, pH 8.0 and applied to a Superdex 200 10/300GL preparative grade gel filtration column (Cytiva). Purification of DsyB was monitored by electronic absorbance spectroscopy and SDS-PAGE; for the latter, proteins were visualized by Coomassie stain. Aliquots of the purified protein were flash-frozen in liquid nitrogen and stored at -80°C until required.

### **DsyB enzyme assays *in vitro***

*Labrenzia aggregata dsyB* J571 (10) was grown to late exponential phase in MBM. Cell lysates were prepared by centrifuging 100 ml of culture for 10 min at 2,500 g. The pellet was washed and resuspended with 20 ml 20 mM HEPES, 150 mM NaCl, pH 7.5 before cell lysis via French press (16,000 psi). The cell lysate was heat-treated at 80°C for 10 min to denature proteins, then applied to a 10 ml PD10 column, eluted over 10 ml, and collected in 1 ml fractions. DsyB activity was monitored by performing *in vitro* enzyme assays with 50 µl of the separate heat-killed extract fractions, 1 mM SAM (Sigma-Aldrich), 1 mM DL-MTHB and 5 mM DsyB or no protein (control) in 400 µl total volume. Experiments were done as above with purified DsyB without addition of heat-killed extracts, but these gave no activity (data not shown). Reactions were incubated at 28°C for 30 mins and then 800 µl of activated charcoal (38 mg ml<sup>-1</sup> in 0.1 M acetic acid) was added to the samples and mixed. Samples were centrifuged for 15 mins at 14,000 g and the supernatant was retained. For GC analysis, 200 µl of the supernatant was added to a 2 ml vials with 100 µl 10 M NaOH and vials were immediately crimped. Crimped vials were then heated to 80 °C for 10 minutes (to release DMS) and incubated at 22 °C for 16 h in the dark. These samples were subsequently used for quantification of DMS by GC analysis (as above) and reported as nmol DMSHB mg protein<sup>-1</sup> min<sup>-1</sup>. No DMS was produced from the no DsyB protein controls.

For kinetics analysis of *N. denitrificans* DsyB, the as-isolated protein was activated by addition of 400 µl heat-killed fraction 5 (see Fig. S1).  $K_m$  and  $V_{max}$  values were determined by nonlinear analysis based on the initial rates using 1.97 µM DsyB and 0-2 mM SAM, or 0-2 mM DL-MTHB (Fig. 2). The reaction mixture was incubated at 28°C for 30 min before detection DMSHB, as above. The enzyme activities were linear with respect to incubation time and enzyme concentration. Origin version 8.5 was used to calculate  $K_m$ .

### **Mass spectrometry analysis**

Denaturing LC-MS was conducted using a Bruker microQToF-QIII electrospray ionisation time of flight mass spectrometer, operating in positive mode. The spectrometer was calibrated with ESI-L Low concentration tuning mix (Agilent technologies). Samples were prepared by ten-fold dilution of ~100 µM

protein solution with 2% (v/v) acetonitrile and 0.1% (v/v) formic acid to 0.5 ml. Samples were chromatographically separated by an UltiMate 3000 HPLC system (Dionex) fitted with a ProSwift reversed phase RP-1S column (4.6 × 50 mm; Dionex). Hystar (Bruker Daltonics) was used to coordinate mass spectrometer and HPLC operations. Bound proteins were eluted using an isocratic gradient (2 – 100% B) at a flow rate of 0.2 ml min<sup>-1</sup> using the following solvents: solvent A (water, 0.1% (v/v) formic acid); and, solvent B (acetonitrile, 0.1% (v/v) formic acid). The eluant was continuously infused into the source of the mass spectrometer operating with the following parameters: dry gas flow 8 L min<sup>-1</sup>; nebuliser gas pressure 1.8 bar; dry gas 240 °C; capillary voltage 4,500 V; offset 500 V; collision RF 650 Vpp.

Hydrophilic interaction chromatography (HILIC)-MS (60) experiments were performed using the same mass spectrometer and HPLC system fitted with a Luna NH<sub>2</sub> column (2 x 100 mm) (Phenomenex). For HILIC chromatography the following solvents were freshly prepared: solvent A (95% (v/v) aqueous 5 mM ammonium formate pH 3.75, 5% (v/v) acetonitrile); solvent B (95% (v/v) acetonitrile, 5% (v/v) aqueous 100 mM ammonium formate pH 3.75). Standard compounds (SAM, DL-MTHB, SAH (Sigma-Aldrich), DMSHB (10)) were used to calibrate the elution profile of the HILIC column. Samples were brought to 92% (v/v) acetonitrile and loaded onto a column pre-equilibrated with solvent B. An optimized HILIC gradient was applied and compounds eluted (0.6 ml min<sup>-1</sup>) using the HILIC gradient between solvent A and solvent B, as previously described (10). The eluant was continuously infused into the source of the mass spectrometer (optimized for 50 – 600 *m/z*) with the following parameters: dry gas flow 8.5 L/min; dry gas 200 °C; nebulizer pressure 1.2 bar; capillary voltage 4500 V; offset 500 V; collision RF 400 Vpp. Each HILIC-MS run contained an internal sodium formate calibration segment at the end of the run.

For non-denaturing mass spectrometry, an aliquot of proteins was exchanged into 250 mM ammonium acetate pH 8.0 using Zeba spin (Thermo Scientific) or PD mini trap (Cytiva) desalting columns and infused (0.3 ml/h) directly into the ESI source of the mass spectrometer with the following parameters: dry gas flow 4 L min<sup>-1</sup>; nebulizer gas pressure 0.8 bar; dry gas temperature 190 °C; capillary voltage 3000 V; capillary offset 500 V; ion energy 8eV; collisional RF 1500 Vpp; collision cell voltage 5 V; and, ion transmission range 1500 – 5500 *m/z*.

Processing, isotope pattern simulation and analysis of denaturing LC-MS, HILIC-MS and non-denaturing MS data were carried out using Compass Data Analysis version 4.1. For denaturing LC-MS and non-denaturing MS, neutral mass spectra were generated using ESI compass Maximum Entropy deconvolution algorithm version 1.3. Proteins masses are reported from peak centroids representing the isotope average neutral mass and compared to predicted masses (Expasy) (61).

### **DsyB activity detection using HILIC-MS**

Heat-killed J571 fractions that restored MTHB S-methyltransferase were added to pure DsyB as above to yield activated samples for analysis here. Samples of as-isolated DsyB prepared in HEPES buffer or activated DsyB, were immediately desalted (PD10, Cytiva) into 25 mM Tris, 100 mM NaCl pH 8.0 prior to

conducting HILIC or non-denaturing MS experiments. The methyltransferase activity of DsyB was measured using DL-MTHB and SAM as substrates, as previously described (17), with a slight modification. The reaction mixture (20 µl) consisted of 14 µl water, 2 µl of buffer (100 mM Tris-HCl, pH 7.5), 0.5 µl of 20 mM DL-MTHB, 1.5 µl of 32 mM SAM as co-substrate, and 2 µl of enzyme solution (DsyB in range 7 – 30 µM depending on particular experiment). The reaction was incubated at 25 °C, over-night and quenched by the addition of 230 µl acetonitrile. Samples were analysed immediately by mass spectrometry.

### Crystallization and data collection

The purified DsyB protein was concentrated to ~8 mg/ml in 10 mM Tris-HCl (pH 8.0) and 100 mM NaCl. Initial crystallization trials for DsyB were performed at 20°C using the sitting drop vapor diffusion method. To obtain crystals of DsyB-SAM complex, the purified DsyB protein was mixed with 1 mM SAM at 4°C for 30 min. Diffraction-quality crystals of DsyB-SAM complex were obtained in hanging drops containing 0.2 M NaCl, 0.1 M Hepes (pH 7.5) and 25% (wt/vol) polyethylene glycol (PEG) 3350 at 20°C after 1-week incubation. Crystals of the DsyB-SAM complex Se derivative were obtained in hanging drops containing 0.2 M sodium acetate, 0.1 M Bis-Tris propane (pH 7.5) and 20% (wt/vol) PEG 3350 at 20°C after 1-week incubation. To obtain crystals of DsyB-SAH-MTHB complex, the purified DsyB protein was mixed with SAH (1 mM) and DL-MTHB (1 mM) at 4°C for 30 min. Crystals of DsyB-SAH-MTHB complex were obtained in hanging drops containing 0.2 M NaCl, 0.1 M Tris (pH 8.0) and 20% PEG 4000 at 20°C after 1-week incubation. X-ray diffraction data were collected on the BL18U1&BL19U1 beamlines at the Shanghai Synchrotron Radiation Facility. The initial diffraction data sets were processed by the HKL3000 program (62).

