DSYB catalyses the key step of dimethylsulfoniopropionate biosynthesis in many phytoplankton

Andrew R. J. Curson¹, Beth T. Williams¹, Benjamin J. Pinchbeck¹, Leanne P. Sims¹, Ana Bermejo Martinez¹, Peter Paolo L. Rivera¹, Deepak Kumaresan², Elena Mercadé³, Lewis G. Spurgin¹, Ornella Carrión¹, Simon Moxon¹, Rose Ann Cattolico⁴, Unnikrishnan Kuzhiumpambili⁵, Paul Guagliardo⁶, Peta L. Clode⁶,⁷, Jean-Baptiste Raina⁵, Jonathan D. Todd¹*

¹School of Biological Sciences, University of East Anglia, Norwich Research Park, Norwich, NR4 7TJ, UK. ²School of Biological Sciences and Institute for Global Food Security, Queen’s University Belfast, Belfast BT9 7BL, UK. ³Laboratori de Microbiologia, Facultat de Farmàcia, Universitat de Barcelona, Av. Joan XXIII s/n, 08028 Barcelona, Spain. ⁴Department of Biology, University of Washington, Seattle, Washington, USA. ⁵Climate Change Cluster (C3), Faculty of Science, University of Technology, Sydney, NSW 2007 Australia. ⁶The Centre for Microscopy Characterisation and Analysis, University of Western Australia, Crawley, Australia. ⁷Oceans Institute, University of Western Australia, Crawley, Australia.

*corresponding author

Dimethylsulfoniopropionate (DMSP) is a globally important organosulfur molecule, and the major precursor for dimethyl sulfide (DMS). These compounds are important infochemicals, key nutrients for marine microorganisms, and are involved in global sulfur cycling, atmospheric chemistry and cloud formation¹-³. DMSP production was thought to be confined to eukaryotes, but heterotrophic bacteria can also produce DMSP, via the pathway used by most phytoplankton⁴, and the DsyB enzyme catalysing the key step
of this pathway in bacteria was recently identified. However, eukaryotic phytoplankton likely produce most of Earth’s DMSP, yet no DMSP biosynthesis genes have been identified in any such organisms. Here we identify functional \( dsyB \) homologues, termed \( DSYB \), in many phytoplankton and corals. \( DSYB \) is a methylthiohydroxybutyrate (MTHB) methyltransferase enzyme localised in the chloroplasts and mitochondria of the haptophyte \( Prymnesium parvum \), and stable isotope tracking experiments support these organelles as sites of DMSP synthesis. \( DSYB \) transcription levels increased with DMSP concentrations in different phytoplankton and were indicative of intracellular DMSP. The identification of the eukaryotic \( DSYB \) sequences, along with bacterial \( dsyB \), provide the first molecular tools to predict the relative contributions of eukaryotes and prokaryotes to global DMSP production. Furthermore, evolutionary analysis suggests that eukaryotic \( DSYB \) originated in bacteria and was passed to eukaryotes early in their evolution.

Not all phytoplankton produce DMSP, and in those that do, intracellular DMSP concentrations vary considerably across groups and within genera. Previous studies identified candidate genes involved in DMSP synthesis via the transamination pathway (Fig. 1a), which is common to DMSP-producing bacteria and algae. A proteomic study of the diatom \( Fragilariopsis cylindrus \) identified putative DMSP synthesis enzymes, including the MTHB methyltransferase reaction catalysed by DsyB in bacteria. Another study on corals identified homologues of two of the \( F. cylindrus \) enzymes in \( Acropora millepora \), one being a candidate MTHB methyltransferase. None of these enzymes have been functionally ratified, and the putative MTHB methyltransferases share no significant sequence similarity to DsyB. When we cloned and expressed the \( F. cylindrus \) and \( A. millepora \) putative MTHB methyltransferase genes they had no such enzyme activity (Supplementary Table 1), suggesting that the identity of an algal MTHB methyltransferase was still unknown.
We identified homologues to the bacterial MTHB methyltransferase gene $dsyB^5$ in available genomes and/or transcriptomes of all marine prymnesiophytes; most dinoflagellates, some corals, and ~20% of diatoms and Ochrophyta (Fig. 1b, Supplementary Table 2, Supplementary Table 3 and Supplementary Data 1). The only dinoflagellate transcriptomes lacking $dsyB$ were from *Oxyrrhis marina*, a heterotroph which produces no detectable DMSP$^6,9$. Furthermore, many dinoflagellates, and some haptophytes, diatoms and corals, have multiple $dsyB$ homologs. The grouping of these multiple homologues across the phylogeny was consistent with multiple gene duplication and gene loss events over the evolutionary history of eukaryotes$^{10}$ (Fig. 1b, Supplementary Table 2, and Supplementary Table 3). These $dsyB$-like genes, termed $DSYB$, from representatives of the corals (*Acropora cervicornis*), diatoms (*F. cylindrus*), dinoflagellates (*Alexandrium tamarense, Lingulodinium polyedrum, Symbiodinium microadriaticum*) and prymnesiophytes (*Chrysochromulina tobin, Prymnesium parvum*) were cloned and shown to have MTHB methyltransferase activity, at similar levels to bacterial DsyB from *Labrenzia* (Supplementary Table 1). These algal DSYB enzymes fully complement bacterial $dsyB$ mutants, defective in DMSP production. Furthermore, enzyme assays with purified DSYB and MTHB substrate alone showed no activity, but *in vitro* $S$-adenosyl methionine (SAM)-dependent MTHB methyltransferase activity was observed when the same assays were incubated with heat-denatured *P. parvum* cell lysates (Supplementary Table 4). This suggests that a co-factor(s) present in *P. parvum* lysates might be required for activity. The $K_M$ values of DSYB for MTHB and SAM were 88.2 $\mu$M and 60.1 $\mu$M respectively (Supplementary Table 4, Supplementary Fig. 1). DSYB showed no detectable methyltransferase activity with other potential substrates (including methionine (Met), 4-methylthio-2-oxobutyrate (MTOB) and methylmercaptopropionate (MMPA); Supplementary Table 4). Thus, $DSYB$ encodes the first DMSP synthesis enzyme to be identified and functionally ratified from any eukaryotic algae.
DSYB is found across many, but by no means all, major groups of eukaryotes, and eukaryotes are monophyletic in the DsyB/DSYB phylogeny, suggesting either i) that DSYB was present in the last eukaryotic common ancestor (LECA) and has been lost across many eukaryotic groups, or ii) that dsyB has been transferred to eukaryotes multiple times. Homology and phylogenetic analyses place Alphaproteobacteria as the sister clade to the eukaryotes for this gene (Fig. 1b); we note that Alphaproteobacterial genes make up a significant proportion of eukaryotic genomes, due to endosymbiotic events with the ancestor of mitochondria. We suggest that DMSP production originated in prokaryotes, and was transferred to the eukaryotes, either via endosymbiosis at the time of mitochondrial origin, or more recently via horizontal gene transfer (HGT). Interestingly, coral DSYB paralogs grouped with dinoflagellate sequences from coral symbionts of the genus Symbiodinium (Fig. 1b). This is consistent with HGT between corals and their symbionts, as documented for other genes, and suggests that DMSP production in corals may be a result of recent HGT of DSYB from dinoflagellates. However, we cannot discount the possibility that coral DSYB sequences might be contaminant sequences unintentionally extracted from their symbionts.

