Release of Ecologically Relevant Metabolites by the Cyanobacterium, *Synechococcus elongatus* CCMP 1631

Cara L Fiore, Krista Longnecker, Melissa C Kido Soule, and Elizabeth B Kujawinski*

Department of Marine Chemistry & Geochemistry, Woods Hole Oceanographic Institution, Woods Hole, MA, USA

*Corresponding author

360 Woods Hole Rd. MS#4
Woods Hole, MA, USA 02543
Phone: 508-289-3493
Fax: 508-289-2164
ekujawinski@whoi.edu

Running title: Metabolomics of Synechococcus

Abstract

Photoautotrophic plankton in the surface ocean release organic compounds that fuel secondary production by heterotrophic bacteria. Here we show that an abundant marine cyanobacterium, *Synechococcus elongatus*, contributes a variety of nitrogen-rich and sulfur-containing compounds to dissolved organic matter. A combination of targeted and untargeted metabolomics and genomic tools was used to characterize the intracellular and extracellular metabolites of *S. elongatus*. Aromatic compounds such as 4-hydroxybenzoic acid and phenylalanine, as well as nucleosides (e.g., thymidine, 5'-methylthioadenosine, xanthosine), the organosulfur compound 3-mercaptopropionate, and the plant auxin indole 3-acetic acid, were released by *S. elongatus* at multiple time points during its growth. Further, the amino acid kynurenine was found to accumulate in the media even though it was not present in the predicted metabolome of *S. elongatus*. This indicates that some metabolites, including those not predicted by an organism's genome, are likely excreted into the environment as waste; however, these molecules may have broader ecological relevance if they are labile to nearby microbes. The compounds described herein provide excellent targets for quantitative analysis in field settings to assess the source and lability of dissolved organic matter *in situ*.

Introduction

Dissolved organic matter (DOM) comprises one of the largest reservoirs of reduced carbon on the planet (Hedges, 2002). This organic carbon includes a complex mixture of compounds ranging from high molecular weight colloidal material (Amon and Benner, 1994) to free amino acids and simple carbohydrates (Granum *et al.*, 2002). It is within this dissolved milieu that microbial interactions (Grossart and Simon, 2007), acquisition of specific compounds (Ito and Butler, 2005), and microbial growth (Romera-Castillo *et al.*, 2011) all take place (reviewed by Kujawinski (2011)). In order to constrain the flux of organic nutrients through the dissolved pool, it is critical to characterize this material at the molecular level, and to assess the influence of different microbes on DOM composition.

Phytoplankton release DOM as a byproduct of photosynthetic carbon fixation and this material contributes significantly to the biologically-available organic carbon and nitrogen in the ocean (Nagata and Kirchman, 1992; Kaiser and Benner, 2008). Phytoplankton-derived DOM is used as an energy source and for growth by microbial communities (Moreira et al., 2011; Nelson and Carlson, 2012), particularly heterotrophic bacteria (Cole *et al*., 1982; Kaiser and Benner, 2008). Further, the composition of phytoplankton exudates influences microbial community structure over spatial and temporal gradients (Cottrell and Kirchman, 2000; Nelson and Carlson, 2012; Landa *et al*., 2013). Investigations of DOM released by phytoplankton have shown that its composition is quite variable and is influenced by phylogeny (Romera-Castillo *et al.*, 2011; Becker *et al.*, 2014) and growth conditions (Grossart and Simon, 2007; Baran *et al.*, 2011).

Studies with diatoms and flagellated eukaryotes have documented the exudation of macromolecules (Nagata and Kirchman, 1992) and polysaccharides (Aluwihare and Repeta, 1999) within the high-molecular-weight fraction of DOM (>1000 Da). In contrast, the composition of labile low-molecular weight (LMW) components of DOM remains poorly constrained due to its heterogeneous nature and the challenges associated with extracting this material from high salt solutions. The pool of labile LMW DOM is a small fraction of the total LMW DOM pool (Amon and Benner, 1996; Nagata *et al*., 2003; Hama *et al*., 2004). Nevertheless, this fraction is of particular interest because these 'small' molecules comprise metabolites that are transferred among members of microbial consortia (e.g., Durham *et al*., 2015). Biochemical classes of labile LMW DOM such as hydrolyzed amino acids and sugars have been quantified in seawater using several techniques (e.g., Kaiser and Benner, 2000; Kaiser and Benner 2009), providing insight into source and turnover of these components. There are still gaps in our knowledge, however, regarding the molecular composition of labile LMW DOM derived from ecologically relevant taxa of phytoplankton and the factors that influence the turnover of this material. In particular, more information is needed on the identities of DOM molecules that are important for microbe-microbe interactions (e.g., signaling molecules, exudates that are utilized by other microbes) and microbe-DOM interactions (e.g., labile metabolites). Identification of such ecologically relevant components of labile phytoplankton-derived DOM will then allow for quantification of the turnover of these components by microbial communities, which is critical for building predictive global carbon cycle models (Follows *et al*., 2007).

Recent advances in mass spectrometry and computational tools are providing an avenue for exploring phytoplankton-derived DOM on a molecular level. Generally termed 'environmental metabolomics,' the goals of these studies are typically the

quantification of known metabolites ("targeted metabolomics") or the discovery of new metabolites ("untargeted metabolomics"; e.g., Patti *et al*., 2012). In untargeted metabolomics studies, complex datasets of thousands of mass features are generated, but statistical analyses (Rosselló-Mora *et al*., 2008; Kujawinski *et al*., 2009), stable isotope labeling (Baran *et al.*, 2010), or comparative approaches (Becker *et al.*, 2014) can be applied to uncover patterns in the data and to obtain features of interest for metabolite identification. After identification, these features can be quantified using targeted metabolomics methods in either intracellular or extracellular metabolite mixtures to discern their metabolic role and/or their temporal dynamics (e.g., Ankrah *et al*., 2014). Here, we use targeted and untargeted metabolomics to characterize DOM released from cyanobacterial cultures.

Cyanobacteria such as *Synechococcus* spp. are abundant primary producers in the photic areas of the ocean and are important in global carbon cycling (Li, 1994; Scanlan and West, 2002). *Synechococcus elongatus* in particular, is found in pelagic and coastal waters (Seymour *et al*., 2010; Muralitharan and Thajuddin, 2011) and comprises approximately 10% of the *Synechococcus* spp. genes recovered in the global ocean sampling data set (Rusch *et al.*, 2007; Seymour *et al.*, 2010). The prevalence of *S. elongatus* in surface waters implies that this organism constitutes a reliable source of labile DOM in the environment and thus is an excellent candidate for culture-based experiments aimed at characterizing excreted ecologically relevant metabolites.

We paired targeted and untargeted metabolomics techniques to characterize LMW DOM released from *Synechococcus elongatus* CCMP 1631 during different stages of growth. Intracellular and extracellular metabolites were extracted from cultures of *S. elongatus* and analyzed using high-performance liquid chromatography (LC) coupled to triple stage quadrupole (QQQ; targeted) and Fourier transform ion cyclotron resonance

(FT-ICR; untargeted) mass spectrometers. Biologically relevant mass features were identified from untargeted data and compared to predicted metabolomes generated from several genomes representing *Synechococcus* spp. and other bacterioplankton and phytoplankton species. With these methods, we have identified and quantified compounds not previously associated with the metabolic profile of *S. elongatus*. The dynamics of these components have important implications for the turnover of DOM and for microbial interactions *in situ*. The metabolites highlighted here represent compounds that should be prioritized for further functional characterization.