### Structure determination and refinement

The crystals of DsyB-SAM complex belong to the  $P2_12_12_1$  space group, while the crystals of DsyB-SAH-MTHB complex belong to the  $P2_1$  space group. The structure of DsyB-SAM complex Se derivative was determined by SAD phasing. The crystal structures of DsyB-SAM complex and DsyB-SAH-MTHB complex were determined by molecular replacement using the CCP4 program Phaser (63) with the structure of the Se derivative as the search model. The refinements of these structures were performed using Coot (64) and Phenix (65). All the structure figures were produced with the PyMOL (<http://www.pymol.org/>).

### Analyses of DMSP synthesis genes in cultured microorganisms

The presence or absence of DMSP synthesis genes in 111 cultured DMSP-producing bacteria (published since the discovery of bacterial DMSP synthesis (7, 10, 11, 32, 33) was analysed (Table S2). This is based previously published work that analysed their sequenced genomes and/or used degenerate primers to detect the presence of *burB*, *dsyB* and/or *mmtN*. Percentage abundances were calculated for all three DMSP synthesis genes within these cultured organisms, as well as those containing both *dsyB* and *mmtN*.

Eukaryotic transcriptomes from the MMETSP (66) (Table S3 and S4) were analysed for the presence of *DSYB* and/or *TpMMT* through tblastn searches against ratified *DSYB* (12) and *TpMMT* (17) sequences. These were manually curated to confirm identity (E value cutoff of  $1e^{-30}$  for *DSYB*), although since *TpMMT* has only been shown to be functional in *T. pseudonana*, we did not assume any sequences below 70% identity to *T. pseudonana* *TpMMT* to be functional. Strains confirmed to contain *DSYB* and/or *TpMMT* are listed in Table S3 and summarised in Table S4, alongside literature reporting the presence of DMSP synthesis in that particular strain (if tested).

### Metagenome and metatranscriptome analyses.

Verified sequences (7, 10-12, 67) were aligned using ClustalOmega (68), and profile hidden Markov models (hmms) of *dysB*, *DSYB*, and *mmtN* were constructed using the hmmbuild function of hmmer 3.3 (69). Tara metagenome (OM-RGC\_v2\_metaG/MATOU\_v1\_metaG) (prokaryotic/eukaryotic, respectively), and metatranscriptome (OM-RGC\_v2\_metaT/ MATOU\_v1\_metaT) sequences together with their abundances and taxonomic assignments were downloaded from the Ocean Gene Atlas site (70) using an hmmsearch e-value threshold of  $1e^{-70}$  (*dysB*),  $1e^{-80}$  (*DSYB*) or  $1e^{-60}$  (*mmtN*). A blastp search (e-value threshold of  $1e^{-80}$ ) was used for *TpMMT*, using the *Thalassiosira pseudonana* *TpMMT* sequence as query. Environmental *dysB*/*DSYB* sequences were aligned with *Nisaea denitrificans dysB* using ClustalOmega (68), and sequences that did not possess all six essential residues were excluded from further analysis. Environmental *TpMMT* sequences greater than 400 amino acids in length were also excluded from further analysis. Prokaryotic sequence abundances were normalized using the median abundance of 10 single copy marker genes/transcripts (71). This gave abundance as a percentage of single copy gene abundance (equivalent to the percentage of cells containing a copy) in the metagenome, and transcription as a percentage of single copy gene transcription in the metatranscriptome. These marker gene/transcript abundances were downloaded from the Ocean Gene Atlas using the hmm profiles developed by Milanese *et al.* (71) with an e-value threshold of  $1e^{-80}$ . The MATOU\_v1\_metaG (metagenomic) database featured few MIX and MES sampling sites (2 and 7, respectively), limiting the power of comparative analysis between sampling depths, thus, these sites were excluded from analysis. Statistical analysis was performed in R (version 4.02) using RStudio.

### Accession numbers

The structures of DsyB-SAM complex and DsyB-SAH-MTHB complex have been deposited in the Protein Data Bank (PDB) under the accession codes 7DFD and 7DFF, respectively.

## Declarations

### Acknowledgements

We thank the staff from BL18U1&BL19U1 beamlines of the National Facility for Protein Sciences Shanghai (NFPS) and Shanghai Synchrotron Radiation Facility, for assistance during data collection. This work was supported by: the National Science Foundation of China (grants 31630012, U1706207,

91851205, 42076229, 31800107, and 41706152), the National Key Research and Development Program of China (2016YFA0601303, 2018YFC1406700, 2018YFC1406504), Major Scientific and Technological Innovation Project (MSTIP) of Shandong Province (2019JZZY010817), the Program of Shandong for Taishan Scholars (tspd20181203), AoShan Talents Cultivation Program Supported by Qingdao National Laboratory for Marine Science and Technology (2017ASTCP-OS14), the grant of Laboratory for Marine Biology and Biotechnology (OF2019NO02), Pilot National Laboratory for Marine Science and Technology (Qingdao), the Fund of Key Laboratory of Global Change and Marine-Atmospheric Chemistry, MNR, (GCMAC1908), the Fundamental Research Funds for the Central Universities, the United Kingdom's Natural and Environmental Research Council (NERC, NE/P012671/1 and NE/N002385/1), and the United Kingdom's Biotechnology and Biological Sciences Research Council (BBSRC, BB/P006140/1).

## References

1. Gali M, Devred E, Levasseur M, Royer SJ, & Babin M (2015) A remote sensing algorithm for planktonic dimethylsulfoniopropionate (DMSP) and an analysis of global patterns. *Remote Sens Environ* 171:171-184.
2. Simo R, Archer SD, Pedros-Alio C, Gilpin L, & Stelfox-Widdicombe CE (2002) Coupled dynamics of dimethylsulfoniopropionate and dimethylsulfide cycling and the microbial food web in surface waters of the North Atlantic. *Limnol Oceanogr* 47(1):53-61.
3. Zhang XH, *et al.* (2019) Biogenic production of DMSP and its degradation to DMS-their roles in the global sulfur cycle. *Sci China Life Sci* 62(10):1296-1319.
4. Stefels J (2000) Physiological aspects of the production and conversion of DMSP in marine algae and higher plants. *J Sea Res* 43(3):183-197.
5. Wolfe GV & Steinke M (1996) Grazing-activated production of dimethyl sulfide (DMS) by two clones of *Emiliania huxleyi*. *Limnol Oceanogr* 41(6):1151-1160.
6. Sunda W, Kieber DJ, Kiene RP, & Huntsman S (2002) An antioxidant function for DMSP and DMS in marine algae. *Nature* 418(6895):317-320.
7. Zheng Y, *et al.* (2020) Bacteria are important dimethylsulfoniopropionate producers in marine aphotic and high-pressure environments. *Nat Commun* 11(1):4658.
8. Vallina SM & Simó R (2007) Strong relationship between DMS and the solar radiation dose over the global surface ocean. *Science* 315(5811):506-508.
9. Curson AR, Todd JD, Sullivan MJ, & Johnston AW (2011) Catabolism of dimethylsulphonopropionate: microorganisms, enzymes and genes. *Nat Rev Microbiol* 9(12):849-859.
10. Curson AR, *et al.* (2017) Dimethylsulfoniopropionate biosynthesis in marine bacteria and identification of the key gene in this process. *Nat Microbiol* 2:17009.
11. Williams BT, *et al.* (2019) Bacteria are important dimethylsulfoniopropionate producers in coastal sediments. *Nat Microbiol* 4(11):1815-1825.

