

1 **An extended siderophore suite from *Synechococcus* sp. PCC 7002 revealed by LC-**  
2 **ICPMS-ESIMS.**

3  
4 Rene M. Boiteau<sup>a,b</sup> and Daniel J. Repeta<sup>a</sup>

5  
6 <sup>a</sup>Department of Marine Chemistry and Geochemistry, Woods Hole Oceanographic  
7 Institution, Woods Hole, MA 02543

8  
9 <sup>b</sup>Department of Earth, Atmospheric and Planetary Sciences, Massachusetts Institute of  
10 Technology, Cambridge, MA 02139

11  
12 **Abstract**

13  
14 Siderophores are thought to play an important role in iron cycling in the ocean, but  
15 relatively few marine siderophores have been identified. Sensitive, high throughput  
16 methods hold promise for expediting the discovery and characterization of new  
17 siderophores produced by marine microbes. We developed a methodology for  
18 siderophore characterization that combines liquid chromatography (LC) inductively  
19 coupled plasma mass spectrometry (ICPMS) with high resolution electrospray ionization  
20 mass spectrometry (ESIMS). To demonstrate this approach, we investigated siderophore  
21 production by the marine cyanobacteria *Synechococcus* sp. PCC 7002. Three  
22 hydroxamate siderophores, synechobactin A-C, have been previously isolated and  
23 characterized from this strain. These compounds consist of an iron binding head group  
24 attached to a fatty acid side chain of variable length (C<sub>12</sub>, C<sub>10</sub>, and C<sub>8</sub> respectively). In  
25 this study, we detected six iron-containing compounds in *Synechococcus* sp. PCC 7002  
26 media by LC-ICPMS. To identify the molecular ions of these siderophores, we aligned  
27 the chromatographic retention times of peaks from the LC-ICPMS chromatogram with  
28 features detected from LC-ESIMS spectra using an algorithm designed to recognize  
29 metal isotope patterns. Three of these compounds corresponded to synechobactins A  
30 (614 m/z), B (586m/z), and C (558m/z). The MS<sub>2</sub> spectra of these compounds revealed  
31 diagnostic synechobactin fragmentation patterns which were used to confirm the identity  
32 of the three unknown compounds (600, 628, and 642 m/z) as new members of the  
33 synechobactin suite with side chain lengths of 11, 13, and 14 carbons. These results  
34 demonstrate the potential of combined LCMS techniques for the identification of novel  
35 iron-organic complexes.

36  
37 **Introduction**

38  
39 Iron availability can influence the biological productivity and ecosystem community  
40 composition of the ocean<sup>1-4</sup>. In oxic seawater, dissolved iron (III) rapidly forms  
41 insoluble oxyhydroxides and precipitates from solution<sup>5</sup>. This poses a challenge to  
42 microbes that require iron for photosynthesis, nitrogen fixation, respiration, protection  
43 against oxidative stress, and other essential biological processes. In regions where  
44 available iron is scarce, microbes may gain a competitive advantage by producing  
45 siderophores; biomolecules that strongly and specifically complex iron in a form that can  
46 be recovered using dedicated membrane transporters<sup>6,7</sup>.

1  
2 Isolating and identifying siderophores is important for understanding the competition for  
3 iron in regions of the ocean where iron scarcity exerts a selective pressure on the  
4 microbial community. Indeed, siderophores appear to be present at low concentrations in  
5 seawater collected from open ocean sites<sup>8</sup>, and metagenomic analysis has suggested that  
6 siderophore production and uptake in the ocean may be common<sup>9-12</sup>. Some siderophores  
7 have been discovered in the extracts of a number of marine bacteria grown in laboratory  
8 cultures<sup>13</sup>. However their characterization in marine environmental samples remains a  
9 challenge due to their low concentrations. Of the known marine siderophores, many are  
10 amphiphilic, consisting of a polar, iron-binding head group and a nonpolar fatty acid  
11 tail<sup>14-19</sup>, which may allow the siderophore to be tethered to the outer cell membrane or  
12 form micelles in an environment where free siderophores might otherwise diffuse away  
13 too quickly to facilitate iron uptake<sup>20,21</sup>. It is common for amphiphilic siderophores to  
14 occur as homologous series with fatty acid tails that differ by CH<sub>2</sub>, such that a single  
15 bacteria may produce a suite of structurally related siderophores.

16  
17 Liquid chromatography (LC) coupled to mass spectrometry offers a sensitive, high  
18 throughput means of detecting and characterizing siderophores. Two complementary  
19 methods have been used to detect organic-metal complexes: inductively coupled plasma  
20 mass spectrometry (ICPMS) and electrospray ionization mass spectrometry (ESIMS)<sup>22-26</sup>.  
21 In the case of ICPMS, as the metal complexes elute from the chromatography column,  
22 the metal is atomized and detected directly as the elemental ion. This is a convenient  
23 method for sensitively detecting metal complexes and estimating their abundances in  
24 biological and environmental samples<sup>8,27-30</sup>. However, ICPMS provides no information  
25 on the structure of the organic complex. Electrospray ionization mass spectrometry  
26 (ESIMS) can be used to detect isotopologues of the intact metal complexes<sup>31-35</sup>. For  
27 metals such as iron that have multiple isotopes, computer algorithms can search LC-  
28 ESIMS mass spectra and identify features that match a metal's characteristic isotope  
29 pattern<sup>33-35</sup>. Furthermore, ESI instruments capable of generating MS<sub>2</sub> spectra of the  
30 parent ion can reveal diagnostic fragmentation patterns that facilitate compound  
31 identification. However, it is difficult to quantify organic-metal complexes using ESIMS,  
32 and search algorithms often miss isotopologues that occur in low abundance or report  
33 false positives due to isobaric interferences from co-eluting compounds<sup>36</sup>.

34  
35 We developed a methodology that combines both LCMS techniques and used it to  
36 investigate the production of siderophores by *Synechococcus* sp. PCC 7002 (Fig. 1). The  
37 metabolites produced by this fast growing model marine cyanobacteria have been the  
38 focus of numerous studies<sup>37-42</sup>. *Synechococcus* sp. PCC 7002 is known to produce a suite  
39 of three siderophores known as synechobactins A, B and C<sup>43-45</sup>. Synechobactins consist  
40 of a citrate head group attached to two 1,3-diaminopropane moieties. The terminal  
41 amines of the diaminopropane moieties are hydroxylated and one is linked to acetic acid  
42 while the other is linked to a fatty acid (Fig. 2). Synechobactins A-C differ by the carbon  
43 number of their fatty acid chains, with synechobactin A having octanoic acid (C<sub>8</sub>), B  
44 having decanoic acid (C<sub>10</sub>), and C having dodecanoic acid (C<sub>12</sub>)<sup>45</sup>. Additional iron  
45 complexes were later detected in the media of *Synechococcus* sp. PCC 7002 with LC-  
46 ICPMS<sup>29</sup>. Here we report the improved separation of *Synechococcus* sp. PCC 7002

1 siderophores and the use of an isotope matching algorithm to assign parent ion masses to  
2 three previously undescribed complexes. By comparing the MS2 fragmentation spectra  
3 of known synechobactins to the unknown complexes, we were able to identify them as  
4 new members of the synechobactin family of siderophores.

