Biogeosciences Discuss., https://doi.org/10.5194/bg-2017-558
Manuscript under review for journal Biogeosciences
Discussion started: 26 January 2018

Biogeosciences Discussions Open Access EGU

Iron triggers colony formation in *Phaeocystis antarctica*: connecting molecular mechanisms with iron biogeochemistry

Sara J. Bender[a,b], Dawn M. Moran[a], Matthew R. McIlvin[a], Hong Zheng[c], John P. McCrow[c], Jonathan Badger[c,f], Giacomo R. DiTullio[e], Andrew E. Allen[c,d], Mak A. Saito[a,*]

[a]Marine Chemistry and Geochemistry Department, Woods Hole Oceanographic Institution, Woods Hole, Massachusetts 02543 USA

[b]Current address: Gordon and Betty Moore Foundation, Palo Alto, California 94304 USA

[c]Microbial and Environmental Genomics, J. Craig Venter Institute, La Jolla, California 92037 USA

[d]Integrative Oceanography Division, Scripps Institution of Oceanography, UC San Diego, La Jolla, California 92037 USA

[e]College of Charleston, Charleston South Carolina 29412, USA

[f]Current address: Center for Cancer Research, Bethesda, Maryland 20892, USA

[*]Correspondence to M. Saito (msaito@whoi.edu)

Submitted to Biogeosciences

December 25, 2017

Biogeosciences Discuss., https://doi.org/10.5194/bg-2017-558
Manuscript under review for journal Biogeosciences
Discussion started: 26 January 2018

Abstract.

Phaeocystis antarctica is an important phytoplankter of the Ross Sea where it dominates the early season bloom after sea ice retreat and is a major contributor to carbon export. The factors that influence *Phaeocystis* colony formation and the resultant Ross Sea bloom initiation have been of great scientific interest, yet there is little known about the underlying mechanisms responsible for these phenomena. Here, we present laboratory and field studies on *Phaeocystis antarctica* grown under multiple iron conditions using a coupled proteomic and transcriptomic approach. *P. antarctica* had a lower iron limitation threshold than a Ross Sea diatom *Chaetoceros* sp., and at increased iron nutrition (>120 pM Fe') a shift from flagellate cells to a majority of colonial cells in *P. antarctica* was observed, implying a role for iron as a trigger for colony formation. Proteome analysis revealed an extensive and coordinated shift in proteome structure linked to iron availability and life cycle transitions with 327 and 436 proteins significantly different between low and high iron in strains 1871 and 1374, respectively. The enzymes flavodoxin and plastocyanin that can functionally replace iron metalloenzymes were observed at low iron treatments consistent with cellular iron sparing strategies, with plastocyanin being more dynamic in range. The numerous isoforms of the putative iron-starvation induced protein ISIP group (ISIP2A and ISIP3) had abundance patterns coincided with that of either low or high iron (and coincident flagellate or the colonial cell types in strain 1871), implying that there may be specific iron acquisition systems for each life cycle type. The proteome analysis also revealed numerous structural proteins associated with each cell type: within flagellate cells actin and tubulin from flagella and haptonema structures as well as a suite of calcium-binding proteins with EF domains were observed. In the colony-dominated samples a variety of structural proteins were observed that are also often found in multicellular organisms including spondins, lectins, fibrillins, and glycoproteins with von

Willebrand domains. A large number of proteins of unknown function were identified that became abundant at either high and low iron availability. These results were compared to the first metaproteomic analysis of a Ross Sea *Phaeocystis* bloom to connect the mechanistic information to the *in situ* ecology and biogeochemistry. Proteins associated with both flagellate and colonial cells were observed in the bloom sample consistent with the need for both cell types within a growing bloom. Bacterial iron storage and B_{12} biosynthesis proteins were also observed consistent with chemical synergies within the colony microbiome to cope with the biogeochemical conditions. Together these responses reveal a complex, highly coordinated effort by *P. antarctica* to regulate its phenotype at the molecular level in response to iron and provide a window into the biology, ecology, and biogeochemistry of this group.

1. Introduction

The genus *Phaeocystis* is a cosmopolitan marine phytoplankton group that plays a key role in global carbon and sulfur cycles (Hamm et al., 1999; Matrai et al., 1995; Rousseau et al., 2007; Schoemann et al., 2005; Smith et al., 1991; Solomon et al., 2003; Thingstad and Billen, 1994; Verity et al., 2007). Because of their large cell concentrations during bloom formation, *Phaeocystis* have a significant impact on the ocean biogeochemistry through carbon fixation (Arrigo et al., 1999; Hamm et al., 1999; Matrai et al., 1995; Rousseau et al., 2007; Schoemann et al., 2005; Smith et al., 1991; Solomon et al., 2003; Thingstad and Billen, 1994; Verity et al., 2007), the release of large concentrations of organic carbon upon grazing/viral lysis (Alderkamp et al., 2007; Hamm et al., 1999; Lagerheim, 1896; Verity et al., 2007), and export as aggregates out of the photic zone (DiTullio et al., 2000). Through the production of dimethylsulfide (DMS), they also directly connect ocean and atmospheric processes and carbon and sulfur cycling (Smith et al., 2003).

Biogeosciences Discuss., https://doi.org/10.5194/bg-2017-558
Manuscript under review for journal Biogeosciences
Discussion started: 26 January 2018

Some *Phaeocystis* species, including *Phaeocystis antarctica*, undergo multiple morphotypes and can occur as flagellated single-cells or in gelatinous colonies consisting of thousands of non-motile cells (Fig. 1). Microscopic and chemical analyses have found that *Phaeocystis* colonies are filled with a mucilaginous matrix surrounded by a thin, but strong, hydrophobic skin(Hamm, 2000; Hamm et al., 1999). Once formed, cells typically associate with this outer layer of the colony (Smith et al., 2003). Colony formation involves the exudation of (muco)polysaccharides and carbohydrate-rich dissolved organic matter, as well as amino sugars and amino acids; it is estimated that approximately 50 – 80% of *Phaeocystis* carbon is allocated to this extracellular matrix (Hamm et al., 1999; Matrai et al., 1995; Rousseau et al., 2007; Solomon et al., 2003; Thingstad and Billen, 1994). Thus, not only does the colony increase the size of *Phaeocystis* by several orders of magnitude, but the extracellular matrix material also constitutes the majority of measured algal (carbon) biomass (Rousseau et al., 1990). The colonial form of *Phaeocystis* has been suggested as a defense mechanism against grazers (Hamm et al., 1999), a means to sequester micronutrients such as iron and manganese (Noble et al., 2013; Schoemann et al., 2001), as a means of protection from pathogens (Hamm, 2000; Jacobsen et al., 2007), and as a microbiome vitamin B_{12} source (Bertrand et al., 2007). Colony formation of *Phaeocystis* species, including *P. antarctica* and *P. globosa*, has been linked to numerous physiological triggers including the synergistic effects of iron and irradiance (Feng et al., 2010), grazer-induced chemical cues (Long et al., 2007), phosphate concentrations (Riegman et al., 1992), and the presence of different nitrogen species (Riegman and van Boekel, 1996; Smith et al., 2003).

The Ross Sea is one of the most productive regions of the Southern Ocean (Arrigo et al., 1999; 1998; Feng et al., 2010; Garcia et al., 2009; Sedwick and DiTullio, 1997), and the latter is an important contributor to the cycling of carbon in the oceans (Lovenduski et al., 2008; Sarmiento

et al., 1998). In the early spring when the sea ice retreats and polynyas form, phytoplankton blooms and regional phytoplankton productivity are fed by the residual winter iron inventory and perhaps iron-rich sea ice melt (Noble et al., 2013; Sedwick and DiTullio, 1997); blooms have also been linked to changes in irradiance and mixed layer depth (Arrigo et al., 1999; Coale et al., 2003; Martin et al., 1990; Sedwick and DiTullio, 1997; Sedwick et al., 2000). In the Ross Sea Polynya (RSP), *P. antarctica* colonial cells form almost mono-specific blooms until the austral summer season begins, comprising > 98% of cell abundance at the peak of the bloom (Smith et al., 2003). Although diatom abundance dominates in the summer, the RSP typically harbors the co-existence of flagellated single cells of *P. antarctica* along with diatoms (Garrison et al., 2003). During blooms *P. antarctica* can draw down more than twice as much carbon relative to phosphate as diatoms and contribute to rapid carbon export, leaving a lasting biogeochemical imprint on surrounding waters (Arrigo et al., 1999; 2000; DiTullio et al., 2000; Dunbar et al., 1998). Recent *in vitro* iron addition experiments provide evidence that iron nutrition influences *P. antarctica* growth in this region, with increasing *P. antarctica* biomass in incubation experiments (Bertrand et al., 2007; Feng et al., 2010). Moreover, laboratory experiments with *P. antarctica* have observed a high cellular iron requirement and evidence for use of strong organic iron complexes (Sedwick et al., 2007; Strzepek et al., 2011).

The multiphasic lifecycle of *P. antarctica* in the Ross Sea gives it a spectrum of nutrient drawdown phenotypes and trophic interactions, dependent on the presence of flagellated versus colonial cells (Smith et al., 2003). Given its prominence during early spring sea ice retreat, it has been hypothesized that the triggers of colony formation for *Phaeocystis* cells are also the triggers of the spring phytoplankton bloom. Yet experimental and molecular analyses of potential environmental triggers and how they manifest in changes in cellular morphology have remained

elusive. Little is known about the mechanisms responsible for colony formation in *P. antarctica*, nor how these mechanisms respond to an environmental stimulus such as iron, both of which appear to be integral to the ecology and biogeochemistry of *P. antarctica.*

2. Materials and methods

2.1 Culture experiments

Two strains of *Phaeocystis antarctica* (treated with Provasoli's antibiotics), CCMP 1871 and CCMP 1374 (Provasoli-Guillard National Center for Culture of Marine Phytoplankton), and a Ross Sea centric diatom isolate *Chaetoceros* sp. RS-19 (collected by M. Dennett at 76.5° S, 177.1° W in December 1997 and isolated by D. Moran) were grown in F/2 media with a trace metal stock (minus $FeCl_3$) according to Sunda and Huntsman (Sunda and Huntsman, 2003; 1995), using a modified 10 µM EDTA concentration, and an oligotrophic seawater base. Strains were chosen because they were culturable representatives from two distinct regions in the Southern Ocean.

Semi-continuous batch cultures were grown at 4 ºC under 200 µmol photons m^{-2} s^{-1} continuous light. Each strain was acclimated to growth on one of six growth condition concentrations: the concentration of dissolved inorganic iron within each treatment was 2 pM, 41 pM, 120 pM, 740 pM, 1200 pM, and 3900 pM Fe' as set by the metal buffer EDTA (where $Fe'/Fe_{Total} = 0.039$) (Sunda and Huntsman, 2003). During the experiment, cultures were maintained in 250 mL polycarbonate bottles; and, subsamples were collected every 1-2 days in 5 mL 13x100 mm borosilicate tubes to measure relative fluorescence units (RFUs) and cell counts in the treatments. Mid-to-late exponential phase cultures were harvested for transcriptome and proteome analysis and cell size was measured for both strains; cell pellets were stored at -80 ºC (see Supplementary Information for additional methods). Cell counts were conducted using a Palmer-Maloney counting chamber and a Zeiss Axio Plan microscope on 400x magnification; cell

numbers were used to determine the final growth rate of each strain/treatment. During mid-to-late exponential phase (time-of-harvest), cell size was determined for both strains (n=20 cells were counted for each strain), calculated using the Zeiss 4.8.2 software and a calibrated scale bar. The number of cells in colonies (versus as single cells) was determined for strain 1871 only. Briefly, counts (number of cells associated with colonies versus unassociated) were averaged from 10 fields of view at five distinct time points (50 fields of view total).

2.2 Protein extraction, digestion, and mass spectrometry analyses

Proteins from cell pellets (one pellet per treatment, two strains and six iron treatments for a total of 12 proteomes) was extracted using the detergent B-PER (Thermo Scientific), quantified, purified by immobilization within an acrylamide tube gel, trypsin digested, alkylated and reduced, and analyzed by liquid chromatography-mass spectrometry (LC-MS) using a Michrom Advance HPLC with a reverse phase C18 column (0.3 x 10 mm ID, 3 µm particle size, 200 Å pore size, SGE Protocol C18G; flow rate of 1 µL min^{-1}, nonlinear 210 min gradient from 5% to 95% buffer B, where A was 0.1% formic acid in water and B was 0.1% formic acid in CAN, all solvents were Fisher Optima grade) coupled to a Thermo Scientific Q-Exactive Orbitrap mass spectrometer with a Michrom Advance CaptiveSpray source. The mass spectrometer was set to perform MS/MS on the top 15 ions using data-dependent settings (dynamic exclusion 30 s, excluding unassigned and singly charged ions), and ions were monitored over a range of 380-2000 m/z (see Supplementary Information for detailed protocol). Peptide-to-spectrum matching was conducted using the SEQUEST algorithm within Proteome Discoverer 1.4 (Thermo Scientific) using the translated transcriptomes for *P. antarctica* strain 1871 and strain 1374 (see below). Normalized spectral counts were generated using Scaffold 4.0 (Proteome Software Inc.), with a protein false discovery

rate (FDR) of 1.0%, a minimum peptide score of 2, and a peptide probability threshold of 95%. Spectral counts refer to the number of peptide-to-spectrum matches that are attributed to each predicted protein from the transcriptome analysis, and the Scaffold normalization scheme involves a small correction normalizing the total number of spectra counts across all samples to correct for run-to-run variability and improve comparisons between treatments. The R package "FactoMineR" (Lê et al., 2008) was used for the PCA analysis; for heatmaps, the package "gplots" was used (Warnes et al., 2009). Proteomic samples taken from each laboratory condition were not pooled downstream as part of the analyses; replicates shown for each treatment are technical replicates.