Results

Growth of Synechococcus elongatus *CCMP 1631*

Cultures of *S. elongatus* exhibited lag, exponential, and decline growth phases, and sampling was performed during each of the different growth stages (Figure S2). Fluorescence microscopy showed an occasional small non-auto-fluorescent cell in some cultures of *S. elongatus* (average of 0.002 non-auto-fluorescent putative cells per *S. elongatus* cell; see SI); however, these were in low abundance relative to the cyanobacteria and thus were unlikely to significantly alter the outcome of the experiment. No-cell controls were clear of cells except for day 15, which had approximately 9×10^6 cells ml^{-1}, which is at least 100-fold less than any of the cultures. The no-cell control from day 15 was more similar to the other control samples than to the culture samples in terms of metabolite composition and concentration across all quantified metabolites, lending support for the use of the day-15 control in the analysis. Inorganic nutrients were not limiting for *S. elongatus* at the start of the experiment and decreased over time (Table S2). Total organic carbon (TOC) ranged from approximately 170 to 300 μM in the

cultures. The extraction efficiencies of organic carbon from the extracellular media by solid-phase extraction (SPE) were lower than expected (3-13%; Table S2), but are similar to previous studies using SPE of phytoplankton-released DOM (Landa *et al*., 2014; Becker *et al*., 2014). Additionally, the extraction efficiencies of individual compounds within the fraction of LMW DOM targeted in this study are likely much higher based on nearly 100% recovery of several deuterated organic compounds from seawater using this SPE method (Johnson, Kido Soule, and Kujawinski, unpublished).

Metabolic profile of Synechococcus elongatus *CCMP 1631*

We generated metabolic profiles for *S. elongatus* along a growth curve. At each time point, we quantified intra- and extracellular concentrations of pre-chosen metabolites. Simultaneously, we analyzed intra- and extracellular profiles of all detectable metabolites. In these untargeted datasets, data was organized into mass features, where a feature is defined by a unique combination of mass-to-charge (*m*/*z)* ratio and a retention time. The focus of this report is the composition and relative concentrations of extracellular metabolites. However, results of the intracellular metabolite analysis provide valuable complementary insights into the physiology of *S. elongatus*, and are highlighted where they relate to the production or utilization of extracellular compounds or for comparative analysis.

In the untargeted analysis, metabolic profiles of intra- and extracellular fractions were significantly different based on analysis of similarity (ANOSIM; $R = 1$, $p = 0.001$ (positive and negative); Figure 1). There was a large overlap of mass features, however, with 98% (positive mode) and 90% (negative mode) of features detected in both intra- and extracellular factions. The intracellular profiles showed some separation by time in

positive mode (ANOSIM; R=0.509, p=0.006; Figure 1A), but no trend in negative mode (ANOSIM, R=0.124, p=0.16; Figure 1C).

Cluster analysis of the targeted intracellular data showed minimal clustering of samples by time. For example, one sample each from days 0 and 8 did not cluster with their corresponding replicates, although the rest of the early time points clustered together (Figure 2). Additionally, variability in the composition and concentrations of intracellular metabolites was observed across time points (Figure 2). Diverse metabolites were detected at relatively high concentrations within the quantified intracellular compounds including several amino acids, 5'-adenosine monophosphate (AMP), and several carbohydrate derivatives (Table 1). The polyamine, spermidine, was the dominant and most consistent metabolite within the quantified intracellular compounds (Figure 2), and the amino acids involved in polyamine synthesis clustered together with spermidine (Figure 2, cluster A). Some compounds of interest were detected in the intracellular fraction only at low concentrations including *N*-acetylchitobiose (chitobiose) and dimethylsulfoniopropionate (DMSP).

Within the targeted extracellular metabolic dataset, thymidine, tryptophan, phenylalanine, and 4-hydroxybenzoic acid comprised the dominant features (Figure 3; Table 2). The pyrimidine nucleoside, thymidine, increased in concentration over time when normalized to volume and to TOC (Figure S3). Succinic acid and inosine were present at early time points but below detection after day 8 or 10. Some compounds including 3-mercaptopropionate, indole 3-acetic acid (IAA), 5'-methylthioadenosine (MTA), and *N*-acetylglutamic acid were detected at relatively low but consistent levels throughout the experiment. IAA is typically considered a catabolic product of tryptophan and extracellular concentrations of tryptophan and IAA were significantly correlated (Pearson's r = 0.94, p-value <0.01; Figure S4). Metabolites of interest in the dissolved

fraction could be grouped as to whether they increased throughout the experiment, or were variable in TOC-normalized concentration over time (Table 2; Figure 3). The removal of metabolites from the media may be due to either uptake by *S. elongatus* or abiotic degradation. More experiments are needed to determine the sink for these metabolites.

Characterization of DOM by untargeted analysis

Several dissolved metabolites were observed in both the targeted and untargeted datasets. These metabolites exhibited a similar trend in each dataset (i.e., increased or decreased over time) as shown by the correlation between tryptophan concentration determined in the targeted method and the peak area of the mass feature identified as tryptophan in the untargeted method (Figure 4, Spearman's $\rho = 0.81$, p-value < 0.01). Untargeted metabolic profiles from extracellular fractions showed a shift over time and this was supported by ANOSIM (Figure 1A (positive): R=0.702, p=0.001; Figure 1C (negative): R=0.788, p=0.002). Approximately 25% (646 peaks (negative) and 413 peaks (positive)) of all extracellular features (Table 3) increased over time when normalized to TOC and had an average peak area greater than the average peak area for the cell-free controls. Of the features that increased over time, approximately 10% (67 (negative) and 40 (positive)) had corresponding MS/MS (fragmentation) spectra, which facilitate putative identification and thus were selected for further investigation. Additional features were selected if the mass of the feature matched the exact mass of a metabolite in one of the predicted metabolomes (Table 3). Altogether, 21 features of interest, which generally increased in peak area over time, were putatively identified or confirmed in the untargeted dataset (Table S3; see SI).

One such feature was identified as kynurenine, a tryptophan oxidation product and a predicted metabolite of *Ruegeria pomeroyi*, a heterotrophic α-proteobacterium in the Roseobacter clade, and *Thalassiosira pseudonana,* a centric diatom. The presence of kynurenine in the *S. elongatus* dataset was confirmed with an authentic standard analyzed with our methods. Kynurenine accumulated over time in the extracellular fraction even when normalized to TOC (as a proxy for biomass; Figure 5). It was also observed intracellularly at variable concentrations over time (data not shown). Comparative genomics analysis was used to investigate whether sequenced *Synechococcus* spp. contain the enzymes responsible for oxidation of tryptophan to kynurenine (Table S4, SI), but this analysis did not yield any potential orthologous genes.

Discussion

Intracellular metabolites

Phytoplankton are considered the major producers of DOM in the surface ocean, the composition of which can vary (Myklestad, 1995; Biersmith and Benner, 1998; Becker *et al*., 2014) due to species-based differences in metabolism. In order to determine the impact of the metabolism of the cosmopolitan cyanobacterium, *Synechococcus elongatus,* on DOM composition, as well as to quantify diverse and unexpected metabolites associated with *S. elongatus,* we applied both targeted and untargeted metabolomics techniques to analyze intra- and extracellular metabolites derived from cultures of *S. elongatus*. Few metabolomics studies have examined both intra- and extracellular metabolite profiles of marine microbes (Baran *et al*., 2010; Rosselló-Mora *et al*., 2008; Longnecker *et al*., 2015). The intracellular dataset described here thus provided unique insight into the metabolism of *S. elongatus* and highlighted specific compounds

with undefined functional roles as targets for future physiological and biochemical studies.

We expected the intracellular metabolites to be more variable in concentration and composition over time reflecting shifts in metabolism at different stages of growth. This overall variability was observed in the intracellular profile derived from the untargeted analyses and to some extent in the targeted intracellular analyses. The targeted intracellular analyses, however, also showed that a few metabolites were consistently detected at relatively high concentrations or steadily increased or decreased over time (Table 1). For example, the most prominent intracellular metabolite in the targeted method was the polyamine spermidine, sometimes comprising ~40% of the total quantified metabolites in a sample. Polyamines are known to have a wide range of functions in the cell (Tabor and Tabor, 1984; Panagiotidis *et al*., 1995), and the relatively high and stable concentrations in the intracellular metabolite samples suggests that spermidine has a central role in the metabolic functioning of *S. elongatus*. While spermidine could not be quantified in the extracellular fraction for analytical reasons, a side product of spermidine production, 5'-methylthioadenosine (MTA) was observed to accumulate in the media. This suggests that MTA is not entirely recycled back into central metabolism (i.e., *via* a methionine salvage pathway) and may be released through a regulatory mechanism. Furthermore, many of the amino acids involved in spermidine biosynthesis clustered together (Figure 2), which suggests that this pathway affects concentrations of all these compounds.