## 5 6 **Experimental**

### 7 8 **Materials and reagents**

9  
10 Ultrahigh purity water (18.2 MΩ cm), and LCMS grade methanol (MeOH), ammonium  
11 formate, and formic acid (Optima, Fisher scientific) were used in this study. The  
12 methanol was further purified by sub-boiling-point distillation in a  
13 polytetrafluoroethylene (PTFE) still to reduce Fe contamination<sup>29</sup>. Nutrient salts and  
14 vitamins for culture media were obtained from Sigma Aldrich. Polycarbonate plastic  
15 bottles used for culturing and PTFE vials for sample storage were soaked overnight in  
16 0.1% detergent (Citranox), rinsed 5x with H<sub>2</sub>O, and then soaked in 1 N hydrochloric acid  
17 (J.T. Baker) for 2 days followed by a final 5x rinse with H<sub>2</sub>O. PTFE and platinized  
18 silicone tubing (Cole Parmer) and tube adapters (Visiprep, Sigma Aldrich) used for solid  
19 phase extraction were cleaned by rinsing with 1 N HCl through the tubing for 12 hrs  
20 followed by rinsing with H<sub>2</sub>O for another 12 hours using a peristaltic pump (Cole  
21 Parmer). All samples for LCMS analysis were placed in certified 2 mL amber glass  
22 autosampler vials or 250 μL vial inserts (Agilent).

### 23 24 ***Synechococcus* sp. PCC 7002 culture**

25  
26 Cultures of *Synechococcus* sp. PCC 7002 were grown in polycarbonate bottles in  
27 continuous light at 23°C. A 10 mL inoculum was used to inoculate 500mL of sterile SN  
28 medium<sup>46</sup> containing only 50 nM FeCl<sub>3</sub>\*6H<sub>2</sub>O plus 1 μg/L cyanocobalamin. After seven  
29 days of growth, the culture media was centrifuged to remove cells, filtered (0.2 μm  
30 polyethersulfone sterivex, Millipore), and pumped through an ENV+ resin column (1 g, 6  
31 mL, Biotage) at a flow rate of 10 mL/min. A 500 mL volume of sterile media was  
32 processed as a procedural blank to monitor contamination. The columns were rinsed with  
33 H<sub>2</sub>O to remove salts and eluted with 6 mL of distilled MeOH into PTFE vials. This  
34 organic extract was concentrated to 0.5 mL by evaporation under a stream of nitrogen gas  
35 and stored in the dark at -20°C until analysis by LC-MS. Effort was made to minimize  
36 exposure of the sample to light and avoid photodegradation.

### 37 38 **Liquid chromatography**

39  
40 Organic extracts were separated using an Agilent 1260 series bioinert high pressure  
41 chromatography pump and autosampler fitted with a C18 column (Kinetex 2.1x100mm,  
42 1.7 μm particle size) and polyetheretherketone (PEEK) tubing and connectors. The  
43 mobile phase consisted of (A) 5 mM aqueous ammonium formate or 0.1% formic acid in  
44 H<sub>2</sub>O (B) 5 mM ammonium formate or 0.1% formic acid in distilled MeOH. Optimal  
45 separation and electrospray ionization were achieved using ammonium formate buffer  
46 and a 30 minute solvent gradient from 5% to 100% B followed by a 10 minute isocratic

1 elution in 100% B at a flow rate of 100  $\mu$ L/min. A post column PEEK flow splitter  
2 directed 50% of the flow into the ICP-MS or ESI-MS. Reducing the flow entering the  
3 ICP-MS to 50  $\mu$ L/min eliminated the need for post-column desolvation, even when  
4 eluting with 100% organic phase. On the same column, chromatographic retention times  
5 were reproducible within 0.1 min over several months of analysis.

### 7 **LC-ICPMS conditions**

9 The flow of the LC column was coupled directly to a quadrupole ICPMS (Thermo  
10 ICAPq) using a teflon STD micronebulizer (ESI) and a cyclonic spray chamber cooled to  
11 0°C. Oxygen gas was introduced to the plasma at 25 mL/min to combust organic  
12 solvents and buffers to CO<sub>2</sub> thereby preventing the formation of reduced carbon deposits  
13 that would otherwise accumulate on the cones. The ICPMS was equipped with platinum  
14 sampler and skimmer cones, and was tuned each day using an automated tuning feature  
15 and the 'Tune B' solution purchased from Thermo Scientific. <sup>54</sup>Fe, <sup>56</sup>Fe, <sup>57</sup>Fe, and <sup>59</sup>Co  
16 were monitored with an integration time of 0.05 seconds each. The instrument was run in  
17 KED mode with a He collision gas introduced at a rate of 4.2 mL/min to minimize ArO<sup>+</sup>  
18 interferences on <sup>56</sup>Fe.

### 20 **LC-ESIMS conditions**

22 For LC-ESIMS analysis, the flow from the LC was coupled to a Thermo Scientific  
23 Orbitrap Fusion mass spectrometer equipped with a heated electrospray ionization  
24 source. ESI source parameters were set to a capillary voltage of 3500 V, sheath, auxiliary  
25 and sweep gas flow rates of 12, 6, and 2 (arbitrary units), and ion transfer tube and  
26 vaporizer temperatures of 300°C and 75°C. MS1 scans were collected in high resolution  
27 (450K) positive and negative mode. High energy collision induced dissociation (HCD)  
28 MS2 spectra for the most abundant compounds in each orbitrap scan were collected  
29 simultaneously on the ion trap mass analyzer. Ions were trapped using quadrupole  
30 isolation of a  $\pm 2$  m/z mass window and were then fragmented using an HCD collision  
31 energy of 35%. For the six iron containing compounds investigated in this study, targeted  
32 high resolution MS2 spectra were collected with the orbitrap mass analyzer during a  
33 second analytical run to obtain accurate fragment masses.

### 35 **Data processing**

37 The LC-ESIMS data was converted to mzXML file format (MSconvert, proteowizard)  
38 and iron isotopologue features were identified using a data-mining algorithm (script  
39 written in R using xcmsRaw class<sup>47</sup>). This algorithm searches through each ms scan and  
40 compiles a list of peak sets (retention time, masses, and intensities) that fit a specified  
41 isotope pattern. Both low tolerance and high tolerance (described below) peak ratio filters  
42 were used. The compiled peak sets are then binned by mass in 0.01 m/z increments. Two  
43 criteria are used to automatically remove mass sets that result from instrumental noise  
44 rather than chromatographic peaks: (1) Mass sets that do not appear at least twice within  
45 a 10 second interval are discarded. (2) Mass sets that are found in 8 or more 30 second  
46 intervals are discarded if they do not contain any points above the intensity baseline. The

1 baseline is calculated based on the maximum intensity points from each 30 second time  
2 interval. The highest 25% of intensities are discarded, and the baseline is calculated as  
3 3x the standard deviation plus the mean of the remaining values.