No DSYB homologs were identified in available transcriptomes from marine ascomycota, cercozoa, chlorophyta, ciliophoran, cryptophyta, euglenozoa, glaucophyta, labyrinthista, perkinsozoa, or rhodophyta (Supplementary Table 3), although some members of these taxa, such as chlorophyta and rhodophyta, are known to produce DMSP. DSYB homologs were also absent in the genomes of the DMSP-producing diatoms Phaeodactylum tricornutum and Thalassiosira pseudonana. Some marine eukaryotes lack DSYB simply because they do not produce DMSP. Others may (i) have DSYB but not express it under the tested transcriptome conditions, (ii) contain a MTHB methyltransferase isoform, or (iii) produce DMSP via a different synthesis pathway.
Intracellular DMSP concentrations are generally high in dinoflagellates (reported up to 3.4 M, but unlikely to be this high given seawater osmolarity is ~1 Osm l⁻¹) and haptophytes (up to 413 mM), but significant intra-group variance exists, with some representatives not producing DMSP at detectable levels⁶,⁹. Since eukaryotic DSYB enzymes had MTHB methyltransferase rates similar to bacterial DsyB enzymes (Supplementary Table 1), it is unlikely that variation in DsyB and DSYB amino acid sequences is responsible for the differing intracellular DMSP concentrations in these organisms (Supplementary Table 1). To understand this variance, we studied model DMSP-producing phytoplankton, starting with *Chrysochromulina tobin* CCMP291 and *Chrysochromulina* sp. PCC307, two haptophytes adapted to different salinity levels (fresh-brackish and marine waters¹⁶,¹⁷, respectively). Both *Chrysochromulina* strains produced very low intracellular DMSP concentrations (Supplementary Table 1, Supplementary Fig. 2), which were unaffected by variation in salinity and nitrogen availability, conditions that have been shown previously to affect DMSP production in bacteria⁵ and phytoplankton⁷,¹⁸. Consistent with these findings, *C. tobin* CCMP291 DSYB was transcribed at very low levels (Supplementary Fig. 2), perhaps indicating a DMSP function in these haptophytes that only requires low concentrations. Many haptophytes produce high DMSP concentrations, consistent with an osmoregulatory function, but this contrasts the low *C. tobin* DMSP concentrations and highlights the variability in the process and requirement for a methodology to predict which phytoplankton are high and low DMSP producers. Perhaps other compatible solutes, possibly sugars or amino acids, are the major osmolytes in CCMP291 and PCC307. Consistent with this, the osmolyte glycine betaine (551 ± 6 nmol) was present in ~10-fold higher amounts than DMSP (52 ± 6 nmol) in CCMP291.

Next, we investigated DMSP production in six *Prymnesium* strains, from brackish/marine sources, and found they had similar intracellular DMSP concentrations, which were much
higher than those for *C. tobin* (Supplementary Fig 2). *P. parvum* CCAP946/6 *DSYB* transcription was also higher than that for *C. tobin* *DSYB* under standard conditions (Supplementary Fig 2). Interestingly, *DSYB* transcription, *DSYB* protein levels and DMSP concentration in *P. parvum* were all enhanced by increased salinity but unaffected by other environmental conditions, including nitrogen availability or temperature (Supplementary Fig. 2; Supplementary Fig. 3). Increased salinity enhances DMSP production in many phytoplankton, notably *P. parvum*, where DMSP is thought to be a significant osmolyte\(^1\). Our findings, and those of Dickson and Kirst\(^1\), are consistent with DMSP playing an osmoregulatory role in this haptophyte. However, *dsyB* transcription and DMSP production is regulated by salinity in bacteria, yet no detrimental effect on growth was observed in a bacterial *dsyB* mutant when grown in saline conditions\(^5\). Thus, increased *DSYB* expression and DMSP production with raised salinity does not necessarily indicate a major role for DMSP in osmoprotection.

**P. parvum** *DSYB* protein was concentrated to the chloroplasts and mitochondria (Fig. 2; Supplementary Fig. 4). We propose these organelles as sites of DMSP synthesis in *P. parvum* and perhaps other eukaryotic phytoplankton. Although DMSP production in mitochondria has not been reported, DMSP is produced in the chloroplasts of the higher plant *Wollastonia*, albeit using a different pathway\(^2\). Based on *in silico* sequence analysis (see Methods), *DSYB* from *P. parvum* and some other phytoplankton are predicted to be targeted to the mitochondria and/or chloroplasts (Supplementary Table 5). However, chromophyte algae, such as haptophytes and diatoms, have complex plastids\(^2\), which may render such *in silico* predictions less reliable.

Nanoscale secondary-ion mass spectrometry (NanoSIMS), with a cryopreservation method previously shown to preserve cytosolic DMSP\(^2\), was used to identify potential sub-cellular
sites of DMSP production and storage in *P. parvum* through tracking $^{34}$SO$_4$ uptake. After 48 h incubation, ~23.2 % ± 0.2 of the *P. parvum* intracellular DMSP pool was labelled with $^{34}$S ($^{34}$S-DMSP; Supplementary Fig. 5). Within the cells, the $^{34}$S appeared to be localized in subcellular compartments, with increasing levels appearing over time in the chloroplasts ($^{34}$S/$^{32}$S: 3.4 ± 0.17 after 48 h) and in submicrometer hotspots ($^{34}$S/$^{32}$S: 3.1 ± 0.15 after 48 h) (Fig. 3). Given the size and location of these hotspots, they are likely to be mitochondria or small lipid vesicles (Fig. 3). Although many sulfur compounds are present in algal cells, DMSP represents more than 50% of the total organosulfur compounds in marine phytoplankton and it is expected to account for a significant fraction of the $^{34}$S signal detected by NanoSIMS. However, it cannot be discounted that the increased $^{34}$S content in the chloroplast could be due to transport of sulfur and subsequent assimilation via plastid-located enzymes, such as ATP sulfurylase, APS reductase and sulfite reductase. Nonetheless, the simultaneous increase in $^{34}$S in the chloroplasts and potentially mitochondria supports our hypothesis that these organelles are indeed sites of DMSP synthesis and storage in *P. parvum* and likely other phytoplankton. Given the role of these organelles in energy production, it is perhaps not surprising that DMSP production, an energy-demanding process, may occur at these sites. With DMSP being far less concentrated in the cytosol, it is less likely that its primary function in *P. parvum* is as a typical cytosolic osmolyte, but it may be a key osmolyte in the chloroplasts and/or mitochondria, as proposed in *Wollastonia* chloroplasts. Also, considering reactive oxygen species (ROS) are generated in the mitochondria and chloroplasts, and that DMSP is an effective scavenger of ROS, the production of DMSP in these organelles is in line with its putative role in oxidative stress protection.

Diatoms are thought to produce the lowest intracellular DMSP levels (typically < 50 mM). We studied DMSP production in the polar ice diatom *Fragilariopsis cylindrus*, one of the few diatoms with a functional *DSYB* (Supplementary Table 2), finding that, under standard
conditions, intracellular DMSP levels and DSYB transcription were relatively low, when compared to (e.g.) P. parvum (Supplementary Fig. 2). However, consistent with work in other diatoms\(^{18}\), both F. cylindrus DMSP production and DSYB transcription increased with nitrogen limitation and increased salinity (Supplementary Fig. 2). The latter supports a role for DMSP in osmoregulation and salinity-induced oxidative stress protection in F. cylindrus, as suggested by Lyon et al.\(^{7}\). DSYB was not detected as one of the salinity-induced proteins in Lyon et al.\(^{7}\), despite using the same salinity conditions for our experiments, reflecting the nature of 2D gel electrophoresis studies, whereby not all proteins are identified.

Given the trend of intracellular DMSP concentration increasing with DSYB transcription, we studied Symbiodinium microadriaticum CCMP2467, a dinoflagellate from a genus producing high DMSP concentrations\(^{6}\). S. microadriaticum gave the highest intracellular DMSP (282 mM) and cumulative DSYB transcription of the tested phytoplankton (Supplementary Fig. 2). Similarly, available transcriptomic data showed that high DMSP-producing dinoflagellate and haptophyte phytoplankton (see above) had the highest average DSYB transcription, which was ~3 and 8-fold higher, respectively, than that in diatoms (Supplementary Table 2).