Other notable intracellular metabolites include the sulfur-containing osmolyte dimethylsulfoniopropionate (DMSP), typically associated with eukaryotic algae. Interestingly, the DMSP degradation intermediate, 3-mercaptopropionate, which also contains sulfur, was detected at low concentrations intracellularly and extracellularly.

Synechococcus elongatus might be catabolizing DMSP to 3-mercaptopropionate (Visscher *et al*., 1994; Yoch, 2002). Although this is the most likely metabolic source for 3-mercaptopropionate, other metabolic pathways may generate 3-mercaptopropionate (Kiene *et al*., 1990) and further experiments are needed to determine the source of 3-mercaptopropionate in *S. elongatus*. Cyanobacteria have been shown to produce DMSP (Jonkers *et al.*, 1998), which can function as both an osmolyte and a scavenger of reactive oxygen species in marine algae (ROS; Sunda *et al*., 2002). While release of DMSP could not be quantified here (due to method constraints), other studies have indicated that cyanobacteria are likely to be an insignificant source of DMSP relative to other marine algae (Keller, 1989). Nonetheless, the release of 3-mercaptopropionate, and potentially DMSP, even at low concentrations could be a consistent source of organosulfur metabolites for microbes inhabiting the same environment as *S. elongatus*.

Another interesting metabolite in *S. elongatus* was chitobiose in the intracellular fraction. As the samples were grown in artificial seawater and this molecule was not detected in the cell-free controls, we must conclude that *S. elongatus* is making chitobiose. The functional role of chitobiose in *S. elongatus* is not known but it may be exported to the outer membrane and/or function in some structural capacity.

Extracellular metabolites

Characterization of the molecules released by phytoplankton can be difficult due to the complex nature of DOM (Nagata, 2008) and the rapid turnover of small compounds (Rich *et al.*, 1996). However, these small labile compounds, which are a quantitatively small fraction of LMW DOM, are important in the flux of organic nutrients through the dissolved pool and in supporting bacterioplankton populations. Here, we

identified several ecologically and biochemically relevant metabolites released by *Synechococcus elongatus*.

The fact that there is a large overlap in features between the intracellular and extracellular metabolic profiles suggests that the differences identified by ANOSIM are driven by the relative concentrations of the overlapping metabolites. These differences are influenced by 1) the amount of metabolite released and 2) the degree to which each metabolite is subsequently utilized by *S. elongatus*. Thus, while most of the metabolites produced by *S. elongatus* are released into the environment, only a subset accumulates to significant concentrations.

Major features observed in the targeted analysis of the extracellular metabolite fractions can, in part, define the influence of *S. elongatus* on DOM composition. One example is the nucleoside thymidine, a dominant feature in the extracellular fraction. The genomes of *Synechococcus* species do not contain the gene that encodes for thymidine kinase, the enzyme that phosphorylates thymidine to thymidine monophosphate (TMP) for incorporation into DNA. Instead, the gene encoding for the enzyme thymidylate synthase, which converts uridine monophosphate (UMP) to TMP, was present in the investigated genomes of *Synechococcus*. Therefore, any thymidine that is released by DNA degradation, or by hydrolysis of TMP, is excreted from *S. elongatus* cells and may provide a valuable source of thymidine for heterotrophic bacteria. Thymidine incorporation ([^{3}H]-thymidine) has been used to quantify bacterial production, as many heterotrophic bacteria are known to transport and incorporate thymidine into DNA (Fuhrman and Azam, 1982; Zweifel *et al*., 1993). Nucleic acids, as well as triphosphate nucleotides (i.e., ATP, GTP) have been identified as an important component in DOM (Karl and Bailiff, 1984; Björkman and Karl, 2001); however, to our knowledge there are no published reports of thymidine concentrations in seawater.

The aromatic amino acids phenylalanine and tryptophan were also major components in the extracellular fraction. These compounds appeared to be both released and taken up by *S. elongatus* to different degrees. Phenylalanine concentrations remained relatively steady, only increasing with increasing TOC. This indicates that there is some degradation or uptake of phenylalanine that prevents this amino acid from accumulating in the media. Tryptophan increased in concentration initially but was not detected in the later time points, indicating this amino acid was released to the extracellular fraction but was then degraded or utilized by *S. elongatus*. Tryptophan is energetically costly to synthesize (Barton *et al*., 2010) and the release of tryptophan and other aromatic compounds may constitute an important source of these molecules for nearby microbes. Whether there was uptake of amino acids by *S. elongatus* or abiotic degradation of these amino acids in the media cannot be resolved using our data. However, uptake of amino acids by *Synechococcus* spp. has been documented previously in culture (Baran *et al*., 2011) and during *in situ* incubation experiments (Zubkov *et al*., 2003; Michelou *et al*., 2007). Taken together, this information emphasizes the need for quantifying the autotrophic contribution to DOM turnover in order to improve models of the ocean carbon cycle.

Tryptophan and phenylalanine have been quantified previously in seawater (0 – 54 and 4 – 27 nM, respectively in surface water (< 100 m)) and their concentrations and distribution with depth indicate a link to surface water production (Yamashita and Tanoue, 2003; Jørgensen *et al*., 2011). These metabolites are also among a subset of amino acids considered to be important labile components of DOM (Yamashita and Tanoue, 2003), but the source of these amino acids has not been identified within the surface planktonic community. The work presented here provides an important step

toward identifying ecologically relevant components of labile DOM and elucidating the microbial sources of these components in the ocean.

Other minor components in the extracellular media may be ecologically important even at low concentrations, particularly because they were detected at most or all time points. For example, indole 3-acetic acid (IAA) has been characterized in symbiotic cyanobacteria associated with plants, although its functional role in free-living bacteria is not known (Sergeeva *et al.*, 2002). IAA from *S. elongatus* appears to be derived from tryptophan, based on the strong correlation between these two metabolites in our data. Additionally, while the full biochemical pathway linking the two metabolites remains uncharacterized, a gene encoding for a pyridoxal-dependent decarboxylase family protein was found in two *Synechococcus* genomes. This enzyme could potentially use tryptophan as a substrate (see SI), and highlights the potential of metabolomics studies for improving annotation of marine microbial genomes.

The distribution and functional roles of IAA have been investigated in soil systems (Martens & Frankenberger, 1993 b). However, few studies have quantified IAA in seawater (Bentley, 1960; Maruyama et al., 1989; Mazur and Homme 1993) and these had limited spatial distributions for sampling. The released IAA may function as a signaling molecule or may simply be a waste product. There are several lines of evidence for a connection between the occurrence and concentration of IAA in seawater and mucus exudation by diatoms (Homme and Mazur, 1992; Mazur and Homme, 1993). The production of mucus has broader ecological implications as it can influence the flux of organic matter to the deep ocean (Riebesell et al., 1995; Turk et al., 2010) and the mucus itself consists of a rich matrix of phytoplankton exudates and polysaccharides (Myklestad, 1995; Brussaard et al., 1996). The fact that *S. elongatus* releases IAA may mean that this

organism can influence mucus production by other phytoplankton, a hypothesis that should be addressed by future experimental studies.

Several mass features in the untargeted extracellular dataset were identified and confirmed with authentic standards (see SI). Some of these compounds, such as xanthosine and tyrosine, may be released as metabolic overflow during photosynthesis (Behrends *et al.*, 2009). Of particular interest was the mass feature identified as the amino acid, kynurenine, which clearly accumulated over time in the media. When normalized to TOC, the peak areas of the mass feature still increased over time, suggesting that there is continual release of kynurenine over time and that there is little, if any, degradation. This would be expected if the compound were a waste product of a continuous process or of a constitutively active pathway. In fact, a similar trend is produced when thymidine concentration is normalized to TOC, and as discussed above, thymidine is a waste product with no salvage pathway.

