environmental microbiology

Environmental Microbiology (2022) **24**(5), 2380–2403 doi:10.1111/1462-2920.16020

Glycine betaine uptake and metabolism in marine microbial communities

Angela K. Boysen,[1†] **Bryndan P. Durham**,[2] **William Kumler**,[1] **Rebecca S. Key**,[2] **Katherine R. Heal**,[1‡] **Laura T. Carlson**,[1] **Ryan D. Groussman**,[1] **E. Virginia Armbrust**[1] and **Anitra E. Ingalls**[1*]

[1]*School of Oceanography, University of Washington, Seattle, WA, 98195.*

[2]*Department of Biology, Genetics Institute, University of Florida, Gainesville, FL, 32610.*

Summary

Glycine betaine (GBT) is a compatible solute in high concentrations in marine microorganisms. As a component of labile organic matter, GBT has complex biochemical potential as a substrate for microbial use that is unconstrained in the environment. Here we determine the uptake kinetics and metabolic fate of GBT in two natural microbial communities in the North Pacific characterized by different nitrate concentrations. Dissolved GBT had maximum uptake rates of 0.36 and 0.56 nM h^{-1} with half-saturation constants of 79 and 11 nM in the high nitrate and low nitrate stations respectively. During multiday incubations, most GBT taken into cells was retained as a compatible solute. Stable isotopes derived from the added GBT were also observed in other metabolites, including choline, carnitine and sarcosine, suggesting that GBT was used for biosynthesis and for catabolism to pyruvate and ammonium. Where nitrate was scarce, GBT was primarily metabolized via demethylation to glycine. Gene transcript data were consistent with SAR11 using GBT as a source of methyl groups to fuel the methionine cycle. Where nitrate concentrations were higher, more GBT was partitioned for lipid biosynthesis by both bacteria and eukaryotic phytoplankton. Our data highlight unexpected metabolic pathways and potential routes of microbial metabolite exchange.

Received 7 December, 2021; revised 24 March, 2022; accepted 15 April, 2022. *For correspondence. E-mail aingalls@uw.edu; Tel. 206-221-6748. †Present address: Department of Geophysical Sciences, University of Chicago, Chicago, IL 60637, USA. ‡Present address: Integral Consulting, Seattle, WA 98104, USA.

Introduction

In the ocean, the flow of resources and information carried by organic molecules provides the foundation for complex and globally important microbial ecosystems. The quantity and identity of molecules produced by phytoplankton help shape community structure and influence the balance between organic carbon production and degradation, which together influence carbon storage in the ocean (Gómez-Consarnau *et al*., 2012; Moran, 2015). Small, polar metabolites released into seawater are exchanged and used by cells in a variety of ways (Poretsky *et al*., 2010; Vorobev *et al*., 2018). A molecule might be incorporated directly into biomass as a building block for macromolecules, respired for energy, converted into biosynthetic intermediates, or used for specific biochemical roles as osmolytes, cofactors, cellular signals, antioxidants, or chemical defences. Here we focus on a single compound, the compatible solute glycine betaine (GBT). We elucidate the uptake kinetics and metabolic fate of GBT in two natural microbial communities with different nitrate concentrations.

Once released into seawater through excretion or cell death, compatible solutes are a quantitatively important pool of labile dissolved organic matter (DOM) (Bronk *et al*., 1994; Welsh, 2000; Bronk, 2002). GBT, a quaternary amine, is one of the most effective compatible solutes and is found in high concentrations inside cells in saline environments (Airs and Archer, 2010; Klähn and Hagemann, 2011; Johnson *et al*., 2019). GBT comprises a significant portion of intracellular metabolites in the North Pacific subtropical gyre and in the North Pacific transition zone (NPTZ) (Heal *et al*., 2021) and exhibits diel oscillations indicative of active production and consumption (Boysen *et al*., 2021). Incubation experiments with ^{14}C-GBT show that natural microbial communities both respire and assimilate carbon from GBT (Kiene *et al*., 1998; Kiene and Williams, 1998; Mausz *et al*., 2022). However, GBT has complex biochemical potential, and previous work has left the flow and fate for GBT-derived carbon and nitrogen mostly unresolved. Furthermore, few studies have investigated the role that

environmental conditions have on the biochemical fate of GBT.

Across vast swaths of the ocean, inorganic nitrogen availability limits phytoplankton growth (Moore *et al*., 2013). In inorganic nitrogen-depleted waters, dissolved organic nitrogen (DON) is the dominant form of nitrogen, sometimes comprising 100% of measurable total dissolved nitrogen (Berman and Bronk, 2003). Microorganisms that can take up and use DON have an advantage in these environments. However, the specific structures of molecules within the DON pool determine which organisms can take it up and how they can use it. Incubations of DOM amended seawater show that GBT transporters are transcribed in response to DOM additions, presumably to take up GBT (Poretsky *et al*., 2010; Vorobev *et al*., 2018), but these studies did not directly measure uptake or how GBT was subsequently used.

There are three pathways of GBT metabolism with well-characterized enzymes (Fig. 1 pathways 1–3). The demethylation pathway removes methyl groups from GBT one at a time (1 in Fig. 1). Betaine-homocysteine methyltransferase (BHMT) converts GBT and homocysteine to dimethylglycine and methionine, which can fuel methylation throughout the cell via the methionine cycle (Barra *et al*., 2006). Alternatively, the genes *gbcA* and *gbcB* encode a Rieske-family oxidase and flavin-linked oxidoreductase that convert GBT to dimethylglycine with the production of formaldehyde (Wargo *et al*., 2008; Shao *et al*., 2018). Further oxidation of dimethylglycine to sarcosine and glycine occurs through subsequent demethylations in which methyl groups are transferred to tetrahydrofolate and ultimately oxidized to CO_2 in an energy-yielding reaction (Sun *et al*., 2011). The first of these demethylations is catalysed by either dimethylglycine dehydrogenase (DMGDH) or dimethylglycine oxidase (DMGO) and then the second by sarcosine dehydrogenase (SARDH), monomeric sarcosine oxidase (MSOX), or tetrameric sarcosine oxidase (SoxBDAG) (Wagner *et al*., 2000; Wagner and Jorns, 2000; Wargo *et al*., 2008; Lahham *et al*., 2021). Organisms with the full demethylation pathway, including the ubiquitous marine *Alphaproteobacteria* clade SAR11, can convert GBT to glycine and use the glycine for respiration, biosynthesis and protein synthesis (Sun *et al*., 2011; Wargo, 2013; Noell and Giovannoni, 2019). Through this pathway organisms can use GBT to fulfil cellular methylation, glycine, carbon, nitrogen and energy requirements for growth (Lidbury *et al*., 2015a).

The second pathway with characterized enzymes is the transformation of GBT to choline (2 in Fig. 1). A non-ribosomal peptide synthetase-like glycine betaine reductase (ATRR) responsible for this transformation was

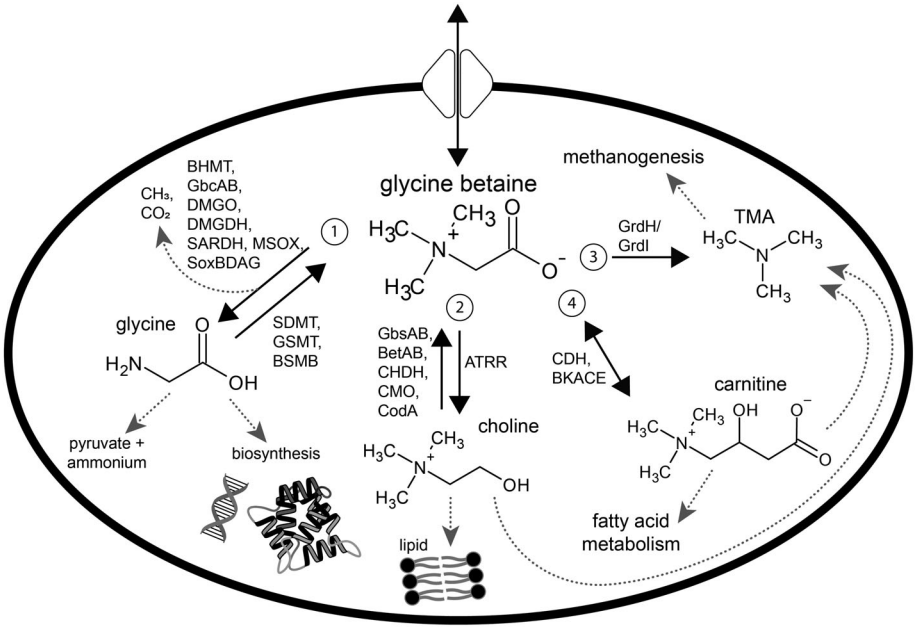

Fig. 1. Schematic of formation and degradation pathways for GBT. Circled numbers are GBT metabolism pathways referenced in the text. Solid black arrows indicate pathways of GBT synthesis or degradation that have one or more characterized enzymes. Grey dotted arrows represent select potential fates of GBT-related molecules. Enzyme names are abbreviated as follows: BHMT = betaine-homocysteine methyltransferase; GbcAB = glycine betaine demethylase; DMGO = dimethylglycine oxidase; DMGDH = dimethylglycine dehydrogenase; SARDH = sarcosine dehydrogenase; MSOX = monomeric sarcosine oxidase; SoxBDAG = tetrameric sarcosine oxidase; GSMT = glycine-sarcosine methyltransferase; BSMB = dimethylglycine N-methyltransferase; GbsAB, BetAB, CHDH = choline dehydrogenase; CMO = choline monooxygenase; CodA = choline oxidase; ATRR = non-ribosomal peptide synthetase-like glycine betaine reductase; CDH = carnitine dehydrogenase; BKACE = β-keto acid cleavage enzyme; GrdHI = betaine reductase.

identified in fungi (Hai *et al*., 2019), though the presence of this pathway in marine microbes is unexplored. The third characterized pathway of GBT metabolism is GBT reduction to trimethylamine (TMA) via betaine reductase (GrdHI) (3 in Fig. 1). TMA can then fuel anaerobic methanogenesis or be oxidized to trimethylamine-*N*-oxide (TMAO) (Sun *et al*., 2019). TMA and TMAO can both be demethylated to ultimately form formaldehyde and ammonium, which can feed back into central carbon and nitrogen metabolism (Sun *et al*., 2019). Betaine reductase is considered an anaerobic pathway, but *grdH* homologues have been identified in proteobacteria from open ocean marine metagenomes (Jameson *et al*., 2016). A fourth GBT metabolism pathway that produces carnitine through GBT-CoA and dehydrocarnitine (4 in Fig. 1) is possible since the enzymes that produce GBT from carnitine [CDH and β-keto acid cleavage enzyme (BKACE)] are theoretically reversible (Sellés Vidal *et al*., 2018).

Cyanobacteria, heterotrophic bacteria and eukaryotic algae have a variety of abilities to synthesize, transport and metabolize GBT (Keller *et al*., 1999a; Welsh, 2000; Mou *et al*., 2007; Torstensson *et al*., 2019; McParland *et al*., 2021). As described above, some heterotrophic bacteria can transport and metabolize GBT. Some diatoms, dinoflagellates and haptophytes that produce GBT can take up extracellular GBT under salinity stress or when extracellular concentrations of GBT are elevated, but there is no evidence that these organisms metabolize GBT (Keller *et al*., 1999b; Torstensson *et al*., 2019). There are also marine microorganisms that lack both the known synthesis and degradation genes for GBT but still have GBT transporters (McParland *et al*., 2021). Thus, there are several potential fates for extracellular GBT once taken up by marine microbes. It could be retained as a compatible solute, catabolized to provide cellular energy, used as a methyl donor (1 in Fig. 1), used for biomolecule synthesis including lipids (through choline, 2 in Fig. 1) and proteins and nucleobases (through glycine), or used for the synthesis of other metabolites (3 in Fig. 1). Intermediates along these pathways may also be exchanged between organisms, implicating GBT as a foundation for complex microbial interactions.

Here we explore the fate of GBT in marine surface waters using transcriptomics and stable isotope tracing untargeted metabolomics. We incubated natural seawater collected at two stations in the NPTZ with $^{13}C_5$, $^{15}N_1$-GBT to measure the kinetics of GBT uptake and determine the fate of the carbon and nitrogen. The NPTZ is a dynamic region characterized by seasonally migrating fronts of temperature, salinity, nutrients and chlorophyll (Ayers and Lozier, 2010; Polovina *et al*., 2017), and thus a useful region to investigate the effects of environmental conditions on GBT uptake and metabolism. We show that microbial communities are poised to take up and retain high concentrations of GBT. Once GBT is in cells, we show that the proportion of carbon and nitrogen that flow through each GBT pathway differs between the two sites. The observed differences are consistent with more intense recycling of photosynthate, and, in particular, the nitrogen in GBT, in the southern, more oligotrophic, environment.

Results and discussion

Environmental description

Experiments to determine the uptake kinetics and cellular fate of GBT were performed in spring 2019 at two stations in the NPTZ. The northern station (41.68 °N, −158 °W) was characterized by 0.85 ± 0.08 μg chlorophyll L^{-1} and 6.42 μM dissolved inorganic nitrogen (DIN, NO_3^- +-NO_2^-). The southern station (37 °N, −158 °W) was characterized by 0.57 ± 0.03 μg chlorophyll L^{-1} and DIN below our limit of detection (near 0.02 μM, Table 1). There was more particulate carbon, heterotrophic bacterial biomass and picophytoplankton biomass in the southern station (Table 1). The elevated chlorophyll in the northern station was likely due to larger phytoplankton (>4 μm diameter) outside the analytical window of the flow cytometer.

Glycine betaine uptake kinetics

GBT uptake was well described by Michaelis–Menten kinetics at both locations (Fig. 2 top row). A non-linear least-squares fit was used to determine two uptake parameters: maximum uptake rate, V_{max}, and half saturation concentration plus *in situ* GBT concentration, $K_t + S$. The southern station had a V_{max} of 0.56 ± 0.06 nM h^{-1} and $K_t + S$ of 11.06 ± 5.53 nM. The northern station had lower V_{max} of 0.36 ± 0.06 nM h^{-1} and higher $K_t + S$ of 79.46 ± 48.26 nM, indicating a lower capacity to take up GBT at low concentrations (Table 2, Supplemental Table S1). The GBT uptake kinetic parameters measured here are within the wide range of published values for amino acids, the DON compounds best described in the literature (Sipler and Bronk, 2015). The half saturation constants of GBT uptake that we estimate are on the high end of those previously measured for free-living bacteria which had $K_t + S$ values of <1 to 17.2 nM, (Kiene *et al*., 1998; Mausz *et al*., 2022) and for whole seawater communities in coastal environments (1.8–49 nM for waters with salinity >10 ppt) (Kiene and Williams, 1998). We estimated a turnover time for dissolved GBT of 8.7–37.5 h in the southern station and >100 h in the northern station (Table 2). Compared to the <1–11 h GBT turnover time of coastal and estuarine environments (Kiene

Table 1. Summary of the environmental conditions where samples were collected in this study.

	Date	Lat	Lon	Depth	PC	PN	C:N	TON	DIN	Temperature		Chl		Bacteria biomass		*Prochlorococcus* biomass		Picoeukaryote biomass		*Synechococcus* biomass	
										Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
North experiment	4/16/2019	41.68	−158.05	15	4.24	0.71	5.97	4.4	6.42	9.59	0.05	0.85	0.08	6.21	0.42	NA	NA	6.02	0.29	0.52	0.03
South experiment	4/20/2019	37	−158	15	6.69	0.65	10.29	4.29	NA	13.5	0.09	0.57	0.03	12.8	0.5	1.43	0.04	9.91	0.06	3.73	0.01

Depth is in meters. Particulate carbon (PC), particulate nitrogen (PN), total organic nitrogen (TON) and dissolved inorganic nitrogen (DIN) are in μM. Chlorophyll is in μg L^{-1}. Temperature is in °C. Microbial biomass is in μg C L^{-1}.

et al., 1998; Kiene and Williams, 1998), these data suggest that open ocean communities are slower to turn over the available pool of GBT.