12. Curson ARJ, *et al.* (2018) DSYB catalyses the key step of dimethylsulfoniopropionate biosynthesis in many phytoplankton. *Nat Microbiol* 3(4):430-439.
13. Summers PS, *et al.* (1998) Identification and stereospecificity of the first three enzymes of 3-dimethylsulfoniopropionate biosynthesis in a chlorophyte alga. *Plant Physiol* 116(1):369-378.
14. Gage DA, *et al.* (1997) A new route for synthesis of dimethylsulphoniopropionate in marine algae. *Nature* 387(6636):891-894.
15. Rousseau H, *et al.* (2017) Evolution of DMSP (dimethylsulfoniopropionate) biosynthesis pathway: Origin and phylogenetic distribution in polyploid *Spartina* (Poaceae, Chloridoideae). *Mol Phylogenet Evol* 114:401-414.
16. Ito T, Asano Y, Tanaka Y, & Takabe T (2011) Regulation of Biosynthesis of Dimethylsulfoniopropionate and Its Uptake in Sterile Mutant of *Ulva Pertusa* (Chlorophyta)(1). *J Phycol* 47(3):517-523.
17. Kageyama H, Tanaka Y, Shibata A, Waditee-Sirisattha R, & Takabe T (2018) Dimethylsulfoniopropionate biosynthesis in a diatom *Thalassiosira pseudonana*: Identification of a gene encoding MTHB-methyltransferase. *Arch Biochem Biophys* 645:100-106.
18. Cantoni GL (1975) Biological methylation: selected aspects. *Annu Rev Biochem* 44:435-451.
19. Liscombe DK, Louie GV, & Noel JP (2012) Architectures, mechanisms and molecular evolution of natural product methyltransferases. *Nat Prod Rep* 29(10):1238-1250.
20. O'Hagan D & Schmidberger JW (2010) Enzymes that catalyse S(N)2 reaction mechanisms. *Nat Prod Rep* 27(6):900-918.
21. Urios L, Michotey V, Intertaglia L, Lesongeur F, & Lebaron P (2008) *Nisaea denitrificans* gen. nov., sp nov and *Nisaea nitritireducens* sp nov., two novel members of the class Alphaproteobacteria from the Mediterranean Sea. *Int J Syst Evol Micr* 58:2336-2341.
22. Meinnel T, Mechulam Y, & Blanquet S (1993) Methionine as Translation Start Signal - a Review of the Enzymes of the Pathway in *Escherichia-Coli*. *Biochimie* 75(12):1061-1075.
23. Pellicer Martinez MT, *et al.* (2019) Mechanisms of iron- and O<sub>2</sub>-sensing by the [4Fe-4S] cluster of the global iron regulator RirA. *eLife* 8, e47804.
24. Volbeda A, *et al.* (2017) Crystal structures of the NO sensor NsrR reveal how its iron-sulfur cluster modulates DNA binding. *Nat Commun* 8, 10152.
25. Crack JC, Thomson AJ, & Le Brun NE (2017) Mass spectrometric identification of intermediates in the O<sub>2</sub>-driven [4Fe-4S] to [2Fe-2S] cluster conversion in FNR. *P Natl Acad Sci U S A* 114(16):E3215-E3223.
26. Chen SC, *et al.* (2015) Structure and mechanism of an antibiotics-synthesizing 3-hydroxykynurenine C-methyltransferase. *Sci Rep* 5, 10100.
27. Martin JL & McMillan FM (2002) SAM (dependent) I AM: the S-adenosylmethionine-dependent methyltransferase fold. *Curr Opin Struct Biol* 12(6):783-793.