Genes homologous to known protein-coding genes involved in the kynurenine pathway of tryptophan catabolism (Kurnasov *et al.*, 2003) were not detected in *Synechococcus* genomes. However, several studies have reported modification of tryptophan residues in proteins of photosystem II (PSII), where the first step of the light reactions for photosynthesis take place (Sharma *et al.*, 1997; Anderson *et al.*, 2002; reviewed by Dreaden Kasson and Barry (2012)). Anderson *et al.*, (2002) detected kynurenine and other tryptophan oxidation products in the place of tryptophan residues in one of the PSII proteins (CP43) suggesting that these modifications are a result of reactive oxygen species generated during the light reactions. Oxidation of tryptophan residues to *N*-formylkynurenine in the D1 protein by ROS generated during photosynthesis was also recently observed (Dreaden *et al*., 2011; Dreaden Kasson *et al*., 2012). The PSII proteins, particularly D1, are known to turn over rapidly relative to other

cellular proteins (Ohad *et al.*, 1984) and the presence of these tryptophan oxidation products may be an indicator of ROS-damaged proteins (Anderson *et al*., 2002). We hypothesize that tryptophan residues in the PSII proteins are oxidized to kynurenine by ROS, and that this kynurenine is released into the extracellular pool.

While *Synechococcus* do not have the ability to utilize kynurenine as a nitrogen or carbon source (see SI), the resulting release of kynurenine from the cell may make *S. elongatus,* and more generally *Synechoccoccus* spp., an important source of this amino acid to nearby heterotrophic bacteria. Genetic evidence for the presence of the kynurenine pathway (Figure S8) in several different genera of bacteria was recently obtained using comparative genetics (Kurnasov *et al*., 2003). In addition, the kynurenine pathway was found to be the major route of catabolism for added carbon in the form of L-tryptophan in an experiment with soil microbes (Martens and Frankenberger, 1993 a). The extent to which marine bacterioplankton utilize the kynurenine pathway as a catabolic source of carbon and nitrogen has not been investigated. Nevertheless, it may be important in some marine microbes and represents an intriguing area for further research.

We have shown that *S. elongatus* is a consistent source of nucleosides, amino acids, and organosulfur compounds to the dissolved organic pool. Interestingly, many of the compounds released by *S. elongatus* in this study are not typically quantified in seawater, such as IAA, kynurenine, MTA, thymidine, 4-hydroxybenzoic acid, and xanthosine. It is not clear, from these data, whether these compounds are excreted intentionally or leak passively through the semipermeable cell membrane. Our observations of kynurenine and thymidine accumulation in the growth media, however, implicate intentional release for at least some metabolites, particularly waste products that cannot be salvaged. Our results provide evidence for metabolites that are generated

during central metabolism and are released by the organism, but are not predicted by its genome and therefore would not normally be considered in the ecology of this organism. Nevertheless, compounds such as kynurenine could play important roles in the metabolism of other marine microbes expected to occur in the same communities as our target organism, *S. elongatus*, and thus constitute possible points of synergy within microbial communities. In short, continually-released waste products may have an important role in microbial interactions with DOM. All of the compounds described here are high-potential targets for *in situ* metabolomics analyses to further explore the source and turnover of metabolites within marine DOM. Similar metabolomics studies with metabolically-diverse microbial taxa are needed to discover additional metabolites of biological importance in microbial consortia, including compounds that are generated by gene-independent processes. In this way, environmental metabolomics studies, such as described here, can provide needed complementary data to genomics and transcriptomics studies in order to elucidate the functional roles of diverse marine microbes.

Experimental procedures

Culture of Synechococcus elongatus *CCMP 1631*

All glassware was acid-cleaned and combusted at 450°C for at least 4 h. *Synechococcus elongatus* CCMP 1631 was purchased from the National Center for Marine Algae and Microbiota (Boothbay, ME, USA) and grown in 300 ml of L1-Si media (https://ncma.bigelow.org/algal-recipes) using Turks Island Salts as a base. A starter culture in exponential phase was used to inoculate 12 flasks of media with 10% of the volume (30 ml). Another six flasks contained only the media as cell-free controls.

The cultures were maintained at room temperature under a 12:12 light:dark regime (56 μmol m^{-2} s^{-1}) and were sampled at approximately 30 min into the light cycle, at each time point. Cultures were grown for three weeks and samples were taken on days: 0, 2, 8, 10, 15, and 17. The initial samples (day 0) were taken at the time of the experiment setup. At each sampling point 750 μl, 40 ml, and 60 ml were removed for cell counts, TOC analysis, and nutrient analysis, respectively (SI). Potential contamination by heterotrophic bacteria was monitored with DAPI-stained cells of each time point. Briefly, 750 μl of culture was fixed with 75 μl buffered 37% formaldehyde (buffered with sodium borate) and frozen at -80°C. The fixed cells were stained with 4',6-diamidino-2-phenylindole dihydrochloride (DAPI, 70 μg ml^{-1} final concentration) on Isopore membrane filters (0.2 μm, Millipore). The filters were viewed with a Zeiss Axiostar plus microscope, where 25-35 fields or at least 150 cells were counted (SI). *Synechococcus elongatus* abundance (cells ml^{-1}) of each sample was calculated from DAPI-derived counts by following the formula from (Wetzel and Likens, 1991). The abundance estimates based on cell counts were used to calculate a growth curve of *S. elongatus* during the experiment and we could clearly differentiate separate stages of growth (Figure S2). Cell counts were not used in data analysis (i.e., data normalization), however, due to the presence of cell clumps that made obtaining accurate counts difficult.

Metabolite extractions

Extraction of metabolites from the *S. elongatus* cultures followed the overview given in Figure S1. Each culture was filtered through a 0.2 μm PTFE filter (Omnipore, Millipore, MA, UA). Filters were stored at -80°C until extraction using a method adapted from (Rabinowitz and Kimball, 2007). Intracellular metabolites from the filter

were extracted in 500 μl of cold 40:40:20 acetonitrile:methanol:water with 0.1 M formic acid. The solvent and filter mix was sonicated for 10 minutes and the extract was centrifuged at 20,000 × *g* for 5 minutes. The supernatant was transferred to a new vial and neutralized with 6 M ammonium hydroxide. Intracellular extracts were dried in a vacufuge to near dryness and redissolved in 450 μl water and 50 μl acetonitrile. Deuterated biotin was added to each sample as an HPLC injection standard (biotin-(*ring*-6,6-d_2), final concentration 0.05 μg ml^{-1}). At this stage, 100 μL of the intracellular extract was used for targeted metabolomics analysis.

A portion of the remaining intracellular extract (250 μl) was processed using solid phase extraction (SPE) with 50 mg/ 1 cc PPL cartridges (BondElut, Agilent, Santa Clara, CA, USA) as described previously (Dittmar *et al.*, 2008) to remove salt. The eluted metabolite extract was dried down and redissolved in 250 μl of 95:5 water:acetonitrile and 200 μl of the extract was removed for untargeted metabolomics analysis. A pooled sample was also made consisting of equal amounts of each intracellular extract.

The filtrate for each sample (229 ml), representing the extracellular metabolite fraction, was processed using SPE as described above but with 1 g / 6 cc cartridges. The dried extracellular metabolite extract was redissolved in 500 μl of 95:5 water:acetonitrile for both targeted and untargeted metabolomics analysis. The PPL protocol is the most popular method for DOM extraction (e.g., Zhang *et al*., 2014; Perminova *et al*., 2014) and while there is bias in the components that are retained by any extraction method, the chosen method allows for comparison of metabolite data with other studies (Kido Soule *et al*., submitted).

Instrument methods

Intracellular and extracellular metabolites for the targeted metabolomics method were analyzed within 24-48 hours of extraction by high-performance liquid chromatography (Thermo PAL autosampler and Accela pump) coupled to a triple stage quadrupole mass spectrometer (TSQ Vantage, Thermo Fisher Scientific, MA, USA) via a heated electrospray ionization (H-ESI) source operated in both positive and negative ion modes. Selected reaction monitoring (SRM) conditions were optimized for each metabolite standard (n = 84, Table S1). Compound separation was achieved on a Synergi Fusion-RP column (4 μm, 2.0 x 150 mm; Phenomenex) with a guard column, heated to 35°C. The extracts were gradient-eluted at 250 μl min^{-1} using (solvent A) water with 0.1% formic acid and (solvent B) acetonitrile with 0.1% formic acid (hold 5% B for 2 min, ramp 5-65% B for 18 min, ramp 65-100% B for 5 min, and hold 100% B for 8 min; total run time = 40 min). The column was equilibrated for 7.5 min prior to each 10-μl injection. Milli-Q water blanks and a standard mix of metabolites were included in each analytical run. Five-to-seven point manually-curated calibration curves for each standard (0.5 – 500 ng ml^{-1}) were used to determine metabolite concentrations (XCalibur 2.0). Peak integrations for all samples were also manually curated and the resulting quality-checked metabolite concentrations were then exported to Microsoft Excel. Absolute quantification of metabolites in our samples is not possible without internal standards for each metabolite. Instead, external standard curves were used to quantify the metabolites, resulting in potential underestimates of metabolite concentrations in the samples due to matrix effects (Taylor, 2005). Nevertheless, our method allows us to compare metabolite concentrations over time.