The presence of more transporters for a given substrate is one way to increase the maximum uptake rate, V_{max}, measured at the community scale, and thus biomass is one of the parameters that can directly influence V_{max} (Button, 1998). At the community scale, a low half saturation constant, K_t, is likely due to high-affinity transporters making up a substantial proportion of total transporters, regardless of transporter type. Though the community level K_t value does not necessarily reflect any individual organism or transporter K_t value, it does inform us about the general affinity level that transporters have for a substrate (Kiene *et al*., 1998). Free-living heterotrophic bacteria were more abundant in the southern station than the northern station (Table 1) and likely contributed to the high V_{max} and low $K_t + S$ values which were near the SAR11 GBT transporter K_t values of 0.89 and 1.85 nM (Noell and Giovannoni, 2019). The microbial community in the northern station had more large eukaryotic phytoplankton, which contributed to the high chlorophyll levels at that location (Table 1). The only characterized GBT transporters in phytoplankton are in sea-ice diatoms, which have K_t values of 189–315 nM (Torstensson *et al*., 2019) that are near the value measured at the northern station. The high community K_t value in the north indicates that more of the organisms taking up GBT have lower affinity transporters, and we suggest that phytoplankton are substantially contributing to uptake at that station. Our results support the hypothesis that community composition, bacterial abundance and nutrient concentrations influence GBT uptake kinetics (Mausz *et al*., 2022).

Glycine betaine over time

To assess the fate of GBT once it was taken up into cells, we added $^{13}C_5$, $^{15}N_1$-GBT to natural seawater communities in excess of the half saturation constants (500 nM) and traced carbon and nitrogen isotopes through metabolic pathways over several days. Samples for the initial time point were filtered directly after the addition of labelled GBT, resulting in an effective incubation time of 15–30 min. Though we expect ambient dissolved GBT concentrations to be lower than 500 nM, these concentrations may be reached in microenvironments such as the phycosphere, where nutrients are exchanged between phytoplankton and their associated bacteria (Stocker, 2012).

In the first 4–6 h, the total particulate GBT concentrations increased from 0.60 to 1.03 nM in the south and from 0.61 to 0.84 nM in the north (Fig. 2 bottom row). Though the particulate GBT concentrations continued to

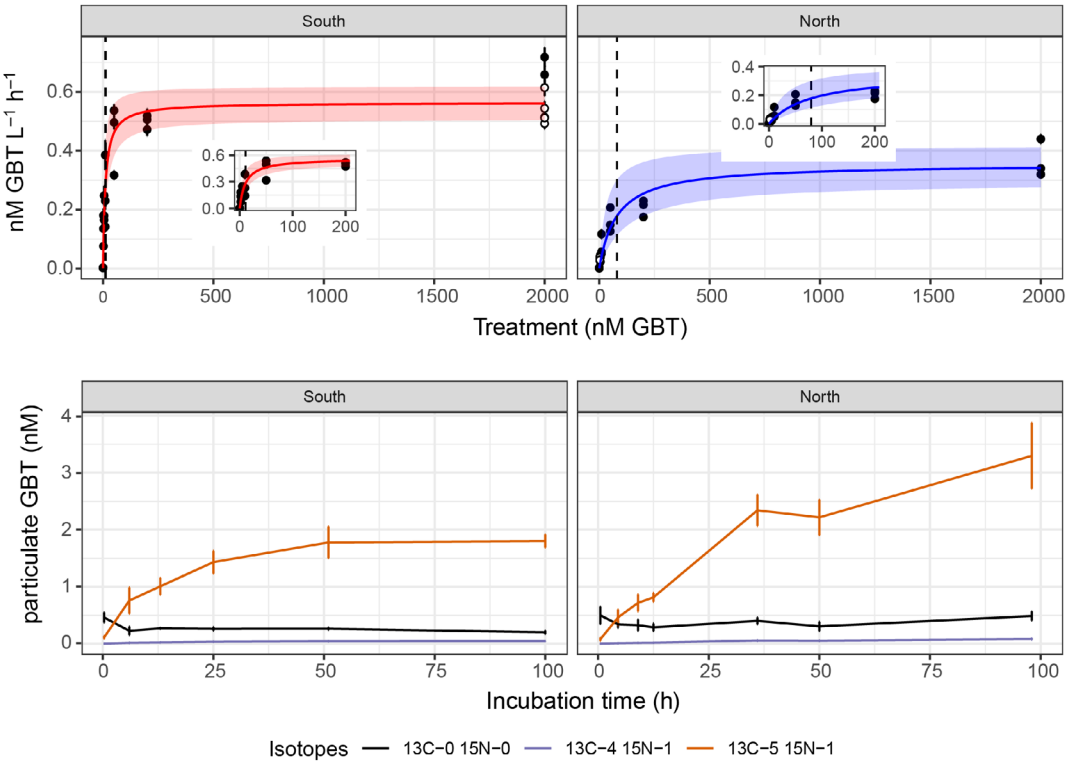

Fig. 2. Top row: Uptake of $^{13}C_5$, ^{15}N-GBT in natural seawater. Black points are rates calculated from individual bottles, standard error propagated from the uncertainty associated with quantification is shown as the bars, sometimes obscured by the size of the points. Open circles are replicate injections of the same sample. Uptake kinetics curve fit to the data with a non-linear least squares and Monte Carlo randomization, as described in the Experimental Procedures section, is shown in the solid line, with one standard deviation of the best fit $K_t + S$ and V_{max} parameters shown in shading. Dashed vertical line is the best fit $K_t + S$ value. Inset shows the shape of the curve for the 0–200 nM treatments. Bottom row: Particulate GBT over time in the incubation experiments with 500 nM $^{13}C_5$, ^{15}N-GBT. The three most abundant isotopologues are shown here. Others were measured and are shown in Supplemental Fig. S2. Solid lines indicate the mean of triplicate bottles and the error bars represent one standard deviation.

increase over the course of the experiments due to uptake of the added GBT, the composition of the compatible solute pool overall did not change substantially, since homarine dominated the total compatible solute pool at both stations at all time points (Supplemental Fig. S1). Though particulate GBT concentrations continued to increase in the last 2 days of the northern incubation, the fraction of total particulate GBT that was in the $^{13}C_5$, $^{15}N_1$-GBT isotope form (a value represented by the mass isotopologue distribution, MID) stabilized after 2 days of incubation. $^{13}C_5$, $^{15}N_1$-GBT, the isotopologue of GBT that was added, reached approximately 85% of the total particulate GBT at both locations (Supplemental Fig. S2).

By searching metabolomics data for compounds enriched in ^{13}C and ^{15}N we were able to identify multiple active pathways of GBT transformation. Unmetabolized GBT dominated ^{13}C and ^{15}N isotopes in aqueous intracellular metabolites (Fig. 3, Supplemental Fig. S3). In the final time point of the northern experiment, GBT isotopologues occupied 85 $\pm$ 19% of ^{15}N and 80 $\pm$ 17% of ^{13}C within the quantifiable particulate metabolites. In the final time point of the southern experiment, GBT isotopologues occupied 88 $\pm$ 7% of ^{15}N and 86 $\pm$ 7% of ^{13}C within the quantifiable particulate metabolites. Homarine represented a substantial portion of the total ^{13}C pool at all time points due to its high concentration and the natural abundance of ^{13}C, not due to enrichment from GBT (Supplemental Fig. S1). We underestimate the total isotope uptake and transformation because our analytical window does not allow us to observe every metabolite, not all metabolites we observe can be quantified, and we did not measure metabolic products, including carbon dioxide and ammonium, in the dissolved pool. Nevertheless, our data indicate that a substantial portion of total GBT taken up is retained as a compatible solute.

Our observations contrast studies in coastal environments where 48-h incubations of ^{14}C-GBT showed that much of the carbon taken up as GBT was incorporated into non-soluble molecules interpreted to be macromolecules (Kiene and Williams, 1998). These previous studies focused on the <0.8 μm size fraction and concluded that most GBT uptake and metabolism occurred within the

Table 2. Summary of uptake kinetic parameters for various metabolites in the marine environment.

Compound	Location	V_{max} (nmol L^{-1} h^{-1})[a]	Kt + S (nM)[a]	Turnover time (h)[a]	References
Glycine betaine	N. Pacific (41 N)	0.3–0.42	31.2–127.72	106–169	This study
Glycine betaine	N. Pacific (37 N)	0.5–0.63	4.81–21.49	8.7–37.5	This study
Glycine betaine	Mobile Bay	0.39–44	1.8–49	0.52–11	Kiene and Williams (1998)
Glycine betaine	Mobile Bay, 0.7 μm filtrate	2.4–27	0.86–4.5	0.15–1.8	Kiene *et al*. (1998)
Glycine betaine	Western English Channel, 1.2 μm filtrate	0.6–40.8	1.4–17.2	0.19–6.2	Mausz *et al*. (2022)
DMSP	Mobile Bay, 0.7 μm filtrate	8.94	3.77		Kiene *et al*. (1998)
DMSP	NPSG			4.1–7.4	Del Valle *et al*. (2012)
Choline	Mobile Bay estuarine water	0.5–3.3	1.7–4.1	1.3–3.2	Kiene *et al*. (1998)
Choline	Western English Channel, 1.2 μm filtrate	0.2–3.4	1.9–5.3	<1–25	Mausz *et al*. (2022)
Individual DFAA	Sargasso Sea (N. Atlantic)	0.003–0.2	0.1–5.5	1.6–181.9	Suttle *et al*. (1991)
Alanine	Sargasso Sea (N. Atlantic)	0.015–0.15	0.9–1.8	1.6–13.2	Suttle *et al*. (1991)
Ornithine	Sargasso Sea (N. Atlantic)			48.7–181.9	Suttle *et al*. (1991)
Taurine	Off the Iberian Peninsula			6–240	Clifford *et al*. (2019)
Leucine	Off the Iberian Peninsula			21–6000	Clifford *et al*. (2019)

[a]Ranges represent the range of the values in the publication or the best estimate plus and minus one standard deviation, the values for this study include the full range of estimates from the linear transformation and non-linear least squares fit.

bacterial community. We hypothesize that the uptake and retention of GBT, which was greater in the northern experiment than the south experiment, is driven in part by eukaryotic phytoplankton uptake. Cultures of diatoms, haptophytes and dinoflagellates take up GBT when provided with exogenous GBT (Keller *et al*., 1999b; Torstensson *et al*., 2019) but lack the known genes for GBT demethylation (McParland *et al*., 2021). Taking up and retaining compatible solutes has a cost associated with maintaining membrane transporters, but those costs may well be offset by saving the energy, carbon and nitrogen required to synthesize compatible solutes, especially if the transporter can be used for multiple substrates as is the case for some characterized GBT transporters (Kiene *et al*., 1998; Noell and Giovannoni, 2019). Excess uptake and storage of GBT may be a strategy for surviving patchy and nutrient-limited regions of the ocean. Similar luxury uptake strategies have been proposed for inorganic nutrients and other organic nutrients, including uptake of polyamines by SAR11 (Gaubert *et al*., 2020; Noell *et al*., 2021). Future studies of organisms in culture should explore this potential use of GBT and other nitrogen-containing osmolytes as a mechanism for increased fitness in low and fluctuating nitrogen regimes.

Primary pathways of GBT transformation

Isotope tracing metabolomics revealed four potential pathways of GBT metabolism in both incubation experiments, as summarized in Fig. 1: (i) the demethylation pathway into dimethylglycine, sarcosine, glycine and creatine, (ii) the choline lipid pathway into choline and glycerophosphocholine, (iii) the reduction pathway into TMA and TMAO and (iv) the carnitine synthesis pathway into carnitine and trimethylammoniobutanoate (TMAB). In our environment, the third and fourth potential pathways may be a single pathway, since TMAO can also be produced from carnitine, rather than from GBT directly, as discussed below.

Our stable isotope probing experimental results led us to search for which organisms might be responsible for these various transformations by searching metatranscriptomes collected from the NPTZ on a cruise in April–May 2016. Unless otherwise specified, we present the gene transcription data as an average of the five stations in the NPTZ, as defined by the second derivative of salinity, following Gradoville *et al*. (2020). We use the regional average because the stations where metatranscriptomes were collected were not concurrently sampled with those used for the GBT experiments. This analysis reveals which organisms within this region are likely carrying out the observed metabolisms *in situ*, and thus lets us infer how various microbial community members may rely on GBT and the immediate products of its metabolism. Metatranscriptomes collected from the same region in May–June 2017 suggest that the organisms expressing GBT-related genes are relatively stable from year to year (Supplemental Fig. S4). We identified taxonomic patterns in the transcription of the genes involved

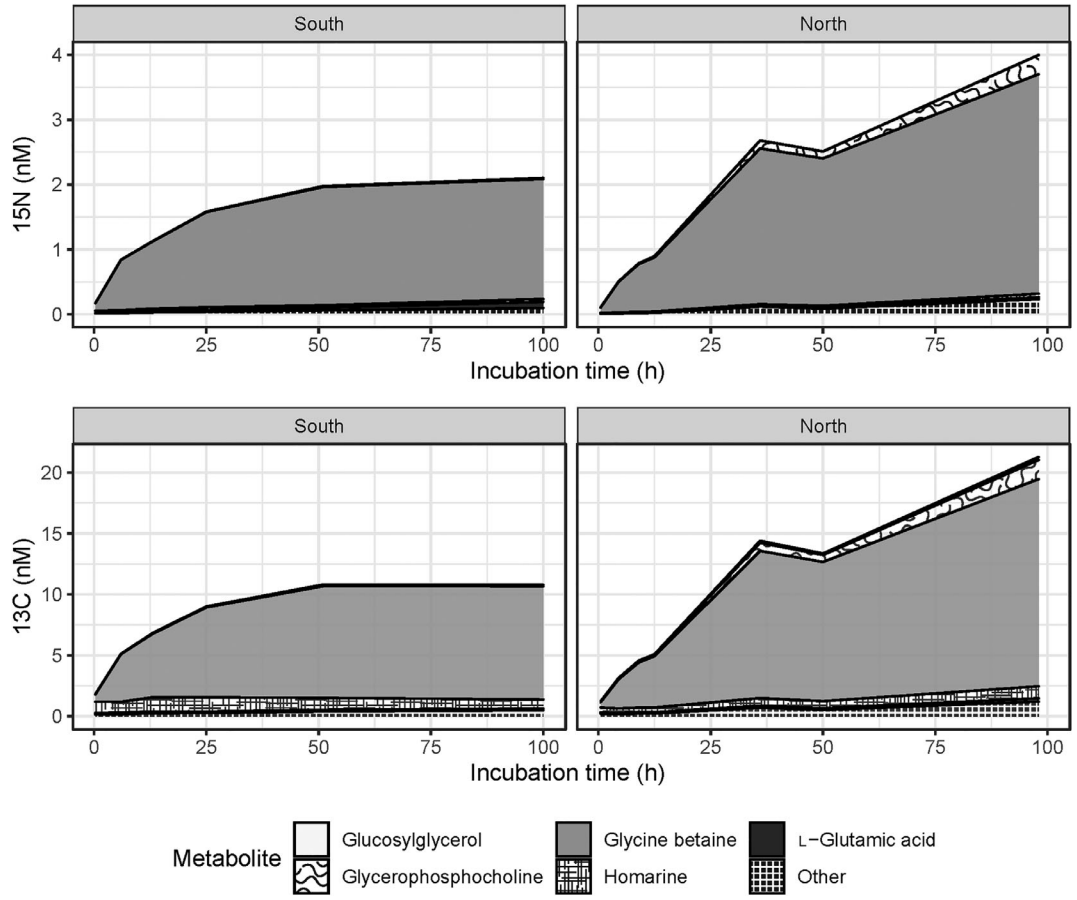

Fig. 3. The total particulate ^{15}N (top) and ^{13}C (bottom) isotope concentrations in the south (left) and north (right) incubation experiments in known, quantifiable metabolites. 'Other' includes all compounds quantified that are not otherwise specified, as listed in Supplemental Table S4. Note that natural abundance isotopes are also included in this total. ^{13}C did not accumulate over time in homarine, but the concentration of homarine was high enough that its natural abundance ^{13}C content contributes substantially to the total.

in the three steps of the demethylation pathway (*bhmt*, *gbcAB*, *dmg*, *dmgdh*, *sardh*, *sox*, *soxAB*) and the choline synthesis pathway (*atrr*). There were no obvious taxa transcribing homologues of *grdI*, the first step of the reduction pathway, and we therefore hypothesize that TMA and TMAO are formed from carnitine.