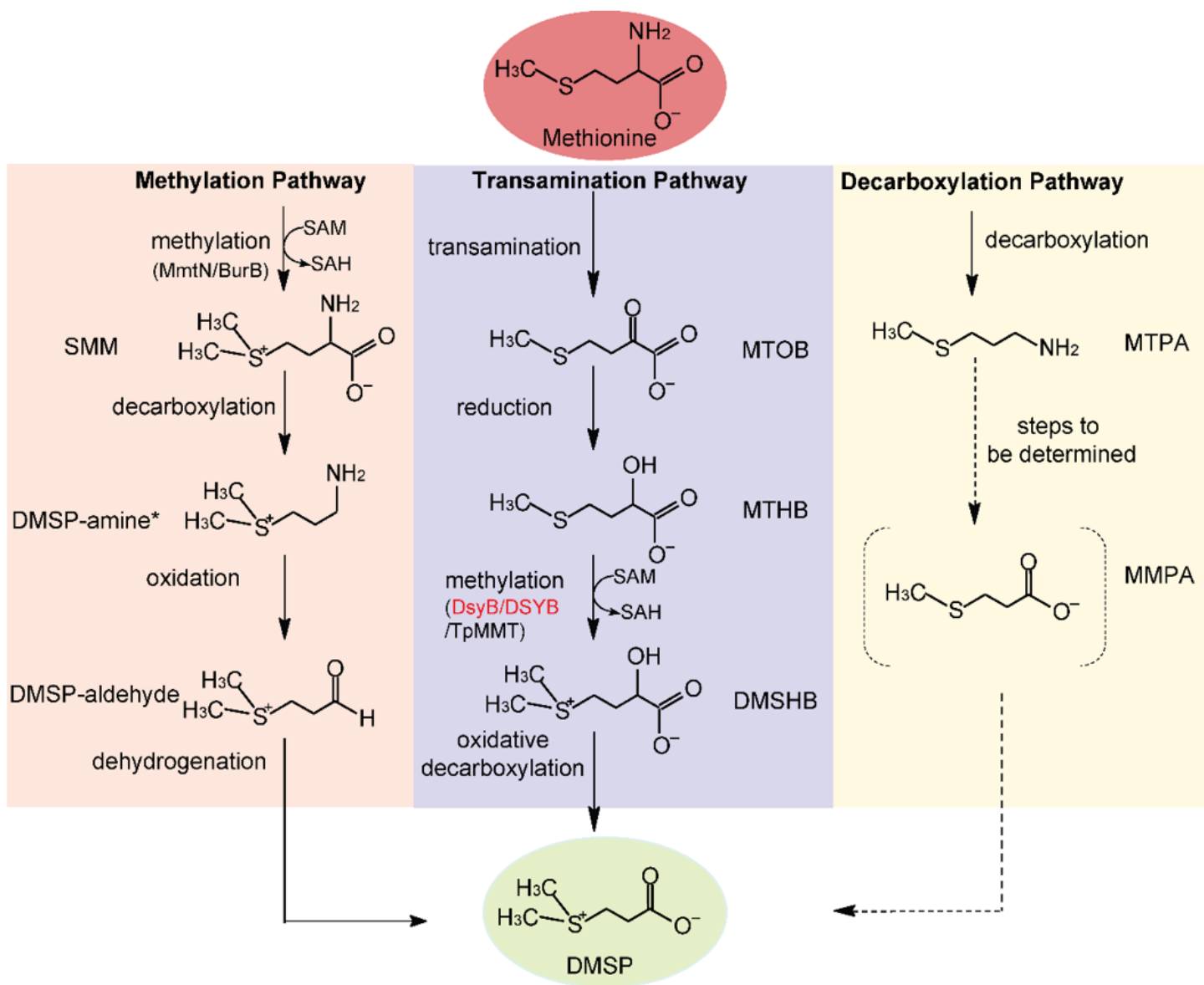


28. Peng Y, *et al.* (2008) Structural basis of substrate recognition in thiopurine S-methyltransferase. *Biochemistry* 47(23):6216-6225.
29. Coirier H, *et al.* (2006) Methylation of sulfhydryl groups: a new function for a family of small molecule plant O-methyltransferases. *Plant J* 46(2):193-205.
30. Dang HY, Li TG, Chen MN, & Huang GQ (2008) Cross-Ocean distribution of Rhodobacterales bacteria as primary surface colonizers in temperate coastal marine waters. *Appl Environ Microb* 74(1):52-60.
31. Waterhouse A, *et al.* (2018) SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res* 46(W1):W296-W303.
32. Trottmann F, *et al.* (2020) Sulfonium Acids Loaded onto an Unusual Thiotemplate Assembly Line Construct the Cyclopropanol Warhead of a *Burkholderia* Virulence Factor. *Angew Chem Int Edit* 59(32):13511-13515.
33. Liu J, *et al.* (2021) Bacterial Dimethylsulfoniopropionate Biosynthesis in the East China Sea. *Microorganisms* 9(3), 657.
34. Tully BJ, Graham ED, & Heidelberg JF (2018) The reconstruction of 2,631 draft metagenome-assembled genomes from the global oceans. *Sci Data* 5:170203.
35. Keeling PJ, *et al.* (2014) The Marine Microbial Eukaryote Transcriptome Sequencing Project (MMETSP): illuminating the functional diversity of eukaryotic life in the oceans through transcriptome sequencing. *PLoS Biol* 12(6):e1001889.
36. Bucciarelli E, *et al.* (2013) Increased intracellular concentrations of DMSP and DMSO in iron-limited oceanic phytoplankton *Thalassiosira oceanica* and *Trichodesmium erythraeum*. *Limnol Oceanogr* 58(5):1667-1679.
37. Keller MD, Bellows WK, & Guillard RRL (1989) Dimethyl Sulfide Production in Marine Phytoplankton. *Biogenic Sulfur in the Environment*, ACS Symposium Series, (American Chemical Society), Vol 393, pp 167-182.
38. Keller MD & Korjef-Bellows W (1996) Physiological Aspects of the Production of Dimethylsulfoniopropionate (DMSP) by Marine Phytoplankton. *Biological and Environmental Chemistry of DMSP and Related Sulfonium Compounds*, eds Kiene RP, Visscher PT, Keller MD, & Kirst GO (Springer US, Boston, MA), pp 131-142.
39. Meng A, *et al.* (2018) Analysis of the genomic basis of functional diversity in dinoflagellates using a transcriptome-based sequence similarity network. *Mol Ecol* 27(10):2365-2380.
40. Caruana AMN, *et al.* (2020) *Alexandrium pacificum* and *Alexandrium minutum*: Harmful or environmentally friendly? *Mar Environ Res* 160:105014.
41. Wolfe GV, Strom SL, Holmes JL, Radzio T, & Olson MB (2002) Dimethylsulfoniopropionate cleavage by marine phytoplankton in response to mechanical, chemical, or dark stress. *J Phycol* 38(5):948-960.
42. Royer C, Borges AV, Lapeyra Martin J, & Gypens N (2021) Drivers of the variability of dimethylsulfoniopropionate (DMSP) and dimethylsulfoxide (DMSO) in the Southern North Sea. *Cont Shelf Res* 216:104360.

43. Caruana AMN, Steinke M, Turner SM, & Malin G (2012) Concentrations of dimethylsulphonioacetate and activities of dimethylsulphide-producing enzymes in batch cultures of nine dinoflagellate species. *Biogeochemistry* 110(1-3):87-107.
44. Caruana AMN & Malin G (2014) The variability in DMSP content and DMSP lyase activity in marine dinoflagellates. *Prog Oceanogr* 120:410-424.
45. Gutierrez-Rodriguez A, *et al.* (2017) Dimethylated sulfur compounds in symbiotic protists: A potentially significant source for marine DMS(P). *Limnol Oceanogr* 62(3):1139-1154.
46. Yost DM & Mitchelmore CL (2009) Dimethylsulfonylacetate (DMSP) lyase activity in different strains of the symbiotic alga *Symbiodinium microadriaticum*. *Mar Ecol Prog Ser* 386:61-70.
47. Franklin DJ, Steinke M, Young J, Probert I, & Malin G (2010) Dimethylsulphonioacetate (DMSP), DMSP-lyase activity (DLA) and dimethylsulphide (DMS) in 10 species of coccolithophore. *Mar Ecol Prog Ser* 410:13-23.
48. Steinke M, Wolfe GV, & Kirst GO (1998) Partial characterisation of dimethylsulfonylacetate (DMSP) lyase isozymes in 6 strains of *Emiliania huxleyi*. *Mar Ecol Prog Ser* 175:215-225.
49. Spielmeyer A, Gebser B, & Pohnert G (2011) Dimethylsulphide sources from microalgae: Improvement and application of a derivatization-based method for the determination of dimethylsulfonylacetate and other zwitterionic osmolytes in phytoplankton. *Mar Chem* 124(1-4):48-56.
50. Kinsey JD, Kieber DJ, & Neale PJ (2016) Effects of iron limitation and UV radiation on *Phaeocystis antarctica* growth and dimethylsulfonylacetate, dimethylsulfoxide and acrylate concentrations. *Environ Chem* 13(2):195-211.
51. Jean N, Bogé G, Jamet J-L, Richard S, & Jamet D (2005) Annual contribution of different plankton size classes to particulate dimethylsulfonylacetate in a marine perturbed ecosystem. *J Marine Syst* 53(1):235-247.
52. Beringer JE (1974) R Factor Transfer in *Rhizobium-Leguminosarum*. *J Gen Microbiol* 84(Sep):188-198.
53. Gonzalez JM, Whitman WB, Hodson RE, & Moran MA (1996) Identifying numerically abundant culturable bacteria from complex communities: an example from a lignin enrichment culture. *Appl Environ Microbiol* 62(12):4433-4440.
54. Baumann P, Baumann, L. (1981) *The Prokaryotes: A Handbook on Habitats, Isolation and Identification of Bacteria* (Springer-Verlag Berlin Heidelberg) 1st edn Ed.
55. Figurski DH & Helinski DR (1979) Replication of an origin-containing derivative of plasmid RK2 dependent on a plasmid function provided in trans. *P Natl Acad Sci U S A* 76(4):1648-1652.
56. Carrion O, *et al.* (2015) A novel pathway producing dimethylsulphide in bacteria is widespread in soil environments. *Nat Commun* 6, 6579.
57. Tett AJ, Rudder SJ, Bourdes A, Karunakaran R, & Poole PS (2012) Regulatable vectors for environmental gene expression in Alphaproteobacteria. *Appl Environ Microbiol* 78(19):7137-7140.