Extracted metabolites for untargeted analysis were analyzed with high-performance liquid chromatography (Micro AS autosampler and Surveyor MS pump plus) coupled *via* ESI to a hybrid linear ion trap-FT-ICR mass spectrometer (7T LTQ FT

Ultra, Thermo Fisher Scientific). Separate autosampler injections of 20 µl each were performed for positive and negative mode analyses. The pooled sample was analyzed every six injections in both the targeted and untargeted metabolomics methods for quality control. The column, gradient, and flow rate were the same as described for the targeted method. In parallel to MS acquisition in the FT-ICR cell (resolving power 100,000 at 400 *m/z*), MS/MS data were collected at nominal mass resolution in the ion trap from the four features with the highest peak intensities in each scan. Data were collected in profile mode using XCalibur 2.0.7 and converted to centroid mode with msConvert (Chambers *et al.*, 2012).

Metabolomics data analysis

Quality-checked targeted metabolomics data were normalized to total metabolite biomass for the intracellular metabolites, and to filtrate volume or TOC for the extracellular metabolites. Figures were generated with R statistical software (v3.0.2; R Core Team; see SI) using TOC- or volume-normalized concentrations. Extracellular mass features were normalized to TOC as an estimate of biomass and additionally to the volume filtered as these provide two views on the behavior of the metabolites in the extracellular media. For example, if the concentrations of a metabolite were correlated to biomass then we would expect it to be stable over time when normalized to TOC. On the other hand, normalizing to volume shows the total amount released to the media over time. Normalized intracellular metabolite values used to generate the heatmap were standardized to the average concentration for each metabolite. The Pearson product-moment correlation coefficient was calculated to assess correlation between concentrations of tryptophan and indole 3-acetic acid from the extracellular metabolite

fraction. The non-parametric Spearman's rank correlation coefficient (Spearman's rho) was used to estimate correlation between extracellular tryptophan concentrations from the targeted dataset and the peak area of the mass feature identified as tryptophan from the untargeted dataset. Only a subset of targeted metabolites (n=18) could be quantified in the extracellular fraction due to the others' incompatibility with our solid-phase extraction resin (Table S1). This limits direct comparison between the intracellular and extracellular data due to variability in extraction efficiency among compounds; however, our analysis is focused on changes in concentrations of metabolites over time rather than on their absolute values.

Untargeted metabolomics data were processed with XMCS 1.38.0 (Smith *et al.*, 2006) and CAMERA 1.18.0 programs (Kuhl *et al.*, 2012) in R. The resulting peak table from XCMS and CAMERA was quality checked using several parameters. For each mass feature of the intracellular metabolite data, the coefficient of variation was calculated for the samples (using peak area) and for the pooled samples described above. The coefficient of variation within the samples was required to be greater than that within the pooled samples (Vinaixa *et al*., 2014). The average peak area of samples also had to be at least 35% greater than the average peak area of the corresponding media control samples. Additionally, each mass feature was required to be present in at least three samples in positive mode and four samples in negative mode (including replicate samples), as there was more noise in the negative mode data. Peak areas of mass features are proportional to metabolite concentration and thus were normalized to the filtrate volume or to TOC for analysis. Quality-checked mass feature data (normalized peak areas) were used in nMDS and analysis of similarity (ANOSIM) analyses to characterize changes in metabolic profile over time and between samples (see SI). Draft predicted metabolomes were generated (Karp *et al*., 2011) and compared to experimental mass

features to obtain features of interest (see SI). Mass features of interest with an associated fragmentation spectrum were processed using the xcmsFragment function in XCMS and the spectra were queried against the METLIN metabolomics database (Smith *et al.*, 2005) and in the program MetFrag (Wolf *et al.*, 2010). MetFrag searches chemical and biological databases such as KEGG (Kanehisa *et al*., 2014) using the parent ion mass of the feature of interest and the associated fragmentation spectrum. For each putative match in the database, MetFrag generates all possible fragments (*in silico*) of the molecule and compares this fragmentation spectrum to that of the mass feature of interest. Putative identifications were confirmed with commercial standards analyzed by our methods (i.e., kynurenine, tryptophan and other standards from the targeted method, see SI). The processed untargeted and targeted metabolomics data from the current project have been submitted to MetaboLights (Haug *et al*., 2013) and are available as MTBLS155.

Acknowledgements

This project was funded by the Gordon and Betty Moore Foundation through Grant #3304 to E. Kujawinski. The authors have no conflict of interest to declare. We thank Crystal Breier for assistance with experiment setup and sampling, Gretchen Swarr for assistance with metabolite extractions and Paul Henderson for nutrient analyses. We are grateful to Paul Carini for helpful discussions on metabolic pathways and microbial physiology, and to Bryndan Durham for helpful comments on this manuscript.

Literature Cited

Aluwihare, L.I. and Repeta, D.J. (1999) A comparison of the chemical characteristics of ocean DOM and extracellular DOM produced by marine algae. *Mar Ecol Prog Ser* **186**: 105–117.

Amon, R.M. and Benner, R. (1994) Rapid cycling of high-molecular-weight dissolved organic matter in the ocean. *Lett Nature* **369**: 549–552.

Amon, R.M. and Benner, R. (1996). Bacterial utilization of different size classes of dissolved organic matter. *Limnol Oceanogr* **41**:41-51.

Anderson, L.B., Maderia, M., Ouellette, A.J., Putnam-Evans, C., Higgins, L., Krick, T., et al. (2002) Posttranslational modifications in the CP43 subunit of photosystem II. Proc *Natl Acad Sci USA* **99**: 14676–14681.

Ankrah, N.Y.D., May, A.L., Middleton, J.L., Jones, D.R., Hadden, M.K., Gooding, J.R., et al. (2014) Phage infection of an environmentally relevant marine bacterium alters host metabolism and lysate composition. *ISME J* **8**: 1089-1100.

Baran, R., Bowen, B.P., Bouskill, N.J., Brodie, E.L., Yannone, S.M., and Northen, T.R. (2010) Metabolite Identification in Synechococcus sp. PCC 7002Using Untargeted Stable Isotope Assisted Metabolite Profiling. *Anal Chem* **82**: 9034–9042.

Baran, R., Bown, B.P., and Northen, T.R. (2011) Untargeted metabolic footprinting reveals a surprising breadth of metabolite uptake and release by *Synechococcus* sp. PCC 7002. *Molec BioSystems* **7**: 3200–3206.

Becker, J.W., Berube, P.M., Follett, C.L., Waterbury, J.B., Chrisholm, S.W., DeLong, E.F., and Repeta, D.J. (2014) Closely related phytoplankton species produce similar suites of dissolved organic matter. *Front Microbiol* **5**: 1–14.

Biersmith, A., and Benner, R. (1998). Carbohydrates in phytoplankton and freshly produced dissolved organic matter. *Mar Chem* **63**:131-144.

Brussaard, C.P.D., Gast, G.J., van Duyl, F.C., and Riegman, R. (1996). Impact of phytoplankton bloom magnitude on a pelagic microbial food web. *Mar Ecol Prog Ser* **144**:211-221.

Chambers, M.C., Maclean, B., Burke, R., Amode, D., Ruderman, D.L., Neumann, S., et al. (2012) A cross-platform toolkit for mass spectrometry and proteomics. *Nature Biotech* **30**: 918–920.

Cole, J.J., Likens, G.E., and Strayer, D.L. (1982) Photosynthetically produced dissolved organic carbon: An important carbon source for planktonic bacteria. *Limnol Oceanogr* **27**: 1080–1090.