Transcriptomic data was also congruent with high variability in intracellular DMSP levels within dinoflagellates and haptophytes\(^{6,9}\). While additional factors, such as DSYB protein levels, DMSP excretion, DMSP catabolism and cell volume, will affect an organism’s intracellular DMSP concentration, the data presented here on a small number of phytoplankton supports the hypothesis that DSYB transcription is a reasonably good indicator of DMSP concentration. Some DSYB-containing phytoplankton may also contain MTHB methyltransferase isoform enzymes or utilise other DMSP synthesis pathways, in which case such predictions may be inaccurate. Further work is required to substantiate this hypothesis.
The prominence of environmental DMSP-producing bacteria and eukaryotes was examined in the ocean microbial reference gene catalogue (OM-RGC) metagenomic dataset, generated from samples fractionated to < 3 µm (Supplementary Table 6 and Supplementary Fig. 6). The dsyB gene was predicted to be present in 0.35% of total bacteria in these samples. For comparison, DMSP lyase genes (dddD, dddL, dddK, dddP, dddQ, dddW, dddY and Alma1) were also used. The dsyB gene was more abundant than dddL, dddW, dddY, and the algal DMSP lyase gene Alma1, but was less abundant than dddD, dddK, dddP and dddQ in the OM-RGC dataset. Despite only 3% of the OM-RGC microorganisms likely being eukaryotes, DSYB genes were detected and were ~25-fold less abundant than bacterial dsyB. Since no DSYB sequences have been identified in bacteria, we conclude that picoeukaryotes in these samples contain DSYB and thus, the genetic potential to make DMSP. The production of DMSP by DSYB-containing picoeukaryotes could contribute, along with DMSP-producing bacteria, to the DMSP measured from particles <2 µm in size in seawater samples.

We also investigated the occurrence of dsyB and DSYB in marine metatranscriptomes (Supplementary Table 7). dsyB transcripts were detected in all tested Tara oceans metatranscriptomic datasets apportioned to marine bacteria (Supplementary Table 8 and Supplementary Fig. 6). dsyB transcript abundance (normalised to total sequence numbers) was similar to dddD and greater than dddL, dddW, dddY and Alma1, but was far less than dddK, dddP and dddQ. Although these datasets do not consider phytoplankton >3 µm, DSYB transcripts, likely from picoeukaryotes, were detected at levels only 3-fold lower than the bacterial dsyB gene, again suggesting that these smaller eukaryotes, like bacteria, should be considered as potentially significant DMSP producers (Supplementary Table 8).
We also analysed the North Pacific Ocean metatranscriptomes (GeoMICS) which used appropriate fractionation methods for bacteria and larger phytoplankton. As expected, eukaryotic DSYB transcript numbers were higher than those of bacterial dsyB in all of the 2-53 µm fractions, which should contain relatively more phytoplankton than bacteria, and the opposite was true in most of the 0.2-2 µm fractions, which should have relatively more bacteria but not contain the larger phytoplankton (Supplementary Table 9). Analysing data from both the large and small size fractions at different sites allowed us to gauge the relative total transcript numbers of DSYB and dsyB in these samples, as well as those of the DMSP lyase genes. Prokaryotic dsyB transcripts (normalised to the recovery of an internal standard) were more abundant than those for the bacterial DMSP lyase genes dddK, dddL, dddQ, dddY and dddW, 3-fold less than dddP and Alma1 and 27-fold less than dddD (Supplementary Table 9). Eukaryotic DSYB transcripts were slightly less abundant than those for the eukaryotic DMSP lyase (Alma1), but, were ~2-fold more abundant than those for bacterial dsyB. With similar DsyB and DSYB enzyme rates (Supplementary Table 1), this metatranscriptomic data suggests that eukaryotic phytoplankton may be the major contributors to DMSP production via the DsyB/DSYB pathway in these samples. However, direct extrapolation from these data to predict eukaryotic versus bacterial DMSP production (via DsyB/DSYB) is not likely accurate since other factors, such as DsyB/DSYB protein stability or the differing expression and activities of other enzymes in the pathway, may also affect DMSP production. Nonetheless, dsyB and DSYB sequences provide invaluable tools for future, in-depth studies to investigate the relative contribution of bacterial and algal DMSP production in varied marine environments. Molecular studies are also required to identify DMSP synthesis genes in DMSP-producing organisms which lack dsyB or DSYB.
Methods

Media and general growth of algae and bacteria

*Prymnesium parvum* CCAP941/1A, *Prymnesium parvum* CCAP941/6, *Prymnesium parvum* CCAP946/1B, *Prymnesium parvum* CCAP946/1D, *Prymnesium parvum* CCAP946/6, *Prymnesium patelliferum* CCAP946/4, *Chrysochromulina* sp. PCC307 and *Symbiodinium microadriaticum* CCMP2467 were grown in F/2 medium made with ESAW artificial seawater\(^3\) and without any added Na\(_2\)SiO\(_3\). Axenic *Fragilariopsis cylindrus* CCMP1102 was supplied by Mock et al.\(^3\)\(^2\) and grown in F/2 medium made with ESAW artificial seawater at 4 °C with a light intensity of 120 µE m\(^{-2}\) s\(^{-1}\) and constant illumination. *Chrysochromulina tobin CCMP291* was grown in the proprietary medium RAC-5\(^3\)\(^3\). All algal cultures (except *F. cylindrus*) were grown at 22 °C with a light intensity of 120 µE m\(^{-2}\) s\(^{-1}\) and a light dark cycle of 16 h light/8 h dark, unless otherwise stated. Where necessary, media for algal growth were modified according to the requirements of the experimental conditions being tested. Where strains were not already known to be axenic, cultures were treated with multiple rounds of antibiotic treatment prior to experiments. Test cultures with and without antibiotic treatments showed no significant difference in total DMSP in samples. For *P. parvum* CCAP946/6, and *Chrysochromulina* sp. PCC307 cultures, streptomycin (400 µg ml\(^{-1}\)), chloramphenicol (50 µg ml\(^{-1}\)), gentamicin (20 µg ml\(^{-1}\)) and ampicillin (100 µg ml\(^{-1}\)) were added, and for *S. microadriaticum* cultures, streptomycin (100 µg ml\(^{-1}\)) and neomycin (100 µg ml\(^{-1}\)) were added. *E. coli* was grown in LB\(^3\)\(^4\) complete medium at 37 °C. *R. leguminosarum* was grown in TY\(^3\)\(^5\) complete medium or Y\(^3\)\(^5\) minimal medium (with 10 mM succinate as carbon source and 10 mM NH\(_4\)Cl as nitrogen source) at 28 °C. *L. aggregata* J571 was grown in YTSS\(^3\)\(^6\) complete medium or MBM\(^3\)\(^7\) minimal medium (with 10 mM succinate as carbon source and 10 mM NH\(_4\)Cl as nitrogen source) at 30 °C. Where necessary, antibiotics were added to
bacterial cultures at the following concentrations: streptomycin (400 µg ml\(^{-1}\)), kanamycin (20 µg ml\(^{-1}\)), spectinomycin (200 µg ml\(^{-1}\)), gentamicin (20 µg ml\(^{-1}\)), ampicillin (100 µg ml\(^{-1}\)).

Strains used in this study are listed in Supplementary Table 10.

---

**Staining with 4’,6-diamidino-2-phenylindole (DAPI)**

The absence of bacterial contamination was confirmed by epifluorescence microscopy of culture samples stained with DAPI\(^{38}\). Briefly, 13 ml of culture was removed and fixed with 765 µl paraformaldehyde, then 130 µl of DAPI stain (1 mg ml\(^{-1}\) in H\(_2\)O) was added and samples were stored in the dark at 4 °C for 16 h. After staining, 3 ml of the stained cells were removed and filtered onto a Whatman Nuclepore track-etched membrane (25 mm, 0.2 µm, polycarbonate). To prepare slides, one drop of immersion oil was added onto the slide then the sample filter was placed on the oil and another drop of immersion oil was added onto the filter. A cover slip was then placed on top of the filter and pressed down with forceps to remove air bubbles. The slide was then tilted and left on absorbent paper towel to allow any excess oil to drain/wick away. Slides were examined using an Olympus BX40 microscope equipped with an Olympus Camedia C-7070 digital camera.

---

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

Plasmids (Supplementary Table 10) were transferred to *E. coli* by transformation, and *Rhizobium leguminosarum* J391 or *Labrenzia aggregata* J571 by conjugation in a triparental mating using the helper plasmid pRK2013\(^{39}\). Routine restriction digestions and ligations for cloning were performed essentially as in Downie et al.\(^{40}\). The oligonucleotide primers used
for molecular cloning were synthesised by Eurofins Genomics and are detailed in Supplementary Table 11. Sequencing of plasmids and PCR products was performed by Eurofins Genomics.