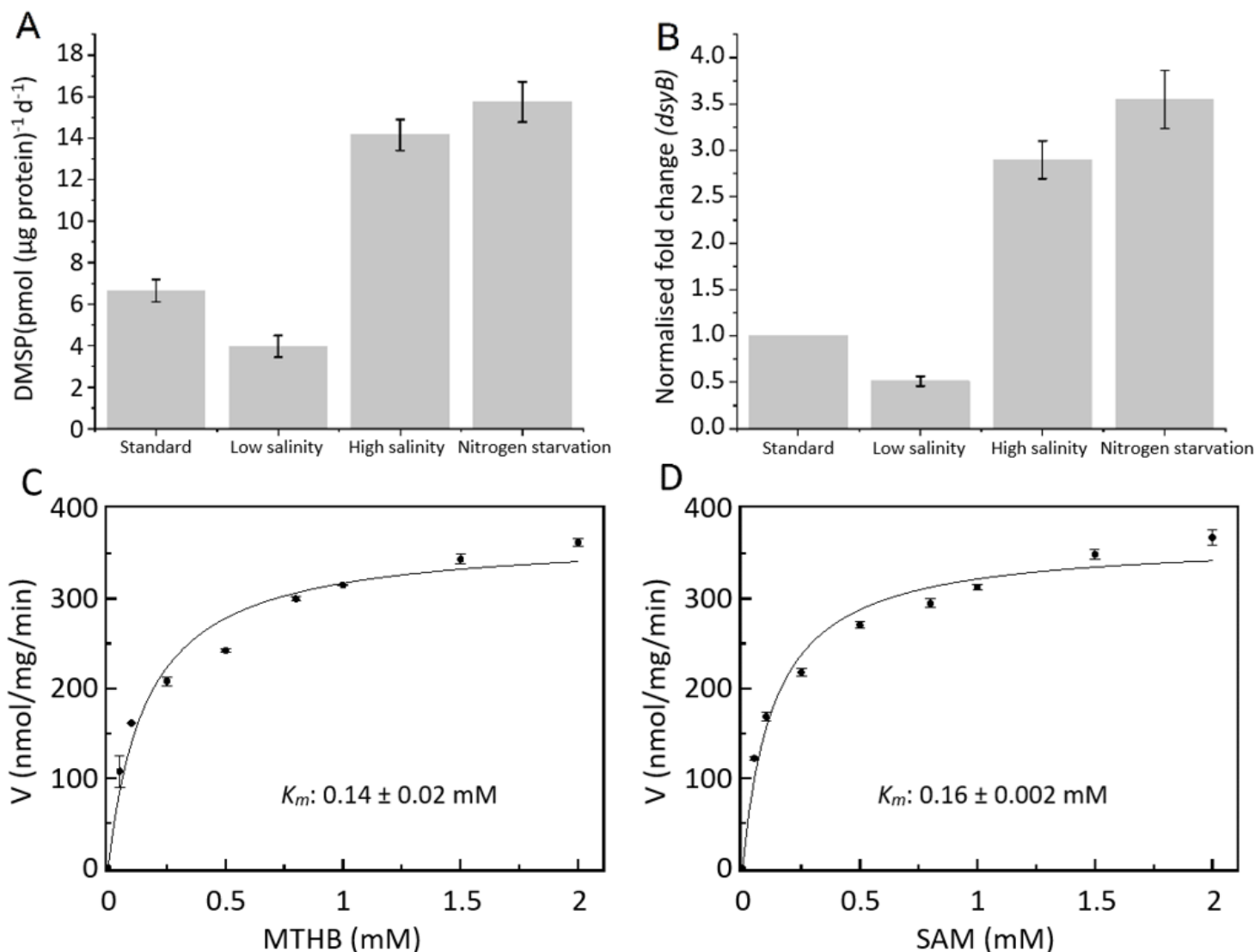
58. Lalitha S (2000) Primer Premier 5. *Biotech Software & Internet Report* 1(6):270-272.
59. Pfaffl MW (2001) A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res* 29(9).
60. Tang DQ, Zou L, Yin XX, & Ong CN (2016) HILIC-MS for metabolomics: An attractive and complementary approach to RPLC-MS. *Mass Spectrom Rev* 35(5):574-600.
61. Wilkins MR, *et al.* (1999) Protein identification and analysis tools in the ExPASy server. *Method Mol Biol* 112:531-552.
62. Minor W, Cymborowski M, Otwinowski Z, & Chruszcz M (2006) HKL-3000: the integration of data reduction and structure solution—from diffraction images to an initial model in minutes. *Acta Crystallogr D Biol Crystallogr* 62(Pt 8):859-866.
63. Winn MD, *et al.* (2011) Overview of the CCP4 suite and current developments. *Acta Crystallogr D Biol Crystallogr* 67:235-242.
64. Emsley P, Lohkamp B, Scott WG, & Cowtan K (2010) Features and development of Coot. *Acta Crystallogr D Biol Crystallogr* 66:486-501.
65. Adams PD, *et al.* (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* 66:213-221.
66. Keeling PJ, *et al.* (2014) The Marine Microbial Eukaryote Transcriptome Sequencing Project (MMETSP): Illuminating the Functional Diversity of Eukaryotic Life in the Oceans through Transcriptome Sequencing. *PLoS Biol* 12(6).
67. Liao CS & Seebeck FP (2019) In Vitro Reconstitution of Bacterial DMSP Biosynthesis. *Angew Chem Int Edit* 58(11):3553-3556.
68. Sievers F, *et al.* (2011) Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol Syst Biol* 7:539-539.
69. Finn RD, Clements J, & Eddy SR (2011) HMMER web server: interactive sequence similarity searching. *Nucleic Acids Res* 39(Web Server issue):W29-37.
70. Villar E, *et al.* (2018) The Ocean Gene Atlas: exploring the biogeography of plankton genes online. *Nucleic Acids Res* 46(W1):W289-W295.
71. Milanese A, *et al.* (2019) Microbial abundance, activity and population genomic profiling with mOTUs2. *Nat Commun* 10(1):1014.

## Figures



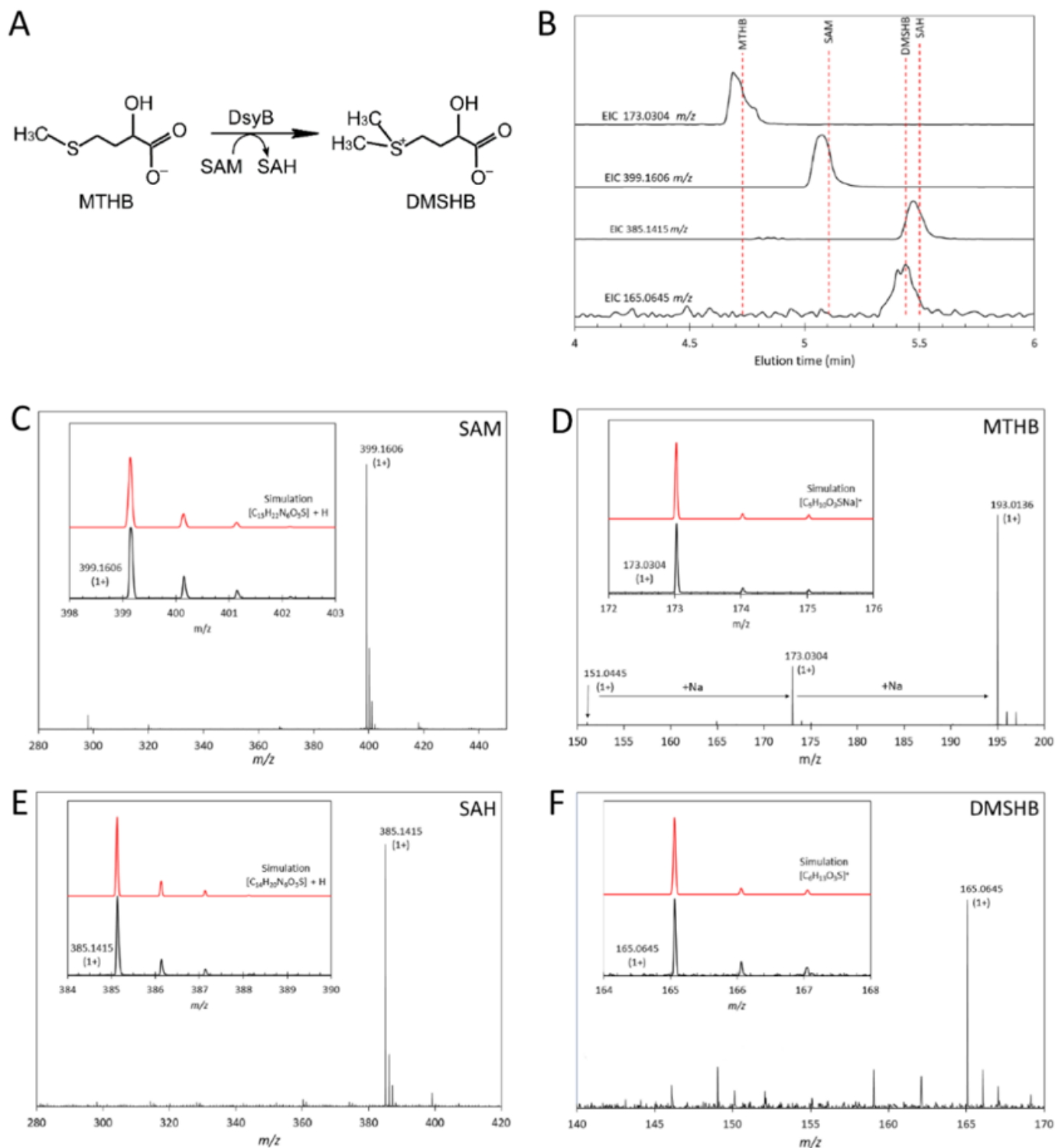
**Figure 1**

Predicted DMSP biosynthesis pathways (10). Different pathways are shown in different colours. Enzymes of interest in this study (DsyB/DSYB) are in red. \*SMM is converted to DMSP-aldehyde directly in Wollastonia. Dotted lines represent unconfirmed steps of the decarboxylation pathway. Abbreviations: SMM, S-methylmethionine; MTOB, 4-methylthio-2-oxobutyrate; MTHB, 4-methylthio-2-hydroxybutyrate; DMSHB, 4-dimethylsulfonio-2-hydroxybutyrate; MTPA, 3-methylthiopropylamine; MMPA, methylmercaptopropionate.