GBT demethylation. The pathway of GBT metabolism to dimethylglycine and sarcosine is well characterized in marine heterotrophic bacteria, and this was a major pathway of GBT transformation observed in our experiments (Fig. 4). Unlabelled dimethylglycine was below our limit of detection due to elevated noise in the mass spectrometer signal at the relevant retention time (RT) and mass-to-charge ratio; however, $^{13}C_4$, $^{15}N_1$-dimethylglycine, the expected product of demethylation of the isotopologue of GBT we added, was detected and accumulated over time in both experiments (Supplemental Fig. S5). $^{13}C_3$, $^{15}N_1$-sarcosine, the product of $^{13}C_4$, $^{15}N_1$-dimethylglycine demethylation, also appeared in the second time point of each incubation experiment and reached a maximum of 30% $\pm$ 8% and 78% $\pm$ 4% of the total sarcosine pool in the north and south experiments respectively (Fig. 4). The accumulation of this sarcosine isotopologue is evidence that organisms were using part of the GBT molecule as an energy source, since characterized enzymes producing sarcosine from dimethylglycine, DMGO and DMGDH, result in complete oxidation of the removed methyl group (Sun *et al*., 2011; Lidbury *et al*., 2015a). Future incubations with radiocarbon labelled GBT, such as those conducted by Kiene and Williams, could quantify the proportion of GBT carbon oxidized for energy by open ocean microbial communities (Kiene and Williams, 1998).

In the northern incubation, sarcosine isotopologues did not accumulate to the same extent as the GBT isotopologues despite the general increase in sarcosine concentration in the bottles (Supplementary Figs S6 and S7). This suggests that sarcosine production in the northern incubation was not dominated by the GBT

demethylation pathway. Sarcosine can also be produced by the methylation of glycine, a pathway that has been identified in marine bacteria, phytoplankton and invertebrates (Nyyssölä *et al*., 2000; Lu *et al*., 2006; Kageyama *et al*., 2018; Ngugi *et al*., 2020). This difference of sarcosine production in our two incubations through dominantly a degradative pathway (in the south) and a synthesis pathway (in the north) suggests that the community in the south may be acting in a predominately catabolic mode and an anabolic mode in the north. This difference may be due to the availability of nitrate to support more phytoplankton growth in the northern incubations (Table 1).

Glycine, the product of sarcosine demethylation and a potential precursor to sarcosine synthesis, was measured in the samples but no ^{13}C or ^{15}N isotopologues were detected. Based on our limit of quantification, the concentrations of unlabelled glycine measured in each experiment, and our inability to measure the $^{13}C_1$ glycine isotopologue, we estimate that the isotopologues could have been enriched up to 5% of the unlabelled glycine concentration and remain undetectable.

Transcription of *bhmt*, which encodes the first step of GBT demethylation, was dominated by SAR11, with a smaller contribution from Roseobacteria and other Alphaproteobacteria (Fig. 4). Interpretation of the *bhmt* gene expression should be conservative, as homologues identified in *Pseudomonas aeruginosa*, which also has *gbcAB*, do not participate in GBT metabolism (Wargo *et al*., 2008). Gene expression of *gbcA* homologues were detected at three of the five NPTZ stations sampled in 2016, while gene expression of *gbcB* homologues were detected at all NPTZ stations. Expression of *gbcA* and *gbcB* was dominated by Gammaproteobacteria and Alphaproteobacteria. The demethylation of dimethylglycine to sarcosine via dimethylglycine oxidase (encoded by *dmg*) was dominated by transcripts belonging to Gammaproteobacteria with some SAR11 contribution. The same transformation encoded by dimethylglycine dehydrogenase (*dmgdh*) had transcription in SAR11, SAR116, Gammaproteobacteria and Roseobacters. The third step of the demethylation pathway, encoded by *sardh*, monomeric *sox (msox)*, or tetrameric *soxBDAG*, was transcribed by a variety of both eukaryotic and prokaryotic organisms (Fig. 4). It is possible that sarcosine oxidases catalyse the demethylation of other products in addition to or instead of sarcosine, potentially explaining the two paralogous tetrameric sarcosine oxidase operons and a MSOX in SAR11 (Chlumsky *et al*., 1995; Sun *et al*., 2011). Few organisms transcribed all three steps of the demethylation pathway, but exchange of intermediates may have enabled the downstream transformations caused by a wide diversity of organisms.

Creatine production from sarcosine. The $^{13}C_3$, $^{15}N_1$-creatine isotopologue appeared in the second time point in both experiments (Fig. 4, Supplemental Fig. S6). Creatine is primarily thought to be synthesized from glycine and arginine, via glycine aminotransferase and guanidinoacetate *N*-methyltransferase, and with these enzymes can be produced as a byproduct of the urea cycle (Allen *et al*., 2006; Joncquel-Chevalier Curt *et al*., 2015). We expect the accumulation of the $^{13}C_2$, $^{15}N_1$-creatine isotopologue because those are the atoms derived from glycine. However, we instead observed accumulation of the $^{13}C_3$, $^{15}N_1$-creatine isotopologue (Fig. 4). This indicates that the active production of creatine in these communities was from sarcosine rather than glycine (Fig. 4). These data suggest that either there is an unidentified enzyme that catalyses the reaction from sarcosine to creatine or that creatinase, the enzyme which produces sarcosine from creatine, is reversible (Wyss and Kaddurah-Daouk, 2000).

GBT as a methyl donor. BHMT, one of the enzymes responsible for the first step of the demethylation pathway, generates methionine from homocysteine (Fig. 4) (Lidbury *et al*., 2015a). Other pathways for methionine synthesis are via methionine synthase (MetH), which uses cobalamin (Vitamin B_{12}) and methyltetrahydrofolate as cofactors, and MetE which directly transfers the methyl from methyltetrahydrofolate without the use of cobalamin. Methionine is used to make the universal methyl donor S-adenosyl methionine (SAM) from S-adenosyl homocysteine via methionine adenosyltransferase. Genomes of both prokaryotes and eukaryotes contain *bhmt* homologues (Lidbury *et al*., 2015a; Figueroa-Soto and Valenzuela-Soto, 2018), though no homologues have been functionally verified in eukaryotic phytoplankton. Our analysis only retrieved sequences with homology to experimentally confirmed genes in marine bacteria to avoid including sequences that may encode other types of closely related methyltransferases. Therefore, we only detected *bhmt* transcripts from heterotrophic bacteria, dominated by SAR11 (Fig. 4). SAR11, originally thought to have a strict glycine and reduced sulfur requirement, can use the methyl donor GBT to satisfy their requirement for glycine through complete GBT demethylation. This pathway also lessens their reliance on exogenous methionine and helps explain the lack of both MetE and the Vitamin B_{12} dependent MetH pathway for methionine synthesis in SAR11 (Tripp *et al*., 2008; Sun *et al*., 2011; Carini *et al*., 2013). Our data suggest that SAR11 uses GBT as a source of both glycine and methionine *in situ*.

We did not detect the $^{13}C_1$ natural abundance isotopologues nor label accumulation in methionine or SAM. Instead, we observed the accumulation of ^{13}C in the methyl-acceptors β-alanine betaine and dimethylsulfoniopropionate

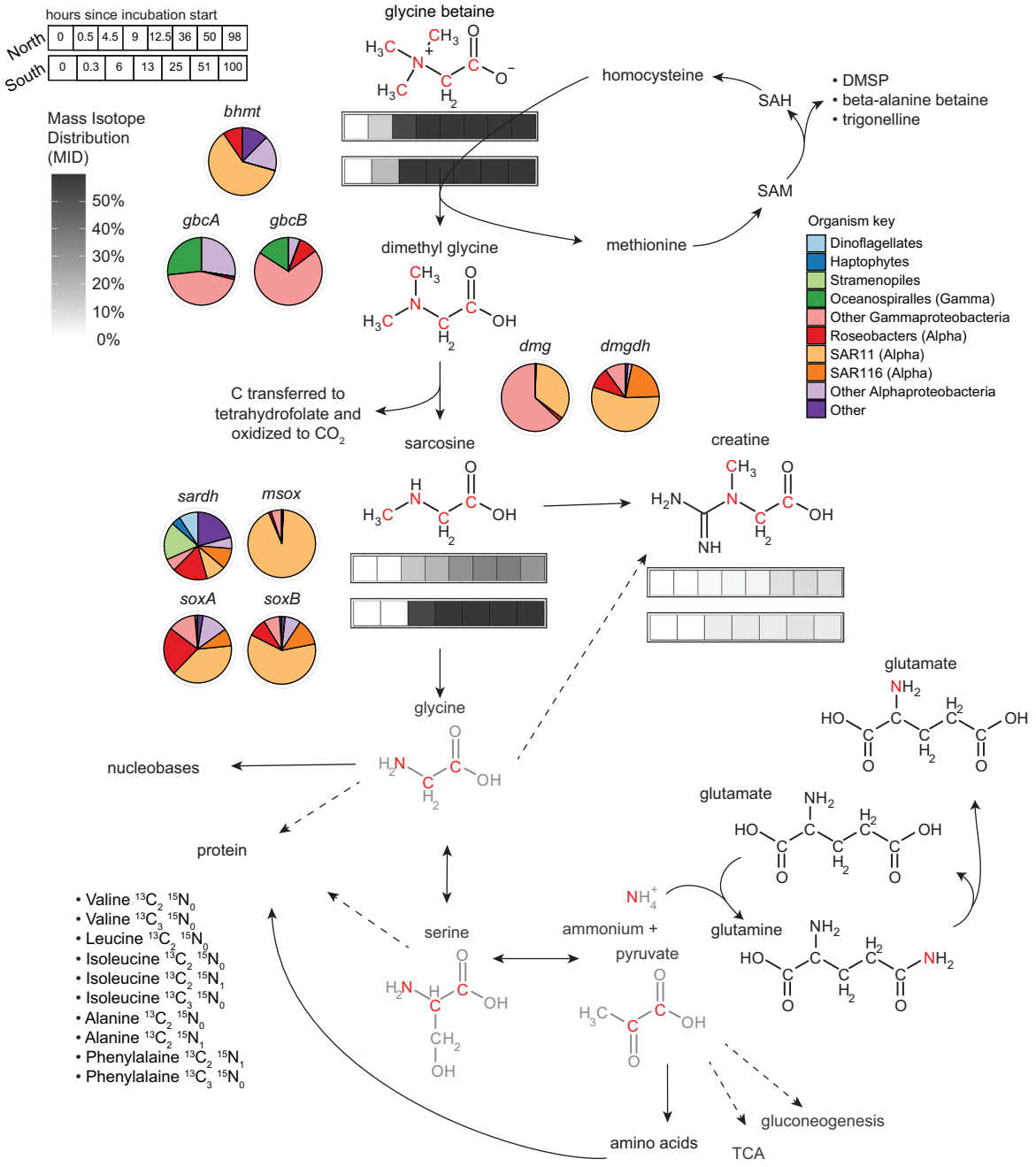

Fig. 4. The GBT demethylation pathway and associated metabolic pathways. Shaded squares represent the MID, which is the proportion of a given metabolite that was in the isotopologue shown, where red atoms in molecular structures indicate ^{13}C or ^{15}N derived from GBT. For each molecule, the top row represents the northern experiment and the bottom row represents the southern experiment, with the initial *in situ* conditions as the first square and the final timepoint as the last square. Molecules in grey were not detected in our data but we infer their presence based on the downstream metabolites. Dashed arrows are known metabolic pathways that we did not see evidence for in our isotope-enriched metabolomics data. The listed amino acids are those which increased in their relative isotopologue abundance in the THAAs in the southern experiment. DMSP, β-alanine betaine and trigonelline are compounds that are methylated through SAM-dependent methyltransferases and had significant increases in their ^{13}C isotopologues. The pie charts show the proportion of organisms transcribing the genes associated with the demethylation pathway in this region in spring 2016: GBT to dimethylglycine (betaine-homocysteine methyltransferase, *bhmt*, or glycine betaine oxidase, *gbcAB*), dimethylglycine to sarcosine (dimethylglycine dehydrogenase, *dmgdh*, or dimethylglycine oxidase, *dmg*), sarcosine to glycine (sarcosine dehydrogenase, *sardh*, monomeric sarcosine oxidase, *msox*, or tetrameric sarcosine oxidase, *soxBDAG* with *soxA* and sox*B* homologues displayed here). Pie charts show an average across the five stations in the NPTZ, with the exception of *gbcA* which was only detected at three stations. 'Other alphaproteobacteria' includes non-Roseobacter Rhodobacterales, Rhizobiales and Sphingomonadales.

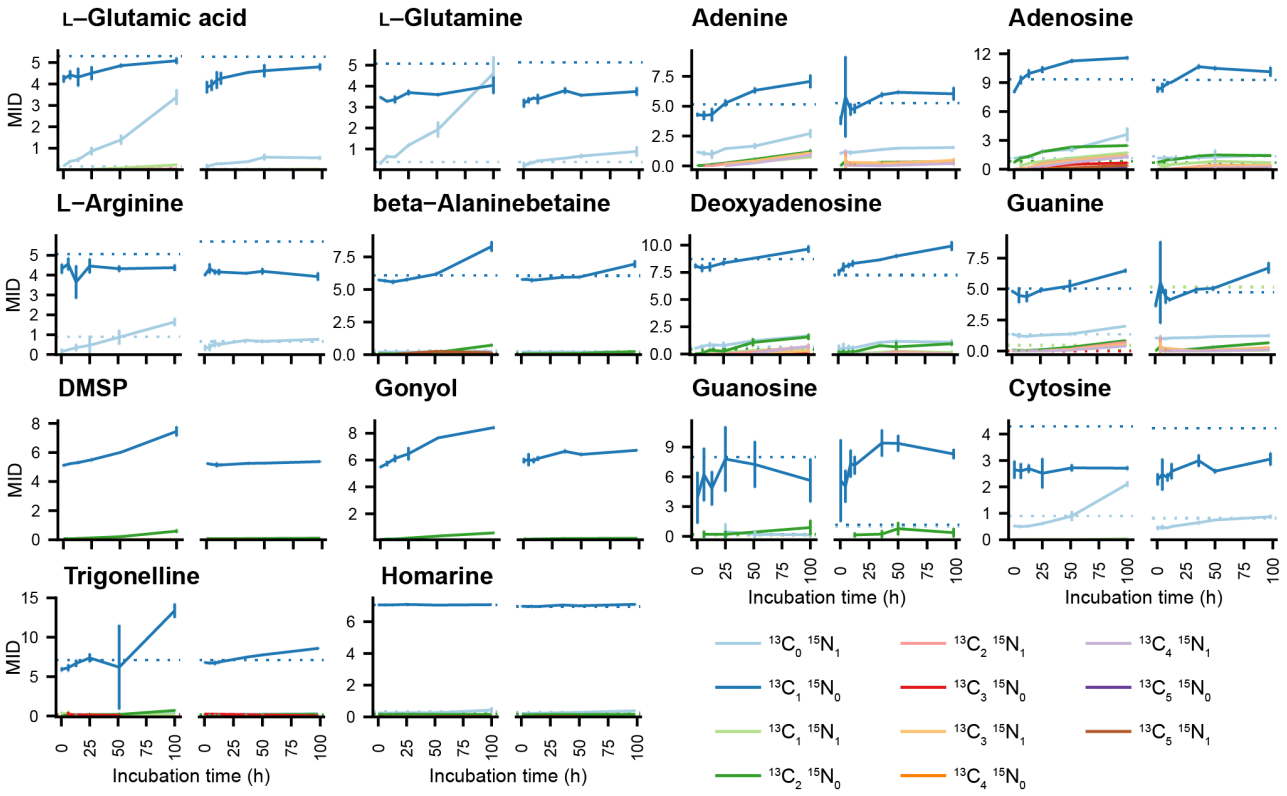

Fig. 5. MIDs (in percent) over time for various isotopologues detected in the metabolites that show evidence of C and N recycling through cellular metabolism. MID is the percent of a given metabolite that was in a particular isotopologue. Isotopologues with only ^{12}C and ^{14}N atoms that make up the majority of the MID are not shown so that the isotopologues with atoms derived from the added $^{13}C_5$, $^{15}N_1$-GBT can be better resolved on the *y*-axis. The *x*-axis is incubation time in hours. For each compound, the northern experiment is on the right and the southern experiment is on the left. Solid lines indicate the mean of triplicate bottles and the error bars represent one standard deviation. Horizontal dashed lines indicate the MID for samples taken at the stations where each experiment was performed.