4  
5 The resulting mass list was manually curated by inspecting the extracted ion  
6 chromatograms (EICs) of the putative isotopologues. A constant time offset was applied  
7 to the LC-ESIMS chromatogram in order to align the retention time of the  
8 cyanocobalamin  $[M+H]^{2+}$  peak (EIC of  $m/z = 678$ ) with the cyanocobalamin  $^{59}\text{Co}$  peak at  
9 20.7 min in the LC-ICPMS chromatogram. Iron-containing ions were considered valid if  
10 peaks in the EICs: (1) Have the same retention time and peak shape compared to each  
11 other when the intensities are scaled relative to the expected isotope ratio. (2) Have the  
12 same retention time as the associated peak in the LC-ICPMS spectrum.

## 13 14 **Results and discussion**

### 15 16 **LC-ICP-MS**

17  
18 Tandem LC-ICPMS enables rapid detection of metal organic complexes. Over the course  
19 of the chromatographic separation, siderophores elute from the column and enter the  
20 plasma where the iron is atomized, ionized, and detected by the mass spectrometer as  
21  $^{56}\text{Fe}^+$ . The chromatogram that results from this analysis indicates the retention time and  
22 abundance of each compound that contains iron.

23  
24 LC-ICPMS was used to detect iron-binding compounds extracted from *Synechococcus*  
25 sp. PCC 7002 culture (Fig. 3a). Chromatographic conditions were optimized to achieve  
26 baseline resolution of the compounds which appear as six distinct peaks of iron, with  
27 characteristic retention times of 29.4, 33.9, 35.8, 37.1, 38.5 and 39.6 minutes. This  
28 separation was achieved using a 30-minute gradient from 0-100% methanol using a 5  
29 mM aqueous ammonium formate buffer.

### 30 31 **LC-ESIMS**

32  
33 While ICPMS provides information on the quantity and number of different ligands in  
34 the samples along with their retention times, ESIMS provides complimentary information  
35 on the parent ion mass and fragmentation pattern. Using the same chromatography as  
36 described above, the *Synechococcus* media extract was analyzed by LC-ESIMS. As a soft  
37 ionization technique, ESIMS measures the mass of the intact metal-ligand complexes.

38  
39 Samples were analyzed in both positive and negative ionization modes, with either 0.1%  
40 formic acid or 5 mM ammonium formate as a mobile phase buffer. For the previously  
41 characterized siderophores (synechobactins A, B, and C), the use of 5 mM ammonium  
42 formate in positive mode resulted in the greatest parent ion signal intensity. The apo  
43 (metal-free) form of the siderophore was also detected under these conditions, although  
44 the intensity was <2% relative to the iron-bound form. When 0.1% formic acid was used  
45 in the mobile phase, the parent ion signal decreased by a factor of 10 and the apo form  
46 intensity was 8-11% relative to the iron-bound form. While the intensity ratio between a

1 ligand and its complex in ESI MS spectra does not necessarily reflect concentration ratio  
2 as the two species may ionize differently in an ESI source, these results demonstrate that  
3 ammonium formate is a preferred buffer for detecting iron-bound hydroxamate  
4 siderophores such as synechobactins.

5  
6 To determine the masses of the three major unknown iron containing compounds, an R  
7 based algorithm was used to extract mass spectral features that matched the natural  
8 abundance pattern of iron stable isotopes. Results were manually curated to determine  
9 which matches aligned with one of the six iron peaks detected by LC-ICPMS (Fig. 3).  
10 The choice of isotope pattern criteria balances flexibility and specificity. Ideally, the  
11 selection criteria are flexible enough to account for imprecision in the instrumental  
12 measurement but specific enough to exclude other common isotope patterns and to  
13 reduce the risk of false positives that result from co-eluting ions or from instrumental  
14 noise that coincidentally match the specified isotope pattern.

15  
16 The algorithm used in this study searched each spectral scan for pairs of peaks with a  
17 mass difference of 1.995 m/z and a light/heavy intensity ratio of 0.06, corresponding to  
18 isotopologues containing  $^{54}\text{Fe}$  and  $^{56}\text{Fe}$  respectively. A mass window of  $\pm 3$  mDa was  
19 used to account for uncertainty in the accuracy of the orbitrap analyzer at 450,000 mass  
20 resolution. Since the uncertainty in the isotopologue intensity ratio depends on signal  
21 intensity, detecting low abundance isotope pairs requires a larger range in the intensity  
22 ratio criteria than is required for more abundant compounds. When a narrow ratio  
23 window ( $0.06 \pm 0.015$ ) was used, 40 isotope pairs were identified by the algorithm and 17  
24 of these peaks passed the manual curation step. Using a wide ratio window ( $0.06 \pm$   
25  $0.036$ ), 92 isotope sets were initially identified, and 29 passed the manual curation step.  
26 The most abundant isotope pair found at each retention time was assigned as the parent  
27 ion (Fig. 3, Table 1). Other peak pairs corresponded to  $^{13}\text{C}$  isotopologues as well as  
28 adducts with other abundant coeluting compounds (a complete list is included in the  
29 supplementary information, Fig. SI-1, Table SI-1).

## 30 31 **MS2 Fragmentation**

32  
33 MS2 fragmentation patterns were used to structurally characterize the six synechobactins  
34 detected in this study (Fig. 4). Once the parent ion masses were assigned, the sample was  
35 reanalyzed by LC-ESIMS for targeted high resolution MS2 analysis. An isolation  
36 window of  $\pm 2$  m/z was used to retain the iron isotope patterns of fragments that contain  
37 iron.