2.3 RNA extraction, Illumina sequencing, and annotation

For *P. antarctica* cultures total RNA was isolated from cell pellets (one pellet per treatment, two strains and three iron concentrations for a total of six transcriptome) following the TRIzol Reagent (Life Technologies, manufacturer's protocol). RNeasy Mini kit (Qiagen) was used for RNA cleanup, and DNase I (Qiagen) treatment was applied to remove genomic DNA. Libraries, from polyA enrichment mRNA, were constructed using a TruSeq RNA Sample Preparation Kit V2 (IlluminaTM), following the manufacturer's TruSeq RNA Sample Preparation Guide. Sequencing was performed using the Illumina HiSeq platform. Downstream, reads were trimmed for quality and filtered. CLC Assembly Cell (CLCbio) was used to assemble contigs, open reading frames (ORFs) were predicted from the assembled contigs using FragGeneScan(Rho et al., 2010), and additional rRNA sequences were removed. The remaining ORFs were annotated de novo via KEGG, KO, KOG, Pfam, and TigrFam assignments. Taxonomic classification was assigned to each ORF and the Lineage Probability Index (LPI, as calculated in (Podell and Gaasterland, 2007).

ORFs classified as Haptophytes were retained for downstream analyses. Analysis of sequence counts ("ASC") was used to assign normalized fold change and determine which ORFs were significantly differentially expressed in pairwise comparisons between treatments. The ASC approach offers a robust analysis of differential gene expression data for non-replicated samples (Wu et al., 2010).

For metatranscriptomes, RNA was extracted from frozen cell pellets using the TRIzol reagent manufacturer's protocol (Thermo Fisher Scientific) (see Supplementary Information for additional details on metatranscriptome processing).

2.4 Ross Sea *Phaeocystis* bloom: sample collection and protein extraction and analysis

The meta 'omics samples were collected in the Ross Sea (170.76° E, 76.82° S) during the CORSACS expedition (Controls on Ross Sea Algal Community Structure) on December 30, 2005 (near pigment station 137; http://www.bco-dmo.org/dataset-deployment/453377) (Saito et al., 2010; Sedwick et al., 2011). Surface water was concentrated via a plankton net tow (20 µm mesh), gently decanted of extra seawater, then split into multiple replicate cryovials and frozen in RNAlater at -80 °C for metatranscriptome and metaproteome analysis. The pore size of the net tow would have preferentially captured the colony form of *Phaeocystis*, although filtration with small pore size membrane filters was particularly challenging during this time period due to the abundance of *Phaeocystis* colonies and the clogging effect of their mucilage. Moreover, the physical process of deploying the net tow appears to have entrained some smaller cells including the *Phaeocystis* flagellate cells by adsorption to partially broken colonies and associated mucilage as observed in the metaproteome results. In the lab, two of these replicate bloom samples were

frozen for proteome analysis. A third replicate sample from this field site was extracted for metatranscriptome analysis as described above.

Proteins were extracted, digested, and purified following the lab methods above, and then identified on a Q-Exactive Orbitrap mass spectrometer using a Michrom Advance CaptiveSpray source. Proteins were then identified within the mass spectra using three databases: the translated transcriptome database for both *Phaeocystis* strains (Database #1), a Ross Sea metatranscriptome generated in parallel from this metaproteome sample (Database #2; this transcriptome is a combination of eukaryotic and prokaryotic communities derived from total RNA and poly(A) enriched RNA sequencing), and a compilation of five bacterial metagenomes from the Amundsen Sea polynya (Database #3) (Delmont et al., 2014), using SEQUEST within Proteome Discoverer 1.4 (Thermo Scientific)(Eng et al., 1994) and collated with normalized spectral counts in Scaffold 4.0 (Proteome Software Inc.) (see Supplementary Information for additional details).

2.5 Data availability

Phaeocystis antarctica RNA sequence data reported in this paper have been deposited in the NCBI sequence read archive under BioProject accession no. PRJNA339150, BioSample accession nos. SAMN05580299 – SAMN05580303. Ross Sea metatranscriptomes have been deposited under BioProject accession no. PRJNA339151, BioSample accession nos. SAMN05580312 – SAMN05580313. Proteomic data from the lab and field components was submitted to the Pride database (Project Name: Phaeocystis antarctica CCMP 1871 and CCMP 1374, Ross Sea Phaeocystis bloom, LC-MSMS; Project accession: PXD005341; Project DOI: 10.6019/PXD005341).

Biogeosciences Discuss., https://doi.org/10.5194/bg-2017-558
Manuscript under review for journal Biogeosciences
Discussion started: 26 January 2018

3. Results and discussion

3.1 Physiological response to iron availability: Growth limitation and colony formation

The two strains of *P. antarctica* (1374 and 1871 hereon) were acclimated to six iron concentrations to capture the metabolic response under different iron regimes (Fig. 2*a* and *b*). A biphasic response in *P. antarctica* strain 1871 was observed; cultures exhibited a clear single-cell versus colony response to low and high iron, respectively: the three low iron treatments (2 pM, 41 pM, and 120 pM Fe') cultures contained only single, flagellated cells, whereas the three higher iron treatments (740 pM, 1200 pM, and 3900 pM Fe') had a majority of colonial cells, based on detailed microscopy counts shown in Fig. 2*c*. The presence of both colony and flagellate cells is expected in actively growing populations with colonies since reproduction is thought to require revisiting the flagellate life cycle stage. Single cells and colonies were not counted in experiments with strain 1374, as these experiments were conducted prior to those of 1871 and the iron-induced colony formation observations therein. However, strain 1374 was observed to become "clumpy" at high iron. This clumping observation may reflect the loss of a specific factor needed for the colony completion lost during long-term maintenance in culture. This interpretation is consistent with the overall similar structural protein expression patterns observed in both strains described below. Strzepek et al. also observed co-varying of iron concentration and colony formation in some strains of *P. antarctica* (Strzepek et al., 2011).

The two strains of *P. antarctica* were able to maintain growth rates for all but the lowest of iron concentrations used here, similar to prior studies of *P. antarctica* strain AA1 that observed no effect of scarce iron on growth rates (Strzepek et al., 2011). Parallel experiments with polar diatoms such as *Chaetoceros* (Fig. 2*d*) observed growth limitation at moderate iron abundances using an identical media composition, indicating 1) that *P. antarctica* has an impressive capability

for tolerating low iron compared to *Chaetoceros* and other diatoms (e.g. a Ross Sea *Pseudo-nitzschia* sp. isolate, data not shown), and 2) demonstrating an absence of iron contamination in these experiments. Growth rates for 1871 were significantly different between the 2 pM Fe' treatment and all other treatments (student's t-test with Bonferroni correction, $p < 0.05$; Fig. 2*a*); there were no significant differences among growth rates for strain 1374 (Fig. 2*b*). Cell size (including both flagellate and colonial cells) decreased with lower iron concentration, a trend that was statistically significant (ANOVA with TukeyHSD test, $p < 0.05$) for both strains when cell sizes from each high iron treatment (740 pM, 1200 pM, and 3900 pM Fe') were compared to cell sizes from each low iron treatment (2 pM, 41 pM, and 120 pM Fe') (Fig. 2*e* and *f*).

3.2 Molecular response to low and high iron concentrations

Global proteomics enabled by peptide-to-spectra matching to transcriptome analyses, revealed a clear statistically significant molecular transition across the iron gradient for each strain (Fig. 3). The global proteome consisted of 536 proteins identified in strain 1871 and 1085 proteins identified in strain 1374 (Table 1; Supplementary Data 1), after summing unique proteins across the six iron treatments. There were 55 proteins identified in strain 1871 and 64 proteins in strain 1374 (Fig. 3) that drove the statistical separation of proteomes across iron treatments using principle component analysis (PCA, Axis 1 PCA correlation coefficient ≥ 0.9 or ≤ -0.9). Axis one accounted for 49% variance for 1871 and 36% variance for 1374. Moreover, using a Fisher Test (P-value ≤ 0.05), 327 proteins (strain 1871) and 436 proteins (strain 1374) were identified as significantly different in relative protein abundance between representative "low" (41 pM Fe') and "high" (3900 pM Fe') iron treatments. This significant change in the proteome composition paralleled observations of a shift from flagellate to colonial cells. Iron-starvation responses and

iron metabolism were detected within the high and low iron PCA protein subsets, including iron-starvation induced proteins (ISIPs), flavodoxin, and plastocyanin, demonstrating a multi-faceted cellular response to iron scarcity (Fig. 4). Surprisingly, there was also a highly pronounced signal in the proteome that appeared to reflect the structural changes occurring in *P. antarctica*. These structural proteins included multiple proteins with protein family (PFam) domains suggestive of extracellular function, adhesion, and/or ligand binding, including putative glycoprotein domains (for example, spondin) that were present in the high iron PCA subset in both strains (Fig. 4); the appearance of these proteins also corresponded to the occurrence of colonies in strain 1871 (Fig. 1). Similarly, a distinct suite of proteins was more abundant in the low iron PCA subset (Fig. 4), including proteins relating to cell signaling (for example, calmodulin/EF-hand, PHD zinc ring finger). A number of proteins with unknown function were also detected in the PCA subsets: 71% unknown for strain 1871 and 42% unknown for strain 1374. Outside of the PCA analyses, additional iron and adhesion-related proteins were identified that demonstrated a similar expression profile to the PCA subset (Supplementary Fig. 1).

Identification and characterization of proteins and transcripts induced by iron scarcity are valuable in improving an understanding of the adaptive biochemical function of these complex phytoplankton as well as for their potential utility for development as environmental stress biomarkers (Roche et al., 1996; Saito et al., 2014). The enzymes flavodoxin and plastocyanin, which require no metal and copper, respectively and that functionally replace iron metalloenzymes counterparts ferredoxin and cytochrome c6, had isoforms that increased in concentrations at the lower iron treatments consistent with cellular iron sparing strategies (Fig. 5, Supplementary Fig. 2) (Peers and Price, 2006; Whitney et al., 2011; Zurbriggen et al., 2008). In strain 1374 however, there was an increase in both of these iron-sparing systems at the highest iron concentration (Fig.

295 5d and 5f, Supplementary Table 1). While both experiments were in exponential growth at the
296 time of harvest, those of strain 1374 were as much as 7.6 fold denser in cell number than those of
297 strain 1874 (based on cell counts from treatments specifically used for transcriptome analyses),
298 and as a result the denser 1374 strain appears to have also experienced iron stress even at this
299 highest iron concentration as the high biomass depleted iron within the medium. Of these two iron
300 sparing enzymes, plastocyanin appeared to show a clearer increase in abundance at lower
301 environmental iron concentrations (Fig 5c and 5f). In contrast, some flavodoxin isoforms could be
302 interpreted as being constitutive, two of the three isoforms were still present in reasonable spectral
303 counts at higher iron concentrations (Figs. 5a and 5d). Prior measurements during a Ross Sea
304 colonial *P. antarctica* spring bloom in 1998 were consistent with this where ferredoxin
305 concentrations were below detection and flavodoxin present (Maucher and DiTullio, 2003). A
306 constitutive flavodoxin could help explain *P. antarctica's* ability to tolerate all but the lowest iron
307 treatment observed in the physiological experiments (Fig 2a and 2b), and implies that the careful
308 selection of isoforms, or better, the inclusion of all isoforms of a protein biomarker of interest may
309 be valuable in interpreting complex field results.

310 There were also numerous isoforms of the iron-starvation induced proteins (ISIP) group
311 identified within the proteome of each *P. antarctica* strain: specifically 9 ISIP2A's and 3 ISIP3's
312 in strain 1871 and 3 ISIP2A's and 4 ISIP3's in strain 1374 (Supplementary Fig. 1; Supplementary
313 Table 1). These ISIPs were identified based on their transcriptome response to iron stress in
314 diatoms and most recently have been implicated in a diatom cell surface iron concentrating
315 mechanism (Allen et al., 2008; Morrissey et al., 2015). Interestingly in this *P. antarctica*
316 experiment, these ISIPs exhibited both "high" or "low" iron responses, where specific isoforms
317 were more abundant only under one of those respective conditions (Fig. 5.). Given the

metamorphosis of *P. antarctica* between flagellate and colonial cell types observed by microscopy and the proteome across the gradient in iron, we hypothesize that this diversity of iron stress responses in the ISIP proteins may reflect the complexity associated with *P. antarctica*'s life cycle. As the abundant winter iron and sloughed basal sea ice reserves are depleted, newly formed colonial cells will inevitably find themselves in the iron-depleted environments that have been characterized in the Ross Sea almost immediately upon bloom formation due to iron's small dissolved inventory (Bertrand et al., 2015; Sedwick et al., 2011). As a result, *P. antarctica* may have distinct iron stress protein isoforms associated specifically with the colonial cell type (such as the high iron/colonial ISIP proteins, Figs. 4 and 5) in order to acquire scarce iron during blooms, in addition to a distinct suite of iron stress proteins produced within the flagellate cells (low iron/flagellate ISIP proteins, also Figs. 4 and 5). Given the rapid depletion of iron during Ross Sea blooms, it is also conceivable that these iron acquisition proteins are constitutive within the colony morphotype, rather than being connected to an iron-sensing and regulatory response system. Future short-term iron perturbation studies that would complement the steady-state experiments presented here could further investigate this hypothesis. The multiplicity of ISIP proteins produced within each strain also is consistent with the observation that both *P. antarctica* strains maintained high growth rates even at the lower 41 and 120 pM Fe' concentrations, compared to the diatom *Chaetoceros sp.* whose growth rate is less than 50% of maximal growth in similar media (Fig. 2).