(DMSP, Figs 4 and 5), providing evidence that GBT fuelled the methionine cycle in these experiments. For example, β-alanine betaine is made through successive methylations of β-alanine via a SAM-dependent methyltransferase (Parthasarathy *et al*., 2019). β-alanine betaine showed changes in ^{13}C isotopologues even though β-alanine did not (Fig. 5, Supplemental Tables S3 and S4), indicating that the ^{13}C was transferred as a methyl group derived from SAM. The methylation of SAM from GBT, as demonstrated by these downstream products, may be an important pathway in marine microbial communities. Vitamin B_{12} (the cofactor for MetH) can be a limiting nutrient for growth in euphotic marine environments (Sañudo-Wilhelmy *et al*., 2014; Bertrand *et al*., 2015; Browning *et al*., 2017). For organisms with BHMT and MetH, storage of GBT could provide a safeguard for methylation potential during temporary Vitamin B_{12} scarcity.

Because only a subset of organisms can use GBT for methylation (via BHMT), ^{13}C accumulation patterns in metabolites synthesized with SAM-dependent methyltransferases can provide insights into which microbes produce them *in situ*. DMSP is synthesized via a SAM-dependent methyltransferase by both phytoplankton and bacteria (Curson *et al*., 2017, 2018), and $^{13}C_1$ and $^{13}C_2$ isotopologues of DMSP accumulated over time, particularly in the southern experiment. Gonyol, potentially produced from DMSP (Trottmann *et al*., 2020), showed similar isotopologue accumulation (Fig. 5). Though DMSP can be made by heterotrophic bacteria that expressed *bhmt* in this region (Fig. 4), gonyol synthesis seems to be restricted to eukaryotic plankton (Nakamura *et al*., 1993; Gebser *et al*., 2020; Heal *et al*., 2021, Durham *et al*., submitted). The appearance of labelled gonyol therefore suggests that either there is a bacterial gonyol source or that some of the unverified eukaryotic *bhmt* homologues we excluded from our gene expression searches are functional and may include dinoflagellates, as they are considered the main gonyol producers (Gebser *et al*., 2020; Heal *et al*., 2021).

Homarine and trigonelline are structural isomers made by different biosynthetic pathways that both involve SAM-dependent methyltransferases (Hall and Gurin, 1975; Ashihara, 2008). The $^{13}C_1$ isotopologue of trigonelline increased in both experiments, with a greater increase in

the south than in the north (Fig. 5). We therefore propose that organisms synthesizing trigonelline in the south are heterotrophic bacteria, the organisms dominating *bhmt* transcription we measured, or possibly the eukaryotes that are likely producing gonyol, as discussed above. In contrast, the homarine $^{13}C_1$ isotopologue showed no significant change over time in the southern experiment (Fig. 5), indicating that organisms synthesizing the high concentrations of homarine in the south are not using BHMT. Given the high concentration of homarine, some homarine producers may use BHMT, but these organisms are not actively producing enough homarine to cause an observable trend in homarine isotopes. Laboratory cultures implicate phytoplankton as the primary homarine producers (Heal *et al*., 2021, Durham *et al*., submitted) and we suggest they may be the dominant organisms producing homarine in these experiments but that those phytoplankton cannot access the methylation potential of GBT through BHMT.

GBT into total hydrolyzable amino acids. Over the course of the experiments, the concentration of total hydrolyzable amino acids (THAAs) increased by ~50% in the northern experiment and decreased by ~10% in the southern experiment (Supplemental Fig. S8). ^{13}C and ^{15}N isotopologues of several hydrolyzed amino acids were detectable, but there were no significant differences in the isotopic distributions at the initial and final timepoints [false discovery rate (FDR) corrected *p*-value >0.1], likely due to high bottle to bottle variability. In the southern experiment there were increases in the proportional contributions of the $^{13}C_2$ and $^{13}C_3$ isotopologues of phenylalanine, isoleucine, leucine, and valine and $^{13}C_2$, $^{15}N_1$ isotopologues of phenylalanine, isoleucine, and alanine that were significant at a lower threshold (*p*-value <0.05 without FDR correction, Fig. 6). Some of these isotopologues were below our limit of detection in the initial time point and increased to be above our limit of detection at the final timepoint, despite an overall decrease in the THAA concentrations.

Of the five amino acids that showed isotope enrichment, four (alanine, valine, isoleucine and leucine) are synthesized from pyruvate. Phenylalanine has three carbons derived from phosphoenolpyruvate, which itself can also originate from pyruvate. These synthesis routes and the enrichment of two or three ^{13}C atoms in these amino acids provide further evidence that GBT is demethylated to glycine, which can then be converted to serine and then to pyruvate and ammonium through the glycine cleavage system and serine dehydratase (Fig. 4). Metatranscriptomic (Fig. 4) and genomic evidence (McParland *et al*., 2021) suggest heterotrophic bacteria were metabolizing GBT through the demethylation pathway to glycine, and thus likely used GBT for protein production.

GBT to choline. Choline is canonically thought of as a precursor of GBT (Kiene, 1998; Lidbury *et al*., 2015a), but our observations suggest this process is reversible. The labelling pattern we observed ($^{13}C_5$, $^{15}N_1$) strongly suggests that choline was formed directly from GBT in our experiments (Fig. 7, Supplemental Fig. S6). Choline and glycerophosphocholine were in higher concentrations in the northern experiment than in the south (Supplemental Fig. S7), and $^{13}C_5$, $^{15}N_1$ isotopes accumulated more rapidly in the north, indicating high flux through this pathway. At the end of the incubation experiment in the north, approximately 25% of the total choline and glycerophosphocholine metabolite pools were in $^{13}C_5$, $^{15}N_1$ isotopologues. The additional accumulation of the glycerophosphocholine $^{13}C_6$, $^{15}N_1$ and $^{13}C_7$, $^{15}N_1$ isotopologues necessitates that the glycerol group was labelled with ^{13}C, supporting further metabolic reworking of GBT-derived carbon.

Glycerophosphocholine is made from choline-phospholipids, a lipid class predominantly produced by eukaryotic organisms in marine plankton communities (Van Mooy and Fredricks, 2010; Popendorf *et al*., 2011). A non-ribosomal peptide synthetase-like enzyme (ATRR) that catalyses the reduction of GBT to choline was recently characterized in fungi (Hai *et al*., 2019). In the metatranscriptome data, transcription of *atrr* gene homologues was dominated by members of Dinophyceae, Stramenopiles, Haptophyta and Amoebozoa (Fig. 7). These transcripts further support GBT use as an intermediate in lipid biosynthesis by eukaryotes, thus far an underappreciated role for GBT in marine environments. Additional research into the prevalence of *atrr* in marine ecosystems will highlight where and when this use of GBT may be important.

GBT to carnitine. We observed the $^{13}C_5$, $^{15}N_1$ isotopologue of carnitine in the second time point in both experiments (Fig. 7), suggesting direct synthesis from GBT. Carnitine had a similar total concentration in each experiment (Supplemental Fig. S7) but higher isotope accumulation in the south (Fig. 7, Supplemental Fig. S6). GBT can be formed from carnitine through carnitine dehydrogenase (CDH) and a carnitine-specific BKACE (Meadows and Wargo, 2015). CDH is an oxidoreductase with NAD^+ as an acceptor, a type of enzyme that has the potential to be reversible under the right chemical conditions (Sellés Vidal *et al*., 2018). It is therefore possible that these enzymes are responsible for the formation of carnitine from the GBT we observe. However, to avoid misinterpretation of the transcriptomics data we did not search for expression of these enzymes that have only been experimentally demonstrated to synthesize GBT.

TMAB [also known as (3-Carboxypropyl) trimethylammonium and γ-butyrobetaine], is a quaternary

amine that can be a precursor to carnitine or a product of carnitine metabolism (Bremer, 1983; Meadows and Wargo, 2015). The stable isotope accumulation patterns in our experiments suggest that TMAB is derived from carnitine (Fig. 7, Supplemental Fig. S6). The $^{13}C_5$, $^{15}N_1$-TMAB isotopologue increased throughout the course of both experiments, from 0% to 6% of the total TMAB pool in the north and from 0% to 32% of the total TMAB pool in the south. The isotope accumulation in carnitine and TMAB highlight the relevance of unexplored pathways for compatible solute metabolism in marine microbial communities.

Glycine betaine to trimethylamine-N-*oxide.* TMAO is a compatible solute and can be an intermediate in the formation of other methylamines and further oxidized to provide energy (Yancey and Siebenaller, 1999; Sun *et al*., 2011; Lidbury *et al*., 2014; Lidbury *et al*., 2015b). The $^{13}C_3$, $^{15}N_1$-TMAO we observe in our incubations could have been formed from three different pathways (Fig. 7, Supplemental Fig. S6). The most direct route is GBT reduction via betaine reductase (GrdH/GrdI) to form TMA (pathway 3 in Figs 1 and 7), a pathway characterized in anaerobes (Jones *et al*., 2019). TMAO is then produced from TMA by TMA monooxygenase, which is found in both prokaryotes and eukaryotes (Yancey and Siebenaller, 1999; Welsh, 2000; Yancey, 2005). Another route of TMAO synthesis is from choline via choline TMA-lyase, CutC, an enzyme in anaerobic organisms (Jameson *et al*., 2016). Lastly, TMA and TMAO can also be formed from carnitine via carnitine oxygenase, an enzyme in aerobic organisms (Jameson *et al*., 2016).

In the metatranscriptome sequence data, reads associated with *grdI* were only detected at a single station at low abundance. At that station (35.28 °N), transcription was dominated by sequences taxonomically assigned to a marine Clostridium (Fendrich *et al*., 1990) which was unlikely to have been dominant in the oxic water column, since the organism was originally described in anoxic sediments. The low number of *grdI* annotations and corresponding transcripts along with the fact that GrdI and CutC are characterized in anaerobes point to carnitine as the most likely source of TMA for TMAO production in this environment.

Carnitine showed significant labelling in both the northern and southern experiments, as discussed above, and the transformation of carnitine to TMA, encoded by *cntA* and *cntB*, has been demonstrated in Acinetobacter (Zhu *et al*., 2014) and *cntA* homologues have been found in open ocean metagenomes (Jameson *et al*., 2016). In searching marine reference genomes and transcriptomes (see Experimental Procedures) for *cntA* and *cntB* homologues, we were unable to identify organisms with high-quality homologues for both enzymes and thus refrain from speculating about which organisms might have been responsible for TMAO production in this study.

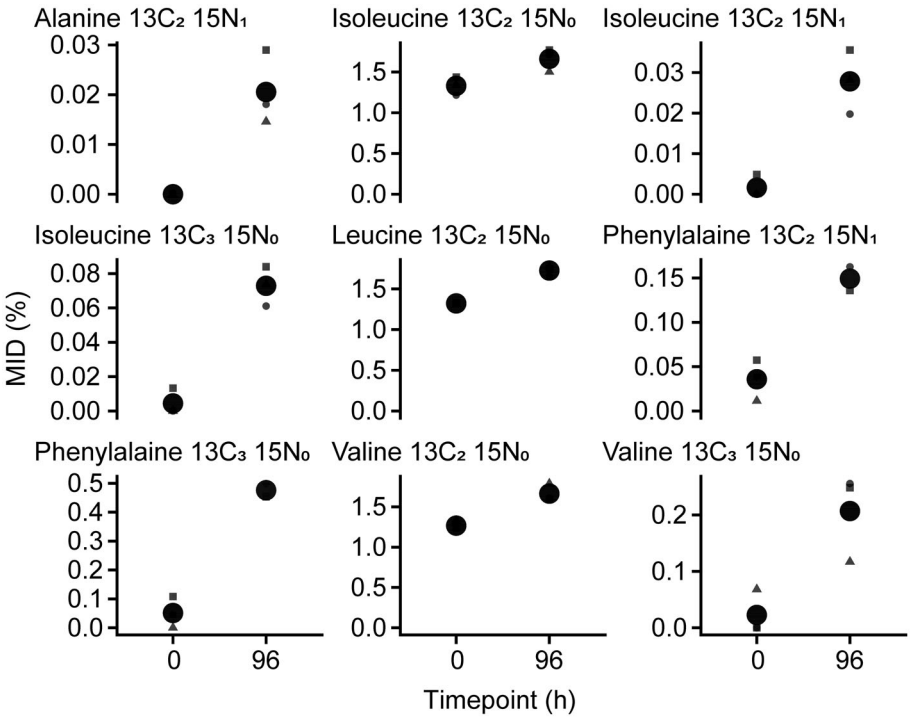

Fig. 6. The isotopologues of individual THAAs with increases in their isotopologue distribution in the southern experiment ($p < 0.05$ without FDR correction). The *x*-axis shows incubation time in hours. Replicate samples ($n = 3$) are in grey and the mean value is in black. No significant increases were observed in the northern experiment.