38  
39 First, the characteristic fragmentation patterns of synechobactins A, B, and C (m/z  
40 614.261, 586.23, and 558.198) were determined. Structurally, synechobactins consists of  
41 an iron binding head group composed of citrate and two aminopropane moieties linked to  
42 a fatty acid or acetic acid. Most of the major fragments that were observed for the known  
43 synechobactins retain iron (based on the appearance of ions corresponding to the  $^{54}\text{Fe}$  and  
44  $^{56}\text{Fe}$  isotopologues). The MS2 analysis of all three synechobactins displayed a major  
45 fragment with neutral losses of 156.006 [ $-\text{C}_6\text{H}_4\text{O}_5$ ], corresponding to cleavage at the  
46 amide linkage and loss of citrate (Fig. 5a). This fragmentation pattern is consistent with

1 those of other citrate and hydroxamate based siderophores<sup>48-50</sup>. Other major neutral mass  
2 losses correspond to fragmentation within the citrate head group including 46.006 [-  
3 CH<sub>2</sub>O<sub>2</sub>], 113.996 [-C<sub>4</sub>H<sub>2</sub>O<sub>4</sub>], and 130.027 [-C<sub>5</sub>H<sub>6</sub>O<sub>4</sub>]. Several additional fragments  
4 indicate head group cleavages that result in the loss of one of the hydroxamate groups  
5 (Fig. 5b). For synechobactin A, for instance, these fragments occur in pairs that differ by  
6 C<sub>10</sub>H<sub>20</sub> (140.157 m/z), corresponding to the loss of the long fatty acid hydroxamate side  
7 chain (including fragments at 206.035, 270.030, 288.040, 298.025, Fig. 4d) and the  
8 acetohydroxamate respectively (fragments at 346.191, 410.186, 428.198, and 438.181,  
9 Fig. 4d). MS2 spectra of the positive apo (iron free) ion of synechobactin A-B and the  
10 negative iron-bound ion of synechobactin A were also measured (supplementary  
11 information, Fig. SI-2).

12  
13 The fragmentation patterns of synechobactins A-C help in the interpretation of MS2  
14 spectra of the three uncharacterized siderophores. The mass difference between the  
15 unknown compounds and synechobactins A-C correspond to the addition or loss of  
16 (CH<sub>2</sub>)<sub>n</sub> ( $\Delta m/z = n*14$ ). All three uncharacterized siderophores exhibit neutral losses that  
17 are characteristic of fragmentation of a citrate head group including 46.006 [-CH<sub>2</sub>O<sub>2</sub>],  
18 113.996 [-C<sub>4</sub>H<sub>2</sub>O<sub>4</sub>], 130.027 [-C<sub>5</sub>H<sub>6</sub>O<sub>4</sub>], and 156.006 [-C<sub>6</sub>H<sub>4</sub>O<sub>5</sub>]. They also contain the  
19 206.035, 270.030, 288.040, and 298.025 fragments that are attributed to cleavage across  
20 the citrate head group that retains the acetohydroxamate group. Thus, these new  
21 compounds vary by the length of the fatty acid side chain, and correspond to the C<sub>13</sub>, C<sub>15</sub>,  
22 and C<sub>16</sub> synechobactins.

23  
24 The MS2 spectra of the low abundance siderophores with m/z of 600.245 and 628.276  
25 contain additional fragments from coeluting compounds (red masses, Fig. 4). By  
26 obtaining off-peak MS2 scans from the interfering parent ions, we were able to identify  
27 the masses that arise from the interfering ion rather than the siderophore (supplementary  
28 information, Fig. SI-3).

### 29 30 **Connection to synechobactin biosynthesis**

31  
32 The six major compounds detected by LC-ICPMS were all identified as homologues of  
33 synechobactin A. A search for ions in the LC-ESIMS spectra that correspond to other  
34 synechobactin homologues revealed low intensity peaks that potentially correspond to  
35 C<sub>9</sub>-synechobactin and C<sub>16</sub>-synechobactin, although the <sup>54</sup>Fe isotopologues for both were  
36 below detection (supplementary information, Fig. SI-4). These results reveal that  
37 *Synechococcus* sp. PCC 7002 is capable of incorporating a wide range of fatty acids into  
38 the hydroxamate side chain of synechobactin. There appears to be a preference for an  
39 even number of carbons over odd carbon numbers. Additional synechobactins with  
40 longer hydrophobic side chains may have been present in the sample, but were not  
41 recovered from the extraction column by the methanol wash.

42  
43 Such a wide range of siderophores produced by *Synechococcus* sp. PCC 7002 suggests  
44 substrate flexibility for the enzymes involved in synechobactin biosynthesis. The operon  
45 responsible for synechobactin synthesis (SYNPCC7002\_G0019-G0024) is located on a  
46 plasmid that is largely dedicated to the synthesis, uptake, and processing of these

1 compounds, and these genes are upregulated in low iron conditions<sup>51</sup>. Synechobactin  
2 synthesis appears to be similar to the synthesis of rhizobactin, which is structurally  
3 equivalent to synechobactin B (C<sub>10</sub>) with a trans double bond between carbons 2,3 of the  
4 fatty acid side chain (Fig. 2). The production of rhizobactin is thought to involve the  
5 synthesis of schizokinen, which contains two acetohydroxamates<sup>52</sup>, followed by the  
6 replacement of one acetic acid with 2-decenoic acid<sup>53</sup>. It also is possible that  
7 *Synechococcus* sp. PCC 7002 uses schizokinen as an intermediate in synechobactin  
8 synthesis. Indeed, in a culture of *Synechococcus* sp. PCC 7002 that was harvested at an  
9 earlier growth phase, an additional chromatographic <sup>56</sup>Fe peak was observed<sup>29</sup> with m/z  
10 474 that corresponds to schizokinen (supplementary information Fig. SI-5). A putative  
11 lipase that is encoded at the end of the synechobactin biosynthesis operon  
12 (SYNPCC7002\_G0018) may be responsible for non-specifically catalyzing the  
13 attachment of the fatty acid side chain to the synechobactin precursor. A better  
14 understanding of the metabolic pathways involved in catalyzing this reaction will help  
15 reveal how amphiphilic marine siderophores are produced, shed light on how this  
16 functionality evolved, and will improve our ability to search for amphiphilic siderophore  
17 biosynthetic potential in marine genomic/metagenomic libraries.

## 18 19 **Conclusions**

20  
21 The methodology described here provides a means to confidently detect and determine  
22 the masses and MS2 spectra of even minor siderophores in a biological extract with no  
23 prior knowledge of the siderophores produced in a culture. Common siderophore  
24 functional groups can have diagnostic fragmentation patterns that facilitate structural  
25 characterization of new siderophores, as illustrated in this study with the citrate head  
26 group fragmentation of the synechobactins. Using this approach, we were able to  
27 characterize three new synechobactins, and provide preliminary evidence for two  
28 additional homologues that may be present at very low concentrations. As MS2 libraries  
29 of known siderophores grow, our ability to characterize new siderophores based on high  
30 resolution mass spectrometry analysis will improve.