3.3 Correspondence between RNA and protein biomolecules

Many of the RNA transcripts of iron-related genes trended with their corresponding proteins: 60% of the iron-related gene transcripts reflected the proteomic response in strain 1871, whereas there was a 30% correspondence between iron-related transcripts and proteins in strain

341 1374 (Supplementary Fig. 1). In total, 47% of expressed proteins in strain 1871 and 26% of
342 proteins in strain 1374 shared expression patterns with associated transcripts (Fig. 6), consistent
343 with recent studies of proteome-transcriptome comparisons that showed limited coordination
344 between inventories of each type of biomolecule (Dyhrman, 2012). As mentioned above, while
345 both experiments were in exponential growth at the time of harvest, strain 1374 was 7.6 fold denser
346 in cell number than those of strain 1874 at that time. Hence, this decrease in transcript-proteome
347 coherence in strain 1374 may be related to harvesting in late-log growth phase, and reflects the
348 challenge of trying to conduct comparisons of these biomolecules that function on different cellular
349 timescales.

350 Examination of the transcriptome revealed a significant increase in transcripts for tonB-
351 like transporters, which can be associated with cross-membrane nutrient transport (e.g. for iron
352 siderophores complexes or vitamin B_{12} (Bertrand et al., 2007; 2013; Morris et al., 2010) under high
353 iron for strain 1871; and, significantly greater transcript abundances for a putative flavodoxin for
354 strain 1374 under low iron consistent with its substitution for ferredoxin due to iron scarcity (Roche
355 et al., 1996).

356

357 **3.4 Observation of an iron-induced switch from single cells to colonies**

358 The strong connection of iron availability to putative structural components of *P.*
359 *antarctica* observed here served as an ideal opportunity to examine the genes and proteins involved
360 in morphological and life cycle transitions and colony construction in this phytoplankter that can
361 otherwise be experimentally difficult to trigger in isolation. *Phaeocystis* colonies have captured
362 the interest of scientists for more than a century (Hamm et al., 1999), yet next to nothing is known
363 about the molecular basis of their construction. Colonies have been considered a collection of

364 loosely connected cells embedded within a gel matrix, and hence described as "balls of jelly" or
365 "bags of water" (Hamm et al., 1999; Lagerheim, 1896; Verity et al., 2007). Results here suggest
366 significant transformations in the cellular proteome that corresponded to solitary and colonial
367 morphological stages, for example, involving structural proteins and proteins known to be post-
368 translationally modified such as glycoproteins or those containing glycoprotein-binding motifs. To
369 our knowledge, such an extensive proteome remodeling has yet to be observed for another colonial
370 organism, nor with the influence of any other environmental stimuli in the genus *Phaeocystis*. As
371 a result the details of this response, while fascinating, are challenging to interpret due to their
372 novelty.

373 A putative spondin protein exhibited one of the largest responses between low and high
374 iron in both strains with a greater than 20-fold increase in relative protein abundance and
375 normalized 11-fold change in transcript abundance in strain 1871, and a greater than 9-fold
376 increase in relative protein abundance and 3-fold change in transcript abundance in strain 1374
377 (Fig. 4*a* and Supplementary Data 1). Spondin proteins are known to be glycosylated, and to be a
378 component of the extracellular matrix (ECM) environment, which may enable multicellularity in
379 metazoans through cell adhesion, and have been found to help coordinate nerve cell development
380 through adhesion and repulsion (Michel et al., 2010; Tzarfati-Majar et al., 2011). Despite this large
381 variation in protein abundance, the function of spondins in eukaryotic phytoplankton, including
382 *Phaeocystis* remains largely unknown. Given their responsiveness to iron availability and
383 associated enrichment in colony rich cultures, these proteins could contribute to ECM-related
384 adhesion of cells, to each other or the colony skin, or even perhaps to the mucilage interior.

385 Additional glycoproteins that exhibited a strong iron response in both strains include those
386 containing von Willebrand factor domains (for example, protein families PF13519, PF00092), and

387 fibrillin and lectin (Fig. 4 and Supplementary Fig. 1). In biomineralizing organisms, such as corals,
388 glycoproteins with von Willebrand domains are hypothesized to play a role in the formation of the
389 extracellular organic matrix through adhesion (Drake et al., 2013; Hayward et al., 2011) laying the
390 scaffolding for calcification. Orthologs of the von Willebrand proteins that contain these domains
391 have also been characterized in humans and have protein-binding capabilities, which are important
392 for coagulation (Ewenstein, 1997). These dynamic von Willebrand proteins appear to contribute
393 to the cell aggregation and colony formation of *P. antarctica* colonies.

394 The suite of structural and modified proteins described above demonstrates a means
395 through which *P. antarctica*'s colonial morphotype could be constructed, and this dataset provides
396 rare molecular evidence for the proteome reconstruction needed to switch between single
397 organisms to a multicellular colony. The evolution of multicellularity in Eukaryotes is an area of
398 significant interest that has mostly focused on model organisms with colonial forms such as
399 Choanoflagellates and *Volvox* (Abedin and King, 2010). Genomic studies of the former identified
400 the presence of protein families involved in cell interactions within metazoans, including C-type
401 lectins, cadherins, and fibrinogen (King et al., 2003). In other lineages of microalgae that form
402 colonial structures, such as *Volvox carteri*, there is supporting evidence for glycoproteins cross-
403 linking within the extracellular matrix of colonies (Hallmann, 2003), as well as serving other
404 important functional roles in cell-cell attachment during colony formation (for example, colony
405 formation in the cyanobacteria *Microcystis aerginosa*) and as an integral component of cell walls
406 (for example, the diatoms *Navicula pelliculosa* and *Craspedostauros australis*) (Chiovitti et al.,
407 2003; Kröger et al., 1994; Zilliges et al., 2008). In this study, environmental isolates of *P.*
408 *antarctica* displayed consistent trends in similar protein families (for example, lectins, fibrillins,
409 and glycoproteins), and they were produced at higher levels under elevated iron conditions when

strain 1871 was primarily in colonial form. Given *P. antarctica*'s environmental importance and an ability to control the transition between flagellates and colony cell types through iron availability, *P. antarctica* may serve as a useful model for studying multicellularity in nature and in the context of environmental change.

In contrast to these putative colonial structural proteins, there were canonical cytoskeletal proteins such as actin and tubulin observed in *P. antarctica* cultures grown under low iron conditions (Supplementary Fig. 1). These proteins were likely associated with the flagella and the haptonema, a shorter organelle containing nine microtubules that is characteristic of Haptophytes (Zingone et al., 1999), found in the solitary *Phaeocystis* cell, and similar to other eukaryotic flagellar systems such as *Chlamydomonas* (Watanabe et al., 2004). Additionally, a suite of proteins with calcium-binding domains (EF-hand protein families) was identified in greater relative abundance under low iron growth conditions in both strains (Fig. 4; Supplementary Fig. 1 and Supplementary Data 1). In diatoms, calcium-signaling mechanisms have been directly linked with how cells respond to bioavailable iron, as well as stress responses (Allen et al., 2008; Vardi, 2008). Calcium (and magnesium) ions also play an integral role in the ability for extracellular mucus to gel (van Boekel, 1992). The greater abundance of putative calcium-binding proteins under low iron conditions suggests an important role for intracellular calcium, either in its involvement in flagellate motility and/or having a role in inhibiting the cells' abilities to form colonies while under iron limitation. This use of calcium signaling is notable given that calcium is a major constituent of seawater (0.01 mol L^{-1}), implying a need for efflux and exclusion of calcium from the cytoplasm.

3.5 *Phaeocystis antarctica* strain-specific responses

433 *Phaeocystis antarctica* is believed to have speciated from warm-water ancestors, and populations
434 within the Antarctic are mixed via the rapid Antarctic Circumpolar Current (ACC, 1-2 years)
435 circulation with the Ross Sea and Weddell Sea, which entrains strains nearly simultaneously
436 (Lange et al., 2002). However, given the original geographic location of the isolates, there may be
437 some differences regarding adaptation and ecological role between strains. In the Ross Sea, *P.*
438 *antarctica* dominates, and cells exhibit seasonal variability between flagellated states (early
439 Spring, late summer) and colonial stage (late Spring/early summer) (Smith et al., 2003). In
440 contrast, in the Western Antarctic Peninsula, near the Weddell Sea where strain 1871 was isolated
441 from (Palmer station), *P. antarctica* is outnumbered by diatoms and cryptomonads in terms of
442 algal biomass, and colonies are generally rare (Ducklow et al., 2007). While global proteomic and
443 transcriptomic analyses revealed differences among strains (Supplementary Data 1), both strains
444 had responses that overwhelmingly supported a concerted effort towards structural changes under
445 high iron versus low iron, consistent with the minor phylogenetic differences previously reported
446 for *P. antarctica* isolates due to rapid ACC circulation (Lange et al., 2002).
447
448 **3.6 Examination of a *Phaeocystis* bloom metaproteome from the Ross Sea**
449 The detailed laboratory studies above can be compared to a first metaproteomic analysis
450 of a Ross Sea *Phaeocystis antarctica* bloom to provide an examination of the *in situ* ecology and
451 biogeochemical and their underlying biochemical signatures. Due to the relative newness of
452 metaproteomic eukaryotic phytoplankton research, some methodological detail has been
453 incorporated into this section. For field analysis a net tow sample was collected north of Ross
454 Island (Fig. 7) on December 30th 2005, in which *Phaeocystis* colonies were visually dominant.
455 Temporal changes in the bloom composition have been described for this summer expedition and

456 an austral spring expedition later that year (NBP06-01 and NBP06-08, respectively), and a shift
457 was observed from a *P. antarctica* dominated ecosystem to a mixture of *P. antarctica* and diatoms
458 (Smith et al., 2013). Surface pigment distributions showed the sampling region to be within a
459 particularly intense bloom dominated by *Phaeocystis* as observed by abundant 19'-
460 hexanoyloxyfucoxanthin pigment (Fig. 7.), reaching concentrations of 1096 ng L^{-1} and total
461 chlorophyll *a* concentrations of 1860 ng L^{-1} on the sampling day. CHEMTAX analysis of these
462 HPLC pigments found that *P. antarctica* populations accounted for approximately 88% of surface
463 water total chlorophyll at this time. Fucoxanthin pigment, characteristic of diatoms, was lower
464 here (136 ng L^{-1}) compared to samples from the western Ross Sea (Fig. 7.), consistent with prior
465 Ross Sea observations. Repeated sampling near the sampling region (~77.5ºS) two weeks after
466 taking the metaproteome sample found lower overall chlorophyll *a* levels (Smith et al., 2013),
467 consistent with bloom decay. Iron measured very near this location (76.82º S, 170.76º E also on
468 December 30, 2005), found a surface dissolved iron concentration of 170pM (6m depth) and an
469 acid-labile particulate iron concentration of 1590 pM (Sedwick et al., 2011), consistent with iron
470 depletion in seawater following drawdown of the accumulated winter iron supply and
471 incorporation of iron into biological particulate material (Noble et al., 2013; Sedwick et al., 2000).
472 The metaproteome analyses of the Ross Sea sample were conducted by bottom-up mass
473 spectrometry analysis of tryptic peptides, followed by peptide-to-spectrum matching of putative
474 peptide masses and their fragment ions to predicted peptides from translated DNA sequences.
475 While this approach is common for model organisms and has been successfully applied to
476 primarily prokaryotic components of natural communities (Morris et al., 2010; Ram et al., 2005;
477 Sowell et al., 2008; Williams et al., 2012), there continue to be challenges in metaproteomics
478 analyses of diverse communities particularly when including an extensive eukaryotic component

such as is present in the Ross Sea phytoplankton bloom. To address these issues, we utilized three sequences databases for peptide-to-spectrum matching (see Methods and Supplementary Information). Analysis of both unique (tryptic) peptides and identified proteins are provided here, where unique peptides are particularly valuable in metaproteome interpretation as a basal unit of protein diversity that can be definitively compared across the three sequence databases (Saito et al., 2015).

The combined *P. antarctica* strain transcriptome database (Database #1) generated the largest number of protein and unique peptide identifications, 912 and 2103, respectively (Table 2, Fig. 8*a*.). This strong relative performance of the strain database was surprising, and likely reflects the depth of the *P. antarctica* isolate transcriptomes and resultant translation into greater metaproteomic depth. Sixty percent of field identifications mapped to strain 1374; a broad synthesis of all proteomes based on KOG annotations also indicated that the metaproteomes appeared most similar to the Ross Sea strain 1374 (Supplementary Fig. 3). The Ross Sea metatranscriptome database (Database #2) resulted in 859 proteins and 1520 unique peptides distributed across a large number of taxa, with 324 of those proteins associated with *P. antarctica.* The Antarctic bacterial metagenome database (Database #3) produced 98 proteins and 186 unique peptides that mapped to bacteria likely associated with the phytoplankton communities, given the use of a net that would not otherwise capture free-living bacteria. Due to the extensive diversity present, there was overlap between the peptide identifications from each database for the 3193 total unique peptides: 544 *P. antarctica* peptides were shared between the *Phaeocystis* strain and Ross Sea metatranscriptome databases, 69 bacterial peptides were in common between the Ross Sea metatranscriptome and the bacterial metagenomic databases, followed by very small numbers shared between bacterial metagenome and the *Phaeocystis* strains database searches (8 peptides),

and all three databases (4 peptides), likely due to a small fraction of tryptic peptides shared between diverse organisms (Saito et al., 2015).