Cottrell, M.T. and Kirchman, D.L. (2000) Natural Assemblages of Marine Proteobacteria and Members of the Cytophaga-Flavobacter Cluster Consuming Low- and High-

Molecular-Weight Dissolved Organic Matter. *Appl Environ Microbiol* **66**: 1692–1697.

Dittmar, T., Koch, B., Hertkorn, N., and Kattner, G. (2008) A simple and efficient method for solid-phase extracion of dissolved organic matter (SPE-DOM) from seawater. *Limnol Oceanogr Methods* **6**: 230–235.

Dreaden Kasson, T.M. and Barry, B.A. (2012) Reactive oxygen and oxidative stress: *N*-formyl kynurenine in photosystem II and non-photosynthetic proteins. *Photosyn Res* **114**: 97–110.

Dreaden Kasson, T.M., Rexroth, S., and Barry, B.A. (2012) Light-Induced Oxidative Stress, N-Formylkynurenine, and Oxygenic Photosynthesis. *PLOS ONE* **7**: e42220.

Dreaden, T.M., Chen, J., Rexroth, S., and Barry, B.A. (2011) N-Formylkynurenine as a Marker of High Light Stress in Photosynthesis. *J Biol Chem* **286**: 22632–22641.

Durham, B.P., Sharma, S., Luo, H., Smith, C.B., Amin, S.A., *et al*. (2015). Cryptic carbon and sulfur cycling between surface ocean plankton. *Proc Natl Acad Sci USA* **112**:453-457.

Follows, M.J., Dutkiewicz, S., Grant, S., and Chisholm, S.W. (2007). Emergent biogeography of microbial communities in a model ocean. *Science* **315**:1843-1846.

Granum, E., Kirkvold, S., and Myklestad, S.M. (2002) Cellular and extracellular production of carbohydrates and amino acids by the marine diatom *Skeletonema costatum*: diel variations and effects of N depletion. *Mar Ecol Prog Ser* **242**: 83–94.

Grossart, H.P. and Simon, M. (2007) Interactions of planktonic algae and bacteria: effects on algal growth and organic matter dynamics. *Aquat Microb Ecol* **47**: 163–176.

Hama, T., Yanagi, K., and Hama, J. (2004). Decrease in molecular weight of photosynthetic products of marine phytoplankton during early diagenesis. *Limnol Oceanogr* **49:**471-481.

Haug, K., Salek, R.M., Conesa, P., Hastings, J., de Matos, P., Rijnbeek, M., *et. al*. (2013) MetaboLights—an open-access general-purpose repository for metabolomics studies and associated meta-data. *Nucl Acids Res* **41**: D781-D786.

Hedges, J.I. (2002) Why dissolved organic matter. In, Hansell,D.A. and Carlson,C.A. (eds), *Biochemistry of Marine Dissolved Organic Matter*. Academic Press, pp. 1–33.

Ito, Y. and Butler, A. (2005) Structure of synechobactins, new siderophores of the marine cyanobacterium *Synechococcus* sp. PCC 7002. *Limnol Oceanogr* **50**: 1918–1923.

Jonkers, H.M., Koopmans, G.F., and van Gemerden, H. (1998) Dynamics of dimethyl sulfide in a marine microbial mat. *Microb Ecol* **36**: 93–100.

Kaiser, K. and Benner, R. (2008) Major bacterial contribution to the ocean reservoir of

detrital organic carbon and nitrogen. *Limnol Oceanogr* **53**: 99–112.

Keller, M.D. (1989) Dimethyl sulfide production and marine phytoplankton: The importance of species composition and cell size. *Biol Oceanogr* **6**: 375–382.

Kiene, R.P., Malloy, K.D., and Taylor, B.F. (1990) Sulfur-containing amino acids as precursors of thiols in anoxic coastal sediments. *Appl Environ Microbiol* **56**:156-161.

Kuhl, C., Tautenhahn, R., Böttcher, C., Larson, T.R., and Neumann, S. (2012) CAMERA: An Integrated Strategy for Compound Spectra Extraction and Annotation of Liquid Chromatography/Mass Spectrometry Data Sets. *Anal Chem* **84**: 283–289.

Kujawinski, E.B. (2011) The Impact of Microbial Metabolism on Marine Dissolved Organic Matter. *Annu Rev Mar Sci* **3**: 567–599.

Kujawinski, E.B., Longnecker, K., Blough, N.V., Del Vecchio, R., Finlay, L., Kitner, J.B., and Giovannoni, S.J. (2009) Identification of possible source markers in marine dissolved organic matter using ultrahigh resolution mass spectrometry. *Geochim Cosmo Acta* **73**: 4384–4399.

Kurnasov, O., Jablonski, L., Polanuyer, B., Dorrestein, P., Begley, T., and Osterman, A. (2003) Aerobic tryptophan degradation pathway in bacteria: novel kynurenine formamidase. *FEMS Microbiol Lett* **227**: 219–227.

Landa, M., Cottrell, M.T., Kirchman, D.L., Blain, S., and Obernosterer, I. (2013) Changes in bacterial diversity in response to dissolved organic matter supply in a continuous culture experiment. *Aquat Microb Ecol* **69**: 157–168.

Landa, M., Cottrell, M.T., Kirchman, D.L., Kaiser, K., Medeiros, P.M., Tremblay, L., *et al*. (2014). Phylogenetic and structural response of heterotrophic bacteria to dissolved organic matter of different chemical composition in a continuous culture study. *Environ Microbiol* **16**:1668-1681.

Li, W.K. (1994) Primary production of prochlorophytes, cyanobacteria, and eucaryotic ultraphytoplankton: Measurements from flow cytometric sorting. *Limnol Oceanogr* **39**: 169–175.

Longnecker K., Kido Soule, M.C., and Kujawinski, E.B. (2015). Dissolved organic matter produced by *Thalassiosira pseudonana*. *Mar Chem* **168**:114-123.

Michelou, V.K., Cottrell, M.T., and Kirchman, D.L. (2007). Light-stimulated bacteria production and amino acid assimilation by cyanobacteria and other microbes in the North Atlantic Ocean. *Appl. Environ. Microbiol.* **73**:5539-5546.

Moreira, I.C., I, B., and Vieira, A. (2011) Decomposition of dissolved organic matter released by an isolate of *Microcystis aeruginosa* and morphological profile of the associated bacterial community. *Braz J Biol* **71**: 57–63.

Muralitharan, G. and Thajuddin, N. (2011) Rapid differentiation of phenotypically and genotypically similar *Synechococcus elongatus* strains by PCR fingerprinting. *Biologia* **66**: 238–243.

Myklestad, S.M. (1995). Release of extracellular products by phytoplankton with special emphasis on polysaccharides. *Sci Tot Environ* **165**:155-164.

Nagata, T. (2008) Organic matter-bacteria interactions in seawater. In, Kirchman,D.L. (ed), *Microbial Ecology of the Oceans*. John Wiley & Sons, Inc., pp. 207–241.

Nagata, T. and Kirchman, D.L. (1992) Release of macromolecular organic complexes by heterotrophic marine flagellates. *Mar Ecol Prog Ser* **83**: 233–240.

Nagata, T., Meon, B., and Kirchman, D.L. (2003). Microbial degradation of peptidoglycan in seawater. *Limnol Oceanogr* **48**:745-754.

Nelson, C.E. and Carlson, C.A. (2012) Tracking differential incorporation of dissolved organic carbon types among diverse lineages of Sargasso Sea bacterioplankton. *Environ Microbiol* **14**: 1500–1516.

Ohad, I., Kyle, D.J., and Arntzen, C.J. (1984) Membrane protein damage and repair: Removal and replacement of inactivated 32-kilodalton polypeptides in chloroplast membranes. *J Cell Biol* **99**: 481–485.

Panagiotidis, C.A., Artandi, S., Calame, K., and Silverstein, S.J. (1995) Polyamines alter sequence-specific DNA-protein interactions. *Nucleic Acids Res* **23**: 1800–1809.

Patti, G.J., Yanes, O., and Siuzdak, G. (2012). Metabolomics: the apogee of the omics trilogy. *Nature Rev Mol Cell Biol* **13**: 263-269.