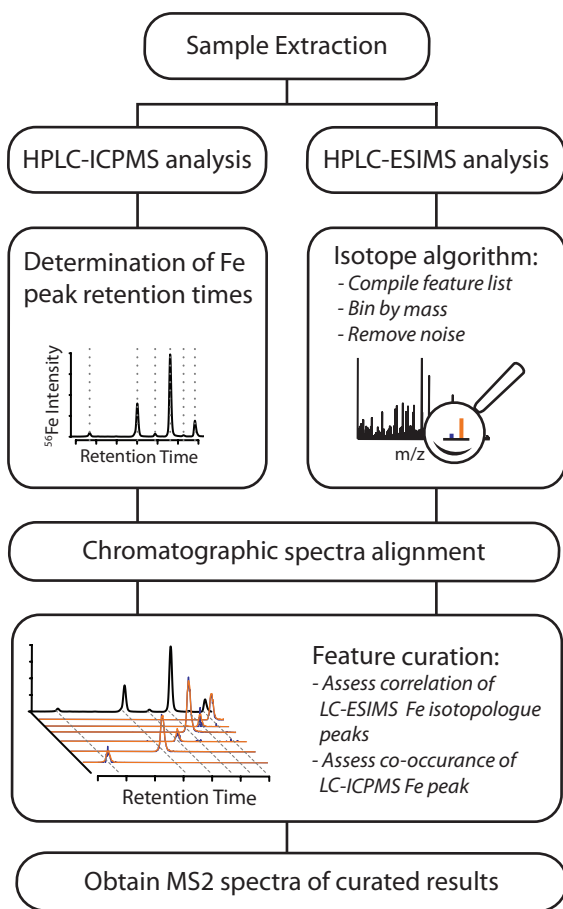
31  
32 The combination of LC-ICPMS and LC-ESIMS has the potential to increase the pace and  
33 depth of siderophore discovery. Looking ahead, this work represents a step towards the  
34 eventual goal of detecting and characterizing metal-binding organic compounds directly  
35 in environmental samples where they impact ecosystems and trace metal cycling.  
36 Organic complexation is thought to largely control iron solubility and bioavailability in  
37 the ocean<sup>54-56</sup>, and understanding the speciation of iron is critical for forecasting  
38 ecosystem changes in low iron regions of the ocean<sup>57-59</sup>. The challenge of characterizing  
39 these marine iron binding molecules (including siderophores<sup>8,60</sup>, heme<sup>61,62</sup>,  
40 polysaccharides<sup>63,64</sup>, and humic substances<sup>65-70</sup>) lies in their low concentrations (often  
41 sub-picomolar for a specific compound) within a very complex organic matrix. Methods  
42 such as the one described in this study have the potential to address these challenges and  
43 reveal the structural diversity of metal organic ligands in environmental samples.

## 44 45 **Acknowledgements**



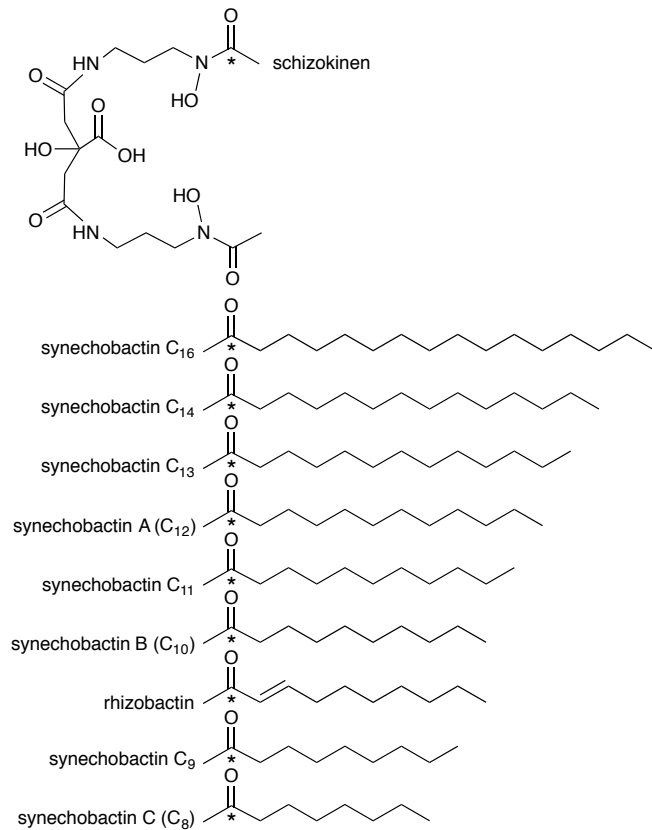
1 We would like to thank Matt McIlvin for his assistance with the orbitrap mass  
2 spectrometer, John Waterbury and Athena Archer for supplying the *Synechococcus* sp.  
3 PCC 7002 strain, and two anonymous reviewers for their constructive comments. This  
4 work was supported by the National Science Foundation program in Chemical  
5 Oceanography (OCE-1356747), and by the National Science Foundation Science and  
6 Technology Center for Microbial Oceanography Research and Education (C-MORE;  
7 DBI-0424599).