**Figure 2**

Characterisation of *N. dentrificans* DMSP production, *dsyB* transcription and the DsyB enzyme. *N. dentrificans* DMSP production (A) and *dsyB* transcription (B) observed under different conditions. Standard conditions were MBM medium at 35 PSU with 10 mM NH<sub>4</sub>Cl compared to low salinity (5 PSU), high salinity (50 PSU) and nitrogen starvation conditions (where cells in standard MBM were resuspended in to standard MBM with no added NH<sub>4</sub>Cl). (C) A non-linear fit curve for MTHB methylation by DsyB. Initial rates were determined with 1.97 μM DsyB (50 mM Tris-HCl, 100 mM NaCl, pH 8.0) and 0 – 2 mM MTHB.  $K_m$  was  $0.14 \pm 0.02$  mM. (D) A non-linear fit curve for SAM demethylation by DsyB. Initial rates were determined with 1.97 μM DsyB and 0 – 2 mM SAM in the same reaction buffer.  $K_m$  was  $0.16 \pm 0.002$  mM. The error bar represents standard deviation of triplicate experiments.

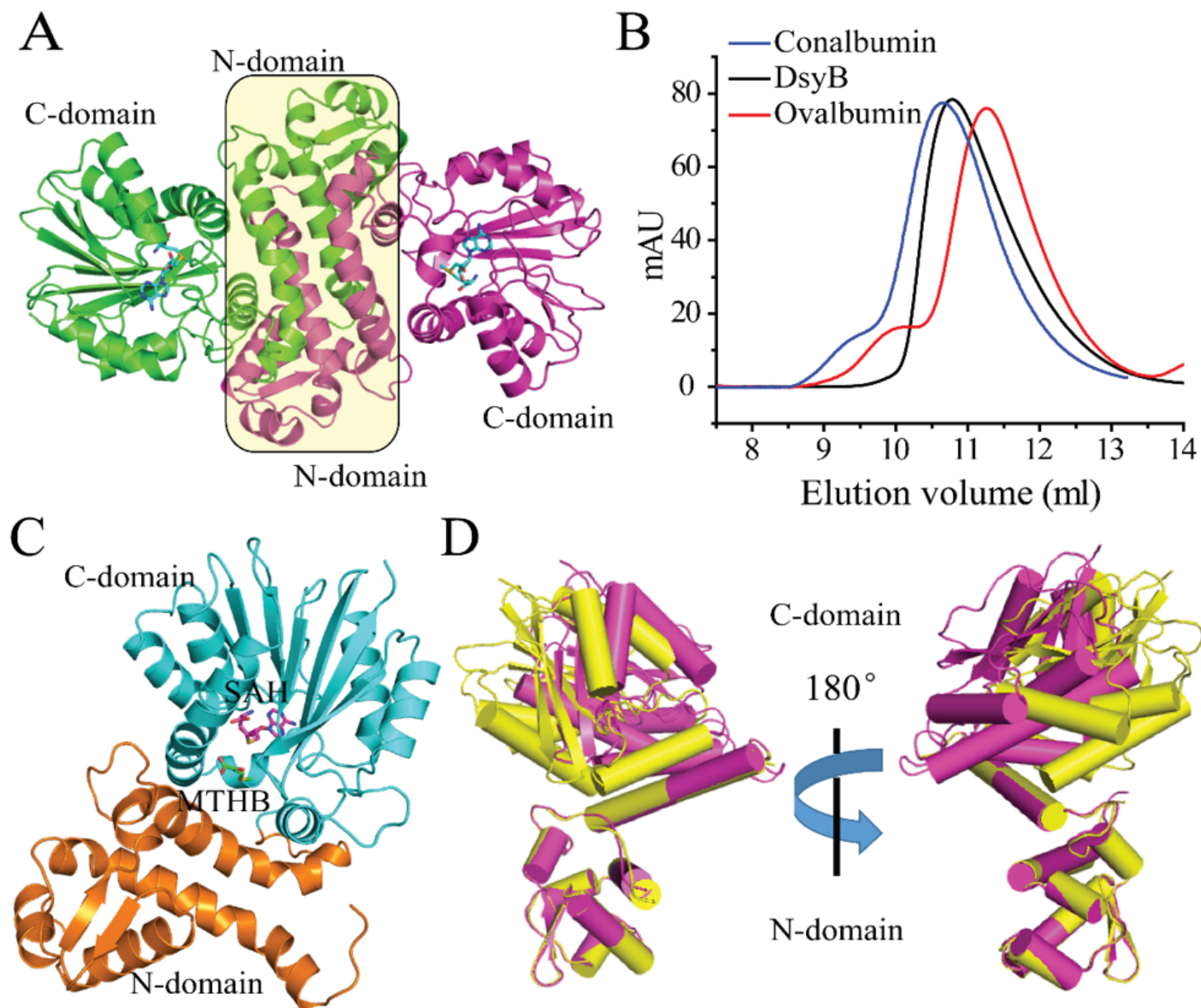


**Figure 3**

Identification of DsyB substrates and products by LC-MS. (A) The chemical equation for MTHB S-methylation into DMSHB. (B) Extracted ion chromatograms for MTHB, SAM, DMSHB and SAH. Here, mass spectrometry data were analysed to extract ion counts as a function of elution volume for the  $m/z$  ions indicated, which correspond to the substrates and products of the DsyB-catalysed reaction. The red broken lines indicate elution volumes of the molecules when run as standards. (C) – (F) mass spectra



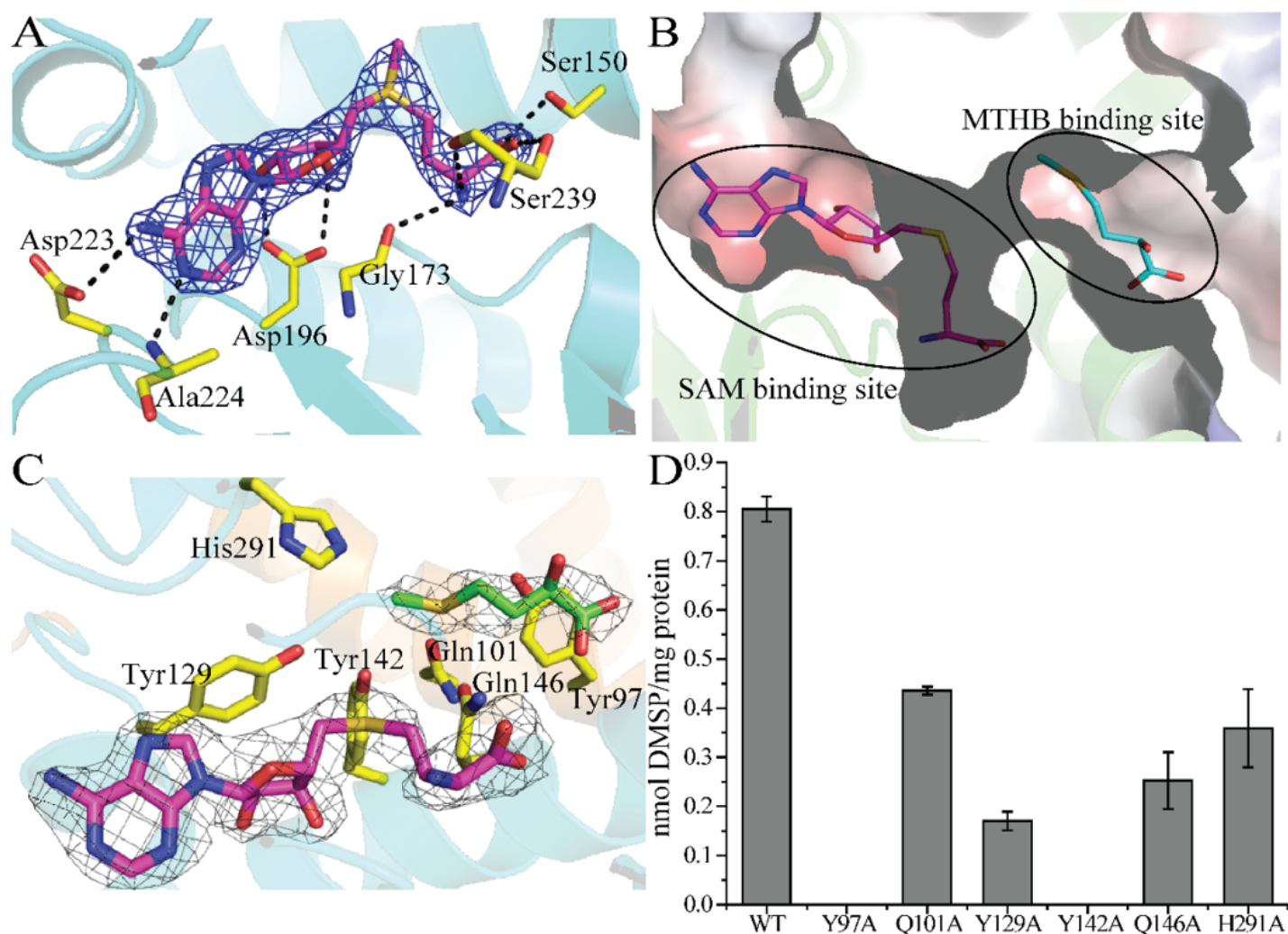
recorded for the eluted species, as indicated. Inset are spectra over a narrower  $m/z$  range (black lines) along with simulated spectra (showing the isotope distribution, red line) for each molecule, providing clear confirmation of identity.