The *DSYB* gene from *P. parvum* CCAP946/6 was PCR-amplified from cDNA and cloned into the IPTG-inducible wide host range expression plasmid pRK41541. All other *DSYB* genes were synthesised by Eurofins Genomics, from sequences codon-optimised (using Invitrogen GeneArt) for expression in *E. coli*, in the vector pEX-K4 (Eurofins Genomics). The synthesised genes were then subcloned into pLMB50942, a taurine-inducible plasmid for the expression of genes in *Rhizobium* and *Labrenzia*, using *Nde*I and *BamH*I or *EcoR*I restriction enzymes. All plasmid clones are described in Supplementary Table 10.

**MTHB methyltransferase (MMT) assays**

To measure MMT activity from pLMB509 clones expressing the *dsyB* or *DSYB* gene 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), 0.1 mM L-methionine and gentamycin, and incubated at 28 °C for 60 h before sampling for gas chromatography (GC) analysis (see ‘Quantification of DMS and DMSP by gas chromatography’) to determine the amount of DMSP product.

To measure MMT activity from pLMB509 clones expressing the *DSYB* gene in the *L. aggregata* *dsyB*-mutant strain J571, cultures were grown (in triplicate) overnight in YTSS complete medium. Following incubation, 1 ml of culture was then centrifuged at 20,000g 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 gentamycin, and incubated at 30 °C for 24 h. Samples were taken for GC analysis and determining protein concentration (t = 0 h timepoint). DL-MTHB (0.5 mM) and L-methionine (0.1 mM) were then added as substrates to the remaining cultures and these were incubated for 4 h at 30 °C before sampling for GC and protein again (t = 4 h timepoint), with activity calculated based on the difference in measured DMSP product between t=0 h and t=4 h.

To measure 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 immediately, incubated at 22 °C for 24 h and monitored by GC assay (see ‘Quantification of DMS and DMSP by gas chromatography’). DsyB/DSYB activity is expressed as pmol DMSP mg protein⁻¹ min⁻¹, assuming that all the DMSP is derived from DMSHB through DDC activity. LC-MS analysis shows no detectable DMSHB in Rhizobium or Labrenzia expressing DsyB/DSYB, presumably due its conversion to DMSP by DDC activity, so DMSP production is used as a proxy for DsyB activity. 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/DSYB activity.

Growth of algae under non-standard conditions

For all P. parvum, F. cylindrus and C. tobin cultures described here, all samples were taken in mid-exponential phase growth before growth rates started to decline (checked by continuing to monitor growth following sampling). To measure DMSP production or DSYB/DSYB expression in P. parvum CCAP946/6 under different conditions, the growth
conditions or F/2 medium were modified as follows. Standard growth conditions were a temperature of 22 °C, light intensity of 120 µE m⁻² s⁻¹, salinity of 35 practical salinity units (PSU) and nitrogen concentration of 882 µM. For increased or decreased salinity, the amount of salts added to the artificial seawater were adjusted to give a salinity of 50 or 10 PSU respectively. For reduced nitrogen concentration cultures, the F/2 medium contained 88.2 µM (10% of standard F/2). For changes in temperature, cultures were grown at 15 °C or 28 °C.

To measure the effect of increased salinity and nitrogen limitation in *F. cylindrus* CCMP1102, this strain was grown in F/2 medium with increased salts in the artificial seawater (to 70 PSU) or reduced nitrogen (88.2 µM, 10% of standard F/2). To measure the effect of increased salinity and nitrogen limitation in *C. tobin* CCMP291, this strain was grown in F/2 medium with sea salts added to the RAC-5 medium (to 5 PSU) or reduced nitrogen (85 µM, 10% of standard RAC-5).

**Sampling methods**

To measure growth of algal cultures, samples were removed, diluted (dependent on level of growth) in artificial seawater and cell counting was done using a Multisizer 3 Coulter counter (Beckman Coulter). The effect of stress on photosystem II was determined by measuring Fv/Fm values using a Phyto-Pam phytoplankton analyzer (Heinz Walz, Germany). To obtain samples for DMSP quantification by GC or liquid chromatography-mass spectrometry (LC-MS), 25 ml of culture was filtered onto 47 mm GF/F glass microfiber filters (Fisher Scientific, UK) using a Welch WOB-L 2534 vacuum pump, and filters were then blotted on paper towel to remove excess liquid and stored at -80 °C in 2 ml centrifuge tubes for particulate DMSP (DMSPp) measurement. To obtain samples for RNA, 50 ml of culture was filtered onto 47 mm 1.2 µm RTTP polycarbonate filters (Fisher Scientific, UK) and filters
were stored in 2 ml centrifuge tubes at -80 °C. To obtain samples for protein for Western
blotting, 50 ml of culture was centrifuged at 600g for 10 min in a 50 ml centrifuge tube, the
supernatant was decanted and cells were transferred in the residual liquid to a 2 ml centrifuge
tube and centrifuged at 600g for 5 mins. All residual liquid was then aspirated and the
pelleted cells were stored at -80 °C.

Quantification of DMS and DMSP by GC

All GC assays involved measurement of headspace DMS, either directly produced or via
alkaline lysis of DMSP or DMSHB, using a flame photometric detector (Agilent 7890A GC
fitted with a 7693 autosampler) and a HP-INNOWax 30 m x 0.320 mm capillary column
(Agilent Technologies J&W Scientific). All GC measurements were performed using 2 ml
glass serum vials containing 0.3 ml liquid samples and sealed with PTFE/rubber crimp caps.
Quantification of DMSP from algal samples filtered on GF/F glass microfiber filters (see
‘Sampling methods’) was performed following methanol extraction. Filters were folded,
placed in a 2 ml centrifuge tube and 1 ml 100% methanol was added. Samples were stored for
24 h at -20 °C to allow the extraction of cellular metabolites, then 200 μl of the methanol
extract was added to a 2 ml vial, 100 μl 10 M NaOH was added, vials were crimped
immediately, incubated at 22 °C for 24 h in the dark and monitored by GC. Control samples
in which DMSP standards were added to algal sample filters prior to methanol extraction
showed that all standard was recovered following our extraction and measurement procedure.
Calibration curves were produced by alkaline lysis of DMSP standards in water (for
Rhizobium/Labrenzia MMT assays) or 100% methanol (for algal methanol extracts), or DL-
DMSHB (chemically synthesised as in Curson et al.5) standards in water with heating at 80
°C for 10 mins (to release DMS from DMSHB) (for assays with purified DSYB protein). The
detection limit for headspace DMS from DMSP was 0.015 nmol in water and 0.15 nmol in methanol, and from DMSHB was 0.3 nmol in water.

Quantification of DMSP by LC-MS

LC-MS was used to confirm that phytoplankton were producing DMSP and at similar levels to that shown by GC, ruling out the possibility that DMS detected by GC was due to some other compound and not DMSP. Samples were extracted as follows: GF/F filters of phytoplankton (see ‘Sampling methods’) were resuspended in 1 ml of 80% LC-MS grade acetonitrile (extraction solvent), and mixed by pipetting and vortexing for 2 min. The resulting mixture was transferred into a fresh 2 ml Eppendorf tube. For a second round of extraction, another 1 ml of the extraction solvent was then added and mixed as previously described. Then the filters were centrifuged at 18,000g for 10 min and the supernatant was collected, giving a total volume of 2 ml of the collected supernatant. The collected supernatant was then centrifuged at 18,000g for 10 min and 1.5 ml of the supernatant was collected for LC-MS analysis. To extract the metabolites from Chrysochromulina sp. CCMP291, 20 ml of sample was centrifuged at 600g for 10 min and the cell pellet was resuspended in a total volume of 0.7 ml of the extraction solvent and mixed by pipetting and vortexing for 2 min. Samples were then centrifuged at 18,000g for 10 min and 0.5 ml of the supernatant was collected for LC-MS analysis.