8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24



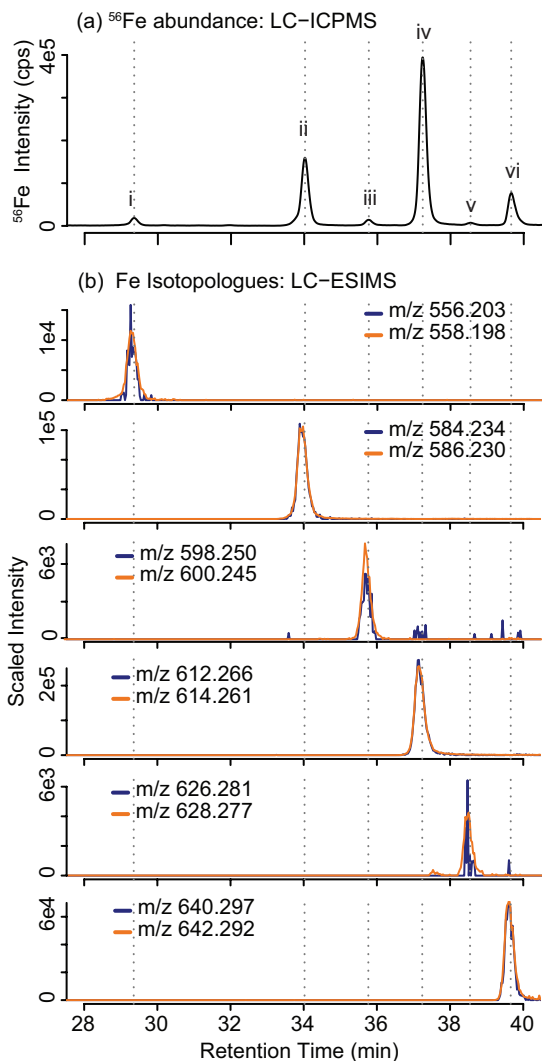
1  
2  
3  
4

**Figure 1:** Workflow for unknown iron compound identification by combined LC-ICPMS-ESIMS.

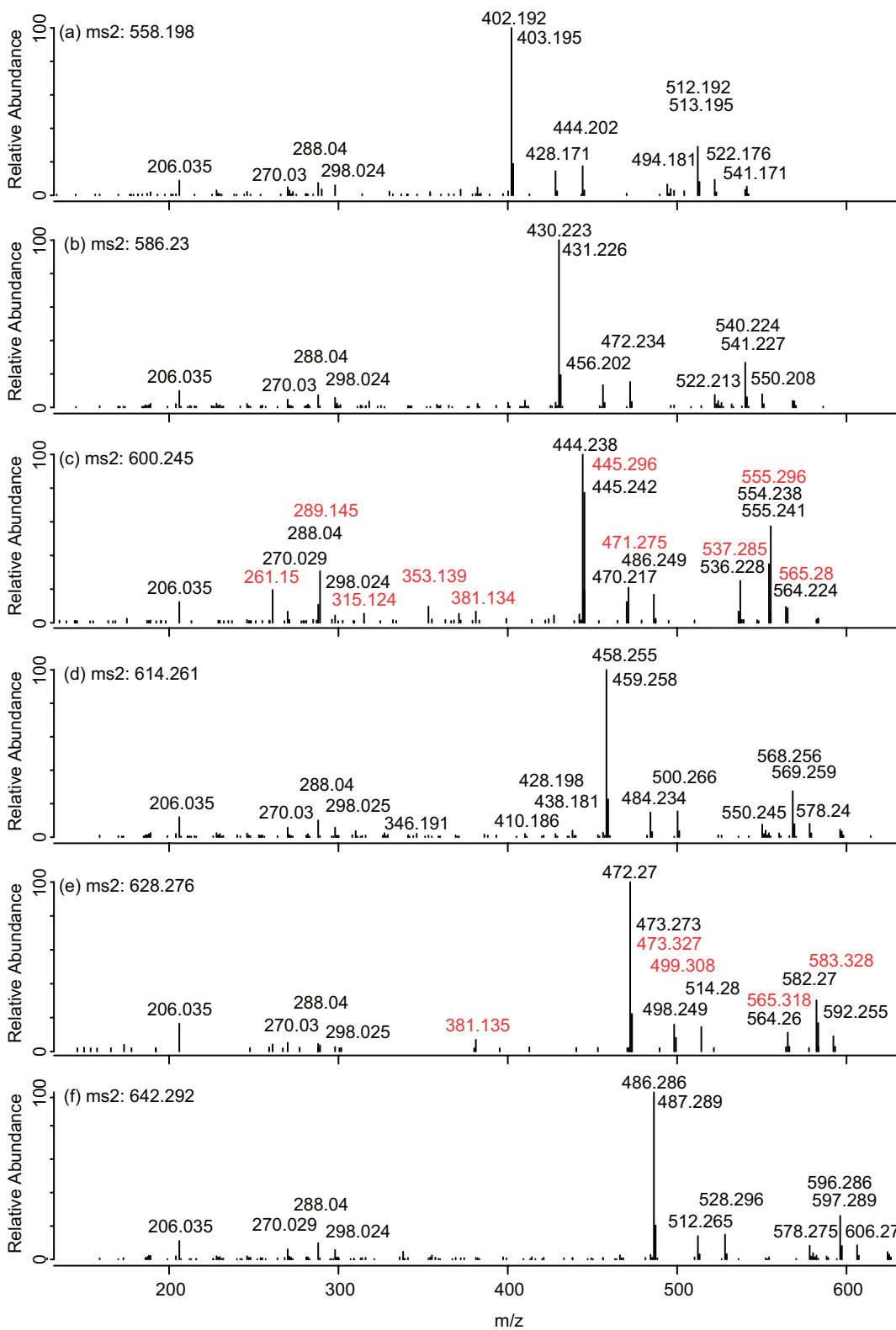


1  
2  
3  
4

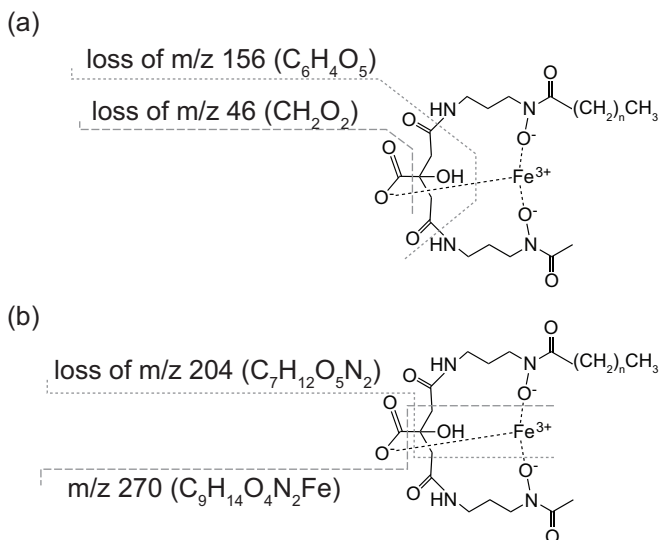
**Figure 2:** Chemical structure of the synechobactins and related compounds. The hydroxamate side chain (indicated by \*) differentiates these compounds.



1  
 2 **Figure 3:** LC-MS chromatograms of *Synechococcus* sp. PCC 7002 media extract. (a)  
 3  $^{56}\text{Fe}$  LC-ICPMS chromatogram. The six labeled peaks correspond to siderophores  
 4 produced by *Synechococcus* sp. PCC 7002 in this experiment (b) extracted ion  
 5 chromatograms from positive mode LC-ESIMS runs. Blue lines correspond to the light  
 6 iron isotopologue  $[\text{M}+^{54}\text{Fe}^{3+}-2\text{H}^+]$  that were identified by the isotope algorithm, and  
 7 orange lines correspond to the heavy iron isotopologue  $[\text{M}+^{56}\text{Fe}^{3+}-2\text{H}^+]$ . The intensity of  
 8 the heavy iron isotopologue has been scaled by the natural abundance ratio of  $^{56}\text{Fe}/^{54}\text{Fe}$   
 9 (divided by 15.7) so that the isotopologues overlap.



1  
2 **Figure 4:** Positive mode MS2 spectra of the six identified Fe containing compounds. (a)  
3 C<sub>8</sub>-Synechobactin, (b) C<sub>10</sub>-synechobactin, (c) C<sub>11</sub>-synechobactin (d) C<sub>12</sub>-synechobactin  
4 (e) C<sub>13</sub>-synechobactin (f) C<sub>14</sub>-synechobactin. Red labels correspond to interferences from  
5 a coeluting ion (see supplementary information Figure 2).



**Figure 5:** Characteristic fragmentation patterns of Fe-synechobactins. (a) Major fragmentations that result in the loss of the citrate head group. (b) Example of fragment pair that results in the symmetric loss of the short and long hydroxamate side chains.