**Figure 4**

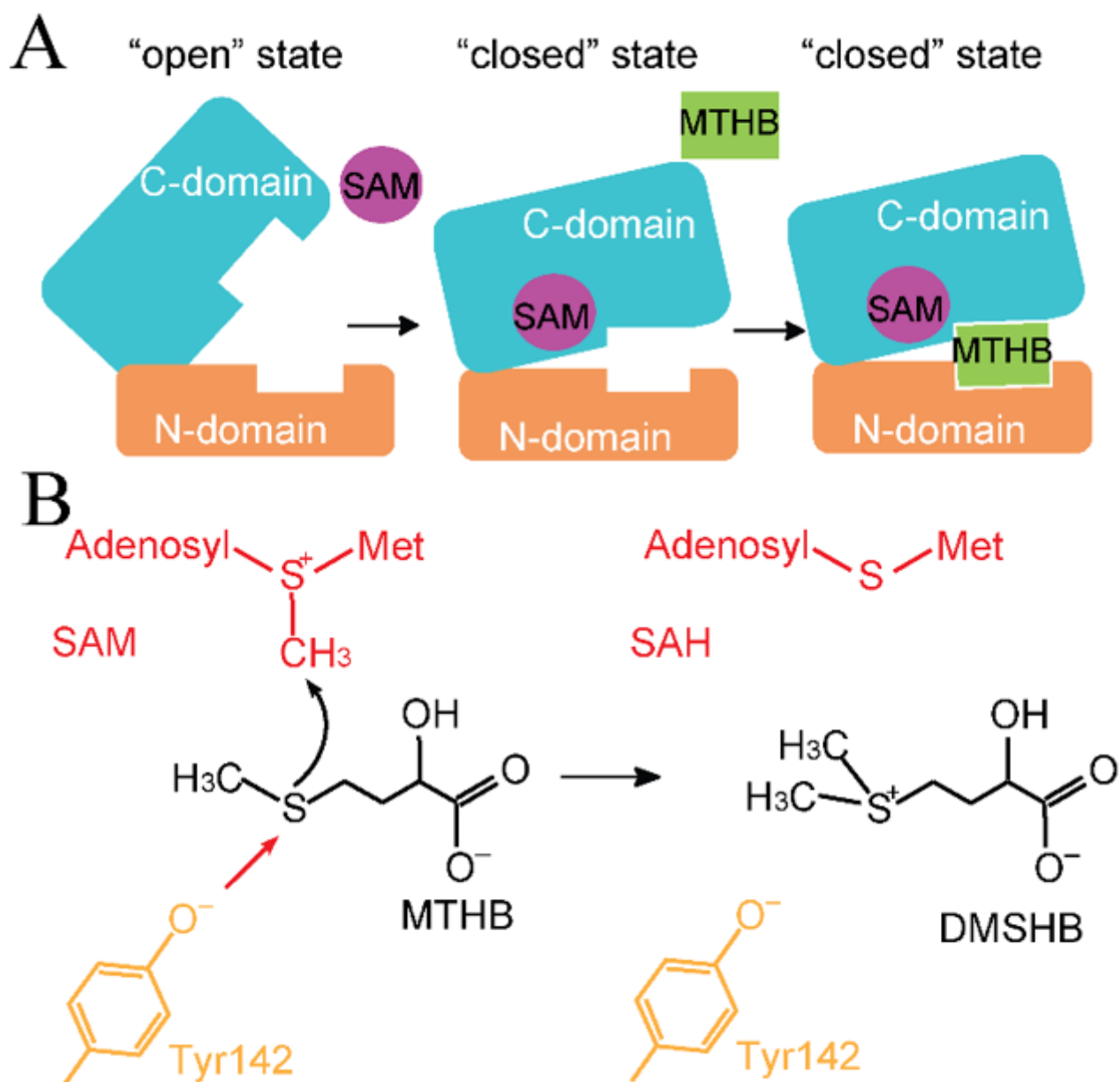
Overall structure analysis of DsyB. (A) Ribbon representation of DsyB dimer. Each DsyB monomer contains an N-domain and a C-domain. SAM molecules are shown as sticks coloured in cyan. (B) Analysis of the form of DsyB in solution by gel filtration. Conalbumin (molecular mass = 75,000 Da; GE Healthcare) and ovalbumin (molecular mass = 43,000 Da; GE Healthcare) were used as markers. The predicted molecular mass of DsyB monomer is 37,215 Da. (C) The overall structure of DsyB-SAH-MTHB complex. The SAH molecule (coloured in purple) and the MTHB molecule (coloured in green) are

represented as sticks. (D) Superimposed structures of DsyB with (coloured in purple) and without (coloured in yellow) binding the SAM molecule.