LC-MS was carried out using a Shimadzu Ultra High Performance Liquid Chromatography (UHPLC) system formed by a Nexera X2 LC-30AD Pump, a Nexera X2 SIL-30AC Autosampler, a Prominence CTO-20AC Column oven, and a Prominence SPD-M20A Diode array detector; and a Shimadzu LCMS-2020 Single Quadrupole Liquid Chromatograph Mass
Spectrometer. Samples were analysed in hydrophilic interaction chromatography (HILIC) mode using a Phenomenex Luna NH2 column (100 x 2 mm with a particle size of 3 µm) at pH 3.75. Mass spectrometry spray chamber conditions were capillary voltage 1.25 kV, oven temperature 30 °C, desolvation temperature 250 °C and nebulising gas flow 1.50 L min⁻¹. Solvent A is 5% acetonitrile + 95% 5 mM ammonium formate in water. Solvent B is 95% acetonitrile + 5% 100 mM ammonium formate in water. Flow rate was 0.6 ml min⁻¹ and gradient (% solvent A/B) was t = 1 min, 100% B; t = 3.5 min, 70% B; t = 4.1 min, 58% B; t = 4.6 min, 50% B; t = 6.5 min, 100% B; t = 10 min, 100% B. The injection volume was 15 µl. All samples were analysed immediately after being extracted. The targeted mass transition corresponded to [M+H]⁺ of DMSP (m/z 135) and of glycine betaine (m/z 118) in positive mode. A calibration curve was performed for quantification of DMSP and glycine betaine using a mixture of DMSP and glycine betaine standards in the extraction solvent.

Reverse transcription quantitative PCR (RT-qPCR)

For each culture, RNA was extracted as follows: 1 ml Trizol reagent (Sigma-Aldrich), prewarmed at 65 °C, was added directly to the frozen phytoplankton filter (see ‘Sampling methods’), followed by 600 mg of < 106 µm glass beads (Sigma-Aldrich). Cells were disrupted using an MP FastPrep®-24 instrument set at maximum speed for 3 x 30 seconds. Following a 5 min recovery time at 22 °C, samples were centrifuged at 13,000g, 4 °C, for 2 min. The supernatant was transferred to a 2 ml screwcap tube containing 1 ml 95% ethanol and RNA was extracted using a Direct-zol™ RNA MiniPrep kit (Zymo Research, R2050), according to the manufacturer’s specifications. Genomic DNA was removed by treating samples with TURBO DNA-free™ DNase (Ambion®) according to the manufacturer’s protocol. The quantity and quality of the RNA
was determined by a NanoDrop 2000 UV-Vis Spectrophotometer (Thermo Scientific) using 1 µl of sample.

Reverse transcription of 1 µg DNA-free RNA was achieved using the QuantiTect® Reverse Transcription Kit (Qiagen). Primers (Supplementary Table 11) were designed, using Primer3Plus\textsuperscript{43}, to amplify ~130 bp region, with an optimum melting temperature of 60 °C. Melting temperature difference between primers in a pair was 2 °C and GC content was kept between 40-60%.

Quantitative PCR was performed with a C1000 Thermal cycler equipped with a CFX96 Real-time PCR detection system (BioRad), using a SensiFAST\textsuperscript{TM} SYBR® Hi-ROX Kit (Bioline) as per the manufacturer’s instructions for a 3-step cycling programme. Reactions (20 µl) contained 50 ng cDNA and a final concentration of 400 nM of each primer, with a 60 °C annealing temperature. Gene expression for each condition was performed upon three biological replicates, each with three technical replicates. Control DNA consisted of pGEMT-Easy (Promega) containing the fragment created by the RT-qPCR primer pair for each gene tested (made through PCR on synthesised cDNA, cloning in \textit{E. coli} 803 and purifying via a Miniprep Kit [Qiagen]).

For each condition and gene, the cycle threshold (Ct) values of the technical and biological replicates were averaged and manually detected outliers were excluded from further analysis. Standard curves of control DNA were calculated from 3 points of 1:10 serial dilutions, starting with 0.01 ng, to absolutely quantify the \textit{DSYB} transcripts for comparison between organisms\textsuperscript{44}. For an individual organism, relative \textit{DSYB} expression was normalised to the β-actin housekeeping gene, and calculated using the 2\textsuperscript{−ΔΔCT} method\textsuperscript{45} to observe changes in response to various conditions.
Analysis of DSYB expression by Western blotting

A polyclonal rabbit IgG was designed against *P. parvum* DSYB using the OptimumAntigen™ software (GenScript Ltd.). The purified IgG was used as a primary antibody in Western blotting and immunogold labelling (see ‘DSYB immunogold labelling’). The specificity of this antibody was ensured by Western blot analysis of DSYB expressed in the heterologous host *R. leguminosarum* J391. J391 strains containing pBIO2275 (positive control) and pRK415 with no cloned insert (negative control) were grown overnight in TY medium with 0.5 mM IPTG. Proteins were extracted by harvesting 1 ml culture, resuspending cell pellet in 200 µl 20 mM HEPES, 150 mM NaCl, pH 7.5 and disrupting with an ultrasonic processor (Cole Palmer) for 2 x 10 s cycles on ice. Cell debris was separated by centrifugation at 18,000g for 10 mins, following which the supernatant was mixed with SDS sample buffer and incubated at 95 °C for 5 min, before resolution on a 15 % (v/v) acrylamide gel.

The specificity of the anti-DSYB antibody was additionally tested on *P. parvum* 946/6, where protein samples were prepared from cell pellets (see ‘Sampling methods’) as for *R. leguminosarum*, without the removal of cell debris. Cell lysate containing 5.5 µg protein was mixed with SDS sample buffer and heat-treated at 95 °C for 20 min, before resolution on a 15 % (v/v) acrylamide gel.

Following SDS-PAGE, proteins were transferred to a PVDF membrane (Amersham Hybond™-P, GE Healthcare) by semi-dry Western blot as outlined by Mahmood and Yang. After 1 hour blocking with 5 % (w/v) skimmed milk powder in TBS (20 mM Tris, 150 mM NaCl, pH 7.5), the anti-DSYB antibody was added at a final concentration of 0.386 µg ml⁻¹. Specific interactions were left to form overnight at 4 °C, before the membrane was washed 4 x 10 min with TBST (TBS + 0.1 % (v/v) Tween 20). TBST (20 ml) was added with 3 µl anti-
rabbit IgG-alkaline phosphatase at 1 mg ml\(^{-1}\) (Sigma). Following 1 h incubation, the membrane was washed as before with two 10 min TBS washes. Colorimetric detection with NBT/BCIP (Thermo Fisher) was used to detect the target protein as per the manufacturer’s instructions. All SDS-PAGE gels were run with Bio-Rad Precision Plus Dual Colour protein size standards and stained with Coomassie using InstantBlue Protein stain (Expedeon).

**Purification of DSYB and *in vitro* catalytic assays**

A 1.1 kb fragment of DNA containing the synthesised coding region of *Chrysochromulina* *tobin* DSYB was subcloned (from pBIO2272) into pET16b as an *Nde*I/*Eco*RI restriction fragment, downstream of a 10-histidine coding sequence, and transformed into *E. coli* BL21 DE3 (New England BioLabs), for protein purification. Batch cultures were grown aerobically in LB medium at 37 °C until reaching an OD\(_{600}\) value of ~0.6 and were then supplemented with 0.2 mM IPTG and incubated at 28 °C overnight to induce recombinant protein expression. Cells were harvested at 5,000\(\times\)g for 20 min and resuspended in buffer A (20 mM HEPES, 150 mM NaCl, 25 mM imidazole, pH 7.5). The mixture was supplemented with protease inhibitor (Roche cOmplete Tablets, Mini EDTA-free, EASYpack (cat. no. 04 693 159 001)), lysed via sonication and separated at 15,000\(\times\)g, 4 °C for 30 min.

DSYB was purified via an immobilized metal affinity chromatography (IMAC, HiTrap Chelating HP, GE Healthcare) column charged with NiSO\(_4\) and equilibrated with buffer A. All steps were performed at 24 °C with a flow rate of 1 ml min\(^{-1}\). Soluble cell lysate was loaded and washed through with 4 column volumes of buffer A. Bound protein was eluted into 1 ml fractions using a stepped gradient of 25 to 150 mM imidazole, applied for 2 column volumes each. Fractions were visualised via SDS-PAGE analysis (Supplementary Fig. 7) and
those containing DSYB were pooled and dialysed at 4 °C overnight against 20 mM HEPES, 150 mM NaCl, pH 7.5.