**Table 1:** Siderophores from *Synechococcus* sp. PCC 7002

| Monoisotopic m/z ( $^{56}Fe$ form) | Retention Time (min) | Parent ion formula       | Compound I.D.                   |
|------------------------------------|----------------------|--------------------------|---------------------------------|
| 558.198                            | 29.4                 | $C_{22}H_{38}O_9N_4Fe^+$ | synechobactin C ( $C_8$ )*      |
| 572.214                            | 31.9                 | $C_{23}H_{40}O_9N_4Fe^+$ | synechobactin C <sub>9</sub> †  |
| 586.230                            | 33.9                 | $C_{24}H_{42}O_9N_4Fe^+$ | synechobactin B ( $C_{10}$ )*   |
| 600.245                            | 35.8                 | $C_{25}H_{44}O_9N_4Fe^+$ | synechobactin C <sub>11</sub>   |
| 614.261                            | 37.1                 | $C_{26}H_{46}O_9N_4Fe^+$ | synechobactin A ( $C_{12}$ )*   |
| 628.276                            | 38.5                 | $C_{27}H_{48}O_9N_4Fe^+$ | synechobactin C <sub>13</sub>   |
| 642.292                            | 39.6                 | $C_{28}H_{50}O_9N_4Fe^+$ | synechobactin C <sub>14</sub>   |
| 670.324                            | 41.8                 | $C_{30}H_{54}O_9N_4Fe^+$ | synechobactin C <sub>16</sub> † |

\*Previously described by *Ito and Butler, 2005*

†Putative ID based on detection of  $^{56}Fe$  monoisotopic mass.  $^{54}Fe$  isotopologue and MS2 peaks were below detection limit.

## 1 References

- 2 1. W. G. Sunda, *Front. Microbiol.*, 2012, **3**, 204.
- 3 2. J. K. Moore and S. C. Doney, *Global Biogeochem. Cycles*, 2007, **21**.
- 4 3. F. M. M. Morel, A. J. Milligan, and M. A. Saito, *Treatise Geochemistry, Vol. 6*,  
5 2003.
- 6 4. H. De Baar, P. Boyd, K. Coale, and M. Landry, *J. Geophys. Res.*, 2005, **110**, 1–24.
- 7 5. X. Liu and F. Millero, *Mar. Chem.*, 2002, **77**, 43–54.
- 8 6. V. Braun and H. Killmann, *Trends Biochem. Sci.*, 1999, **24**, 104–9.
- 9 7. C. Wandersman and P. Delepelaire, *Annu. Rev. Microbiol.*, 2004, **58**, 611–47.
- 10 8. E. Mawji, M. Gledhill, J. A. Milton, G. A. Tarran, S. Ussher, A. Thompson, G. A.  
11 Wolff, P. J. Worsfold, and E. P. Achterberg, *Environ. Sci. Technol.*, 2008, **42**,  
12 8675–80.
- 13 9. E. Toulza, A. Tagliabue, S. Blain, and G. Piganeau, *PLoS One*, 2012, **7**, e30931.
- 14 10. B. M. Hopkinson and K. A. Barbeau, *Environ. Microbiol.*, 2012, **14**, 114–28.
- 15 11. D. K. Desai, F. D. Desai, and J. La Roche, *Front. Microbiol.*, 2012, **3**, 362.
- 16 12. M. Li, B. M. Toner, B. J. Baker, J. A. Breier, C. S. Sheik, and G. J. Dick, *Nat.*  
17 *Commun.*, 2014, **5**, 3192.
- 18 13. A. Butler, *Biometals*, 2005, **18**, 369–374.
- 19 14. J. M. Vraspir, P. D. Holt, and A. Butler, *Biometals*, 2011, **24**, 85–92.
- 20 15. J. S. Martinez, J. N. Carter-Franklin, E. L. Mann, J. D. Martin, M. G. Haygood,  
21 and A. Butler, *Proc. Natl. Acad. Sci. U. S. A.*, 2003, **100**, 3754–9.
- 22 16. V. V Homann, M. Sandy, J. A. Tincu, A. S. Templeton, B. M. Tebo, and A.  
23 Butler, *J. Nat. Prod.*, 2009, **72**, 884–8.
- 24 17. J. M. Gauglitz and A. Butler, *J. Biol. Inorg. Chem.*, 2013, **18**, 489–97.
- 25 18. J. D. Martin, Y. Ito, V. V Homann, M. G. Haygood, and A. Butler, *J. Biol. Inorg.*  
26 *Chem.*, 2006, **11**, 633–41.
- 27 19. J. S. Martinez and A. Butler, *J. Inorg. Biochem.*, 2007, **101**, 1692–1698.

- 1 20. J. Martinez, G. Zhang, P. Holt, and H. Jung, *Science*, 2000, **287**, 1245–1247.
- 2 21. G. Xu, J. S. Martinez, J. T. Groves, and A. Butler, *J. Am. Chem. Soc.*, 2002, **124**,  
3 13408–15.
- 4 22. R. Lobiński, D. Schaumlöffel, and J. Szpunar, *Mass Spectrom. Rev.*, 2006, **25**,  
5 255–89.
- 6 23. S. Mounicou, J. Szpunar, and R. Lobinski, *Chem. Soc. Rev.*, 2009, **38**, 1119–38.
- 7 24. Z. Pedrero, L. Ouerdane, S. Mounicou, R. Lobinski, M. Monperrus, and D.  
8 Amouroux, *Metallomics*, 2012, **4**, 473–9.
- 9 25. T. Grevenstuk, P. Flis, L. Ouerdane, R. Lobinski, and A. Romano, *Metallomics*,  
10 2013, **5**, 1285–93.
- 11 26. L. Grillet, L. Ouerdane, P. Flis, M. T. T. Hoang, M. Isaure, R. Lobinski, C. Curie,  
12 and S. Mari, *J. Biol. Chem.*, 2014, **289**, 2515–25.
- 13 27. D. Pröfrock and A. Prange, *Appl. Spectrosc.*, 2012, **66**, 843–68.
- 14 28. O. J. Lechtenfeld, G. Kattner, R. Flerus, S. L. McCallister, P. Schmitt-Kopplin,  
15 and B. P. Koch, *Geochim. Cosmochim. Acta*, 2014, **126**, 321–337.
- 16 29. R. M. Boiteau, J. N. Fitzsimmons, D. J. Repeta, and E. A. Boyle, *Anal. Chem.*,  
17 2013, **85**, 4357–62.
- 18 30. R. Clough, C. F. Harrington, S. J. Hill, Y. Madrid, and J. F. Tyson, *J. Anal. At.*  
19 *Spectrom.*, 2014, **29**, 1158.
- 20 31. I. Velasquez, B. L. Nunn, E. Ibanmi, D. R. Goodlett, K. A. Hunter, and S. G.  
21 Sander, *Mar. Chem.*, 2011, **126**, 97–107.
- 22 32. H. Waska, A. Koschinsky, M. J. Ruiz Chanco, and T. Dittmar, *Mar. Chem.*, 2014,  
23 DOI: 10.1016/j.marchem.2014.10.001.
- 24 33. S. M. Lehner, L. Atanasova, N. K. N. Neumann, R. Krska, M. Lemmens, I. S.  
25 Druzhinina, and R. Schuhmacher, *Appl. Environ. Microbiol.*, 2013, **79**, 18–31.
- 26 34. M. Deicke, J. F. Mohr, J.P. Bellenger, and T. Wichard, *Analyst*, 2014, DOI:  
27 10.1039/c4an01461h.
- 28 35. O. Baars, F. M. M. Morel, and D. H. Perlman, *Anal. Chem.*, 2014, DOI:  
29 10.1021/ac503000e.
- 30 36. M. J. Keith-Roach, *Anal. Chim. Acta*, 2010, **678**, 140–8.