This multi-database approach and the relatively low overlap illustrates the necessity of employing diverse sequence databases that target distinct components of the biological community, as well as the value in coupling metatranscriptomic and metagenomic sequence databases to metaproteomic functional analysis to capture the extent of natural diversity. This is evident in the taxon group analysis, where the metatranscriptome has a large representation of Dinophyta and diatoms and only a small contribution from Haptophyta that include *Phaeocystis*, likely due to the large genome sizes and transcription rates, particularly of dinoflagellates, and perhaps due to interferences of *Phaeocystis* RNA extraction due to the copious mucilage present (Fig. 8*b*). In contrast the metaproteome derived from the metatranscriptome database is dominated by Haptophyta and Dinophyta, with minor contributions from other groups (Fig. 8*d*), reflecting the dominant organismal composition seen in the pigment analyses (Fig. 7). Due to a coarse net mesh size much larger than a typical bacterial cell, the bacterial community captured by these metatranscriptome and metaproteome analyses most likely reflects the microbiome associated with larger phytoplankton and protists, particularly within the abundant *P. antarctica* colonies . Database #2 and #3 result in 148 and 100 bacterial protein identifications, respectively, including representatives from *Oceanospirillacea*e, *Rhodobacteraceae*, *Cryomorphaceae, Flavobacteria*, and *Gamma proteobacteria* (Fig. 8*c* and *d*).

Together this Ross Sea bloom metaproteome-metatranscriptome analysis provides a window into the complex interactions of this community with its chemical environment. *Phaeocystis antarctica* proteins were abundant in the sample with over 300 proteins identified, yet interestingly, we identified proteins associated with both high and low iron treatments, including

525 those corresponding to flagellate and colonial life stages identified in the culture experiments (Fig.
526 9 and Supplementary Fig. 1). This presence of both life cycle stages of *Phaeocystis* could be
527 interpreted as evidence of an actively growing bloom, with growing flagellate cells coalescing to
528 form new colonies, as well as a standing stock of colonial cells. As mentioned earlier, division and
529 growth of *P. antarctica* colonies is believed to require transitioning back through the flagellate life
530 cycle stage, hence a mixed population of flagellate and colonial stages would be expected of a
531 growing population, consistent with our laboratory observations (Fig. 2*c*).
532 The presence of well-known iron-sparing proteins such as plastocyanin (Fig. 9) was
533 consistent with the depleted dissolved iron concentration (170 pM) in nearby surface waters that
534 are closest to the 120 pM Fe' of the low iron treatments (Peers and Price, 2006; Sedwick et al.,
535 2011), as well as incubation experiments on the same expedition initiated three days prior that
536 demonstrated iron limitation of *P. antarctica* (and iron-B_{12} colimitation of diatom) populations
537 (Bertrand et al., 2007). Notably, the actual Fe' of the Ross Sea was likely considerably lower than
538 this due to the presence of strong organic iron complexes (Boye et al., 2001). Strzepek et al. found
539 evidence for growth of *P. antarctica* and some polar diatoms on strong organic iron complexes at
540 somewhat reduced growth rates in their culture experiments, implying a high-affinity iron
541 acquisition system such as a ferric reductase, although the molecular components of such a system
542 have yet to be identified in *P. antarctica* (Strzepek et al., 2011). As described above, it is likely
543 that both flagellate and colonial cell types have a need to manifest iron stress responses (e.g.
544 distinct ISIP proteins found in the flagellate and colonial dominated cultures, Figs. 4 and 5), and
545 that those distinct responses may be based on the extensive physical differences between life cycle
546 phenotypes. The low contribution of chain-forming diatoms to this metaproteome sample was
547 consistent with the higher sensitivity of some Ross Sea diatom strains to iron stress such as

548 *Chaetoceros* (Fig. 2*d*) and the low iron availability. Careful examination of targeted mass
549 spectrometry results (precursor and fragment ion analysis) for select iron proteins identified in
550 culture studies showed consistently high quality chromatograms within the field sample,
551 demonstrating a capability to measure these potential peptide biomarkers within complex
552 environmental samples in future field studies characterizing bloom and biogeochemical dynamics
553 (Fig. 10 and Supplementary Figs. 4-10).

554 The metaproteome analyses also captured relevant functional elements of the bacterial
555 microbiome associated with the eukaryotic community, based on bacterial proteins identified in
556 both the bacterial databases and the Ross Sea metatranscriptome (Fig. 8*c* and 8*d*). For example,
557 the SAR92 clade of proteorhodopsin-containing heterotrophic bacteria was present (Stingl et al.,
558 2007), and expressed both the iron storage protein bacterioferritin and TonB receptors, the latter
559 of which are involved in siderophore and B_{12} transport. In addition, the Fur iron regulon, iron-
560 requiring ribonucleotide reductase, as well as the vitamin related CobN cobalamin biosynthesis
561 protein, B_{12}-requiring methyl-malonyl CoA, and thiamine ABC transporter were observed from
562 several heterotrophic bacteria species including *Oceanospirillacea*e, *Rhodobacteraceae*, and
563 *Cryomorphaceae* (Supplementary Data 2) (Bertrand et al., 2015; Murray and Grzymski, 2007).
564 These results imply that heterotrophic bacteria known to be associated with the *Phaeocystis*
565 colonies, such as SAR92 and *Oceanospirillaceae*, were also likely responding to micronutrients
566 by concentrating and storing iron, and through biosynthesis of B_{12}. In doing so this bacterial
567 microbiome could have been harboring an "internal" source of the micronutrients, fostering a
568 mutualism with *Phaeocystis* colonies in exchange for a carbon source and consistent with the high
569 particulate iron measured during this station (Sedwick et al., 2011). Together this could create a
570 competitive advantage for *P. antarctica* relative to the iron and B_{12}-stressed diatoms for early

season bloom formation, as previously hypothesized and observed in the Ross Sea in enrichment studies (Bertrand et al., 2007). Although diatoms were less prominent in the dataset, several diatom proteins identified were indicative of the potential for iron stress (e.g., plastocyanin and ISIP3; Supplementary Data 2); however, the diatom CBA1 cobalamin acquisition protein was not identified in the metatranscriptome, and hence would not be detected in the metaproteome using the current methods.

4. Conclusions

Phaeocystis antarctica is a major contributor to Southern Ocean primary productivity, yet arguably is one of the least well understood of key marine phytoplankton species. The multiple life cycle stages of *P. antarctica* add to its ecological and biochemical complexity. Here we have undertaken a detailed combined physiological and proteomic analysis enabled by transcriptomic sequencing under varying conditions of iron nutrition, and compared these to an initial study of the metaproteome of a Ross Sea *Phaeocystis* bloom. These results demonstrate that *P. antarctica* has evolved to utilize elaborate capabilities to confront the widespread iron scarcity that occurs in the Ross Sea and Southern Ocean, including iron metalloenzyme sparing systems and the deployment of transport and other systems that appear to be unique to the flagellate and colonial morphotypes. To our surprise, iron abundance clearly triggered colony formation in one strain in this study, and visual and proteomic evidence implied the second strain was also attempting to do so. Prior studies have invoked light irradiance and mixed layer depth as key factors in colony production and the concurrent Ross Sea *P. antarctica* bloom initiation (Arrigo et al., 1999), and hence there may be other factors that could have this effect as well. These results also provide a first window into the complex cell restructuring process that occurs upon cellular metamorphosis

between life cycle stages in *P. antarctica* as well as identifying numerous dynamic proteins of unknown function for future study. Finally, this study demonstrates the potential for the application of coupled transcriptomic and proteomic biomarker methodologies in studying the ecology of microbial interactions (including iron and B_{12}) and their influence on biogeochemistry in complex polar ecosystems such as the Ross Sea. The improved molecular and biochemical understanding of *P. antarctica* and its response to iron provided here are valuable in the design of future experiments and targeted metaproteomic assays to examine natural populations and to improve understanding of environmental factors that influence the annual bloom formation of an important coastal ecosystem of the Southern Ocean.

Acknowledgements

Support for this study was provided by an Investigator grant to M. Saito from the Gordon and Betty Moore Foundation (GBMF3782), and National Science Foundation grants NSF-PLR 0732665, OCE-1435056, OCE-1220484, the WHOI Coastal Ocean Institute, and a CINAR Postdoctoral Scholar Fellowship provided to S. Bender through the Woods Hole Oceanographic Institution. Support was provided to A. Allen through NSF awards ANT-0732822, ANT-1043671, OCE-1136477, and Gordon and Betty Moore Foundation grant GBMF3828. Additional support was provided to GRD through NSF award, OPP-0338097.We are indebted to Roberta Marinelli for her leadership and vision. We would also like to thank Emily Lorch for her assistance with culturing, Julie Rose for generously sharing a net tow field sample, and Andreas Krupke for manuscript feedback.

617

618 **Author Contributions**

619 S.J.B. contributed to data analysis and writing; D.M.M. conducted the laboratory experiments and
620 (meta)proteome extractions; M.R.M. conducted the mass spectrometry sample preparation and
621 processing; H.Z. conducted RNA extractions; J.P.M. and J.B. contributed to transcriptome
622 sequence analyses; G.R.D. contributed to field measurements and manuscript edits; A.E.A.
623 contributed to the experimental design, data analysis, and writing; M.A.S. contributed to
624 experimental design, data analysis, and writing.

625 **Financial Conflicts:** The authors have no financial conflicts involving the research presented in
626 this manuscript.

Biogeosciences Discuss., https://doi.org/10.5194/bg-2017-558
Manuscript under review for journal Biogeosciences
Discussion started: 26 January 2018

627

628 **Table 1.** Comparison of the total number of proteins and spectra measured in the proteome for
629 each strain/treatment along with the number of differentially expressed transcripts between select
630 conditions for *P. antarctica* strain 1871 and strain 1374. Proteins were identified using a 1% FDR
631 (false discovery rate) threshold, a peptide threshold of 95%, and a minimum of 2 unique peptides
632 per protein. The total number of peptide-to-spectrum matches (PSMs) is given for the total of each
633 strain in parentheses. A threshold of 3 spectral counts in at least one of the treatments was selected
634 for inclusion in the comparative analysis.

635

Strain	Treatment (Fe' pM)	Proteins Identified (PSMs)
1871	2	204
	41	214
	120	234
	740	226
	1200	251
	3900	258
	Total	536 (28887)
1374	2	581
	41	613
	120	600
	740	654
	1200	623
	3900	527
	Total	1085 (72087)

Biogeosciences Discuss., https://doi.org/10.5194/bg-2017-558
Manuscript under review for journal Biogeosciences
Discussion started: 26 January 2018

636 **Table 2.** Comparison of the total number of proteins, peptides, and spectra measured in the
637 Ross Sea metaproteome net tow sample.

Peptide-to-spectrum - matching database	Total proteins	Total Unique Peptides	Total spectra matched	Decoy FDR+ (peptide level)
Phaeocystis strains transcriptomes*	912	2103	8226	0.17%
Ross Sea metatranscriptome**	859	1520	4725	0.7%
Antarctic bacterial metagenomes***	92	186	440	2.33%

638

639 +FDR refers to false discovery rate of a reversed peptide database

640 * Metaproteome annotated using the laboratory-generated transcriptomes for strain 1871
641 and strain 1374.

642 ** Metaproteome annotated using the metatranscriptome generated from sample split of
643 original Ross Sea sample.

644 *** Bacterial metaproteome annotated using bacterial metagenomes from Delmont et al.,
645 2014.

646 **Figure Legends**

647 **Figure 1.** Micrographs of (a) a single *Phaeocystis* in cell culture, and (b) *Phaeocystis* colonies in
648 a Ross Sea bloom.

649

650 **Figure 2.** The effect of iron concentration on colony formation and cell physiology in two strains
651 of *P. antarctica* – 1871 and 1374. Growth rates collected from acclimated culture stocks prior to
652 the start of the experiments (*a*, strain 1871; *b*, strain 1374), calculated using relative fluorescence
653 units from three transfers of acclimated cultures (error bars indicate SD, n=3). Accompanying
654 gray bars represent growth rates calculated based on cell counts made during the course of the
655 proteome-harvest experiments (n=1). (*c*) The number of *P. antarctica* 1871 free-living cells
656 (gray bars) compared to cells associated with colonies (black bars) showed a shift to a majority
657 of colonial cells when Fe' $\geq$ 740 pM. (*d*) Growth rate of Ross Sea diatom isolate *Chaetoceros* sp.
658 strain RS-19 in the same media compositions (n=1), demonstrated a higher sensitivity to iron
659 scarcity and a lack of iron contamination in the media. Cell size for strain 1871 (*e*; black circles)
660 and strain 1374 (*f*; white circles); error bars represent SD of n=20 cell measurements per
661 treatment.

662

663 **Figure 3.** Principle Component Analysis (PCA) of the full proteomes for each iron condition for
664 strain 1871 and strain 1374 and corresponding line graphs highlighting the proteins driving the
665 PCA separation (PCA analyses: $\geq$ 0.9 or $\leq$ -0.9). (*a* and *d*) Iron treatments (pM Fe') are
666 highlighted by color (2, black; 41, red; 120, orange; 740, green; 1200, purple; 3900, blue) and
667 large ellipses indicate confidence ellipses calculated using the R package, FactorMineR. Each
668 small, solid circle represents a technical replicate per iron treatment (n=3); colored, open squares

represent the mean of the iron treatment (empirical variance divided by the number of observations). Proteins with Eigen values ≥ 0.9 or ≤ -0.9 are plotted in graphs b and c for strain 1871 and *e* and *f* for strain 1374 to highlight the subset of proteins driving the variance in Dimension 1. Individual protein spectral counts normalized to total spectral counts for all treatments for a given protein, written as "normalized relative protein abundance" are plotted on the y-axis. The six iron treatments (pM Fe') are plotted from low to high (left to right) on the x-axis.