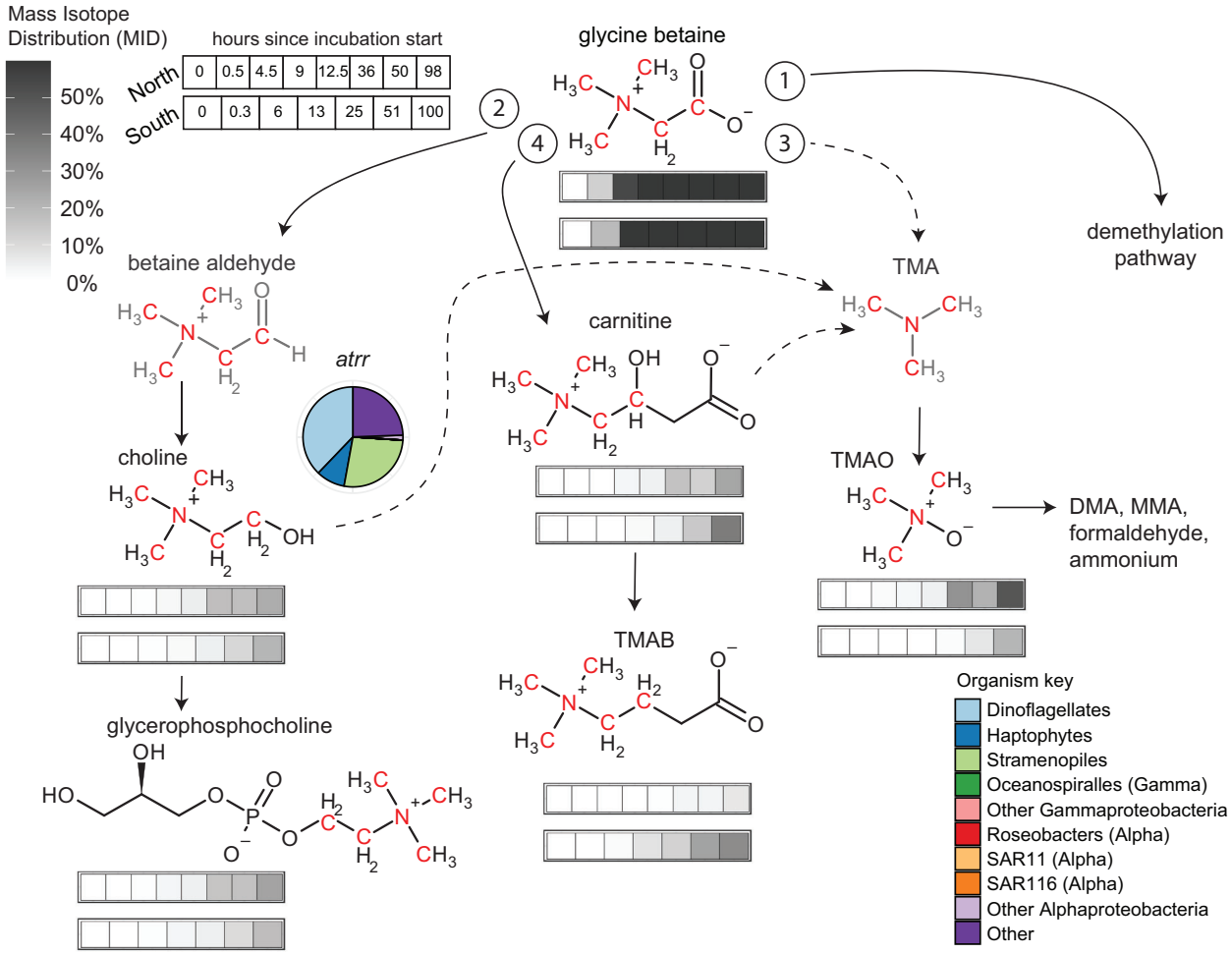

Fig. 7. The transformation of GBT to choline, glycerophosphocholine, TMAO, carnitine and TMAB. Circled numbers refer to pathways as in Fig. 1 and as discussed in the text. Shaded squares represent the MID, which is the proportion of a given metabolite that was in the isotopologue shown, where red atoms in molecular structures indicate ^{13}C or ^{15}N derived from GBT. For each molecule, the top row represents the northern experiment, and the bottom row represents the southern experiment, with the initial *in situ* conditions as the first square and the final timepoint as the last square. Molecules in grey were not detected in our data but we infer their presence based on the downstream metabolites. Dashed arrows indicate the three potential routes of TMA synthesis. The pie charts show the microorganisms transcribing *atrr* in the transition zone in spring 2016. The pie chart for *atrr* is an average of the five stations in the NPTZ.

Biomass production and nitrogen re-mineralization

Increases in the ^{13}C and ^{15}N content of amino acids, nucleosides and nucleobases indicated that GBT-derived carbon and nitrogen were used in macromolecular biosynthesis pathways within cellular metabolism (Fig. 5). Significant increases in isotopologue distributions over time (FDR corrected $p < 0.05$) are discussed below.

GBT nitrogen re-mineralization and reincorporation. The $^{15}N_1$ isotopologues of glutamic acid, glutamine and arginine significantly increased in both experiments (Fig. 5). Glutamine and glutamic acid are the molecules involved in ammonium assimilation, and the appearance of these isotopologues suggests that ammonium was generated from GBT. Other metabolites, including trigonelline, homarine, taurine, proline and alanine, showed ^{15}N incorporation in one or both experiments (Fig. 5, Supplemental Tables S3 and S5). The enhanced ^{15}N isotopologue accumulation trends in many metabolites in the southern station reiterate that, when DIN concentrations are low, organisms are reliant on recycling of organic nitrogen within the photic zone. The small increase in ^{15}N isotopologues in both trigonelline and homarine underscores that nitrogen from GBT made its way through the microbial community, including organisms that may not be able to degrade GBT themselves, as has been observed in co-culture experiments (Lidbury *et al*., 2015a; Zecher *et al*., 2020).

Purine and pyrimidine metabolism. The purines adenine and guanine, and related molecules adenosine,

deoxyadenosine and guanosine, showed enrichment in many isotopologues in both the northern and southern experiments, including $^{15}N_1$ alone, $^{13}C_1$ alone and $^{13}C_2$, $^{15}N_1$ (Fig. 5, Supplemental Table S3). Purines are formed *de novo* from two molecules of formate, glycine, two molecules of glutamine, HCO_3^-, and aspartate. The relatively high accumulation of the $^{13}C_2$, $^{15}N_1$ isotopologues of adenine, adenosine and guanine to 1%, 1.4% and 0.7% in the final time point of the south experiment suggests isotope accumulation from glycine, which contributes two carbons and one nitrogen atom to purines (Fig. 5, Supplemental Table S3, Supplemental Fig. S9). Similar accumulation of the $^{13}C_3$, $^{15}N_1$ isotopologues suggest significant labelling of tetrahydrofolate within these cells, possibly derived from the second two demethylation steps in the demethylation pathway, in which methyl groups are transferred to tetrahydrofolate and potentially ultimately oxidized to CO_2 (Sun *et al*., 2011). These pathways for isotope accumulation in the purines suggest that, in organisms that take up and demethylate GBT, there is significant recycling of carbon and nitrogen from GBT throughout primary metabolites, including possibly formate and CO_2.

In contrast to purines, pyrimidines like cytosine are formed *de novo* from glutamine, HCO_3^- and aspartate, without contribution from glycine. We did not observe significant accumulation of cytosine isotopologues with multiple ^{13}C and ^{15}N atoms, but the $^{15}N_1$ isotopologue of cytosine accumulated more in the south compared to the north (Fig. 5), again emphasizing the relative importance of GBT-derived ammonia being re-assimilated for biomass production in low nitrate waters.

Untargeted metabolites with isotope incorporation

Using untargeted metabolomics, we were able to identify several other metabolites that demonstrate the varied fates of GBT-derived carbon and nitrogen. These metabolites are promising targets for future studies into marine microbial metabolism. By searching the high-resolution metabolomics data for evidence of ^{13}C and ^{15}N incorporation in all mass features (unique combinations of mass-to-charge ratio and RT), we found 74 and 82 high-quality mass features with significant isotope incorporation in the north and south respectively, out of about 2000 initial mass features (see Supplemental Methods). Forty-six of these mass features were shared between the experiments (Supplemental Table S5). In the south, 41% (34/82) of mass features with isotope incorporation were known metabolites, including those discussed above. In the north, 28% (21/74) of mass features with isotope incorporation were known metabolites. The remaining have identities that are either unknown or tentatively identified without an authentic standard for confirmation. Regardless of identification, all mass features represent metabolites that must be direct or indirect products of GBT metabolism. For example, a mass feature in the HILIC-positive mode with an mass-to-charge ratio (*m/z*) of 184.0734 and RT of 12.8 min had a similar increasing $^{13}C_5$, $^{15}N_1$ isotopologue trend as choline and glycerophosphocholine (N-FT0422/S-FT391 in Supplemental Fig. S9, Supplemental Table S2). Based on the exact mass and MS^2 spectra, we identified this metabolite as phosphocholine, which we had not previously investigated in our analysis.

Select unknown metabolites had very similar isotope-accumulation trends as known compounds. For example, a mass feature in the HILIC-positive mode with a *m/z* of 175.1077 and a RT of 11.15 min had a similar ^{13}C isotopologue decrease, $^{13}C_5$ isotopologue increase and $^{13}C_5$, $^{15}N_1$ isotopologue increase as GBT itself (N-FT0365/S-FT328 in Supplemental Fig. S9, Supplemental Table S3). These isotopologue increases and the exact mass suggest this metabolite is GBT-Gly, where a peptide bond has formed between the carboxylic acid of GBT and the amine of glycine. Reports of fungal synthesis of GBT-Ser (Singh *et al*., 2021) support the possibility of a biologically produced GBT-Gly dipeptide.

We tentatively identified several untargeted metabolites as likely mycosporine-like amino acids (MAAs), using exact masses and common MS^2 fragments (as identified in Parailloux *et al*., 2020). Six potential MAAs with isotopologue accumulation in the northern experiment are userjirine (*m/z* 285.144, N-FT0769), mycosporine-glycine (*m/z* 246.09, N-FT0621/S-FT600), palythenic acid (*m/z* 329.134, N-FT0887/S-FT807), porphyra-334 (*m/z* 347.144, N-FT0930/S-FT830), palythine (*m/z* 245.113, N-FT0614/S-FT596) and shinorine (*m/z* 333.129, N-FT0899/S-FT814). All but userjirine also had significant isotopologue accumulation in the southern experiment, with enrichment in multiple isotopologues, including ^{13}C and $^{13}C_2$, $^{15}N_1$ (Supplemental Table S2). This latter isotopologue showed similar accumulation patterns to adenine and guanine and is indicative of the incorporation of a GBT-derived glycine (Supplemental Fig. S9). MAAs are primarily thought to be photoprotective and antioxidant compounds produced by algae, fungi and cyanobacteria (Shick and Dunlap, 2002; Llewellyn and Airs, 2010). Though the abundance of MAAs was generally higher in the northern experiment compared to the southern experiment, the $^{13}C_2$, ^{15}N isotopologue accumulated more in the southern experiment. In marine phytoplankton cultures, molecular features with *m/z* and RTs matching the tentatively identified MAAs here were produced dominantly by dinoflagellates (Heal *et al*., 2021). We did not search for eukaryotic *bhmt* homologues, as discussed above, and therefore it is possible eukaryotic organisms such as dinoflagellates could be directly

synthesizing MAAs from GBT. However, if the demethylation of GBT is primarily a heterotrophic bacterial process and MAA synthesis is primarily done by primary producers, then the labelled glycine incorporation into MAAs is indicative of active metabolite exchange. We speculate that dimethylglycine, sarcosine and glycine may be excreted by SAR11 or other bacteria carrying out the GBT demethylation pathway so that they can use the methyl groups in GBT for methylation and energy but maintain homeostasis, since SAR11 need specific ratios of glycine to pyruvate to maintain optimal growth rates (Carini *et al*., 2013).

Future studies should focus on validating the identities of the molecules tentatively identified here, exploring their cellular roles, and confirming metabolic interactions that they might facilitate. Overall, these results show the promise of using isotope-tracing high-resolution metabolomics to identify metabolites of biological relevance and their biosynthetic origins. We expect that subsequent studies using these techniques with different substrates or microbial communities will lead to more discoveries of ecologically relevant molecules, pathways, and interactions.

Conclusion

Using stable isotope tracing metabolomics, we find that microbial communities in the NPTZ can take up GBT and transform it through both known and previously unrecognized pathways, to ultimately provide methylation capacity, carbon, nitrogen and energy to the community. The uptake kinetics parameters were within the ranges estimated in previous studies of GBT uptake in natural seawater and cultured organisms. The turnover time estimates for dissolved GBT are similar to the turnover time estimates of non-proteinogenic dissolved free amino acids and amino acid-like compounds. The vast majority of the ^{15}N and ^{13}C that we could detect remained in untransformed $^{13}C_5$, $^{15}N_1$-glycine betaine. This luxury uptake of DON may have unexplored consequences for how organisms adapt to and survive chronically nutrient-limited biomes, including nitrogen and cobalamin limitation. We also show that a portion of the GBT is transformed to other metabolites on timescales of hours to days. The difference in uptake and fate of GBT between the two stations emphasize that the DIN availability and microbial community composition influence how DON is used. Though the dominant pathways of GBT metabolism differed, at both sites the carbon and nitrogen from GBT appeared in known and unknown metabolic products including amino acids, nucleobases, lipids and MAAs, suggesting both partial and full recycling through demethylation, pyruvate, the citric acid cycle and ammonium assimilation.

Experimental procedures

Sampling strategy

Samples for GBT experiments were collected aboard the R/V *Kilo Moana* in April 2019. Experiments were conducted at two stations: station 4 at 41° 40.85′ N and 158° 3.01′ W, and station 5 at 37° 0.21′ N and 158° 0.20′ W (Table 1). Water for samples of the *in situ* conditions and for the incubation experiments were collected with Niskin bottles attached to the CTD from 15 m water depth, which was within the surface mixed layer, in the morning. DON, DIN, particulate carbon and particulate nitrogen samples were analysed according to the standard Hawaiian Ocean Time Series program protocols. Temperature and chlorophyll-a concentrations are from the CTD.

GBT uptake kinetics experiments

To determine the kinetics of GBT uptake, whole seawater was spiked with varying concentrations of $^{13}C_5$, $^{15}N_1$-GBT and incubated for 25–42 min. Water was collected into 2 L bottles around 8:00 am local time for both experiments. To minimize the biological transformation of $^{13}C_5$, $^{15}N_1$-GBT into other molecules and to limit induction of enzymatic activity, the incubation time with $^{13}C_5$, $^{15}N_1$-GBT was kept as short as possible given our sampling set up. In order to achieve this, samples were spiked throughout the course of the day since only six samples could be processed at a time. Before and after the addition of $^{13}C_5$, $^{15}N_1$-GBT, bottles were kept in flow-through incubators with blue shading to be at *in situ* temperature and approximately mixed layer light conditions. Samples were spiked to have final concentrations of 0, 2, 5, 10, 50, 200, or 2000 nM $^{13}C_5$, $^{15}N_1$-GBT. Triplicates of each $^{13}C_5$, $^{15}N_1$-GBT concentration were processed for both experiments, with replicates spread throughout the sampling period. After incubation with the spiked molecule, seawater was filtered onto 47 mm diameter, 0.2 μm pore size PTFE (Omnipore) filters using a peristaltic pump, polycarbonate filter holder and Masterflex PharMed BPT tubing (Cole-Parmer). Filtering time was 10–42 min with an average time of 22 min. Experimental blanks were collected for each spike concentration during the northern experiment by collecting filtrate from one replicate of each treatment and re-filtering filtrate onto a new filter. This provided a measure of dissolved organic compounds adsorbed onto the filter during processing. Filters were frozen in liquid nitrogen immediately after filtration and stored at −80°C until the filters were extracted using the metabolite extraction method described in Boysen *et al*. (2018).

GBT and $^{13}C_5$, $^{15}N_1$-GBT were measured using paired liquid chromatography–mass spectrometry (LC–MS) using a Waters Acquity I-Class UPLC (Waters

Corporation, Milford, MA, USA) with a hydrophilic interaction liquid chromatography column (HILIC SeQuant ZIC-pHILIC column, 5 mm particle size, 2.1 mm × 150 mm, from Millipore) paired to a Waters Xevo TQ-S triple quadrupole. Electrospray ionization in selected reaction monitoring mode with polarity switching was used to monitor for the intensity of ions from $^{12}C_5$, $^{14}N_1$-GBT and $^{13}C_5$, $^{15}N_1$-GBT. Standard addition curves were used to quantify GBT as detailed in the supplementary material. Samples were always well above the concentration in the blank, with an average of 63 times greater concentration in the sample than the blank across all treatments. Thus, blank measurements were disregarded for the remainder of the uptake kinetics calculations.