Perminova, I.V., Dubinenkov, I.V., Kononikhin, A.S., Konstantinov, A.I., Zherebker, A.Y., Andzhushev, M.A., et al. (2014). Molecular mapping of sorbent selectivities with respect to isolation of Arctic dissolved organic matter as measured by Fourier transform mass spectrometry. *Environ Sci Technol Lett* **48**: 7461-7468.

Rabinowitz, J.D. and Kimball, E. (2007) Acidic Acetonitrile for Cellular Metabolome Extraction from Escherichia coli. *Anal Chem* **79**: 6167–6173.

Rich, J.H., Ducklow, H.W., and Kirchman, D.L. (1996) Concentrations and uptake of neutral monosaccharides along 140ºW in the equatorial Pacific: Contribution of glucose to heterotrophic bacterial activity and the DOM flux. *Limnol Oceanogr* **41**: 595–604.

Riebesell, U., Reigstad, M., Wassmann, P., Noji, T., and Passow, U. (1995). On the trophic fate of *Phaeocystis pouchetii* (Hariot): VI. Significance of *Phaeocystis*-derived mucus for vertical flux. *Neth J Sea Res* **33**:193-203.

Romera-Castillo, C., Sarmento, H., Álvarez-Salgado, X.A., Gasol, J.M., and Marrase, C. (2011) Net Production and Consumption of Fluorescent Colored Dissolved Organic

Matter by Natural Bacterial Assemblages Growing on Marine Phytoplankton Exudates. *Appl Environ Microbiol* **77**: 7490–7498.

Rosselló-Mora, R., Lucio, M., Peña, A., Brito-Echeverría, J., López-López, A., Valens-Vadell, M., et al. (2008) Metabolic evidence for biogeographic isolation of the extremophilic bacterium *Salinibacter ruber*. *ISME J* **2**: 242–253.

Rusch, D.B., Halpern, A.L., Sutton, G., Heidelberg, K.B., Williamson, S., Yooseph, S., et al. (2007) The Sorcerer II Global Ocean Sampling Expedition: Northwest Atlantic through Eastern Tropical Pacific. *PLoS Biol* **5**: e77.

Scanlan, D.J. and West, N.J. (2002) Molecular ecology of the marine cyanobacterial genera *Prochlorococcus* and *Synechococcus*. *FEMS Microbiol Ecol* **40**: 1–12.

Sergeeva, E., Liaimer, A., and Bergman, B. (2002) Evidence for production of the phytohormone indole-3-acetic acid by cyanobacteria. *Planta* **215**: 229–238.

Seymour, J.R., Ahmed, T., Durham, W.M., and Stocker, R. (2010) Chemotactic response of marine bacteria to the extracellular products of Synechococcus and Prochlorococcus. *Aquat Microb Ecol* **59**: 161–168.

Sharma, J., Panico, M., Shipton, C.A., Nilsson, F., Morris, H.R., and Barber, J. (1997) Primary structure characterization of the photosystem II D1 and D2 subunits. *J Biol Chem* **272**: 33158–33166.

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.

Sunda, W., Kleber, D.J., Kiene, R.P., and Huntsman, S. (2002). An antioxidant function for DMSP and DMS in marine algae. *Lett Nature* **418**: 317-320.

Tabor, C.W. and Tabor, H. (1984) Polyamines. *Ann Rev Biochem* **53**: 749–790.

Turk, V., Hagstrom, A., Kovač, N., and Faganeli, J. (2010). Composition and function of mucilage macroaggregates in the northern Adriatic. *Aquat Microb Ecol* **61**:279-289.

Visscher, P.T., Kiene, R.P., and Taylor, B.F. (1994) Demethylation and cleavage of dimethylsulfoniopropionate in marine intertidal sediments. *FEMS Microbiol Eco*l **14**:179-189.

Wetzel, R.G. and Likens, G.E. (1991) Limnological Analyses 2nd ed. Springer-Verlag.

Wolf, S., Schmidt, S., Müller-Hannemann, M., and Neumann, S. (2010) In silico fragmentation for computer assisted identification of metabolite mass spectra. *BMC Bioinformatics* **11**: 148.

Yoch, D.C. (2002) Dimethylsulfoniopropionate: Its sources, role in the marine food web, and biological degradation to dimethylsulfide. Appl Environ Microbiol 68:5804-5815.

Zhang, F., Harir, M., Moritz, F., Zhang, J., Witting, M, Wu, Y., *et al*. (2014). Molecular and structural characterization of dissolved organic matter during and post cyanobacterial bloom in Taihu by combination of NMR spectroscopy and FTICR mass spectrometry. *Water Res* **57**: 280-294.

Zubkov, M.V., Fuchs, B.M., Tarran, G.A., Burkhill, P.H., and Amann, R. (2003). High rate of uptake of organic nitrogen compounds by Prochlorococcus cyanobacteria as a key to their dominance in oligotrophic oceanic waters. *Appl. Environ. Microbiol*. **69**:1299-1304.

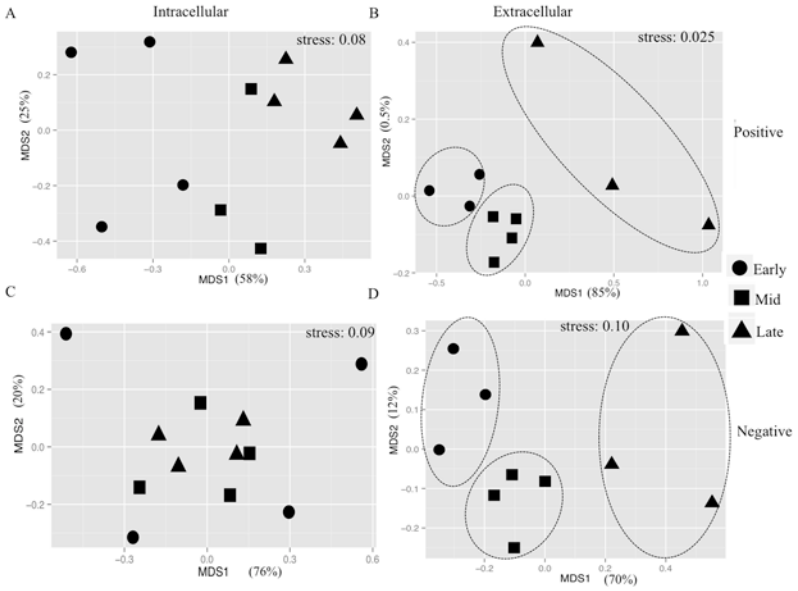

Figure 1. Non-metric multidimensional scaling (nMDS) of intracellular metabolic profiles (A, C) (r^2 = 0.93, 0.85, respectively) and extracellular metabolic profiles (B, D) (r^2 = 0.99, 0.95, respectively) from untargeted analysis in positive ion mode (A, B) and negative ion mode (C, D). Profiles corresponding to time points (Early = day 0, 2; Mid = 8, 10; Late = 15, 17) are indicated by the legend. Extracellular profiles that separate by major time points (B, D) are highlighted by ellipses.