Figure 4. Heatmaps highlighting the relative protein abundance for the six treatments for *P. antarctica* strain 1871 (*a*) and strain 1374 (*b*). The darker green color indicates a greater relative abundance compared to the purple treatments. The "shared abundance patterns" column features a check-mark when a shared response to changes in iron availability between the relative protein abundance and the transcript abundance was observed (for example, both transcripts and proteins have a higher abundance under high iron compared to low iron growth [or] both transcripts and proteins have a higher abundance under low iron compared to high iron growth). The "field presence" column indicates whether or not that protein was detected in the field metaproteome (annotated using Database #1). Protein annotations are based on KEGG, KOG, and PFam descriptions. Annotations in red are associated with iron metabolism and those in blue, cell adhesion/structure.

Figure 5. Examination of iron stress response proteins in *P. antarctica* strain 1871 (top) and 1374 (bottom). Relative protein abundance is shown as normalized spectral counts, where spectral counts have been normalized across experiment treatments for each strain, but not to the

maximum of each protein as used in prior figures to allow comparison of abundance for similar isoforms. Error bars indicate the standard deviation of technical triplicate analyses.

Figure 6. Scatterplots of relative transcript abundance (y-axis) and relative protein abundance (x-axis) for *P. antarctica* strain 1871 (a) and strain 1374 (b) for a high iron treatment (3900 pM Fe') relative to a low iron treatment (41 pM Fe'). Gray circles represent instances where transcript abundance was not significantly different between conditions ($P \geq 0.99$). Quadrants where relative protein and transcript abundances agree (upper right, lower left) and disagree (upper left, lower right) are noted, as are select genes exhibiting the greatest relative protein abundance and/or transcript abundance under a given treatment.

Figure 7. Location of the metaproteome sample and pigment data from a Ross Sea *Phaeocystis* bloom net tow sample. (*a*) Station map of NBP06-01 (December 27, 2005 to January 23, 2006) and the metaproteome sample was taken on December 30th by net tow location (red circle). (*b*) 19'-hexanoyloxyfucoxanthin ("19'-Hex") pigment is associated with *Phaeocystis*, while (c) peridinin and (d) fucoxanthin pigments are typically associated with dinoflagellates and diatoms, respectively (although dinoflagellates living heterotrophically can be lacking in pigment). Comparisons of the spring and summer expeditions (NBP06-08 and NBP06-01, respectively), observed a shift from being dominated by *P. antarctica* to being a mixture of *P. antarctica* and diatoms. See Smith et al., 2013 for further details (Smith et al., 2013).

Figure 8. (a) Venn diagram of the attribution of the 3193 total unique peptides identified in the metaproteome sample to three DNA/RNA sequence databases (Supplementary Table 2). (*b*)

Taxon group composition of genes identified by metatranscriptome analyses (combining Total RNA and PolyA RNA fractions). (c) Taxon group composition of proteins identified by the bacterial metagenomic database (Database #3). (d) Taxon group composition of proteins identified by metatranscriptome database (Database #2).

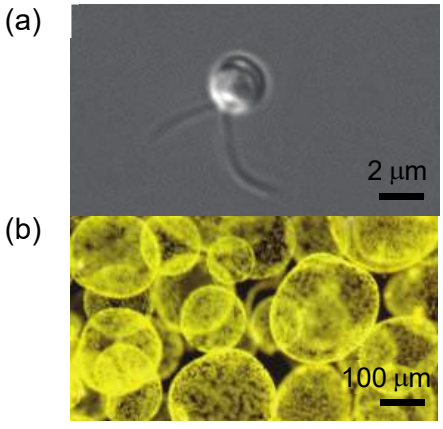
Figure 9. Putative biomarkers identified in the *Phaeocystis* metaproteome annotated using the field metatranscriptome (error bars represent SD of replicate samples; n=2). Green bars indicate putative "low iron" biomarkers; red bars indicate putative "high iron" biomarkers, and correspond to the life cycle stages observed (Fig. 2).

Fig. 10. Example spectra and chromatograms of fragment ions for two peptides corresponding to a *P. antarctica* flavodoxin identified from the Ross Sea metaproteome sample (peptide sequences found within Database #1, 1871, contig_31444_1_606_+, 1374 contig_202625_47_661_+; and, Database #2 contig_175060_39_653_+). Peptide fragmentation spectra are shown in (a) and (c) and example chromatograms of ms1 intensities as well as with +1 and +2 mass addition for isotopic distributions is shown (b) and (d), demonstrating the utility of these iron stress biomarkers in field samples.

Biogeosciences Discuss., https://doi.org/10.5194/bg-2017-558
Manuscript under review for journal Biogeosciences
Discussion started: 26 January 2018

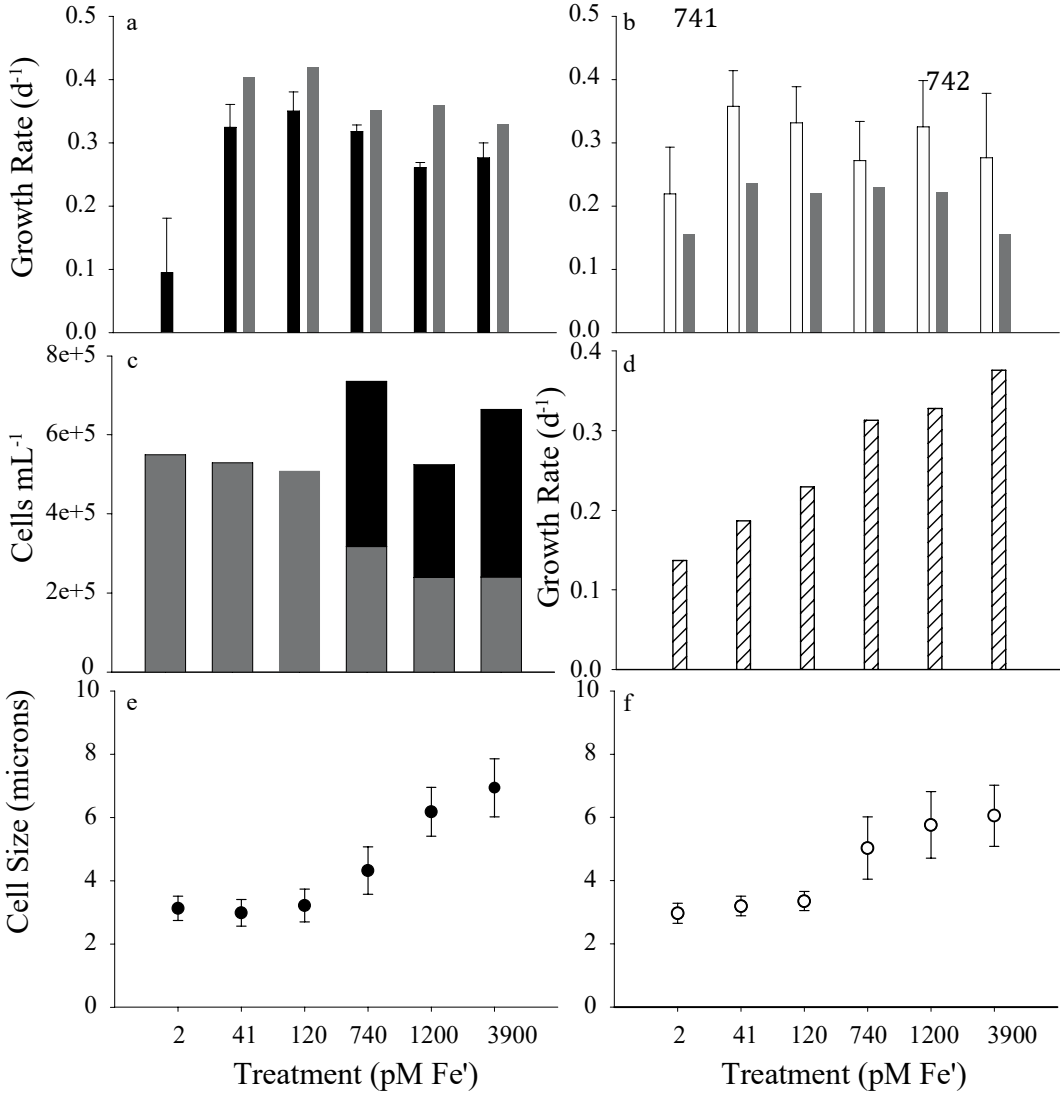
733

734 **Figure 1.**
735
736
737



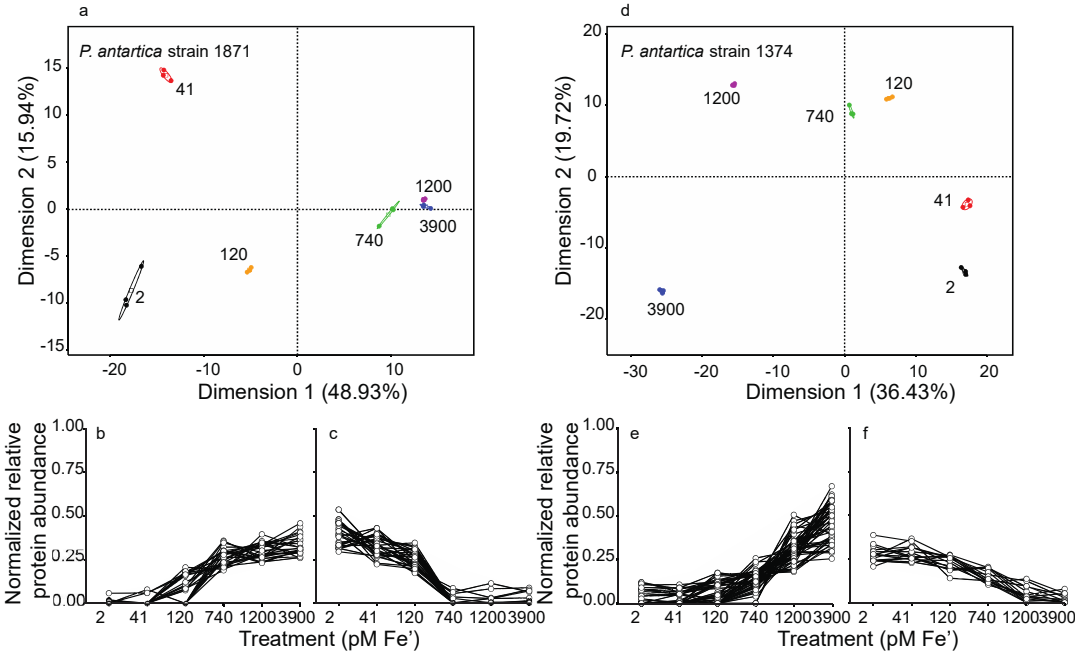
Biogeosciences Discuss., https://doi.org/10.5194/bg-2017-558
Manuscript under review for journal Biogeosciences
Discussion started: 26 January 2018

738 **Figure 2.**
739
740



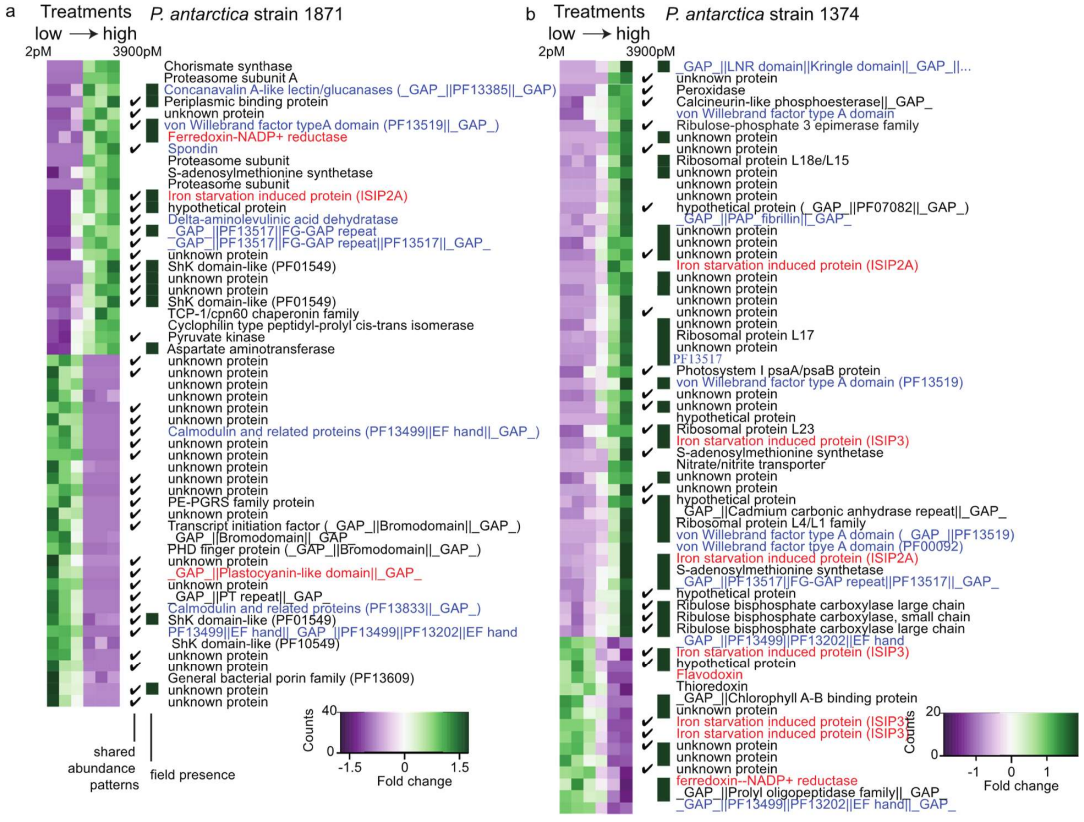
Biogeosciences Discuss., https://doi.org/10.5194/bg-2017-558
Manuscript under review for journal Biogeosciences
Discussion started: 26 January 2018