Rates of $^{13}C_5$, $^{15}N_1$-GBT uptake were calculated by fitting the data to a Michaelis–Menten equation with a non-linear least-squares method and by using a Wright-Hobbie transformation. Details are in the Supplemental Methods. All methods resulted in similar estimates with overlapping uncertainties (Supplemental Table S1). In the results and discussion, parameters from the nonlinear least-squares estimates are used. We note that the uptake rates we calculate may be underestimates if during the incubations the cells became saturated with GBT and slow their uptake rate (Button, 1998).

GBT fate experiments

Two time-course incubation experiments were performed, one at each of the stations described in Table 1. For each experiment, 2 L bottles were filled with seawater collected with the CTD from 15 m depth (within the mixed layer) at approximately 6:00 am local time. All 2 L bottles were spiked with 500 nM $^{13}C_5$, $^{15}N_1$-GBT and incubated in on-deck temperature-controlled incubators at 10°C and 14°C for the north and south experiments respectively. The incubators used blue shading to approximate mixed layer light conditions. Samples for the initial timepoint (T0) were filtered directly after being spiked, resulting in actual incubation times of approximately 20 min. At each time point, triplicate 2 L bottles were sampled for analysis of bacterial and picophytoplankton abundance and biomass via flow cytometry, metabolites and THAAs. Timepoints sampled in the north experiment were 0, 4.5, 9, 12.5, 36, 50 and 98 h. Timepoints sampled in the south experiment were 0, 6, 13, 25, 51 and 100 h. For all samples, incubation time was calculated by taking the difference between the time the bottle was spiked and the midpoint time of sample filtration. In the main text, we report values from these experiments as the mean ± standard deviation of the triplicate bottles.

Flow cytometry samples were taken by aliquoting 2 ml into a cryovial, spiking with glutaraldehyde, incubating in the dark for 20 min, and flash freezing in liquid nitrogen. Samples were stored at −80°C until analysis on an influx flow cytometer. Flow cytometry samples were analysed on a BD Influx Cell Sorter, gating heterotrophic bacteria, *Prochlorococcus*, *Synechococcus* and other phytoplankton (labelled as picoeukaryotes). The light scattering values were converted to cell size and carbon content using Mie theory (more information available at https://github.com/seaflow-uw/fsc-size-calibration and https://github.com/seaflow-uw/fsc-poc-calibration).

Metabolomics samples were collected as described in the kinetics experiment, total volume of 2 L. THAA samples were collected by filtering 200 ml onto glass fibre filters (GF75 grade, 47 mm) using vacuum filtration. Filters were flash-frozen in liquid nitrogen and stored at −80°C until analysis. Experimental blanks for metabolomics samples and THAAs were collected during the northern experiment by re-filtering filtrate from the samples collected at 36 h.

Metabolomics sample processing and data acquisition

Metabolite extractions were performed as in Boysen *et al*. (2018), using a modified Bligh–Dyer extraction that yielded an aqueous and organic fraction. Aqueous soluble metabolites were analysed with two injections for each sample, one with a spike of heavy isotope labelled internal standards and one with an equal volume spike of UPLC grade water. Samples were run using both HILIC and reversed-phase (RP) chromatography on a Waters Acquity I-Class UPLC (Waters Corporation) with injection volumes of 2 and 15 µl respectively, coupled to a Q Exactive Orbitrap MS (Thermo Scientific). Details of the chromatography and mass spectrometry methods are in Boysen *et al*. (2018) and in the Supplemental Methods.

Targeted metabolite data processing

Skyline-daily (Adams *et al*., 2020; Pino *et al*., 2020) was used to integrate the peak areas of the metabolites for which we have authentic standards (Supplemental Table S4), along with the potential isotopologues of those metabolites. Metabolites or their isotopologues were removed from further analysis if they had comparable peak area in the blanks (sample peak area less than two times the blank peak area) or if they were more than 10 ppm mass deviation from expected. MIDs are the percent contribution each individual isotopologue had to the total concentration of all isotopologues for a given metabolite. MIDs were calculated for each metabolite using only the samples that did not receive spikes of isotope labelled internal standards to ensure that all isotopologues were from GBT transformation or ^{15}N and ^{13}C natural abundance and not due to our added internal standards. The significance of MID changes during the

incubations was assessed using a linear regression of MID over time. Samples from the same stations as the incubation experiments but which were not part of the GBT experiments were analysed to determine the natural isotope abundance.

Using the samples that were not spiked with isotope labelled internal standards, the unlabelled metabolites were normalized using best-matched internal standard normalization (Boysen *et al*., 2018). If the coefficient of variation for a given metabolite was less than 0.1 with no normalization or if the internal standards did not reduce the coefficient of variation by more than 0.1 then no internal standard normalization was used for that metabolite. Metabolites with internal standards spiked into the samples prior to analysis were quantified directly (L-alanine, L-histidine, L-proline, L-valine, L-isoleucine, L-methionine, L-phenylalanine, L-tryptophan, adenine, cytosine, guanine, arsenobetaine, isethionic acid, sucrose and trehalose). Except for GBT, other metabolites were quantified using the response factors and response factor ratios of authentic standards spiked into water and into representative pooled samples, as described in Boysen *et al*. (2021).

To calculate the absolute concentration of GBT in the samples, a sample from the cruise that did not come into contact with labelled GBT was used to calculate the response factor in the relevant sample matrix that correlates the peak area of GBT to its known concentration. $^{13}C_5$, $^{15}N_1$-GBT was spiked into this sample at a final concentration of 500 nM just before analysis. This sample was analysed five times over the course of the sample run and thus provided a robust estimate of the response factor and its variability (coefficient of variation = 0.04). This value was used to determine the concentration of all isotopologues of GBT in each vial, which was then converted into the concentration of GBT in each sample.

Untargeted metabolite data processing

Untargeted metabolite data processing was performed as follows for each of the LC–MS analyses performed on aqueous metabolites (HILIC positive mode, HILIC negative mode, aqueous phase RP positive mode). Samples were processed in R [version 4.0.3, (R Core Team, 2013)] with the xcms package [version 3.10.2, (Smith *et al*., 2006)] using parameters in Supplemental Table S6. Custom signal-to-noise ratios were calculated for each peak, as the default implementation has known issues (Myers *et al*., 2017), and were used to filter out likely noise peaks. Briefly, after the initial peak picking step in xcms each peak was fit to an ideal bell curve with varying degrees of skew. For each skew, the standard deviation of the difference between the ideal curve and the real data was calculated, producing a standard deviation of the residuals as a proxy for noise level and peak similarity. The degree of skew with the minimum residual standard deviation was used as a comparison peak. The Pearson correlation between this comparison peak and the real data was combined with the height of the peak and its residual standard deviation to produce a robust proxy for peak quality. This peak quality score was set arbitrarily at a high threshold that produced an acceptable false-discovery rate of noise peaks, and peaks falling below this threshold were removed from the analysis. Default xcms functions were then used to align, correspond and fill in missing peaks. Custom code was used to identify likely ^{13}C and ^{15}N isotopologues based on calculated *m/z* match and chromatogram similarity (code for all the above is available at https://github.com/IngallsLabUW/GBT_uptake_and_use).

Details of mass feature curation and dereplication are in the Supplemental Methods. Metabolites were identified when possible, using an in-house database of authentic standards and MS^2 data for manual comparison against the Metlin database (Smith *et al*., 2005).

For visualization, the MID values for each metabolite isotopologue were then normalized to have a mean of zero and standard deviation of one. Clustering analyses were performed using R's base 'dist' and 'hclust' functions, and the heatmaply package was used to generate interactive heatmaps for ongoing analysis (Galili *et al*., 2018). In both cases, the Manhattan distance metric was used due to its better performance with high-dimensional data.

Total hydrolyzable amino acid analysis and data processing

Amino acids were hydrolyzed as in Fountoulakis and Lahm (1998) with some modifications specified in the Supplemental Methods. Amino acids were derivatized as in Gray *et al*. (2017) using the AccQ Tag Ultra derivatization kit from Waters (Gray *et al*., 2017). A 10 μl aliquot of each sample was transferred into a new glass LC–MS vial and 70 μl of borate buffer was added to each vial. After vortexing, 20 μl of AccQ Tag Ultra derivatizing agent was added to each sample and the vials were heated at 55°C for 10 min. Blank filters that had been collected by filtering 0.2 μM filtrate through the sampling apparatus were collected during the experiment and were analysed alongside the samples.

For analysis of derivatized amino acids, a Waters Acquity UPLC HSS T3 column (1.8 μm particle size, 2.1 mm × 150 mm) was used with 0.1% formic acid in water (Solvent A) and 0.1% formic acid in acetonitrile (Solvent B) at a flow rate of 0.6 ml min^{-1}. LCMS parameters are in the Supplemental Methods.

Skyline was used to integrate peak areas of each derivatized amino acid and the isotopologues of each amino acid with 0 or 1 ^{15}N and 0–5 ^{13}C (Adams *et al*., 2020; Pino *et al*., 2020). The average peak areas of amino acids in the blank samples were subtracted from the sample peak areas. Amino acids that had peak areas in the samples that were smaller than the blank average were discarded and replaced with 0. To quantify amino acids, a standard addition curve was made using a mix of unlabelled amino acid standards. The standard addition curve was run in one representative sample by spiking in amino acid standards for a final concentration of 0.5, 1 and 2.5 μM after hydrolysis and before derivatization. The standard curve samples were then derivatized and analysed alongside the samples. After quality control, MIDs were calculated for the THAAs. A *t*-test was used to determine if there was a significant difference in MID between the initial and final timepoints.

Metatranscriptome sample processing

Metatranscriptome samples were collected from a latitudinal transect (23.5–37.3 °N) at approximately 158 °W in the North Pacific during the R/V *Ka'imikai-O-Kanaloa* Gradients 1 cruise KOK1606 (April/May 2016). These data are used in the main text and figures. To show that springtime gene expression in the NPTZ is relatively consistent from year to year, we also analysed metatranscriptome samples collected from the NPTZ in May and June 2017 aboard the *R/V M.G. Langseth* Gradients 2 cruise MGL1704 (Supplemental Fig. S4). Sample processing and bioinformatics were nearly identical for the 2 years, with the exception that a 200 μm mesh prefilter was used for the 2016 samples and a 100 μm mesh for the 2017 samples. The data are presented as an average across the stations within the NPTZ, as defined by Gradoville *et al*. (2020) by using the second derivative of surface salinity as north of 30.7 °N and 32.8 °N in 2016 and 2017 respectively. Environmental conditions for the 2016 and 2017 cruises are published in Gradoville *et al*. (2020). Plankton samples of ∼10 L of seawater collected in the mixed layer (typically from 15 m) were passed through a 200 μm (2016) or 100 μm (2017) nitex prefilter followed by sequential collection onto a 3 μm followed by a 0.2 μm polycarbonate filter using a peristaltic pump. This collection scheme yielded two size classes (0.2–3 and 3–200 μm) for metatranscriptome samples. Filters were flash frozen in liquid nitrogen and subsequently stored at −80°C until further processing. To generate quantitative transcript inventories, a set of 14 internal mRNA standards were added at known concentrations at the start of RNA extraction. Total RNA was extracted using the ToTALLY RNA kit (Invitrogen) as described previously (Durham *et al*., 2019). Briefly, RNA extracts were treated with DNase, and rRNAs were removed using Illumina's Bacteria and Yeast Ribo-Zero rRNA Removal Kits. rRNA-depleted RNA was cleaned using the Zymo RNA Clean and Concentrator Kit. Purified, depleted RNA was then sheared to ∼225 bp fragments and used to construct TruSeq cDNA libraries according to the Illumina TruSeq®RNA Sample Preparation v2 Guide for paired-end (2 × 150) sequencing using the Illumina NovaSeq 6000 sequencing platform.

Bioinformatics

Raw Illumina data were quality controlled with trimmomatic v0.36 (Bolger *et al*., 2014) using the parameters MAXINFO:135:0.5, LEADING:3, TRAILING:3, MINLEN:60 and AVGQUAL:20, and matching read pairs were merged using flash v1.2.11 (parameters: *-r 150 -f 250 -s 25*) (Magoč and Salzberg, 2011). FASTQ files were converted to FASTA and translated with seqret and transeq vEMBOSS:6.6.0.0 (Rice *et al*., 2000) using standard genetic code. Potential coding frames with ≥40 uninterrupted amino acids were retained for further analysis.

We searched for translated reads corresponding to proteins that carry out demethylation of GBT to dimethylglycine [betaine-homocysteine methyltransferase, BHMT and GBT oxidase, referred to here as GbcAB but sometimes referred to as BmoAB (Shao *et al*., 2018)], the demethylation of dimethylglycine to sarcosine (DMDGH; DMGO), the demethylation from sarcosine to glycine (SARDH, MSOX, and two components of the tetrameric sarcosine oxidase, SoxBDAG), the reduction of GBT to TMA (betaine reductase, GrdI), the reduction to betaine aldehyde and choline (nonribosomal peptide synthetase-like GBT reductase, ATRR), and the transformation of carnitine to TMA (carnitine oxygenase CntAB). Full-length, experimentally verified protein sequences (locus tags given in Supplemental Table S7) from reference genomes were used as queries to identify putative orthologues in translated protein sequences from marine prokaryotic and eukaryotic genomes and transcriptome assemblies complied in a custom marine reference database (Coesel *et al*., 2021) using BLASTp v2.2.31. Genomes and transcriptomes used in this database were derived from Joint Genome Institute (JGI), National Center for Biotechnology Information (NCBI) and the Marine Microbial Eukaryote Transcriptome Sequence Project (MMETSP) (Keeling *et al*., 2014). A complete list is available in Coesel *et al*. (2021). Marine sequences retrieved from the BLASTp search were clustered using usearch (Edgar, 2010) and aligned using MAFFT v7 with the E-ISN-I algorithm (Katoh and Standley, 2013). The alignment was trimmed using trimAl v1.2 using *–gt .05 –*

resoverlap 0.5 –seqoverlap 50 options (Capella-Gutiérrez *et al*., 2009). The trimmed alignment file was converted to Phylip format, and the best-fit amino acid substitution matrix, among-site rate heterogeneity model, and observed amino acid frequency were determined using ProtTest 3 software (Darriba *et al*., 2011). A maximum-likelihood phylogenetic reference tree was built using RAxML v8 (Stamatakis *et al*., 2008), and only those sequences that clustered with experimentally verified enzymes were considered putative homologues. These phylogenetic reference trees served as scaffolds to recruit translated environmental metatranscriptomic reads. An hmm profile was constructed from each reference alignment using hmmbuild, followed by transcript identification and alignment to the reference using hmmsearch (parameters: *-T 40 –incT 40*) and hmmalign respectively, with the HMMER package v2.1 (Eddy, 2011). NCBI taxonomy was assigned to each environmental sequence using pplacer v1.1. alpha19-0-g807f6f3 (Matsen *et al*., 2010) based on the read placement with the best maximum likelihood score to the reference tree (parameters: *–keep-at-most 1 –max-pend 0.7*). Read counts were normalized by recovery of internal mRNA standards to estimate transcript abundance. Size-fractions were summed together for a given biological replicate to yield total transcripts per liter. The NPTZ stations from the 2016 cruise do not match the exact latitudes of those used in the 2019 GBT addition experiments, thus we tried to avoid overinterpretation of detailed latitudinal trends by averaging transcript inventories across the five NPTZ stations, defined as any station with a latitude north of 30.7 °N for the 2016 cruise by Gradoville *et al*. (2020).