*P. parvum* lysate was prepared by centrifuging 100 ml of culture at late exponential phase for 10 min at 2,500g. The pellet was washed with 20 mM HEPES, 150 mM NaCl, pH 7.5 and resuspended in 2 ml buffer supplemented with EDTA-free protease inhibitor (Roche cOmeplete Tablets, Mini EDTA-free, EASYpack (cat. no. 04 693 159 001)). Cells were sonicated 3 x 10 s to lyse, with a 50 s recovery time at 4 °C. Resulting lysate was heat-treated at 80 °C for 10 min to denature proteins (ensuring no activity from native DSYB protein) and centrifuged for 2 min 14,000g. Supernatant was removed to a fresh Eppendorf tube and used for downstream catalytic assays.

DSYB MTHB methyltransferase activity was monitored by performing *in vitro* enzyme assays in 400 µl reactions with 50 µl *P. parvum* lysate and 350 µl purified DSYB (~0.1 mg ml⁻¹) or buffer. All enzyme substrates were added to a final concentration of 1 mM and reactions were incubated at 28 °C for 30 mins. Following this, 800 µl of finely ground charcoal (38 mg ml⁻¹ in 0.1 M acetic acid) was added to the samples and mixed to remove SAM. Samples were centrifuged for 10 mins, 14,000g and the supernatant was retained. For GC analysis, 200 µl of the supernatant was added to a 2 ml vial, 100 µl 10 M NaOH was added, vials were crimped immediately, then heated at 80 °C for 10 minutes (to release DMS from DMSHB) and finally incubated at 22 °C for 24 h in the dark. These samples were subsequently used for quantification of DMSHB by GC analysis as described earlier and activities are reported as nmol DMSHB mg protein⁻¹ min⁻¹. DMS produced from background DMSHB/DMSP present in the *P. parvum* lysate was subtracted from the reported activities.

**DSYB immunogold labelling**
Cells from *P. parvum* 946/6 were cryoimmobilized using a Leica EMPACT High-Pressure Freezer (Leica Microsystems), freeze-substituted in an EM AFS (Leica Microsystems) and embedded in Lowicryl HM20 resin (EMS, Hatfield, USA) as in Perez-Cruz et al.\textsuperscript{47}. Gold grids containing Lowicryl HM20 ultrathin sections were immunolabeled with a specific primary antibody to *P. parvum* DSYB (polyclonal rabbit IgG, GenScript), whose stock concentration was 0.550 mg ml\(^{-1}\) and this was diluted 1:15,000. Secondary antibody was an IgM anti-rabbit coupled to 12 nm diameter colloidal gold particles (Jackson) diluted 1:30. As controls, pre-immune rabbit serum was used as primary antibody, or the gold-conjugated secondary antibody was used without the primary antibody. Sections were observed in a Tecnai Spirit microscope (FEI, Eindhoven, The Netherlands) at 120 kV.

**Prymnesium growth and experimental conditions for NanoSIMS**

*P. parvum* were grown as previously described in F/2 medium (35 PSU)\textsuperscript{30}. Sodium sulfate (Na\(_2\)SO\(_4\), 25 mM) was used as the sole sulfur source, with either \(^{34}\)S (90\% \(^{34}\)S (Sigma-Aldrich, USA; hereafter called \(^{34}\)S-F/2) or natural abundance of \(^{32}\)S (95\% \(^{32}\)S, 0.7\% \(^{33}\)S, 4.2\% \(^{34}\)S; hereafter called natS-F/2). Consequently, the composition of the both the trace metals and vitamin complement had to be slightly modified (with Riboflavin replacing the sulfur-containing Biotin and Thiamine)\textsuperscript{22}. *P. parvum* cells in late exponential phase (grown in natS-F/2) were centrifuged at low speed (1,000 g) for 5 mins, rinsed with \(^{34}\)S-F/2 (to remove potential leftover natS) and transferred in \(^{34}\)S-F/2, whereas a batch incubated only in natS-F/2 acted as a control. Culture were sampled at four time-points: directly after the medium exchange, and after 6 hrs, 24 hrs and 48 hrs. At each timepoint, cultures were sampled for NanoSIMS, mass-spectrometry and cell counts (see below).
Flow cytometry for NanoSIMS samples

Cells were enumerated in triplicate via flow cytometry (BD Accuri C6, Becton Dickinson, USA). For each sample, forward scatter (FSC), side scatter (SSC), and red (chlorophyll) fluorescence were recorded. The samples were analysed at a flow rate of 35 µl min⁻¹.

*Prymnesium* populations were characterized according to SSC and chlorophyll fluorescence and cell abundances were calculated by running a standardized volume of sample (50 µl).

Sample collection for mass spectrometry (NanoSIMS)

At each time point, 1 ml of culture was centrifuged at low speed (1,000 g) for 5 mins, the supernatant was discarded and the cell pellet was extracted with 80% methanol, sonicated on ice for 30 mins and dried.

Dried extracts were reconstituted in methanol to perform LC-MRM-MS analysis. The LC-MS system consisted of an Agilent 1290 series LC interfaced to an Agilent G6490A QQQ mass spectrometer (Agilent, Santa Clara, CA, USA). The MS was equipped with an electrospray ionization source and was controlled by Mass Hunter workstation (version B07) software. A HILIC column (Luna Phenomenex, 150×3 mm, 5 µm, 300 Å) was used for the on-line separations, at a flow rate of 1 ml min⁻¹. The gradient used consisted of a 95 % solvent B (Acetonitrile, 0.1% formic acid), followed by a 2 min linear gradient to 40% solvent A (Milli Q, 0.1 % formic acid), then a 10 min linear gradient to 90% A, and returning to initial conditions at 12.25 min. The injection volume was 2 µl. The MS acquisition parameters were: positive ion mode; capillary voltage, 3,000 V; gas flow 12 l min⁻¹; nebulizer gas, 20 p.s.i.; sheath gas flow rate 7 l/min⁻¹ at a temperature of 250 °C. Acquisition was done in MRM mode with transitions m/z 135- > 63 and m/z 137- > 65 for quantifying $^{32}$DMSP and...
$^{34}$DMSP respectively. The collision energy was optimised as 10 eV to detect the highest possible intensity.

Sample collection and preparation for NanoSIMS

Samples for NanoSIMS were collected and processed following the method described by Raina et al.\textsuperscript{22}. Briefly, samples were snap-frozen, and embedded following by a water-free embedding procedure to effectively prevent the loss of highly soluble compounds such as DMSP from the samples. This method does retain elements in solution by effectively replacing the ‘solution’ with resin, without displacing the ions and osmolytes. \textit{Prymnesium} cultures (20 µl) were dropped onto Theranox strips (Thermo Fisher Scientific, Waltham, USA, 4×18 mm) and placed in humidified chambers. After 20 min, the cells settled onto the strips and the excess medium was carefully removed with filter paper. The strips were then immediately snap-frozen by immersion into liquid nitrogen slush\textsuperscript{22}. Samples were stored in liquid nitrogen until required. Frozen samples for NanoSIMS were freeze-substituted in anhydrous 10% acrolein in diethyl ether, and warmed progressively to room temperature over three weeks in an EM AFS2 automatic freeze-substitution unit (Leica Microsystems, Wetzlar, Germany) as described recently in step-by-step detail by Kilburn and Clode\textsuperscript{48}. The samples were subsequently infiltrated and embedded in anhydrous Araldite 502 resin, after which the Theranox strip was removed and the sample re-embedded and stored in a desiccator. No sulfur was present in processing or resin components. Resin sections (1 mm thick) of embedded \textit{Prymnesium} cells were cut dry using a Diatome-Histo diamond knife on an EM UC6 Ultramicrotome (Leica Microsystems, Wetzlar, Germany), mounted on a silicon wafer and coated with 10 nm of gold.
NanoSIMS analysis