743 **Figure 3.**



Biogeosciences Discuss., https://doi.org/10.5194/bg-2017-558
Manuscript under review for journal Biogeosciences
Discussion started: 26 January 2018

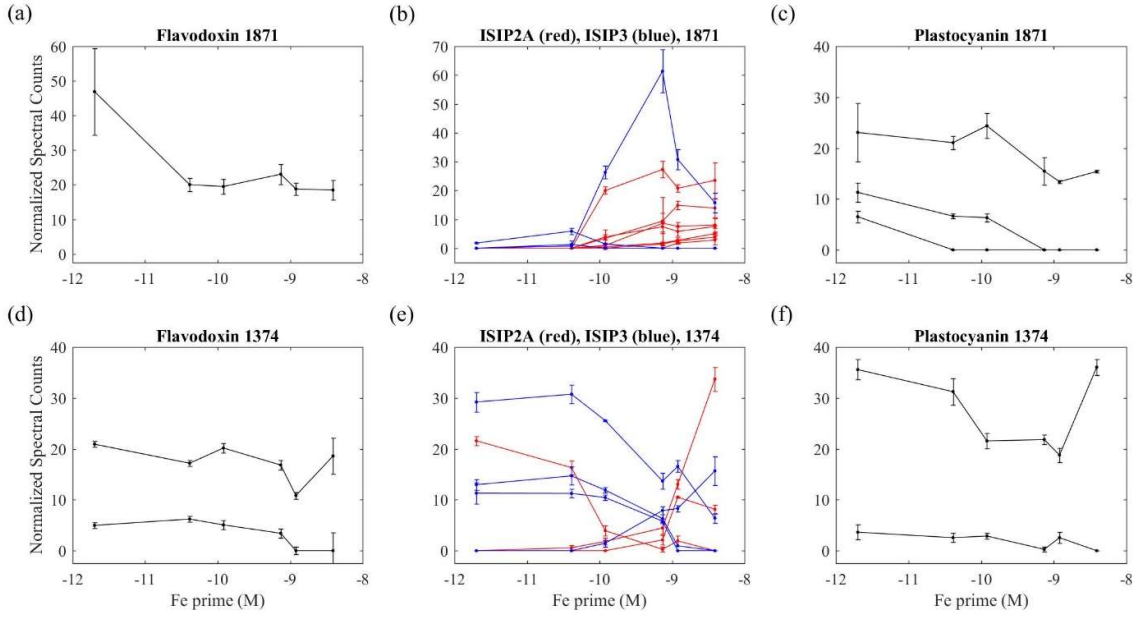
744 **Figure 4.**
745
746



747

Biogeosciences Discuss., https://doi.org/10.5194/bg-2017-558
Manuscript under review for journal Biogeosciences
Discussion started: 26 January 2018

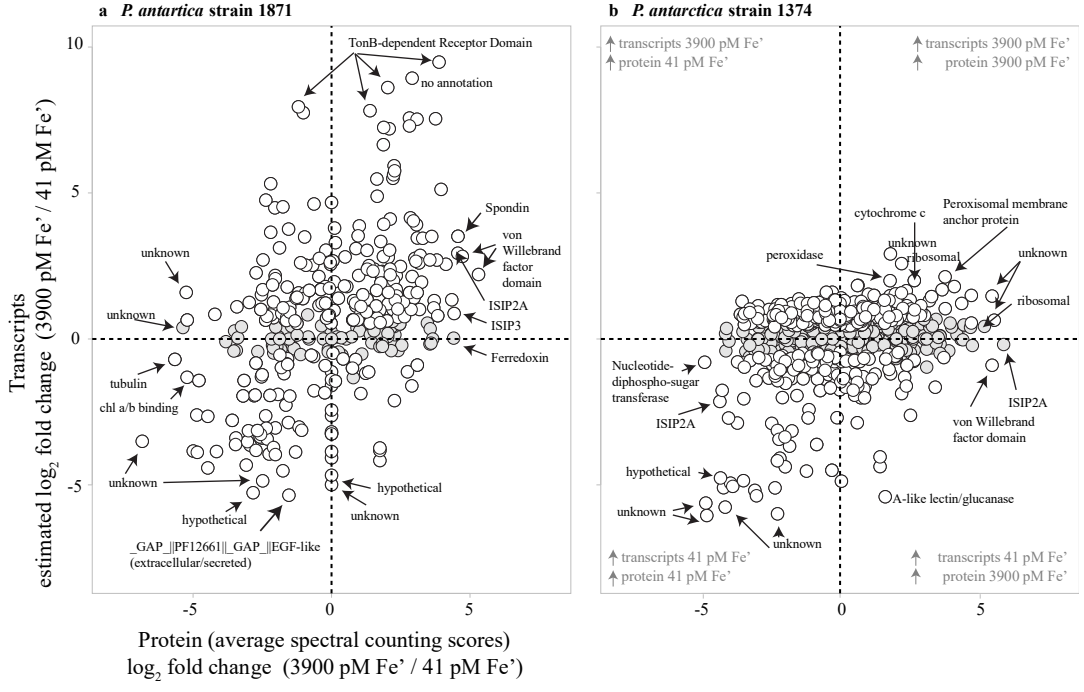
748 **Figure 5.**
749



750

Biogeosciences Discuss., https://doi.org/10.5194/bg-2017-558
Manuscript under review for journal Biogeosciences
Discussion started: 26 January 2018

751 **Figure 6.**
752

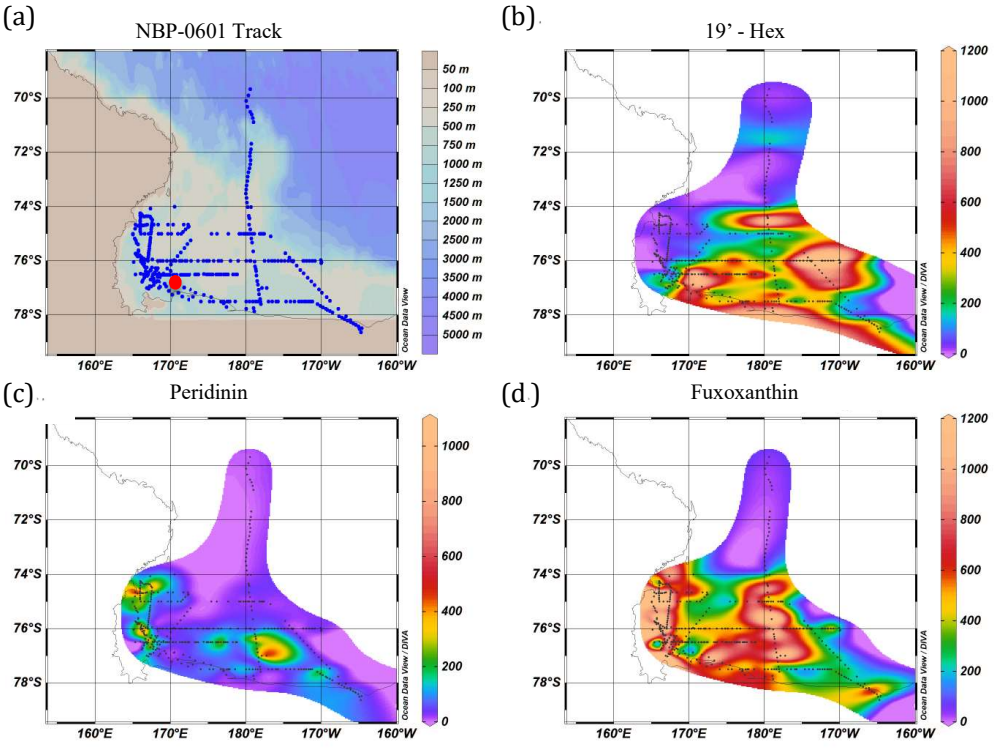


Biogeosciences Discuss., https://doi.org/10.5194/bg-2017-558
Manuscript under review for journal Biogeosciences
Discussion started: 26 January 2018

753 **Figure 7.**

754

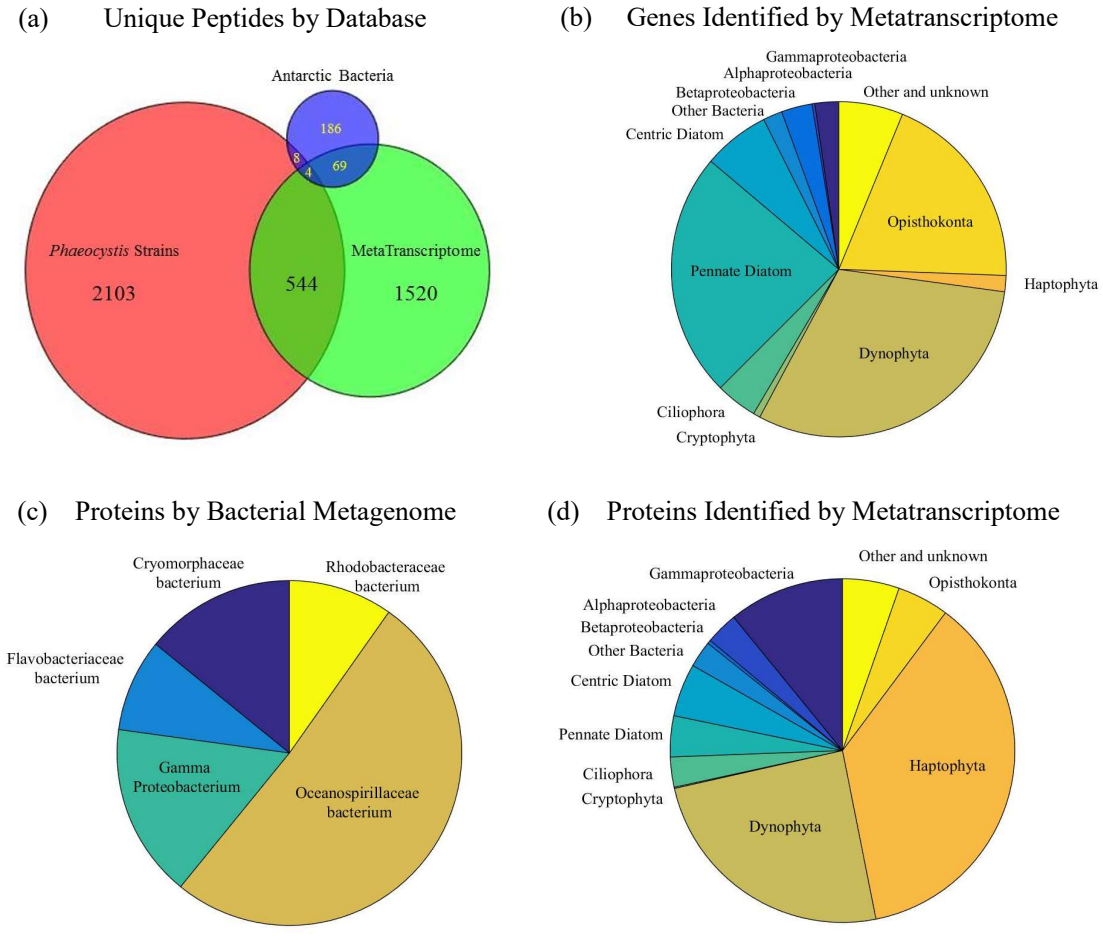
755



756
757
758

Biogeosciences Discuss., https://doi.org/10.5194/bg-2017-558
Manuscript under review for journal Biogeosciences
Discussion started: 26 January 2018

759 **Figure 8.**
760

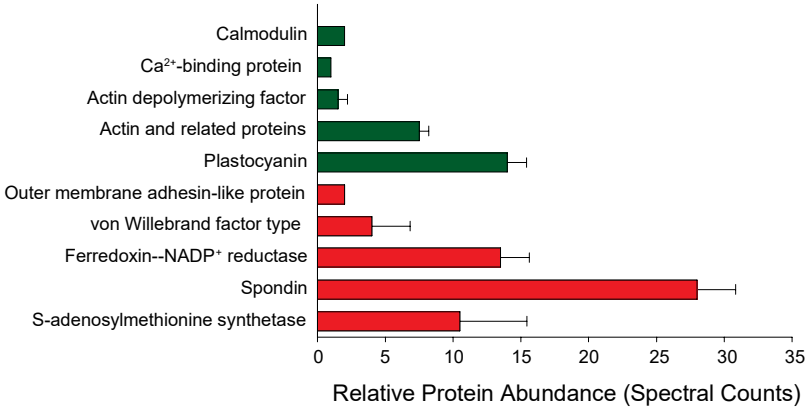


761

Biogeosciences Discuss., https://doi.org/10.5194/bg-2017-558
Manuscript under review for journal Biogeosciences
Discussion started: 26 January 2018

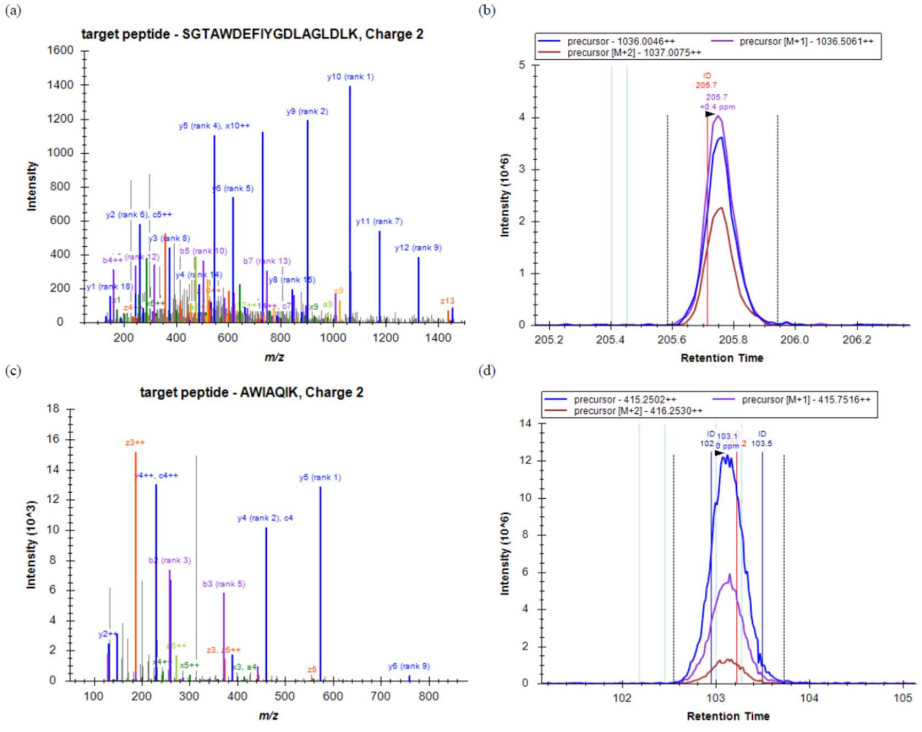
762 **Figure 9.**

763



Biogeosciences Discuss., https://doi.org/10.5194/bg-2017-558
Manuscript under review for journal Biogeosciences
Discussion started: 26 January 2018

764 **Figure 10.**
765
766



767

References

Abedin, M. and King, N.: Diverse and evolutionary paths to cell adhesion, Trends in Cell Biology, 20(12), 734–742, 2010.