Acknowledgements

The authors acknowledge A. Meyers for lab assistance; N. Kellogg, J. Sacks, R. Lionheart for feedback and assistance on data analysis and interpretation; A. Torstensson for input on the experimental design; Z. Bartolek for assistance with making the metatranscriptome data publicly available; K. Cain for the flow cytometry data; the crew and scientific party of the R/V *Kilo Moana* during the Gradients 3 cruise (KM1906); the crew and scientific party of the R/V *Ka'imikai-O-Kanaloa* during the Gradients 1 cruise (KOK1606); the crew and scientific party of the *R/V M.G. Langseth* during the Gradients 2 cruise (MGL1704). This work was supported by grants from the Simons Foundation (LS award ID: 385428, A.E.I.; SCOPE award ID 329108, A.E.I.; SCOPE award ID: 426570SP, E.V.A.; award ID 823165 to B.P.D.; award ID 598819 to K.R.H.); and the National Science Foundation (NSF OCE-1228770 and OCE-1205232 to A.E.I., NSF GRFP to A.K.B.)

Data Availability

Metabolomics data are available through Metabolomics Workbench under Project PR001273. Metatranscriptome data from the 2016 cruise are available through NCBI's SRA under BioProject PRJNA492143. Metatranscriptome data from the 2017 cruise are available through NCBI's SRA under BioProject PRJNA816919.

References

Adams, K.J., Pratt, B., Bose, N., Dubois, L.G., St. John-Williams, L., Perrott, K.M., *et al*. (2020) Skyline for small molecules: a unifying software package for quantitative metabolomics. *J Proteome Res* **19**: 1447–1458.

Airs, R.L., and Archer, S.D. (2010) Analysis of glycine betaine and choline in seawater particulates by liquid chromatography/electrospray ionaization/mass spectrometry. *Limnol Oceanogr Methods* **8**: 499–506.

Allen, A.E., Vardi, A., and Bowler, C. (2006) An ecological and evolutionary context for integrated nitrogen metabolism and related signaling pathways in marine diatoms. *Curr Opin Plant Biol* **9**: 264–273.

Ashihara, H. (2008) Trigonelline (N-methylnicotinic acid) biosynthesis and its biological role in plants. *Nat Prod Commun* **3**: 1423–1428.

Ayers, J.M., and Lozier, M.S. (2010) Physical controls on the seasonal migration of the North Pacific transition zone chlorophyll front. *J Geophys Res Ocean* **115**: C05001.

Barra, L., Fontenelle, C., Ermel, G., Trautwetter, A., Walker, G.C., and Blanco, C. (2006) Interrelations between glycine betaine catabolism and methionine biosynthesis in Sinorhizobium meliloti strain 102F34. *J Bacteriol* **188**: 7195–7204.

Berman, T., and Bronk, D.A. (2003) Dissolved organic nitrogen: a dynamic participant in aquatic ecosystems. *Aquat Microb Ecol* **31**: 279–305.

Bertrand, E.M., McCrow, J.P., Moustafa, A., Zheng, H., McQuaid, J.B., Delmont, T.O., *et al*. (2015) Phytoplankton-bacterial interactions mediate micronutrient colimitation at the coastal Antarctic sea ice edge. *Proc Natl Acad Sci U S A* **112**: 9938–9943.

Bolger, A.M., Lohse, M., and Usadel, B. (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**: 2114–2120.

Boysen, A.K., Carlson, L.T., Durham, B.P., Groussman, R. D., Aylward, F.O., Ribalet, F., *et al*. (2021) Particulate metabolites and transcripts reflect diel oscillations of microbial activity in the surface ocean. *mSystems* **6**: e00896-20.

Boysen, A.K., Heal, K.R., Carlson, L.T., and Ingalls, A.E. (2018) Best-matched internal standard normalization in liquid chromatography-mass spectrometry metabolomics applied to environmental samples. *Anal Chem* **90**: 1363–1369.

Bremer, J. (1983) Carnitine. Metabolism and functions. *Physiol Rev* **63**: 1420–1480.

Bronk, D.A. (2002) Dynamics of DON. In *Biogeochemistry of Marine Dissolved Organic Matter*: New York, NY: Elsevier Science, pp. 153–247.

Bronk, D.A., Gilbert, P.M., and Ward, B.B. (1994) Nitrogen uptake, dissolved organic nitrogen release, and new production. *Science* **265**: 1843–1846.

Browning, T.J., Achterberg, E.P., Rapp, I., Engel, A., Bertrand, E.M., Tagliabue, A., and Moore, C.M. (2017) Nutrient co-limitation at the boundary of an oceanic gyre. *Nature* **551**: 242–246.

Button, D.K. (1998) Nutrient uptake by microorganisms according to kinetic parameters from theory as related to cytoarchitecture. *Microbiol Mol Biol Rev* **62**: 636–645.

Capella-Gutiérrez, S., Silla-Martínez, J.M., and Gabaldón, T. (2009) trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* **25**: 1972–1973.

Carini, P., Steindler, L., Beszteri, S., and Giovannoni, S.J. (2013) Nutrient requirements for growth of the extreme oligotroph "Candidatus Pelagibacter ubique" HTCC1062 on a defined medium. *ISME J* **7**: 592–602.

Chlumsky, L.J., Zhang, L., and Jorns, M.S. (1995) Sequence analysis of sarcosine oxidase and nearby genes reveals homologies with key enzymes of folate one-carbon metabolism. *J Biol Chem* **270**: 18252–18259.

Clifford, E.L., Varela, M.M., De Corte, D., Bode, A., Ortiz, V., Herndl, G.J., and Sintes, E. (2019) Taurine Is a Major Carbon and Energy Source for Marine Prokaryotes in the North Atlantic Ocean off the Iberian Peninsula. *Microbial Ecology* **78**(2): 299–312. https://doi.org/10.1007/s00248-019-01320-y

Coesel, S.N., Durham, B.P., Groussman, R.D., Hu, S.K., Caron, D.A., Morales, R.L., *et al*. (2021) Diel transcriptional oscillations of light-sensitive regulatory elements in open-ocean eukaryotic plankton communities. *Proc Natl Acad Sci U S A* **118**: e2011038118.

Curson, A.R.J., Liu, J., Bermejo Martínez, A., Green, R.T., Chan, Y., Carrión, O., *et al*. (2017) Dimethylsulfoniopropionate biosynthesis in marine bacteria and identification of the key gene in this process. *Nat Microbiol* **2**: 17009.

Curson, A.R.J., Williams, B.T., Pinchbeck, B.J., Sims, L.P., Martínez, A.B., Rivera, P.P.L., *et al*. (2018) DSYB catalyses the key step of dimethylsulfoniopropionate biosynthesis in many phytoplankton. *Nat Microbiol* **3**: 430–439.

Darriba, D., Taboada, G.L., Doallo, R., and Posada, D. (2011) ProtTest 3: fast selection of best-fit models of protein evolution. *Bioinformatics* **27**: 1164–1165.

del Valle, D.A., Kiene, R.P., and Karl, D.M. (2012). Effect of visible light on dimethylsulfoniopropionate assimilation and conversion to dimethylsulfide in the North Pacific Subtropical Gyre. *Aquatic Microbial Ecology* **66**(1): 47–62. https://doi.org/10.3354/ame01557

Durham, B.P., Boysen, A.K., Carlson, L.T., Groussman, R. D., Heal, K.R., Cain, K.R., *et al*. (2019) Sulfonate-based networks between eukaryotic phytoplankton and heterotrophic bacteria in the surface ocean. *Nat Microbiol* **4**: 1706–1715.

Eddy, S.R. (2011) Accelerated profile HMM searches. *PLoS Comput Biol* **7**: e1002195.

Edgar, R.C. (2010) Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26**: 2460–2461.

Fendrich, C., Hippe, H., and Gottschalk, G. (1990) Clostridium halophilium sp. nov. and C. litorale sp. nov., an obligate halophilic and a marine species degrading betaine in the Stickland reaction. *Arch Microbiol* **154**: 127–132.

Figueroa-Soto, C.G., and Valenzuela-Soto, E.M. (2018) Glycine betaine rather than acting only as an osmolyte also plays a role as regulator in cellular metabolism. *Biochimie* **147**: 89–97.

Fountoulakis, M., and Lahm, H.W. (1998) Hydrolysis and amino acid composition analysis of proteins. *J Chromatogr A* **826**: 109–134.

Galili, T., O'Callaghan, A., Sidi, J., and Sievert, C. (2018) heatmaply: an R package for creating interactive cluster heatmaps for online publishing. *Bioinformatics* **34**: 1600–1602.

Gaubert, J., Rodolfo-Metalpa, R., Greff, S., Thomas, O.P., and Payri, C.E. (2020) Impact of ocean acidification on the metabolome of the brown macroalgae Lobophora rosacea from New Caledonia. *Algal Res* **46**: 101783.

Gebser, B., Thume, K., Steinke, M., and Pohnert, G. (2020) Phytoplankton-derived zwitterionic gonyol and dimethylsulfonioacetate interfere with microbial dimethylsulfoniopropionate sulfur cycling. *Microbiology* **9**: e1014.

Gómez-Consarnau, L., Lindh, M.V., Gasol, J.M., and Pinhassi, J. (2012) Structuring of bacterioplankton communities by specific dissolved organic carbon compounds. *Environ Microbiol* **14**: 2361–2378.

Gradoville, M.R., Farnelid, H., White, A.E., Turk-Kubo, K.A., Stewart, B., Ribalet, F., *et al*. (2020) Latitudinal constraints on the abundance and activity of the cyanobacterium UCYN-A and other marine diazotrophs in the North Pacific. *Limnol Oceanogr* **65**: 1858–1875.

Gray, N., Zia, R., King, A., Patel, V.C., Wendon, J., McPhail, M.J.W., *et al*. (2017) High-speed quantitative UPLC-MS analysis of multiple amines in human plasma and serum via precolumn derivatization with 6-aminoquinolyl-N-hydroxysuccinimidyl carbamate: application to acetaminophen-induced liver failure. *Anal Chem* **89**: 2478–2487.

Hai, Y., Huang, A.M., and Tang, Y. (2019) Structure-guided function discovery of an NRPS-like glycine betaine reductase for choline biosynthesis in fungi. *Proc Natl Acad Sci U S A* **116**: 10348–10353.

Hall, E.R., and Gurin, S. (1975) Experiments in marine biochemistry. Homarine metabolism in Penaeus duorarum. *J Biol Chem* **250**: 6943–6946.

Heal, K.R., Durham, B.P., Boysen, A.K., Carlson, L.T., Qin, W., Ribalet, F., *et al*. (2021) Marine community metabolomes carry fingerprints of phytoplankton community composition. *mSystems* **6**: e01334-20.

Jameson, E., Doxey, A.C., Airs, R., Purdy, K.J., Murrell, J. C., and Chen, Y. (2016) Metagenomic data-mining reveals contrasting microbial populations responsible for trimethylamine formation in human gut and marine ecosystems. *Microb Genomics* **2**: e000080.

Johnson, W.M., Longnecker, K., Kido Soule, M.C., Arnold, W.A., Bhatia, M.P., Hallam, S.J., *et al*. (2019) Metabolite composition of sinking particles differs from surface suspended particles across a latitudinal transect in the South Atlantic. *Limnol Oceanogr* **65**: 111–127.

Joncquel-Chevalier Curt, M., Voicu, P.M., Fontaine, M., Dessein, A.F., Porchet, N., Mention-Mulliez, K., *et al*. (2015) Creatine biosynthesis and transport in health and disease. *Biochimie* **119**: 146–165.

Jones, H.J., Kröber, E., Stephenson, J., Mausz, M.A., Jameson, E., Millard, A., *et al*. (2019) A new family of uncultivated bacteria involved in methanogenesis from the ubiquitous osmolyte glycine betaine in coastal saltmarsh sediments. *Microbiome* **7**: 120.

Kageyama, H., Tanaka, Y., and Takabe, T. (2018) Biosynthetic pathways of glycinebetaine in Thalassiosira pseudonana; functional characterization of enzyme catalyzing three-step methylation of glycine. *Plant Physiol Biochem* **127**: 248–255.

Katoh, K., and Standley, D.M. (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol* **30**: 772–780.

Keeling, P.J., Burki, F., Wilcox, H.M., Allam, B., Allen, E.E., Amaral-Zettler, L.A., *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**: e1001889.

Keller, M.D., Kiene, R.P., Matrai, P.A., and Bellows, W.K. (1999a) Production of glycine betaine and dimethylsulfoniopropionate in marine phytoplankton. I. Batch cultures. *Mar Biol* **135**: 237–248.

Keller, M.D., Kiene, R.P., Matrai, P.A., and Bellows, W.K. (1999b) Production of glycine betaine and dimethylsulfoniopropionate in marine phytoplankton. II. N-limited chemostat cultures. *Mar Biol* **135**: 249–257.

Kiene, R.P. (1998) Uptake of choline and its conversion to glycine betaine by bacteria in estuarine waters. *Appl Environ Microbiol* **64**: 1045–1051.

Kiene, R.P., Hoffmann Williams, L.P., and Walker, J.E. (1998) Seawater microorganisms have a high affinity glycine betaine uptake system which also recognizes dimethylsulfoniopropionate. *Aquat Microb Ecol* **15**: 39–51.

Kiene, R.P., and Williams, L.P.H. (1998) Glycine betaine uptake, retention, and degradation by microorganisms in seawater. *Limnol Oceanogr* **43**: 1592–1603.

Klähn, S., and Hagemann, M. (2011) Compatible solute biosynthesis in cyanobacteria. *Environ Microbiol* **13**: 551–562.

Lahham, M., Jha, S., Goj, D., Macheroux, P., and Wallner, S. (2021) The family of sarcosine oxidases: same reaction, different products. *Arch Biochem Biophys* **704**: 108868.

Lidbury, I., Kimberley, G., Scanlan, D.J., Murrell, J.C., and Chen, Y. (2015a) Comparative genomics and mutagenesis analyses of choline metabolism in the marine Roseobacter clade. *Environ Microbiol* **17**: 5048–5062.

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

Lidbury, I.D.E.A., Murrell, J.C., and Chen, Y. (2015b) Trimethylamine and trimethylamine N-oxide are supplementary energy sources for a marine heterotrophic bacterium: implications for marine carbon and nitrogen cycling. *ISME J* **9**: 760–769.

Llewellyn, C.A., and Airs, R.L. (2010) Distribution and abundance of MAAs in 33 species of microalgae across 13 classes. *Mar Drugs* **8**: 1273–1291.

Lu, W.-D., Chi, Z.-M., and Su, C.-D. (2006) Identification of glycine betaine as compatible solute in Synechococcus sp. WH8102 and characterization of its N-methyltransferase genes involved in betaine synthesis. *Arch Microbiol* **186**: 495–506.

Magoč, T., and Salzberg, S.L. (2011) FLASH: fast length adjustment of short reads to improve genome assemblies. *Bioinformatics* **27**: 2957–2963.

Matsen, F.A., Kodner, R.B., and Armbrust, E.V. (2010) pplacer: linear time maximum-likelihood and Bayesian phylogenetic placement of sequences onto a fixed reference tree. *BMC Bioinformatics* **111**: 1–16.

Mausz, M.A., Airs, R.L., Dixon, J.L., Widdicombe, C.E., Tarran, G.A., Polimene, L., *et al*. (2022) Microbial uptake dynamics of choline and glycine betaine in coastal seawater. *Limnol Oceanogr* **9999**: 1–13.

McParland, E.L., Alexander, H., and Johnson, W.M. (2021) The osmolyte ties that bind: genomic insights into synthesis and breakdown of organic osmolytes in marine microbes. *Front Mar Sci* **8**: 732.

Meadows, J.A., and Wargo, M.J. (2015) Carnitine in bacterial physiology and metabolism. *Microbiol (United Kingdom)* **161**: 1161–1174.