- 1 37. R. Baran, B. P. Bowen, N. J. Bouskill, E. L. Brodie, S. M. Yannone, and T. R.  
2 Northen, *Anal. Chem.*, 2010, **82**, 9034–9042.
- 3 38. R. Baran, B. P. Bowen, and T. R. Northen, *Mol. Biosyst.*, 2011, **7**, 3200–6.
- 4 39. R. Baran, N. N. Ivanova, N. Jose, F. Garcia-Pichel, N. C. Kyrpides, M. Gugger,  
5 and T. R. Northen, *Mar. Drugs*, 2013, **11**, 3617–31.
- 6 40. I. Lanekoff, O. Geydebekht, G. E. Pinchuk, A. E. Konopka, and J. Laskin,  
7 *Analyst*, 2013, **138**, 1971–8.
- 8 41. N. Bennette, J. Eng, and G. Dismukes, *Anal. Chem.*, 2011, **83**, 3808–16.
- 9 42. K. McNeely, Y. Xu, N. Bennette, D. a Bryant, and G. C. Dismukes, *Appl. Environ.*  
10 *Microbiol.*, 2010, **76**, 5032–8.
- 11 43. C. Trick and S. Wilhelm, *Mar. Chem.*, 1995, **50**, 207–217.
- 12 44. S. Wilhelm and D. Maxwell, *Limnol. Oceanogr.*, 1996, **41**, 89–97.
- 13 45. Y. Ito and A. Butler, *Limnol. Oceanogr.*, 2005, **50**, 1918–1923.
- 14 46. J. Waterbury, S. Watson, F. Valois, and D. Franks, *Can Bull Fish Aquat Sci*, 1986,  
15 **214**, 71.
- 16 47. C. A. Smith, E. J. Want, G. O’Maille, R. Abagyan, and G. Siuzdak, *Anal. Chem.*,  
17 2006, **78**, 779–87.
- 18 48. M. Gledhill, *Analyst*, 2001, **126**, 1359–1362.
- 19 49. E. Mawji, M. Gledhill, P. J. Worsfold, and E. P. Achterberg, *Rapid Commun. Mass*  
20 *Spectrom.*, 2008, 2195–2202.
- 21 50. J. M. Gauglitz, H. Zhou, and A. Butler, *J. Inorg. Biochem.*, 2012, **107**, 90–5.
- 22 51. M. Ludwig and D. A. Bryant, *Front. Microbiol.*, 2012, **3**, 145.
- 23 52. K. B. Mullis, J. R. Pollack, and J. B. Neilands, *Biochemistry*, 1971, **10**, 4894–  
24 4898.
- 25 53. D. Lynch, J. O’Brien, T. Welch, P. Clarke, P. O. Cuív, J. H. Crosa, and M.  
26 O’Connell, *J. Bacteriol.*, 2001, **183**, 2576–85.
- 27 54. K. A. Hunter and P. W. Boyd, *Environ. Chem.*, 2007, **4**, 221.
- 28 55. M. Gledhill and K. N. Buck, *Front. Microbiol.*, 2012, **3**, 1–17.

- 1 56. S. M. Kraemer, O. W. Duckworth, J. M. Harrington, and W. D. C. Schenkeveld,  
2 *Aquat. Geochemistry*, 2014, DOI: 10.1007/s10498-014-9246-7.
- 3 57. A. Tagliabue, L. Bopp, O. Aumont, and K. R. Arrigo, *Global Biogeochem. Cycles*,  
4 2009, **23**.
- 5 58. A. Tagliabue and C. Völker, *Biogeosciences*, 2011, **8**, 3025–3039.
- 6 59. D. Shi, Y. Xu, B. M. Hopkinson, and F. M. M. Morel, *Science*, 2010, **327**, 676–9.
- 7 60. E. Mawji, M. Gledhill, J. A. Milton, M. V. Zubkov, A. Thompson, G. A. Wolff,  
8 and E. P. Achterberg, *Mar. Chem.*, 2011, **124**, 90–99.
- 9 61. S. L. Hogle, K. A. Barbeau, and M. Gledhill, *Metallomics*, 2014, **6**, 1107–20.
- 10 62. M. Gledhill, *Anal. Chim. Acta*, 2014, **841**, 33–43.
- 11 63. C. S. Hassler, V. Schoemann, C. M. Nichols, E. C. V. Butler, and P. W. Boyd,  
12 *Proc. Natl. Acad. Sci.*, 2011, **108**, 1076–1081.
- 13 64. C. S. Hassler, L. Norman, C. A. Mancuso Nichols, L. A. Clementson, C.  
14 Robinson, V. Schoemann, R. J. Watson, and M. A. Doblin, *Mar. Chem.*, 2014,  
15 DOI: 10.1016/j.marchem.2014.10.002.
- 16 65. L. M. Laglera and C. M. G. van den Berg, *Limnol. Oceanogr.*, 2009, **54**, 610–619.
- 17 66. R. Yang and C. M. G. van den Berg, *Environ. Sci. Technol.*, 2009, **43**, 7192–7.
- 18 67. S. Batchelli, F. L. L. Muller, K.C. Chang, and C. Lee, *Environ. Sci. Technol.*,  
19 2010, **44**, 8485–90.
- 20 68. L. M. Laglera, G. Battaglia, and C. M. G. van den Berg, *Mar. Chem.*, 2011, **127**,  
21 134–143.
- 22 69. K. Misumi, K. Lindsay, J. K. Moore, S. C. Doney, D. Tsumune, and Y. Yoshida,  
23 *Global Biogeochem. Cycles*, 2013, **27**, 450–462.
- 24 70. R. M. Bundy, H. A. N. Abdulla, P. G. Hatcher, D. V. Biller, K. N. Buck, and K. A.  
25 Barbeau, *Mar. Chem.*, 2014, DOI: 10.1016/j.marchem.2014.11.005.

26