Alderkamp, A. C., Buma, A. G. J. and Van Rijssel, M.: The carbohydrates of *Phaeocystis* and their degradation in the microbial food web, Biogeochemistry, 83, 99–118, 2007.

Allen, A. E., LaRoche, J., Maheswari, U., Lommer, M., Schauer, N., Lopez, P. J., Finazzi, G., Fernie, A. R. and Bowler, C.: Whole-cell response of the pennate diatom *Phaeodactylum tricornutum* to iron starvation, vol. 105, pp. 10438–10443. 2008.

Arrigo, K. R., R, D. G., Dunbar, R. B., Robinson, D. H., vanWoert, M. L., Worthen, D. L. and Lizotte, M. P.: Phytoplankton taxonomic variability in nutrient utilization and primary production in the Ross Sea, Journal of Geophysical Research, 105(C4), 8827–8846, doi:10.1029/1998JC000289, 2000.

Arrigo, K. R., Robinson, D. H., Worthen, D. L., Dunbar, R. B., R, D. G., vanWoert, M. L. and Lizotte, M. P.: Phytoplankton community structure and the drawdown of nutrients and CO_2 in the Southern Ocean, Science, 283, 365–367, doi:10.1126/science.283.5400.365, 1999.

Arrigo, K. R., Worthen, D., Schnell, A. and Lizotte, M. P.: Primary production in Southern Ocean waters, Journal of Geophysical Research, 103(C8), 15587–15600, 1998.

Bertrand, E. M., McCrow, J. P., Moustafa, A., Zheng, H., McQuaid, J. B., Delmont, T. O., Post, A. F., Sipler, R. E., Spackeen, J. L., Xu, K., Bronk, D. A., Hutchins, D. A. and Allen, A. E.: Phytoplankton–bacterial interactions mediate micronutrient colimitation at the coastal Antarctic sea ice edge, Proceedings of the National Academy of Sciences, 112(32), 9938–9943, doi:10.1073/pnas.1501615112, 2015.

Bertrand, E. M., Moran, D. M., McIlvin, M. R., Hoffman, J. M., Allen, A. E. and Saito, M. A.: Methionine synthase interreplacement in diatom cultures and communities: Implications for the persistence of B_{12} use by eukaryotic phytoplankton, Limnology and Oceanography, 58(4), 1431–1450, doi:10.4319/lo.2013.58.4.1431, 2013.

Bertrand, E. M., Saito, M. A., Lee, P. A., Dunbar, R. B., Sedwick, P. N. and R, D. G.: Iron limitation of a springtime bacterial and phytoplankton community in the Ross Sea: Implications for Vitamin B_{12} nutrition, Frontiers in Microbiology, 2, 1–12, doi:10.3389/fmicb.2011.00160, 2011.

Bertrand, E. M., Saito, M. A., Rose, J. M., Riesselman, C. R., Lohan, M. C., Noble, A. E., Lee, P. A. and R, D. G.: Vitamin B_{12} and iron colimitation of phytoplankton growth in the Ross Sea, Limnology and Oceanography, 52(3), 1079–1093, 2007.

Boye, M., van den Berg, C., de Jong, J., Leach, H., Croot, P. and de Baar, H. J. W.: Organic complexation of iron in the Southern Ocean, Deep-Sea Research Part I, 48(6), 1477–1497, 2001.

Chiovitti, A., Bacic, A., Burke, J. and Wetherbee, R.: Heterogeneous xylose-rich glycans are

804 associated with extracellular glycoproteins from the biofouling diatom *Craspedostauros*
805 *australis* (Bacillariophyceae), European Journal of Phycology, 38(4), 351–360,
806 doi:10.1080/09670260310001612637, 2003.

807 Coale, K. H., Wang, X., Tanner, S. J. and Johnson, K. S.: Phytoplankton growth and biological
808 response to iron and zinc addition in the Ross Sea and Antarctic Circumpolar Current along
809 170°W, Deep-Sea Research Part II, 50, 635–653, 2003.

810 Delmont, T. O., Hammar, K. M., Ducklow, H. W., Yager, P. L. and Post, A. F.: *Phaeocystis*
811 *antarctica* blooms strongly influence bacterial community structures in the Amundsen Sea
812 polynya, Frontiers in Microbiology, 5, 646, 2014.

813 DiTullio, G. R., Grebmeier, J. M., Arrigo, K. R., Lizotte, M. P., Robinson, D. H., Leventer, A.,
814 Barry, J. P., vanWoert, M. L. and dunbar, R. B.: Rapid and early export of *Phaeocystis*
815 *antarctica* blooms in the Ross Sea, Antarctica, Nature, 404, 595–598, doi:10.1038/35007061,
816 2000.

817 Drake, J. L., Mass, T., Haramaty, L., Zelzion, E., Bhattacharya, D. and Falkowski, P. G.:
818 Proteomic analysis of skeletal organic matrix from the stony coral *Stylophora pistillata*,
819 Proceedings of the National Academy of Sciences, 110(10), 3788–3793, 2013.

820 Ducklow, H. W., Baker, K., Martinson, D. G., Quetin, L. B., Ross, R. M., Smith, R. C.,
821 Stammerjohn, S. E., Vernet, M. and Fraser, W.: Marine pelagic ecosystems: the West Antarctic
822 Peninsula, Philosophical Transactions of the Royal Society B, 362, 67–94,
823 doi:10.1098/rstb.2006.1955, 2007.

824 Dunbar, R. B., Leventer, A. R. and Mucciarone, D. A.: Water column sediment fluxes in the
825 Ross Sea, Antarctica: Atmospheric and sea ice forcing, Journal of Geophysical Research,
826 103(C13), 30741–30759, doi:10.1029/1998JC900001, 1998.

827 Dyhrman, S. T.: Identifying reference genes with stable expression from high throughput
828 sequence data,, 1–10, doi:10.3389/fmicb.2012.00385/abstract, 2012.

829 Eng, J. K., McCormack, A. L. and Yates, J. R.: An approach to correlate tandem mass spectral
830 data of peptides with amino acid sequences in a protein database, J Am Soc Mass Spectrom,
831 5(11), 976–989, doi:10.1016/1044-0305(94)80016-2, 1994.

832 Ewenstein, B. M.: Von Willebrand's disease, Annu. Rev. Med., 48(1), 525–542,
833 doi:10.1146/annurev.med.48.1.525, 1997.

834 Feng, Y., Hare, C. E., Rose, J. M., Handy, S. M., DiTullio, G. R., Lee, P. A., Smith, W. O., Jr,
835 Peloquin, J., Tozzi, S., Sun, J., Zhang, Y., Dunbar, R. B., Long, M. C., Sohst, B., Lohan, M. and
836 Hutchins, D. A.: Interactive effects of iron, irradiance and CO_2 on Ross Sea phytoplankton,
837 Deep-Sea Research Part I, 57, 368–383, doi:10.1016/j.dsr.2009.10.013, 2010.

838 Garcia, N. S., Sedwick, P. N. and DiTullio., G.R.: Influence of irradiance and iron on the growth
839 of colonial *Phaeocystis antarctica*: implications for seasonal bloom dynamics in the Ross Sea,
840 Antarctica, Aquatic Microbial Ecology, 57, 203–220, 2009.

841 Hallmann, A.: Extracellular matrix and sex-inducing pheromone in *Volvox*, in International
842 Review of Cytology, vol. 227, pp. 131–182, International Review of Cytology. 2003.

843 Hamm, C. E.: Architecture, ecology and biogeochemistry of *Phaeocystis* colonies, Journal of Sea
844 Research, 43, 307–315, 2000.

845 Hamm, C. E., Simson, D. A., Merkel, R. and Smetacek, V.: Colonies of *Phaeocystis globosa* are
846 protected by a thin but tough skin, Marine Ecology Progress Series, 187, 101–111, 1999.

847 Hayward, D. C., Hetherington, S., Behm, C. A., Grasso, L. C., Forêt, S., Miller, D. J. and Ball, E.
848 E.: Differential gene expression at coral settlement and metamorphosis - A subtractive
849 hybridization study, PLoS ONE, 6(10), e26411, doi:10.1371/journal.pone.0026411, 2011.

850 Jacobsen, A., Larsen, A., Martínez-Martínez, J., Verity, P. G. and Frischer, M. E.: Susceptibility
851 of colonies and colonial cells of *Phaeocystis pouchetii* (Haptophyta) to viral infection, Aquatic
852 Microbial Ecology, 48, 105–112, 2007.

853 King, N., Hittinger, C. T. and Carroll, S. B.: Evolution of key cell signaling and adhesion protein
854 families predates animal origins, Science, 301(5631), 361–363, doi:10.1126/science.1083853,
855 2003.

856 Kröger, N., Bergsdorf, C. and Sumper, M.: A new calcium binding glycoprotein family
857 constitutes a major diatom cell wall component, The EMBO Journal, 13(19), 4676–4683, 1994.

858 Lagerheim, G.: Ueber Phaeocystis poucheti (Har.) Lagerh., eine Plankton-Flagellate, Oeivers af
859 Vet Akad Foerhandl, 4, 277–288, 1896.

860 Lange, M., Chen, Y.-Q. and Medlin, L. K.: Molecular genetic delineation of *Phaeocystis* species
861 (Prymnesiophyceae) using coding and non-coding regions of nuclear and plastid genomes,
862 European Journal of Phycology, 37(1), 77–92, doi:10.1017/S0967026201003481, 2002.

863 Lê, S., Josse, J. and Husson, F.: FactoMineR: an R package for multivariate analysis, Journal of
864 Statistical Software, 25(1), 2008.

865 Long, J. D., Smalley, G. W., Barsby, T., Anderson, J. T. and Hay, M. E.: Chemical cues induce
866 consumer-specific defenses in a bloom-forming marine phytoplankton, Proceedings of the
867 National Academy of Sciences, 104(25), 10512–10517 [online] Available from:
868 http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=4224216, 2007.

869 Lovenduski, N. S., Gruber, N. and Doney, S. C.: Toward a mechanistic understanding of the
870 decadal trends in the Southern Ocean carbon sink, Global Biogeochemical Cycles, 22(GB3016),
871 doi:10.1029/2007GB003139, 2008.

872 Lubbers, G., Gieskes, W., Del Castilho, P., Salomons, W., and Bril, J.: Manganese accumulation
873 in the high pH microenvironment of *Phaeocystis sp.*(Haptophyceae) colonies from the North
874 Sea, Marine Ecology Progress Series, 1990. 285-293, 1990.

875 Martin, J. H., Fitzwater, S. E. and Gordon, R. M.: Iron deficiency limits phytoplankton growth in

876 Antarctic waters, Global Biogeochemical Cycles, 4(1), 5–12, doi:10.1029/GB004i001p00005,
877 1990.

878 Matrai, P. A., Vernet, M., Hood, R., Jennings, A., Brody, E. and Saemundsdottir, S.: Light-
879 dependence of carbon and sulfur production by polar clones of the genus *Phaeocystis*, Marine
880 Biology, 124, 157–167, 1995.

881 Maucher, J. M. and G. R. DiTullio (2003). Flavodoxin as a Diagnostic Indicator of Chronic Fe-
882 Limitation in the Ross Sea and New Zealand Sector of the Southern Ocean. Biogeochemistry in
883 the Ross Sea. G. R. DiTullio and R. B. Dunbar. Washington DC, AGU**:** 209-220.

884 Michel, G., Tonon, T., Scornet, D., Cock, J. M. and Kloareg, B.: The cell wall polysaccharide
885 metabolism of the brown alga *Ectocarpus siliculosus*. Insights into the evolution of extracellular
886 matrix polysaccharides in Eukaryotes, New Phytologist, 188(1), 82–97, 2010.

887 Morris, R. M., Nunn, B. L., Frazar, C., Goodlett, D. R., Ting, Y. S. and Rocap, G.: Comparative
888 metaproteomics reveals ocean-scale shifts in microbial nutrient utilization and energy
889 transduction, The ISME Journal, 4(5), 673–685, doi:10.1038/ismej.2010.4, 2010.