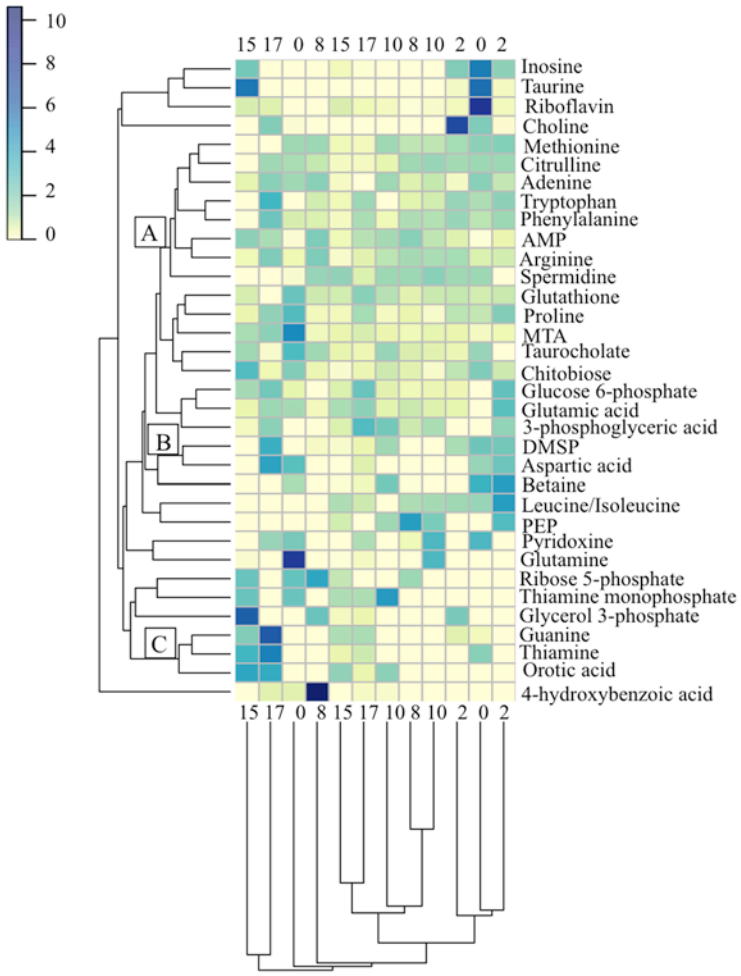

Figure 2. TOC-normalized intracellular metabolites from targeted analysis clustered by sample (columns, days 0-17) and by metabolite (rows) (Bray-Curtis dissimilarity) and expressed as relative concentration to the average for each metabolite. Only metabolites observed in more than one third of samples were used. A value of 1 is equal to the average concentration and is visualized as light green on the color spectrum. Darker colors represent values above the average concentration and lighter colors represent values below the average concentration. Groups of interest are shown as A-C (see SI). MTA= 5'-methylthioadenosine, AMP = 5'-adenosine monophosphate, IMP = 5'-inosine monophosphate, PEP = phosphoenolpyruvate, DMSP = dimethylsulfoniopropionate.

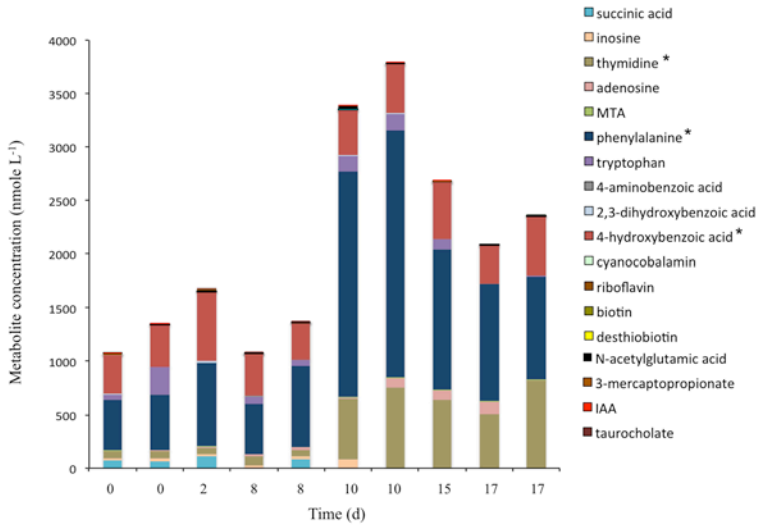

Figure 3. Volume-normalized concentrations of a subset of targeted metabolites in the extracellular fraction over time. Replicate samples are indicated by sampling day. Dominant metabolites are indicated by a (*) in the legend.

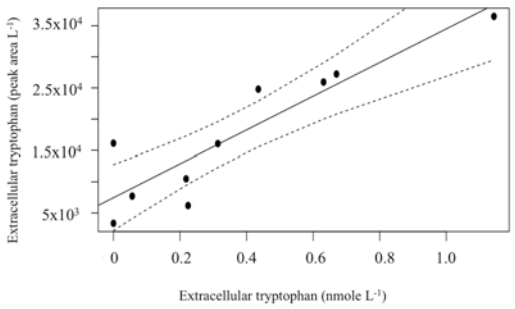

Figure 4. Extracellular tryptophan quantified in the targeted method plotted against the peak area of the mass feature identified as tryptophan (extracellular) in the untargeted method, showing a significant correlation (Spearman's ρ=0.81, $p < 0.01$).

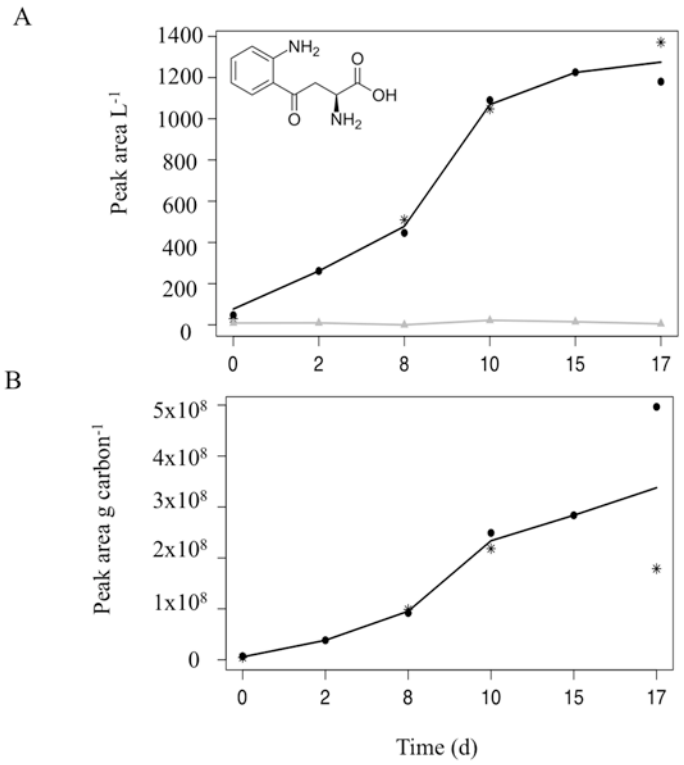

Figure 5. Volume-normalized (A) and TOC-normalized (B) peak area of the mass feature identified as the amino acid kynurenine in the extracellular fraction over time (positive ion mode). The structure of kynurenine is shown as an inset in A. The black lines represent the average between replicate cultures (circles and asterisks), at each time point and the gray represents the media control. For days 2 and 15, data from only one replicate is available.

Table 1. Metabolites of interest in the intracellular fraction from targeted metabolomics analysis. Intracellular metabolites were normalized to total organic carbon (TOC). Metabolites that are shaded light gray increased over time, while metabolites shaded dark gray decreased over time, and non-shaded metabolites were variable or remained steady over time. The fold change from the second time point* to the last time point is shown in the shading.

Most abundant intracellular compounds	**Fold change**
Spermidine	3
Glutathione (reduced)	1
Glutamate	2
5'-adenosine monophosphate (AMP)	4
Dimethylsulfoniopropionate (DMSP)	3
Glucose 6-phosphate	2
Arginine	10**
Citrulline	1
Uracil	
Glutamine	
Thiamin	
N-acetylglucosamine	
Glycine betaine	
Proline	
Chitobiose	

* The second time point rather than the first time point was used to allow cells to be fully acclimated after inoculation.

**This metabolite was detected at the initial time point but not the second time point; therefore, the average of the first two time points was used to calculate the fold change.

Table 2. Metabolites of interest in the extracellular fraction from targeted metabolomics analysis. Extracellular metabolites were normalized to total organic carbon (TOC). Metabolites that are shaded increased over time, while non-shaded metabolites were variable or remained steady over time. The fold change from the second time point* to the last time point is shown in the shading.

Most abundant extracellular compounds	**Fold change**
Adenosine	16
Thymidine	12
N-acetylglutamic acid	2
4-hydroxybenzoic acid	1
Phenylalanine	
Tryptophan	
Succinic acid	
Inosine	
Indole 3-acetic acid	
N-acetyltaurine	
3-mercaptopropionate	
Cyanocobalamin	
Taurocholate	

* The second time point rather than the first time point was used to allow cells to be fully acclimated after inoculation and to mitigate the influence of the inoculating media on extracellular concentrations.