The NanoSIMS-50L (Cameca, Gennevilliers, France) at the Centre for Microscopy, Characterisation and Analysis (CMCA) at the University of Western Australia was used for all subsequent analyses. The NanoSIMS-50L allows simultaneous collection and counting of seven isotopic species, which enables the determination of $^{34}\text{S}/^{32}\text{S}$ ratio. Enrichments of the rare isotope $^{34}\text{S}$ was confirmed by an increase in the sulfur ($^{34}\text{S}/^{32}\text{S}$) ratio above natural abundance values recorded in controls (0.0438). NanoSIMS analysis was undertaken by rastering a 2.5 pA Cs$^+$ beam (~100 nm diameter) across defined 20 µm$^2$ sample areas (256×256 pixels), with a dwell time of 30 ms per pixel. The isotope ratio values are represented hereafter using a colour-coded transform (hue saturation intensity (HSI)) showing natural abundance levels in blue, and grading to high enrichment in pink. Images were processed and analysed using Fiji (http://fiji.sc/Fiji) with the Open-MIMS plug-in (http://nrims.harvard.edu/software). All images were dead-time corrected. Ratio data were tested for QSA (quasi-simultaneous arrivals) by applying different beta values from 0.5 to 162. No differences in the data were observed, indicating that the secondary ion count rates were too low to be affected by QSA. Quantitative data were extracted from the mass images through manually drawn regions of interest, at T0 (whole cells n = 7, hotspot n = 10, chloroplasts n = 3), at T6 (whole cells n = 14, hotspot n = 10, chloroplasts n = 6), at T24 (whole cells n = 12, hotspot n = 10, chloroplasts n = 9), and at T48 (whole cells n = 6, hotspot n = 10, chloroplasts n = 4).

Statistics

Statistical methods for RT-qPCR are described in the relevant section above. All measurements for DMSP production or DSYB/DsyB enzyme activity (in algal strains or...
enzyme assays) are based on the mean of at least three biological replicates per
strain/condition tested, with all experiments performed at least twice. To identify statistically
significant differences between standard and experimental conditions in Supplementary Fig.
2, a two-tailed independent Student’s t-test (P<0.05) was applied to the data, using R51.

Identification of DSYB proteins in eukaryotes

BLASTP and TBLASTN searches52 were used to identify homologues of the Labrenzia DsyB
protein in available eukaryotic genomes and/or transcriptome assemblies at NCBI or JGI.
Any eukaryotic DsyB-like proteins (E values ≤ 1e-30), were aligned to ratified bacterial DsyB
sequences and to non-functional DsyB-like proteins, e.g., in Streptomyces varsoviensis, see
below. Representative DsyB-like proteins, more similar to DsyB than to non-functional S.
varsoviensis DsyB-like proteins, were cloned and assayed for MMT activity (as above).

Ratified eukaryotic DSYB peptide sequences were used in BLASTP searches of 119
eukaryotic transcriptomes (with replicates) downloaded from the Marine Microbial
Eukaryote Transcriptome Sequencing Project (MMETSP)53 via the sequencing repositories
iMicrobe (http://imicrobe.us/project/view/104) and ENA (European Nucleotide Archive)54.
Of these, 45 contained at least one hit to DSYB (E values ≥ 1e-30) (Supplementary Table 3).
Each potential DsyB/DSYB sequence was manually curated by BLASTP analysis against the
RefSeq database and discounted as a true DSYB sequence if the top hits were not to ratified
DSYB sequences detailed in Fig. 1b. DSYB sequences identified from iMicrobe
transcriptomes were aligned to ratified DsyB and DSYB sequences and included in the
evolutionary analysis (Fig. 1b). All DsyB and DSYB protein sequences identified from
genomes or transcriptomes are listed in Supplementary Data 1. Kallisto55 was used to
quantify transcript abundances. Firstly, Kallisto indexes were created for the combined nucleotide assemblies of each organism. Next, Kallisto quant was used to obtain Transcripts Per kilobase Million (TPM) expression values for all datasets using the relevant reference transcriptome index for that organism. Nucleotide sequences corresponding to the DSYB hits were obtained using TBLASTN, and the CAMNT ID number was used to identify the TPM values for each \textit{DSYB} read, giving an estimate of gene expression for organisms grown in standard conditions.

**Phylogenetic analysis of DSYB and DsyB proteins**

All prokaryotic DsyB and eukaryotic DSYB amino acid sequences were aligned in MAFFT\textsuperscript{56,57} version 7 using default settings, then visually checked. Prior to phylogeny construction, model selection was carried out and the best supported model of sequence evolution based on the Bayesian Information Criterion (BIC)\textsuperscript{58} was selected for phylogeny construction (the LG+I+G4 model\textsuperscript{59}). A maximum likelihood phylogeny was then constructed using IQ-TREE\textsuperscript{60} version 1.5.3, implemented in the W-IQ-TREE web interface\textsuperscript{61}, with 1,000 ultrafast bootstrap replicates\textsuperscript{62} used to assess node support. The resulting tree was rooted using a non-DsyB methyltransferase sequence from \textit{Streptomyces varsoviensis}\textsuperscript{5}, and was formatted for publication using the ggtree package\textsuperscript{63} in R\textsuperscript{51}.

**Analysis of DSYB sequences for localisation signals**

Searches for localisation signals in the DSYB protein sequences used the prediction software packages SignalP 4.1 (http://www.cbs.dtu.dk/services/SignalP/), TargetP 1.1
(http://www.cbs.dtu.dk/services/TargetP/) and ChloroP 1.1 (http://www.cbs.dtu.dk/services/ChloroP/).

Analysis of marine metagenomes and metatranscriptomes

Hidden Markov Model (HMM)-based searches for *dysY* and *DSYB* homologs in metagenome and metatranscriptome datasets were performed as described in\(^6^4\) using HMMER tools (version 3.1, http://hmmer.janelia.org/). The DsyB/DSYB protein sequences, shown in Fig. 1b, and ratified DddD\(^\text{65-68}\), DddK\(^\text{69}\), DddL\(^\text{70}\), DddP\(^\text{71}\), DddQ\(^\text{72}\), DddY\(^\text{73}\) DddW\(^\text{74}\) and Alma1\(^\text{75}\) sequences were used as training sequences to create the HMM profiles. Profile HMM-based searches eliminate the bias associated with single sequence BLAST queries\(^7^6\). HMM profiles for the *recA* gene were downloaded from the functional gene pipeline and repository (FunGene\(^7^7\)). The *Ruegeria pomeroyi* DddW\(^\text{74}\) sequence was used to search metagenome and metatranscriptome datasets via BLASTP\(^5^2\) since it is the only ratified DddW. HMM and BLASTP searches were performed against peptide sequences predicted from OM-RGC database assemblies (Supplementary Table 6) and all hits with an E value cut-off of \(1e^{-30}\) were retrieved. In the case of metatranscriptome datasets (*Tara Oceans* and GeoMICS metatranscriptomes), homologs with an E value cutoff of \(1e^{-5}\) were retrieved. Each potential DsyB/DSYB sequence retrieved from the analysis of metagenomes and metatranscriptomes was manually curated by BLASTP analysis against the RefSeq database and discounted as a true DsyB sequence if the top hits were not to DsyB or DSYB sequences detailed in Fig. 1b. If the top hits were to eukaryotic DSYB then the sequence was counted as a true DSYB sequence, and vice versa for bacterial DsyB. Each of the DddD, DddK, DddL, DddP, DddQ, DddW, DddY and Alma1 peptide sequences retrieved were aligned to curated reference sequences using hmmalign and an approximate maximum likelihood tree was constructed.
using FastTree\textsuperscript{78} v2.1. Putative Ddd or Alma1 peptide sequences not aligning most closely to functional Ddd or Alma1 enzymes were removed. To estimate the percentage of bacteria containing \textit{dsyB}, the number of unique hits to DysB in metagenomes was normalised to the number of RecA sequences. Retrieved DsyB/DSYB homolog sequences were aligned to the training sequences using the \textit{dsyB} HMM alignment and this was used to construct an approximately maximum likelihood phylogenetic tree inferred using FastTree\textsuperscript{78} v2.1. The resulting tree (Supplementary Fig. 6) was visualised and annotated using the Interactive Tree Of Life (iTOL)\textsuperscript{79} version 3.2.4.