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

Moran, M.A. (2015). The global ocean microbiome. *Science* **350**(6266): aac8455. https://doi.org/10.1126/science.aac8455

Mou, X., Hodson, R.E., and Moran, M.A. (2007) Bacterioplankton assemblages transforming dissolved organic compounds in coastal seawater. *Environ Microbiol* **9**: 2025–2037.

Myers, O.D., Sumner, S.J., Li, S., Barnes, S., and Du, X. (2017) Detailed investigation and comparison of the XCMS and MZmine 2 chromatogram construction and chromatographic peak detection methods for preprocessing mass spectrometry metabolomics data. *Anal Chem* **89**: 8689–8695.

Nakamura, H., Fujimaki, K., Sampei, O., and Murai, A. (1993) Gonyol: methionine-induced sulfonium accumulation in a dinoflagellate Gonyaulax polyedra. *Tetrahedron Lett* **34**: 8481–8484.

Ngugi, D.K., Ziegler, M., Duarte, C.M., and Voolstra, C.R. (2020) Genomic blueprint of glycine betaine metabolism in coral metaorganisms and their contribution to reef nitrogen budgets. *iScience* **23**: 101120.

Noell, S.E., Barrell, G.E., Suffridge, C., Morré, J., Gable, K.P., Graff, J.R., *et al*. (2021) Sar11 cells rely on enzyme multifunctionality to metabolize a range of polyamine compounds. *MBio* **12**: e0109121.

Noell, S.E., and Giovannoni, S.J. (2019) SAR11 bacteria have a high affinity and multifunctional glycine betaine transporter. *Environ Microbiol* **21**: 2559–2575.

Nyyssölä, A., Kerovuo, J., Kaukinen, P., Von Weymarn, N., and Reinikainen, T. (2000) Extreme halophiles synthesize betaine from glycine by methylation. *J Biol Chem* **275**: 22196–22201.

Parailloux, M., Godin, S., Fernandes, S.C.M., and Lobinski, R. (2020) Untargeted analysis for mycosporines

and mycosporine-like amino acids by hydrophilic interaction liquid chromatography (HILIC)—electrospray orbitrap MS2/MS3. *Antioxidants* **9**: 1–26.

Parthasarathy, A., Savka, M.A., and Hudson, A.O. (2019) The synthesis and role of β-alanine in plants. *Front Plant Sci* **10**: 921.

Pino, L.K., Searle, B.C., Bollinger, J.G., Nunn, B., MacLean, B., and MacCoss, M.J. (2020) The skyline ecosystem: informatics for quantitative mass spectrometry proteomics. *Mass Spectrom Rev* **39**: 229–244.

Polovina, J.J., Howell, E.A., Kobayashi, D.R., and Seki, M.P. (2017) The transition zone chlorophyll front updated: advances from a decade of research. *Prog Oceanogr* **150**: 79–85.

Popendorf, K.J., Lomas, M.W., and Van Mooy, B.A.S. (2011) Microbial sources of intact polar diacylglycerolipids in the Western North Atlantic Ocean. *Org Geochem* **42**: 803–811.

Poretsky, R.S., Sun, S., Mou, X., and Moran, M.A. (2010) Transporter genes expressed by coastal bacterioplankton in response to dissolved organic carbon. *Environ Microbiol* **12**: 616–627.

R Core Team. (2013) *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.

Rice, P., Longden, I., and Bleasby, A. (2000) EMBOSS: the European molecular biology open software suite. *Trends Genet* **16**: 276–277.

Sañudo-Wilhelmy, S.A., Gómez-Consarnau, L., Suffridge, C., and Webb, E.A. (2014) The role of B vitamins in marine biogeochemistry. *Ann Rev Mar Sci* **6**: 339–367.

Sellés Vidal, L., Kelly, C.L., Mordaka, P.M., and Heap, J.T. (2018) Review of NAD(P)H-dependent oxidoreductases: properties, engineering and application. *Biochim Biophys Acta - Proteins Proteomics* **1866**: 327–347.

Shao, Y.H., Guo, L.Z., Zhang, Y.Q., Yu, H., Zhao, B.S., Pang, H.Q., and Lu, W.D. (2018) Glycine betaine monooxygenase, an unusual Riesketype oxygenase system, catalyzes the oxidative N-demethylation of glycine betaine in Chromohalobacter salexigens DSM 3043. *Appl Environ Microbiol* **84**: e00377-18.

Shick, J.M., and Dunlap, W.C. (2002) Mycosporine-like amino acids and related gadusols: biosynthesis, accumulation, and UV-protective functions in aquatic organisms. *Annu Rev Physiol* **64**: 223–262.

Singh, G., Singh, J., Singamaneni, V., Singh, S., Gupta, P., and Katoch, M. (2021) Serine-glycine-betaine, a novel dipeptide from an endophyte Macrophomina phaseolina: isolation, bioactivity and biosynthesis. *J Appl Microbiol* **131**: 756–767.

Sipler, R.E., and Bronk, D.A. (2015) Chapter 4 - dynamics of dissolved organic nitrogen. In *Biogeochemistry of Marine Dissolved Organic Matter*, 2nd ed, Hansell, D.A., and Carlson, C.A. (eds). Boston: Academic Press, pp. 127–232.

Smith, C.A., O'Maille, G., Want, E.J., Qin, C., Trauger, S.A., Brandon, T.R., *et al*. (2005) METLIN: a metabolite mass spectral database. *Ther Drug Monit* **27**: 747–751.

Smith, C.A., Want, E.J., O'Maille, G., Abagyan, R., and Siuzdak, G. (2006) XCMS: processing mass spectrometry data for metabolite profiling using nonlinear peak alignment, matching, and identification. *Anal Chem* **78**: 779–787.

Stamatakis, A., Hoover, P., and Rougemont, J. (2008) A rapid bootstrap algorithm for the RAxML web servers. *Syst Biol* **57**: 758–771.

Stocker, R. (2012) Marine microbes see a sea of gradients. *Science* **338**: 628–633.

Sun, J., Mausz, M.A., Chen, Y., and Giovannoni, S.J. (2019) Microbial trimethylamine metabolism in marine environments. *Environ Microbiol* **21**: 513–520.

Sun, J., Steindler, L., Thrash, J.C., Halsey, K.H., Smith, D. P., Carter, A.E., *et al*. (2011) One carbon metabolism in SAR11 pelagic marine bacteria. *PLoS One* **6**: e23973.

Suttle, C.A., Chan, A.M., and Fuhrman, J.A. (1991) Dissolved free amino acids in the Sargasso Sea: uptake and respiration rates, turnover times, and concentrations. *Marine Ecology Progress Series* **70**: 189–199. https://doi.org/10.3354/meps070189

Torstensson, A., Young, J.N., Carlson, L.T., Ingalls, A.E., and Deming, J.W. (2019) Use of exogenous glycine betaine and its precursor choline as osmoprotectants in Antarctic sea-ice diatoms1. *J Phycol* **55**: 663–675.

Tripp, H.J., Kitner, J.B., Schwalbach, M.S., Dacey, J.W.H., Wilhelm, L.J., and Giovannoni, S.J. (2008) SAR11 marine bacteria require exogenous reduced sulphur for growth. *Nature* **452**: 741–744.

Trottmann, F., Ishida, K., Franke, J., Stanišić, A., Ishida-Ito, M., Kries, H., *et al*. (2020) Sulfonium acids loaded onto an unusual thiotemplate assembly line construct the cyclopropanol warhead of a burkholderia virulence factor. *Angew Chem - Int Ed* **59**: 13511–13515.

Van Mooy, B.A.S., and Fredricks, H.F. (2010) Bacterial and eukaryotic intact polar lipids in the eastern subtropical South Pacific: water-column distribution, planktonic sources, and fatty acid composition. *Geochim Cosmochim Acta* **74**: 6499–6516.

Vorobev, A., Sharma, S., Yu, M., Lee, J., Washington, B.J., Whitman, W.B., *et al*. (2018) Identifying labile DOM components in a coastal ocean through depleted bacterial transcripts and chemical signals. *Environ Microbiol* **20**: 3012–3030.

Wagner, M.A., and Jorns, M.S. (2000) Monomeric sarcosine oxidase: 2. Kinetic studies with sarcosine, alternate substrates, and a substrate analogue. *Biochemistry* **39**: 8825–8829.

Wagner, M.A., Trickey, P., Che, Z.W., Mathews, F.S., and Jorns, M.S. (2000) Monomeric sarcosine oxidase: 1. Flavin reactivity and active site binding determinants. *Biochemistry* **39**: 8813–8824.

Wargo, M.J. (2013) Homeostasis and catabolism of choline and glycine betaine: lessons from Pseudomonas aeruginosa. *Appl Environ Microbiol* **79**: 2112–2120.

Wargo, M.J., Szwergold, B.S., and Hogan, D.A. (2008) Identification of two gene clusters and a transcriptional regulator required for Pseudomonas aeruginosa glycine betaine catabolism. *J Bacteriol* **190**: 2690–2699.

Welsh, D.T. (2000) Ecological significance of compatible solute accumulation by micro-organisms: from single cells to global climate. *FEMS Microbiol Rev* **24**: 263–290.

Wyss, M., and Kaddurah-Daouk, R. (2000) Creatine and creatinine metabolism. *Physiol Rev* **80**: 1107–1213.

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

Yancey, P.H., and Siebenaller, J.F. (1999) Trimethylamine oxide stabilizes teleost and mammalian lactate dehydrogenases against inactivation by hydrostatic pressure and trypsinolysis. *J Exp Biol* **202**: 3597–3603.

Zecher, K., Hayes, K.R., and Philipp, B. (2020) Evidence of interdomain ammonium cross-feeding from methylamine- and glycine betaine-degrading Rhodobacteraceae to diatoms as a widespread interaction in the marine phycosphere. *Front Microbiol* **11**: 2431.

Zhu, Y., Jameson, E., Crosatti, M., Schäfer, H., Rajakumar, K., Bugg, T.D.H., and Chen, Y. (2014) Carnitine metabolism to trimethylamine by an unusual Rieske-type oxygenase from human microbiota. *Proc Natl Acad Sci U S A* **111**: 4268–4273.

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

Appendix S1. Supplementary Information.

Fig. S1. The mean concentration of quantified osmolytes over the course of the GBT fate experiments, in nM C (top) and as a percent of the total C measured in these osmolytes (bottom).

Fig. S2. The particulate concentration of all possible GBT isotopologues over the course of the incubations (top). The mass isotopologue distribution (MID) over the course of the incubations (bottom). Lines are the average of the triplicate bottles, with error bars representing one standard deviation. Dashed lines on the bottom plots indicate the natural isotopic abundance of samples taken from the stations where each experiment took place. Insets show the lower abundance isotopologues.

Fig. S3. The total particulate ^{15}N (top) and ^{13}C (bottom) isotope concentrations in metabolites in the South (left) and North (right) incubation experiments, relative to the total ^{15}N and ^{13}C measured in known, quantifiable metabolites. Other includes all compounds quantified, as listed in Supplemental Table S4.

Fig. S4. Microorganisms expressing GBT-related genes in the NPTZ during springtime cruises in 2016 and 2017. The pie charts show an average of the five stations in the NPTZ for 2016 and the six stations in the NPTZ for 2017, except for *gbcA* which was only detected at three stations in 2016.

Fig. S5. Peak area of $^{13}C_4$, $^{15}N_1$-dimethylglycine over the course of both experiments. This isotopologue of dimethylglycine was not detected in the *in situ* samples from the corresponding stations.

Fig. S6. Mass isotopologue distributions (MID, as a percent) over time for various isotopologues detected in the metabolites that are derived directly from glycine betaine in the southern (left) and northern (right) experiments. The *x*-axis is incubation time in hours. Solid lines indicate the mean of triplicate bottles, and the error bars represent one standard deviation. Horizontal dashed lines indicate the MID for samples taken at the stations where each experiment was performed. Note the different *y*-axis values for each plot.

Fig. S7. The concentration of nitrogen in select metabolites over time in the GBT fate incubation experiments, shown as the concentration of nitrogen in all isotopologues of a given metabolite. The line and error bars are the mean and standard deviation of triplicates.

Fig. S8. The total concentration of carbon in total hydrolyzable amino acids at the initial and final (98 h) time points in the southern and northern GBT fate experiments. Points are the average value of three replicates, with the standard deviation in grey bars. Methionine and histidine were only detected in a few replicates and were therefore excluded from this total.

Fig. S9. Isotopologue accumulation trends for $^{13}C_2$,$^{15}N_1$ and $^{13}C_5$,$^{15}N_1$ compounds identified in the untargeted metabolomics data for both experiments. Triplicate samples for each time-point are shown as independent rows. Blue colours indicate lower mass isotopologue distribution contribution, yellow colours indicate higher MID. White values indicate data removed during quality control. Compounds with identifiers written as N-FT0000/S-FT000 are unknown or tentatively identified metabolites without authentic standards to confirm their identify, with the names corresponding to the mass feature found in the northern experiment (N-FT) and southern experiment (S-FT) as in Supplementary Tables S2 and S3. Known compounds are shown with data processed through the untargeted pipeline as opposed to the targeted pipeline to make the most robust comparisons to the unknown mass features.

Table S1. The K_s, V_{max}, and turnover time. Uptake kinetics values are reported as a mean and standard deviation calculated as discussed in the Experimental Procedures section. Values are reported as: mean (SD).

Table S2. Results of a significance test to determine if isotopologues changed their distribution over time. Each isotopologue that was detected in >2 samples in a given experiment was fit with a linear model to determine if there was a significant trend with time in the MID. The *p*-value and significance (True if $p < 0.05$) results are presented here for each isotopologue tested in each experiment. The ^{13}C-0 ^{15}N-0 isotopologues that have a significant trend had MID decreases over time, while the isotopologues with ^{13}C or ^{15}N that were significant typically had MID increases over time. The average and standard deviation of the MID for each isotopologue in each sample are in Table S4. (XLSX)

Table S3. Table of the estimated concentrations (nM) and isotopologue distributions. The timepoint is the approximate sampling time, Incubation time (h) is the actual sampling time after initial spiking. C.lab and N.lab are the number of ^{13}C and ^{15}N atoms in the isotopologue. Iso.mz is the mass to charge ratio for the isotopologue. Ionization is the ionization we measured that metabolite in. MID is mass isotopologue distribution, where the number 1 would indicate that in a particular sample 100% of the concentration of a given metabolite was in that specific isotopologue and the number 0 would indicate that in a particular sample, 0% of the concentration of a given metabolite was in that specific isotopologue. (XLSX)

Table S4. Table of targeted metabolites measured in the glycine betaine isotope tracing experiments. 'Isotopologue significant change' column is true if in either the north or the south experiment, any isotopologue of that metabolite showed a significant trend in mass isotopologue distribution with time. 'Isotopologue significant change' column is false if we measured the metabolite but did not detect any ^{13}C or ^{15}N isotopologues, or if we did detect the isotopologues but they did not show any significant enrichment over time ($p > 0.05$). The details of which isotopologues were had significant MID changes in which experiment are in Table S3. (XLSX)

Table S5. A summary of the number of mass features identified in each LC-MS configuration and those with isotope incorporation. (XLSX)

Table S6. Parameters used in untargeted metabolomics data processing (XLSX)

Table S7. Reference sequences used for searching metatranscriptomic data. For each gene, there is a list of all reference sequences used (called homologues). The sequence(s) in red were the annotated/experimentally verified sequences used to retrieve the rest of the homologues. All these reference sequences/homologues were used to retrieve short reads from the metatranscriptomes. (XLSX)

Table S8. The K_t (3 parameters) or $K_t + S$ (two parameters), V_{max}, and S values when two versus three parameters (n. params.fit) were fit, reported as a mean and standard deviation as calculated from the Monte Carlo error analysis described in the Experimental Procedures.