890 Morrissey, J., Sutak, R., Paz-Yepes, J., Tanaka, A., Moustafa, A., Veluchamy, A., Thomas, Y.,
891 Botebol, H., Bouget, F.-Y., McQuaid, J. B., Tirichine, L., Allen, A. E., Lesuisse, E. and Bowler,
892 C.: A novel protein, ubiquitous in marine phytoplankton, concentrates iron at the cell surface and
893 facilitates uptake, Current Biology, 25, 364–371, doi:10.1016/j.cub.2014.12.004, 2015.

894 Murray, A. E. and Grzymski, J. J.: Diversity and genomics of Antarctic marine micro-organisms,
895 Philosophical Transactions of the Royal Society B: Biological Sciences, 362(1488), 2259–2271,
896 doi:10.1098/rstb.2006.1944, 2007.

897 Noble, A. E., Moran, D. M., Allen, A. E. and Saito, M. A.: Dissolved and particulate trace metal
898 micronutrients under the McMurdo Sound seasonal sea ice: basal sea ice communities as a
899 capacitor for iron, Frontiers in Chemistry, 1(25), 1–18, doi:10.3389/fchem.2013.00025, 2013.

900 Peers, G. and Price, N. M.: Copper-containing plastocyanin used for electron transport by an
901 oceanic diatom, Nature, 441(7091), 341–344, doi:10.1038/nature04630, 2006.

902 Podell, S. and Gaasterland, T.: DarkHorse: a method for genome-wide prediction of horizontal
903 gene transfer, Genome Biology, 8, R16, doi:10.1186/gb-2007-8-2-r16, 2007.

904 Ram, R. J., VerBerkmoes, N. C., Thelen, M. P., Tyson, G. W., Baker, B. J., Blake, R. C., Shah,
905 M., Hettich, R. and Banfield, J.: Community proteomics of a natural microbial biofilm, Science,
906 308(5730), 1915–1920, doi:10.1126/science, 2005.

907 Rho, M., Tang, H. and Ye, Y.: FragGeneScan: predicting genes in short and error-prone reads,
908 Nucleic Acids Research, 38(20), e191–e191, doi:10.1093/nar/gkq747, 2010.

909 Riegman, R. and van Boekel, W.: The ecophysiology of *Phaeocystis globosa*: a review, Journal
910 of Sea Research, 35(4), 235–242, 1996.

911 Riegman, R., Noordeloos, A. A. M. and Cadee, G. C.: *Phaeocystis* blooms and eutrophication of
912 the continental coastal zones of the North Sea, Marine Biology, 112, 479–484,
913 doi:10.1007/BF00356293, 1992.

914 Roche, J. L., Boyd, P. W., McKay, R. M. L. and Geider, R. J.: Flavodoxin as an *in situ* marker
915 for iron stress in phytoplankton, Nature, 382, 802–805, doi:10.1038/382802a0, 1996.

916 Rousseau, V., Chrétiennot-Dinet, M.-J., Jacobsen, A., Verity, P. G. and Whipple, S.: The life
917 cycle of *Phaeocystis*: state of knowledge and presumptive role in ecology, Biogeochemistry, 83,
918 29–47, doi:10.1007/s10533-007-9085-3, 2007.

919 Rousseau, V., Mathot, S. and Lancelot, C.: Calculating carbon biomass of *Phaeocystis sp.* from
920 microscopic observations, Marine Biology, 107, 305–314, 1990.

921 Saito, M. A., Dorsk, A., Post, A. F., McIlvin, M. R., Rappé, M. S., R, D. G. and Moran, D. M.:
922 Needles in the blue sea: Sub-species specificity in targeted protein biomarker analyses within the
923 vast oceanic microbial metaproteome, Proteomics, 15(20), 3521–3531,
924 doi:10.1002/pmic.201400630, 2015.

925 Saito, M. A., Goepfert, T. J., Noble, A. E., Bertrand, E. M., Sedwick, P. N. and DiTullio, G. R.:
926 A seasonal study of dissolved cobalt in the Ross Sea, Antarctica: micronutrient behavior,
927 absence of scavenging, and relationships with Zn, Cd, and P, Biogeosciences, 7, 4059–4082,
928 doi:10.5194/bg-7-4059-2010, 2010.

929 Saito, M. A., McIlvin, M. R., Moran, D. M., Goepfert, T. J., R, D. G., Post, A. F. and Lamborg,
930 C. H.: Multiple nutrient stresses at intersecting Pacific Ocean biomes detected by protein
931 biomarkers, Science, 345(6201), 1173–1177, 2014.

932 Sarmiento, J. L., Hughes, T. M. C., Stouffer, R. J. and Manabe, S.: Simulated response of the
933 ocean carbon cycle to anthropogenic climate warming, Nature, 393, 245–249,
934 doi:10.1038/30455, 1998.

935 Schoemann, V., Becquevort, S., Stefels, J., Rousseau, V. and Lancelot, C.: *Phaeocystis* blooms
936 in the global ocean and their controlling mechanisms: a review, Journal of Sea Research, 53, 43–
937 66, 2005.

938 Schoemann, V., Wollast, R., Chou, L. and Lancelot, C.: Effects of photosynthesis on the
939 accumulation of Mn and Fe by *Phaeocystis* colonies, Limnology and Oceanography, 46(5),
940 1065–1076, 2001.

941 Sedwick, P. N. and DiTullio, G. R.: Regulation of algal blooms in Antarctic shelf waters by the
942 release of iron from melting sea ice, Geophys. Res. Lett., 24(20), 2515–2518,
943 doi:10.1029/97GL02596, 1997.

944 Sedwick, P. N., DiTullio, G. R. and Mackey, D. J.: Iron and manganese in the Ross Sea,
945 Antarctica: Seasonal iron limitation in Antarctic shelf waters, Journal of Geophysical Research,
946 105(C5), 11321–11336, doi:10.1029/2000JC000256, 2000.

947 Sedwick, P. N., Garcia, N. S., Riseman, S. F., Marsay, C. M. and DiTullio, G. R.: Evidence for
948 high iron requirements of colonial *Phaeocystis antarctica* at low irradiance, Biogeochemistry,
949 83(1-3), 83–97, doi:10.1007/s10533-007-9081-7, 2007.

950 Sedwick, P. N., Marsay, C. M., Sohst, B. M., Aguilar Islas, A. M., Lohan, M. C., Long, M. C.,
951 Arrigo, K. R., Dunbar, R. B., Saito, M. A., Smith, W. O. and DiTullio, G. R.: Early season
952 depletion of dissolved iron in the Ross Sea polynya: Implications for iron dynamics on the
953 Antarctic continental shelf, Journal of Geophysical Research, 116(C12019), 2011.

954 Smith, W. O., Jr, Codispoti, L. A., Nelson, D. M., Manley, T., Buskey, E. J., Niebauer, H. J. and
955 Cota, G. F.: Importance of *Phaeocystis* blooms in the high-latitude ocean carbon cycle, Nature,
956 352, 514–516, 1991.

957 Smith, W. O., Jr, Dennett, M. R., Mathot, S. and Caron, D. A.: The temporal dynamics of the
958 flagellated and colonial stages of *Phaeocystis antarctica* in the Ross Sea, Deep-Sea Research
959 Part II, 50, 605–617, 2003.

960 Smith, W. O., Tozzi, S., Long, M. C., Sedwick, P. N., Peloquin, J. A., Dunbar, R. B., Hutchins,
961 D. A., Kolber, Z. and R, D. G.: Spatial and temporal variations in variable fluorescence in the
962 Ross Sea (Antarctica): Oceanographic correlates and bloom dynamics, Deep Sea Research I, 79,
963 141–155, 2013.

964 Solomon, C. M., Lessard, E. J., Keil, R. G. and Foy, M. S.: Characterization of extracellular
965 polymers of *Phaeocystis globosa* and *P. antarctica*, Marine Ecology Progress Series, 250, 81–
966 89, 2003.

967 Sowell, S. M., Wilhelm, L. J., Norbeck, A. D., Lipton, M. S., Nicora, C. D., Barofsky, D. F., H,
968 C., Smith, R. D. and Giovanonni, S. J.: Transport functions dominate the SAR11 metaproteome
969 at low-nutrient extremes in the Sargasso Sea, The ISME Journal, 3(1), 93–105,
970 doi:10.1038/ismej.2008.83, 2008.

971 Stingl, U., Desiderio, R. A., Cho, J. C., Vergin, K. L. and Giovannoni, S. J.: The SAR92 Clade:
972 an Abundant Coastal Clade of Culturable Marine Bacteria Possessing Proteorhodopsin, Applied
973 and Environmental Microbiology, 73(7), 2290–2296, doi:10.1128/AEM.02559-06, 2007.

974 Strzepek, R. F., Maldonado, M. T., Hunter, K. A., Frew, R. D. and Boyd, P. W.: Adaptive
975 strategies by Southern Ocean phytoplankton to lessen iron limitation: Uptake of organically
976 complexed iron and reduced cellular iron requirements, Limnology and Oceanography, 56(6),
977 1983–2002, doi:10.4319/lo.2011.56.6.1983, 2011.

978 Sunda, W. and Huntsman, S.: Effect of pH, light, and temperature on Fe–EDTA chelation and Fe
979 hydrolysis in seawater, Marine Chemistry, 84(1-2), 35–47, doi:10.1016/S0304-4203(03)00101-4,
980 2003.

981 Sunda, W. G. and Huntsman, S. A.: Iron uptake and growth limitation in oceanic and coastal
982 phytoplankton, Marine Chemistry, 50, 189–206, 1995.

983 Thingstad, F. and Billen, G.: Microbial degradation of *Phaeocystis* material in the water column,

984 Journal of Marine Systems, 5, 55–65, doi:10.1016/0924-7963(94)90016-7, 1994.

985 Tzarfati-Majar, V., Burstyn-Cohen, T. and Klar, A.: F-spondin is a contact-repellent molecule
986 for embryonic motor neurons, Proceedings of the National Academy of Sciences, 98(8), 4722–
987 4727, 2011.

988 van Boekel, W.: *Phaeocystis* colony mucus components and the importance of calcium ions for
989 colony stability, Marine Ecology Progress Series, 87, 301–305, 1992.

990 Vardi, A.: Cell signaling in marine diatoms, Communicative & Integrative Biology, 1(2), 134–
991 136, doi:10.1016/j.cub.2008.05.037, 2008.

992 Verity, P. G., Brussaard, C. P., Nejstgaard, J. C., van Leeuwe, M. A., Lancelot, C. and Medlin,
993 L. K.: Current understanding of *Phaeocystis* ecology and biogeochemistry, and perspectives for
994 future research, Biogeochemistry, 83, 311–330, doi:10.1007/s10533-007-9090-6, 2007.

995 Warnes, G. R., Bolker, B., Bonebakker, L., Gentleman, R., Huber, W., Liaw, A., Lumley, T.,
996 Maechler, M., Magnusson, A., Moeller, S. and Schwartz, M.: gplots: Various R programming
997 tools for plotting data. 2009.

998 Watanabe, Y., Hayashi, M., Yagi, T. and Kamiya, R.: Turnover of actin in *Chlamydomonas*
999 flagella detected by fluorescence recovery after photobleaching (frap), Cell, 29, 67–72,
1000 doi:10.1247/csf.29.67, 2004.

1001 Whitney, L. P., Lins, J. J., Hughes, M. P., Wells, M. L., Chappell, P. D. and Jenkins, B. D.:
1002 Characterization of putative iron responsive genes as species-specific indicators of iron stress in
1003 *Thalassiosiroid* diatoms, Frontiers in Microbiology, 2, 1–14,
1004 doi:10.3389/fmicb.2011.00234/abstract, 2011.

1005 Williams, T. J., Long, E., Evans, F., DeMaere, M. Z., Lauro, F. M., Raftery, M. J., Ducklow, H.,
1006 Grzymski, J. J., Murray, A. E. and Cavicchioli, R.: A metaproteomic assessment of winter and
1007 summer bacterioplankton from Antarctic Peninsula coastal surface waters, The ISME Journal,
1008 6(10), 1883–1900, doi:10.1038/ismej.2012.28, 2012.

1009 Wu, Z., Jenkins, B. D., Rynearson, T. A., Dyhrman, S. T., Saito, M. A., Mercier, M. and
1010 Whitney, L. P.: Empirical bayes analysis of sequencing-based transcriptional profiling without
1011 replicates, BMC Bioinformatics, 11, 564, doi:10.1186/1471-2105-11-564, 2010.

1012 Zilliges, Y., Kehr, J. C., Mikkat, S., Bouchier, C., de Marsac, N. T., Borner, T. and Dittmann, E.:
1013 An extracellular glycoprotein is implicated in cell-cell contacts in the toxic cyanobacterium
1014 *Microcystis aeruginosa* PCC 7806, Journal of Bacteriology, 190(8), 2871–2879,
1015 doi:10.1128/JB.01867-07, 2008.

1016 Zingone, A., Chrétiennot-Dinet, M.-J., Lange, M. and Medlin, L.: Morphological and genetic
1017 characterization of *Phaeocystis cordata* and *P. jahnii* (prymnesiophyceae), two new species from
1018 the Mediterranean Sea, Journal of Phycology, 35(6), 1322–1337, doi:10.1046/j.1529-
1019 8817.1999.3561322.x, 1999.

1020 Zurbriggen, M. D., Tognetti, V. B., Fillat, M. F., Hajirezaei, M.-R., Valle, E. M. and Carrillo, N.:
1021 Combating stress with flavodoxin: a promising route for crop improvement, Trends in
1022 Biotechnology, 26(10), 531–537, 2008.

1023