The GeoMICS metatranscriptome database\textsuperscript{29} generated from North Pacific Ocean samples offered an opportunity to compare prokaryotic and eukaryotic gene expression. Sequences from both the 0.2 µm – 2 µm and 2 µm – 53 µm filtrate fractions for sites P1 and P6 (those samples that had duplicates) were obtained from NCBI (Accession: PRJNA272345) (Supplementary Table 7). Sequences were trimmed using TrimGalore (default parameters, paired-end mode, https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/) and overlapping paired-end reads were joined using PandaSeq\textsuperscript{80}. To create peptide databases, the joined reads were translated using the translate function in Sean Eddy's squid package (http://selab.janelia.org/software.html) to generate all ORFs above 20 amino acids in length. The resulting peptide sequences were used to retrieve \textit{dsyB} and \textit{DSYB} sequences using HMM searches and BLASTP (as above). Read numbers for \textit{dsyB}/\textit{DSYB} were normalised to the read numbers of internal standard\textsuperscript{29} recovered in each sample by dividing the number of reads by the internal standard number and multiplying by 100. Normalised reads from the same site and fraction were averaged (Supplementary Table 9) and ratios of \textit{dsyB}/\textit{DSYB} calculated.
Data availability statement The datasets analysed during the current study are available in the iMicrobe (https://www.imicrobe.us/#/projects/104), European Nucleotide Archive (https://www.ebi.ac.uk/ena), NCBI (https://www.ncbi.nlm.nih.gov/) and Ocean Microbiome (http://ocean-microbiome.embl.de/companion.html) repositories or are available within the paper in Methods section ‘Analysis of marine metagenomes and metatranscriptomes’ and in Supplementary Tables 7, 8 and 9. All data that support the findings of this study are available from the corresponding author upon reasonable request.

References


analysis of a sea-ice diatom: salinity acclimation provides new insight into the


16. Hovde, B. T. *et al.* Genome sequence and transcriptome analyses of *Chrysochromulina*


34. Sambrook, J., Fritsch, E. F., Maniatis, T. & Nolan, C. *Molecular cloning, a laboratory


71. Todd, J. D., Curson, A. R. J., Dupont, C. L., Nicholson, P. & Johnston, A. W. B. The *dddp* gene, encoding a novel enzyme that converts dimethylsulfoniopropionate into dimethyl sulfide, is widespread in ocean metagenomes and marine bacteria and also occurs in some Ascomycete fungi (vol 11, pg 1376, 2009). *Environ. Microbiol.* 11,


Correspondence and requests for materials should be addressed to Jonathan D. Todd (jonathan.todd@uea.ac.uk).

Acknowledgements Funding from the Natural Environment Research Council (NE/J01138X/1, NE/M004449/1, NE/N002385/1 and NE/P012671/1) supported work in J.D.T.’s laboratory. B.T.W. was supported by a NERC EnvEast grant (NE/L002582/1) and A.B.M was supported by a BBSRC Norwich Research Park Biosciences Doctoral Training Partnership grant (BB/M011216/1). The NanoSIMS work was supported by an Australian Research Council Grant (DE160100636) to J-B.R. We thank Pamela Wells and Marco Giardina for general technical support, Thomas Mock for supplying *Fragilariopsis cylindrus*, and Robert Green, Ji Liu and Colin Murrell for advice and discussion of results. We also acknowledge the *Tara* Oceans Consortium for providing metagenomic sequence data, and the facilities at the Australian Microscopy & Microanalysis Research Facility at the Centre for Microscopy, Characterisation & Analysis, University of Western Australia, a facility funded by the University, State and Commonwealth Governments.

Author contributions J.D.T. wrote the paper, designed experiments and performed experiments (gene cloning, enzyme assays, bioinformatics) and analysed data; A.R.J.C. wrote the paper, designed experiments, performed experiments (gene cloning, enzyme assays, gas chromatography to quantify DMSP/DMSHB, phytoplankton growth experiments), analysed data and prepared figures/tables; B.T.W. performed experiments (bioinformatics analysis of DsyB/DSYB in transcriptomes, metagenomes and metatranscriptomes, phylogenetic tree construction), analysed data and prepared figures/tables; B.J.P. performed experiments (gene
cloning, RNA isolation, qRT-PCR experiments, protein purification, in vitro enzyme assays and Western Blots) and analysed data; L.P.S. performed experiments (gene cloning) and analysed data; A.B.M. performed experiments (LC-MS detection of DMSP and glycine betaine) and analysed data; P.P.L.R. performed experiments (phytoplankton growth experiments); D.K. performed experiments (bioinformatic analysis and phylogenetic tree construction); E.M. performed experiments (immunogold labelling, microscopy) and prepared figures; L.G.S. wrote the paper, performed experiments (evolutionary analysis of DsyB and DSYB sequences and phylogenetic tree construction) and prepared figures/tables; J-B.R. wrote the paper, performed experiments (NanoSIMS, LC-MRM-MS) and prepared figures; U.K. performed experiments (LC-MRM-MS); P.L.C. and P.G. performed experiments (NanoSIMS); O.C. designed antibodies and prepared materials for microscopy; S.M. performed experiments (bioinformatic analysis); R.A.C. supplied C. tobin CCMP291 strain. All authors reviewed the manuscript before submission.

Competing interests

The authors declare no competing financial interests.

Additional Information

Supplementary Information is linked to the online version of the paper.

Reprints and permissions information is available at www.nature.com/reprints.
Figure legends

Figure 1. Transamination pathway for DMSP biosynthesis pathway in bacteria and marine algae, and phylogenetic tree of DsyB/DSYB proteins

a, Predicted pathway for DMSP biosynthesis in bacteria (Labrenzia), macroalgae (Ulva, Enteromorpha), diatoms (Thalassiosira, Melosira), prymnesiophytes (Emiliania) and prasinophytes (Tetraselmis). Abbreviations: Met, methionine; MTOB, 4-methylthio-2-oxobutyrate; MTHB, 4-methylthio-2-hydroxybutyrate; DMSHB, 4-dimethylsulphonio-2-hydroxybutyrate. b, Maximum likelihood phylogenetic tree of DsyB/DSYB proteins. Species are colour-coded according to taxonomic class as shown in the key, with proteins shown to be functional marked with an asterisk. Bootstrap support for nodes is marked. Based on 145 protein sequences.

Figure 2. Immunogold localisation of DSYB in Prymnesium parvum CCAP946/6

Representative electron micrographs of P. parvum cells showing location of DSYB by immunogold labelling. a, b, Immunostaining of cell with DSYB antibody and secondary antibody with gold. c, d, Control immunostaining with pre-immune serum. e, f, Control immunostaining with only secondary antibody. Boxes in a, c, and e, correspond to area magnified in b, d, and f respectively. Scale bars are all 500 nm. Abbreviations: ch, chloroplast; g, golgi apparatus; ig, immunogold; m, mitochondrial; nu, nucleus; py, pyrenoid; ri, ribosome; v, vacuole. Experiments were repeated twice and two samples (n=2) were used for each experiment.
Figure 3. Sub-cellular distribution of $^{34}$S in *Prymnesium parvum* CCAP946/6 following sulfur uptake for 48 h. a-d, Representative $^{12}$C$^{14}$N/$^{12}$C$_2$ mass images showing cellular structures of *P. parvum* cells. The cells were imaged straight after the start of the incubation (a), and after 6 h (b), 24 h (c) and 48 h (d). e-h, $^{34}$S/$^{32}$S ratio of the same cells, shown as Hue Saturation Intensity (HSI) images where the colour scale indicates the value of the $^{34}$S/$^{32}$S ratio, with natural abundance in blue, changing to pink with increasing $^{34}$S levels. Each image was only acquired once. i, Isotope ratio of $^{34}$S/$^{32}$S in different cellular regions (biological replicates, number of cells analysed: T0: whole cells n = 7, chloroplasts n = 3, hotspot n = 10; T6: whole cells n = 14, chloroplasts n = 6, hotspot n = 10; T24: whole cells n = 12, chloroplasts n = 9, hotspot n = 10; and T48: whole cells n = 6, chloroplasts n = 4, hotspot n = 10; error bars are shown for standard error). Abbreviations, ch: chloroplast; h: hotspot; py: pyrenoid; v: vacuole. Scale bars: 1 µm.
Sulfur ratio ($^{34}$S/$^{32}$S) over time (hours:
0 6 24 48)

- whole cells
- chloroplasts
- hotspots