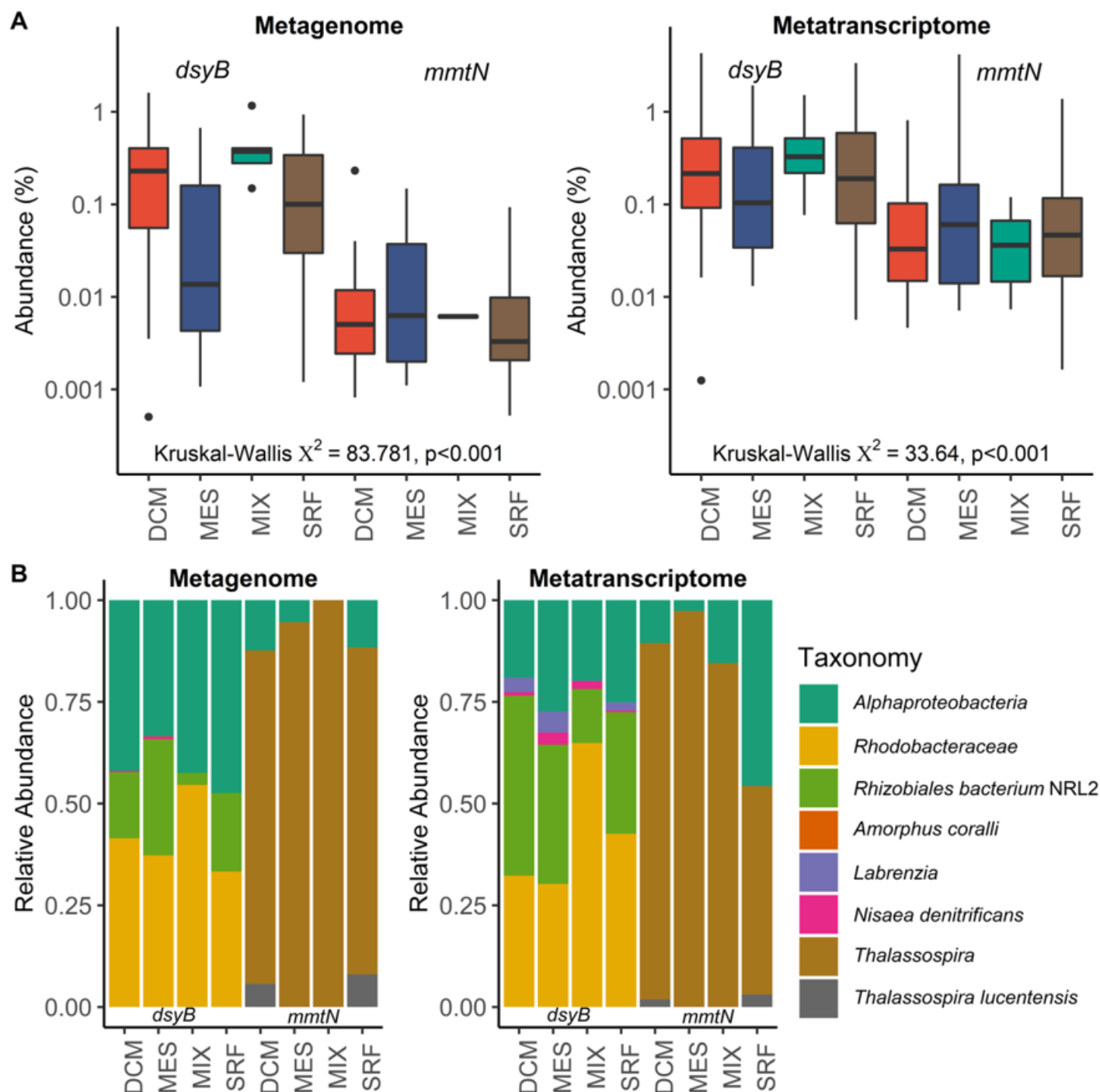
**Figure 5**

Analyses of residues of DsyB involved in binding SAM and MTHB. (A) Interactions between DsyB residues and SAM. SAM is coloured in purple. The possible hydrogen bonds are represented by dashed lines. The 2Fo-Fc density for SAM is contoured in blue at 1.5 $\sigma$ . (B) Electrostatic surface of the crystal structure of DsyB. The SAM binding site and the MTHB binding site can be clearly identified. (C) The binding site of MTHB. Residues of DsyB that may participate in binding MTHB are shown in yellow. The 2Fo-Fc density for SAH (coloured in purple) is contoured in grey at 1.5 $\sigma$ . The 2Fo-Fc density for MTHB (coloured in green) is contoured in grey at 1.0 $\sigma$ . (D) Enzymatic activities of WT DsyB and site-directed mutants. The error bar represents standard deviation of triplicate experiments.



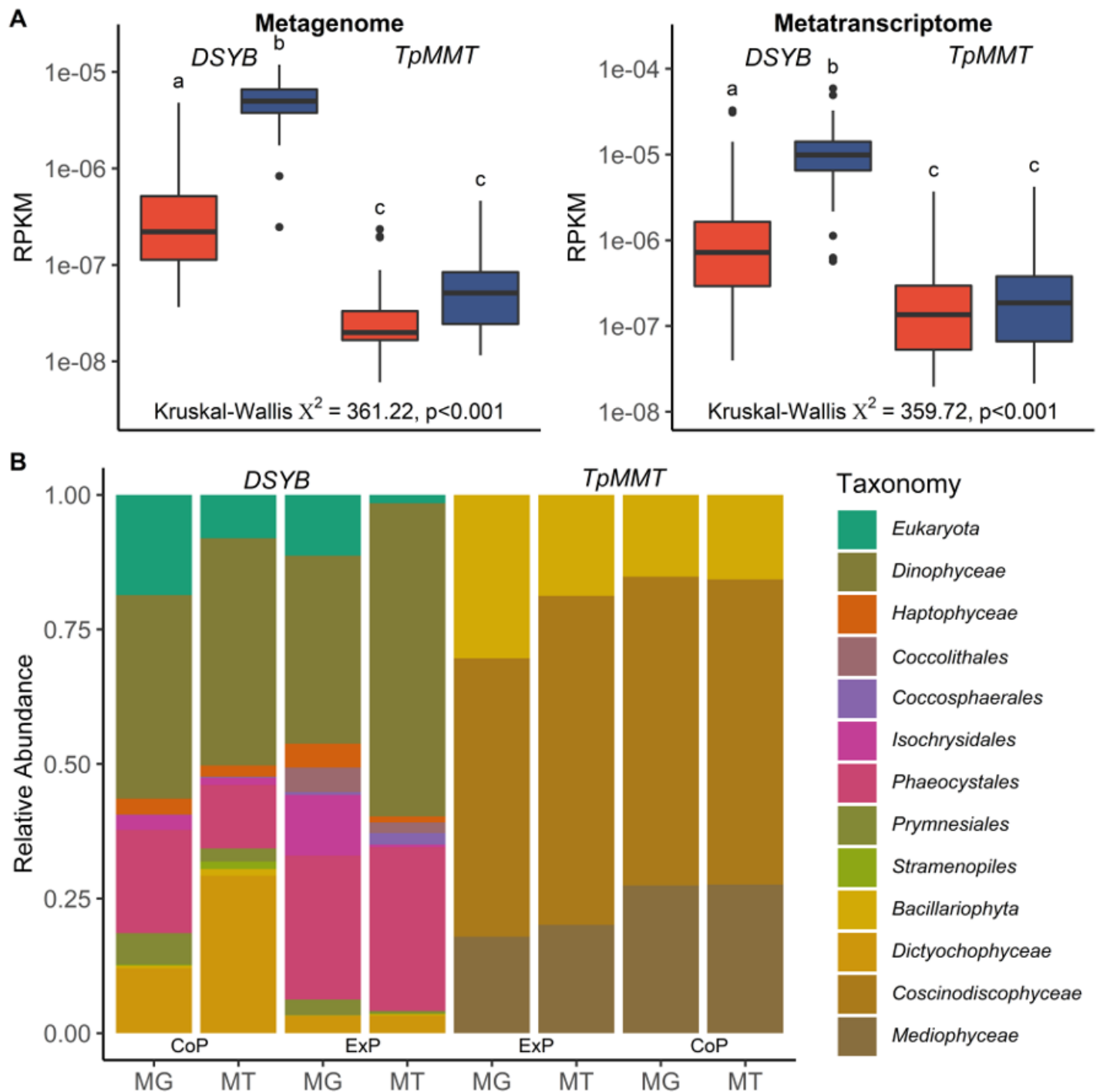
**Figure 6**

A proposed catalytic mechanism of DsyB. (A) The schematic diagrams of the DsyB conformational change triggered by the binding of SAM. (B) The catalytic residue Tyr142 activates the sulfur atom of MTHB for its nucleophilic attack on the methyl group of SAM. The residue Tyr142 is shown in orange. The MTHB molecule and DMSHB molecule are shown in black. And the SAM molecule and the SAH molecule are shown in red.



**Figure 7**

Analysis of MATOU for *dsyB* and *mmtN*. (A) Normalized abundance of *dsyB* and *mmtN* in Tara metagenomes and metatranscriptomes, by sampling depth. Abundances are normalized to the median gene or transcript abundance of 10 single copy marker genes. Kruskal-Wallis  $X^2$  values for comparisons between *dsyB* and *mmtN* (across all depths) are shown. (B) Taxonomic assignment and relative abundance of *dsyB* and *mmtN* sequences in the Tara metagenomes and metatranscriptomes.



**Figure 8**

Analysis of MATOU for DSYB and TpMMT. (A) Normalized abundance of DSYB and TpMMT in MATOU metagenome and metatranscriptomes, by fractions containing picoeukaryotes (blue) and fractions excluding eukaryotes (red). DCM and SRF depths are combined for the purposes of this analysis. Abundances are normalized as reads per kilobase per million mapped reads (RPKM). Kruskal-Wallis  $\chi^2$  values for comparisons between DSYB/TpMMT and fraction abundance are shown. Letters denote gene or transcript/fraction combinations that are significantly different (min.  $p < 0.05$ ) by post-hoc Dunn's test,

using Holm's correction. (B) Taxonomic assignment and relative abundance of DSYB and TpMMT sequences in the MATOU metagenome and metatranscriptome. Taxonomy is reported as Phylum (if available) for DSYB and as Class (if available) for TpMMT

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryMaterials.docx](#)
- [SupTable2.pdf](#)
- [SupTable3.pdf](#)
- [SupTable4.pdf](#)